JMIR Infodemiology

Journal Impact Factor (JIF) (2023): 3.5
Volume 5 (2025) ISSN 2564-1891 Editor-in-Chief: Tim Ken Mackey, MAS, PhD

Contents

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

Viewpoint

XSL•FO
RenderX

The Role of Digital Health Equity Audits in Preventing Harmful Infodemiology

Massimiliano Biondi[1], MSc, MD; Fabio Filippetti[2]; Giorgio Brandi[3], PhD; Elsa Ravaglia[4]; Sofia Filippetti[5]; Pamela Barbadoro[6], PhD

[1]Medical Directorate, Fabriano Hospital Site, World Federation of Public Health Associations GHEDT WG, Ancona Health Authority AST AN, Via Stelluti Scala, 26, Fabriano, Italy

[2]Prevention and Health Promotion Unit in Living and Working Places of the Marche Region, Ancona, Italy

[3]Unit of Hygiene, Department of Biomedical Sciences, University of Urbino Carlo Bo, Urbino, Italy

[4]Pesaro-Urbino Health Authority AST PU, Pesaro, Italy

[5]Department of Public Health and Pediatrics, University of Turin, Turin, Italy

[6]Unit of Hygiene, Preventive Medicine and Public Health, Department of Biomedical Sciences and Public Health, Faculty of Medicine, Marche Polytechnic University, Ancona, Italy

Corresponding Author:
Massimiliano Biondi, MSc, MD
Medical Directorate, Fabriano Hospital Site, World Federation of Public Health Associations GHEDT WG, Ancona Health Authority AST AN, Via Stelluti Scala, 26, Fabriano, Italy

Abstract

Background: Health disparities persist and are influenced by digital transformation. Although digital tools offer opportunities, they can also exacerbate existing inequalities, a problem amplified by the COVID-19 pandemic and the related infodemic. Health equity audit (HEA) tools, such as those developed in the United Kingdom, provide a framework to assess equity but require adaptation for the digital context. Digital determinants of health (DDoH) are increasingly recognized as crucial factors influencing health outcomes in the digital era.

Objective: This editorial proposes an approach to extend HEA principles to create a specific framework, the digital health equity audit (DHEA), designed to systematically assess and address health inequities within the design, implementation, and evaluation of digital health technologies, with a focus on DDoH.

Methods: We propose a cyclical DHEA model based on existing HEA principles, integrating them with digital health equity frameworks. The DHEA cycle comprises six phases: (1) scoping the audit and mobilizing the team (including community members); (2) developing the digital health equity profile and identifying inequities (assessing DDoH at individual, interpersonal, community, and societal levels); (3) identifying high-impact actions to address DDoH and inequities; (4) prioritizing actions for maximum equity impact; (5) implementing and supporting change; and (6) evaluating progress and impact, and refining. This method emphasizes multilevel interventions and stakeholder engagement.

Results: The main result is the articulation of the DHEA framework: a structured, 6-phase cyclical model to guide organizations in the analysis and proactive mitigation of digital health–related disparities. The framework explicitly integrates the assessment of DDoH across multiple levels (individual, interpersonal, community, societal) and promotes the development of targeted interventions to ensure digital solutions promote equity.

Conclusions: The DHEA model offers an integrated approach to consider social, epidemiological, health, and technological variables, aiming to reduce health inequities through the conscious use of new technologies. It is emphasized that digital technologies can be the cause or the solution to inequalities; DHEAs are proposed as a tool to foster equity. Its systematic adoption, along with a collaborative approach (co-design) and trust building, can help ensure that the benefits of health digitization are equitably distributed while strengthening trust in institutions. Continued attention is needed to manage emerging challenges such as infodemiology in the era of big data and artificial intelligence.

(JMIR Infodemiology 2025;5:e75495) doi:10.2196/75495

KEYWORDS

equity; digital; audit; infodemiology; quality of health care

XSL•FO
RenderX

Introduction

The concept of a health equity audit (HEA), as part of guidance issued by Public Health England (PHE) in the context of their Health Equity Assessment Tool, updated in May 2021, helps to articulate a clear framework to address health inequalities and has potential to be extended to the broader field of digital health [1]. Specifically, the PHE guidance defined how the objective of HEAs is to evaluate whether resources are distributed equitably with respect to the health needs of different population groups; they systematically examine health inequalities and access to services for particular groups or areas. Audits also ensure that actions to address health inequalities are incorporated into planning decisions, prioritizing actions to address health inequalities, and addressing how they can evaluate the impact of the actions on reducing inequalities.

These tools, despite having been widely used in the United Kingdom since the 2000s and subsequently neglected due to organizational changes in the British health care system, are currently recommended by PHE, which following the inequalities evidenced during the COVID-19 pandemic, renewed interest in adopting the tool for health care purposes; the tool could also be adapted for various applications of digital health technologies [2,3].

As the field of health disparities has matured to reach a crucial element of health care management and quality, we have simultaneously witnessed the effects of digital transformation on health status as well as health care [4-7]. This necessitates the need for a reexamination of the utility of HEAs in the design, implementation, and evaluation of digital health technologies. These inequalities have amplified during the COVID-19 pandemic, especially due to the misinformation caused by the infodemic, which pushed the World Health Organization to call a conference on the topic. Importantly, the convergence of factors such as volume and speed of information, misinformation, and disinformation flow, combined with political polarization, requires the forging of a community for the evidence-based practice of infodemic management [8].

Different Definitions of Digital Health

According to the National Institutes of Health, digital health refers to the use of information and communication technologies in medicine and other health professions to manage diseases and health risks and promote well-being [9]. For the European Union, digital health and care refers to tools and services that use information and communication technologies to improve the prevention, diagnosis, treatment, monitoring, and management of health-related problems, and to monitor and manage lifestyle habits that affect health. Digital health and care facilitates the use of emerging and innovative technologies, and has the potential to improve access and the quality of care, as well as increase the overall efficiency of the health care sector [10].

The World Health Organization [11] has three key objectives to promote the adoption and expansion of digital health and innovation:

1. Promote data sharing and support the implementation of digital solutions that contribute to informed decision-making
2. Improve knowledge through the best scientific communities
3. Assess and connect countries' needs with the supply of innovations

Access to digital technologies in health, including the internet, technological tools, digital agendas and systems, digital literacy, etc, has also become an increasingly important determinant of health and has a special relationship with social determinants of health. Emerging evidence from the scientific literature recognizes that access to digital technologies is now a determinant of health outcomes [12,13]. As digital determinants of health (DDoH) become increasingly recognized [14,15], a framework for digital health equity audits (DHEAs), including the evaluation of key DDoH, is needed.

Opportunities to Extend HEA Tools

In this editorial, we propose an approach to extend HEA tools to address shared international objectives of synergistically promoting both health equity and digital health adoption and access, framed as a DHEA. This strategy is rooted in the World Health Organization's objectives of equity and digitalization as described above and involves the synthesis of a tool that combines HEA concepts with stated goals of digital health equity, as published in other academic literature, and modeled based on proposals by the Agency for Healthcare Research and Quality in a comprehensive framework [16]. We support the implementation of a similar approach also focusing on improving the health status of the population in relation to the use of technologies and the context of health technology assessments. However, we must consider that new digital "approaches have the potential to address some of the structural challenges for marginalized populations.... Yet the digitalization of health care can also harm health equity if this digitally enabled ecosystem moves forward without proactive engagement, planning, and implementation" [5]. As discussed in the recent literature, the evidence linking inequality in health care to misinformation exposure and mitigating strategies is a complex area where further research is needed [17-19].

The DHEA Cycle

Overview

Building upon standard HEA principles, the DHEA cycle integrates the Framework for Digital Health Equity to specifically address how digital technologies impact health disparities. It emphasizes understanding and acting on DDoH across multiple levels (individual, interpersonal, community, societal) to ensure digital health solutions promote equity rather than widen gaps.

The DHEA cycle phases are shown in Figure 1.

XSL•FO
RenderX

Figure 1. Digital health equity audit (DHEA).

Scope the Audit and Mobilize the Team (Phase 1)

Action

A diverse working group is formed including community members (especially from disparity populations), clinicians, IT specialists, designers, policy makers, and public health professionals.

Integration

The specific digital health tool, service, or system being audited (eg, a new patient portal feature, a telehealth service) is defined. Priority populations and potential equity concerns are agreed on, explicitly considering the Framework for Digital Health Equity [15] and its emphasis on populations adversely affected by health differences (racial/ethnic minorities, those with low-income, those who live in rural areas, sexual and gender minorities, or individuals with disabilities).

Example

When auditing a new telehealth platform for primary care, the team can include patient representatives from low-income neighborhoods, accessibility experts, primary care physicians, and IT developers, prioritizing equitable access and usability for seniors and nonnative speakers.

Develop the Digital Health Equity Profile and Identify Inequities (Phase 2)

Action

Data is gathered to create a profile of the target population's interaction with the specific digital health tool/service and the broader digital environment, using surveys, interviews, use data, population health data, and community assessments.

Integration

This profile must assess relevant DDoH across the four levels:

XSL•FO
RenderX

- Individual: digital literacy, technology access (device/internet), digital self-efficacy, attitudes/trust toward technology
- Interpersonal: implicit technology bias from providers, patient-provider communication via digital tools, family/caregiver technology interdependence
- Community: broadband availability/affordability, local technology support resources (libraries, community centers), health care system's digital infrastructure quality, relevant community technology norms
- Societal: technology policies (reimbursement, privacy), data/design standards (accessibility, language), algorithmic bias, social norms around technology

Action Continued

The profile is analyzed to pinpoint specific inequities—where are there avoidable unfair differences in digital access, use, experience, or outcomes between population groups?

Example

The profile for the telehealth platform reveals that seniors have lower adoption rates. An analysis identifies key DDoH barriers: lower digital literacy and lack of affordable broadband (individual/community), coupled with clinician assumptions about seniors' ability/interest (interpersonal implicit technology bias). An inequity is identified: seniors face avoidable barriers to accessing telehealth compared to younger higher-income groups.

Identify High-Impact Actions to Address DDoH and Inequities (Phase 3)

Action

Potential interventions are suggested to address the specific DDoH barriers and inequities identified in phase 2, reviewing the evidence for effective strategies.

Integration

The focus is on developing multilevel interventions that target "upstream" determinants (community and societal levels) where possible, as these often have a broader and more sustainable impact on equity, as highlighted by the summarized text. Actions addressing individual skills, interpersonal interactions, community resources, and systemic policies/design should be considered.

Example

The following actions can be taken to address the telehealth inequity:

- Individual: offer digital literacy training tailored for seniors; provide loaner tablets
- Interpersonal: train clinicians on identifying and mitigating implicit technology bias
- Community: partner with local libraries or senior centers for technology support hubs; advocate for expanded community broadband initiatives
- Societal: Advocate for policies ensuring telehealth platforms meet high accessibility standards (Web Content Accessibility Guidelines)

Prioritize Actions for Maximum Equity Impact (Phase 4)

Action

The potential actions are evaluated based on criteria such as potential impact on reducing the identified inequity, feasibility, cost, community acceptability, and alignment with organizational goals.

Integration

Actions most likely to address root causes (upstream DDoH) and benefit populations disproportionately are prioritized, ensuring the prioritization process involves stakeholders, especially from affected communities.

Example

Partnering with senior centers for training/support (community/individual: high impact, feasible) and advocating for better broadband (community/societal: high upstream impact, longer term) can be prioritized over simply providing tablets without support (individual: less sustainable).

Implement and Support Change (Phase 5)

Action

An implementation plan is developed, allocating necessary resources (funding, staffing, partnerships) and executing the prioritized actions.

Integration

Resources need to specifically address the DDoH barriers (eg, funding for digital navigators, accessible design implementation, community infrastructure partnerships). Synergy needs to be fostered between the implementation team, decision makers, and the target community through ongoing communication and feedback loops, adapting based on initial rollout experiences.

Example

The following actions can be taken during this phase: secure funding for trainers at senior centers, deploy accessible platform updates, launch clinician training modules, or establish a feedback channel with senior users.

Evaluate Progress and Impact, and Refine (Phase 6)

Action

The implementation process is monitored, and the impact of the actions is evaluated against the initial objectives and the identified inequities.

Integration

Specific, measurable indicators are defined that track changes in DDoH (eg, digital literacy scores, broadband access rates, device ownership) and health equity outcomes related to the digital tool (eg, telehealth use rates stratified by age/income/race, patient satisfaction scores by demographic group, changes in relevant health metrics for disparity groups). Evaluation findings need to be used to refine actions, inform future DHEA cycles, and demonstrate accountability.

XSL•FO
RenderX

Example

Telehealth appointment completion rates can be tracked to analyze seniors versus other groups, pre- and postscores for the digital literacy assessment for participants, qualitative feedback on usability, and number of broadband sign-ups through advocacy efforts. If senior use remains low, phases 2 and 3 can be revisited to identify potentially missed DDoH barriers.

Conclusions

The DHEA model is an integrated model that takes into account social, epidemiological, health, and technological variables. The integration of knowledge and resources, together with the involvement at the institutional and the population levels, should produce a health gain for the majority of the population, reducing health inequities thanks to new technologies and strengthening trust in government institutions and health care. The systematic adoption of integrated and digitized tools for reading the system could certainly contribute to the evaluation of the effectiveness of interventions [20]. However, the digitization of health services in 2025 is increasingly important and its massive implementation is expected in the following years. The large amount of data that we will have to manage—fueled by the continuous flow of information—is converging and will converge with artificial intelligence all over the world, leading to a rapid and proportionally difficult-to-control diffusion of infodemiology. The increased convenience, accessibility, and penetration of internet services have significantly transformed how people obtain information on health-related issues. The rapid proliferation of information and communication technology tools has led to an era of unprecedented accessibility to vast repositories of information, especially through online communication channels and social media platforms [21]. In summary, we could therefore affirm that new technologies can be the cause of inequalities or the solution to health inequalities. We would like the DHEA tool to help increase equity at all levels so that one day everyone can benefit from the advantages of technologies and, through them, be in control of their health and well-being. Facilitators include building trust (eg, providing evidence for health messages), while barriers include user reluctance to accept support. The main recommendations are adopting a collaborative working approach (involving users, developers, health care professionals, policy makers, etc, often through co-design) and using effective advertising to raise awareness of available support.

Being aware of the great advantages of the widespread, adequate, and fair use of continuous technological and digital innovations available to science, we must never forget that they are the tool and not the goal.

Conflicts of Interest

None declared.

References

1. Health Equity Assessment Tool (HEAT): what it is and how to use it. GOV.UK. URL: https://www.gov.uk/government/publications/health-equity-assessment-tool-heat/health-equity-assessment-tool-heat-executive-summary [accessed 2025-05-12]
2. Barr B, Higgerson J, Whitehead M. Investigating the impact of the English health inequalities strategy: time trend analysis. BMJ 2017 Jul 26;358:j3310. [doi: 10.1136/bmj.j3310] [Medline: 28747304]
3. van Daalen KR, Davey F, Norman C, Ford JA. Health equity audits: a systematic review of the effectiveness. BMJ Open 2021 Nov 11;11(11):e053392. [doi: 10.1136/bmjopen-2021-053392] [Medline: 34764176]
4. Crawford A, Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. J Med Internet Res 2020 Jun 2;22(6):e19361. [doi: 10.2196/19361] [Medline: 32452816]
5. Lyles CR, Wachter RM, Sarkar U. Focusing on digital health equity. JAMA 2021 Nov 9;326(18):1795-1796. [doi: 10.1001/jama.2021.18459] [Medline: 34677577]
6. Ramsetty A, Adams C. Impact of the digital divide in the age of COVID-19. J Am Med Inform Assoc 2020 Jul 1;27(7):1147-1148. [doi: 10.1093/jamia/ocaa078] [Medline: 32343813]
7. Sieck CJ, Sheon A, Ancker JS, Castek J, Callahan B, Siefer A. Digital inclusion as a social determinant of health. NPJ Digit Med 2021 Mar 17;4(1):52. [doi: 10.1038/s41746-021-00413-8] [Medline: 33731887]
8. Mackey T, Baur C, Eysenbach G. Advancing infodemiology in a digital intensive era. JMIR Infodemiology 2022 Feb 14;2(1):e37115. [doi: 10.2196/37115] [Medline: 37113802]
9. Digital health. National Library of Medicine. 2024. URL: https://www.ncbi.nlm.nih.gov/mesh/?term=digital%20health [accessed 2025-05-12]
10. Digital health and care. European Commission. 2025. URL: https://health.ec.europa.eu/ehealth-digital-health-and-care/overview_en [accessed 2025-05-12]
11. Digital health. World Health Organization. 2025. URL: https://www.who.int/health-topics/digital-health [accessed 2025-05-12]
12. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. Health Policy 2024 Aug;146:105122. [doi: 10.1016/j.healthpol.2024.105122] [Medline: 38986333]
13. Fraser HS, Marcelo A, Kalla M, Kalua K, Celi LA, Ziegler J. Digital determinants of health: editorial. PLOS Digit Health 2023 Nov 28;2(11):e0000373. [doi: 10.1371/journal.pdig.0000373] [Medline: 38016101]
14. Chidambaram S, Jain B, Jain U, et al. An introduction to digital determinants of health. PLOS Digit Health 2024 Jan 4;3(1):e0000346. [doi: 10.1371/journal.pdig.0000346] [Medline: 38175828]

XSL•FO
RenderX

15. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. NPJ Digit Med 2022 Aug 18;5(1):119. [doi: 10.1038/s41746-022-00663-0] [Medline: 35982146]
16. Evidence- and consensus-based digital healthcare equity framework. Agency for Healthcare Research and Quality. 2024 Feb. URL: https://digital.ahrq.gov/sites/default/files/docs/citation/health-equity-framework.pdf [accessed 2025-05-16]
17. Heley K, Chou WYS, D'Angelo H, et al. Mitigating health and science misinformation: a scoping review of literature from 2017 to 2022. Health Commun 2025 Jan;40(1):79-89. [doi: 10.1080/10410236.2024.2332817] [Medline: 38534199]
18. Hendricks-Sturrup R, Simmons M, Anders S, et al. Developing ethics and equity principles, terms, and engagement tools to advance health equity and researcher diversity in AI and machine learning: modified Delphi approach. JMIR AI 2023 Dec 6;2:e52888. [doi: 10.2196/52888] [Medline: 38875540]
19. Southwell BG, Otero Machuca J, Cherry ST, Burnside M, Barrett NJ. Health misinformation exposure and health disparities: observations and opportunities. Annu Rev Public Health 2023 Apr 3;44:113-130. [doi: 10.1146/annurev-publhealth-071321-031118] [Medline: 36207008]
20. Grimshaw JM, Ivers N, Linklater S, et al. Reinvigorating stagnant science: implementation laboratories and a meta-laboratory to efficiently advance the science of audit and feedback. BMJ Qual Saf 2019 May;28(5):416-423. [doi: 10.1136/bmjqs-2018-008355] [Medline: 30852557]
21. Zhao C, Musonye HA, Wang P, Pan H. Infodemiology as a powerful tool in public health: bridging the digital divide for better well-being. Innov Discov 2024 Jun;1(2):19. [doi: 10.53964/id.2024019]

Abbreviations

DDoH: digital determinants of health
DHEA: digital health equity audit
HEA: health equity audit
PHE: Public Health England

Edited by A Mavragani; submitted 04.04.25; this is a non–peer-reviewed article;accepted 29.04.25; published 30.05.25.

Please cite as:
Biondi M, Filippetti F, Brandi G, Ravaglia E, Filippetti S, Barbadoro P
The Role of Digital Health Equity Audits in Preventing Harmful Infodemiology
JMIR Infodemiology 2025;5:e75495
URL: https://infodemiology.jmir.org/2025/1/e75495
doi:10.2196/75495

© Massimiliano Biondi, Fabio Filippetti, Giorgio Brandi, Elsa Ravaglia, Sofia Filippetti, Pamela Barbadoro. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 30.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of Misinformation on Social Media in the Context of Natural Disasters: Narrative Review

Sonya Hilberts[1], BSc, MAEd; Mark Govers[1], PhD; Elena Petelos[1,2], PhD; Silvia Evers[1], PhD

[1]Department Health Services Research, CAPHRI School at Faculty of Health, Medicine and Life Sciences, Maastricht University, Bonnefantenstraat 2, Maastricht, The Netherlands

[2]Clinic of Social and Family Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece

Corresponding Author:
Sonya Hilberts, BSc, MAEd
Department Health Services Research, CAPHRI School at Faculty of Health, Medicine and Life Sciences, Maastricht University, Bonnefantenstraat 2, Maastricht, The Netherlands

Abstract

Background: Misinformation on social media during natural disasters has become a significant challenge, with the potential to increase public confusion, panic, and distrust. Although individuals rely on social media platforms for timely updates during crises, these platforms also facilitate the rapid spread of unverified and misleading information. Consequently, misinformation can hamper emergency response efforts, misdirect resources, and distort public perception of the disaster's true severity.

Objective: This narrative review aims to (1) critically evaluate the available evidence; (2) unpack the dynamics of misinformation on social media in the context of natural disasters, specifically natural hazards, shedding light on the challenges, implications, and potential solutions; and (3) develop a conceptual model linking misinformation, public impact, and disasters, grounded in sourced evidence.

Methods: The narrative review examines the impact of social media misinformation in the context of natural disasters. The literature search was conducted using the PubMed database and Google Scholar in April 2024. Studies eligible for inclusion were published in English, with no restrictions on publication date, geographic region, or target population. The inclusion criteria focused on the original research that examined social media misinformation related to natural disasters, specifically natural hazards.

Results: From an initial pool of 173 studies, 9 studies met the inclusion criteria for this review. The selected studies revealed consistent patterns in how misinformation spreads during natural disasters, highlighting the role of users, some influencers, and bots in amplified false narratives. The misleading messages disseminated across social media platforms often outpaced official communications, resulting in reduced trust and exacerbating anxiety, stress, and fear among affected populations. This heightened emotional response and erosion of trust in official communications influenced an individual's susceptibility to the misinformation and prompted inappropriate actions. Consequently, such actions led to resource misallocation, overwhelmed emergency services, and diverted attention away from genuine needs. Collectively, these factors negatively impacted public health outcomes and diminished the effectiveness of emergency management efforts, as illustrated in the conceptual model developed to provide a greater understanding of this critical area of study.

Conclusions: This narrative review highlights the significant impact of misinformation in the context of natural disasters, specifically natural hazards. It stresses the urgent need for disaster preparedness and response plans that include targeted interventions such as real-time misinformation detection technologies, public education campaigns focused on digital literacy, and proactive debunking initiatives. Implementing these strategies can help mitigate the harmful effects of misinformation, strengthen public trust in official communications, enhance the effectiveness of disaster response, and improve public health outcomes.

(JMIR Infodemiology 2025;5:e70413) doi:10.2196/70413

KEYWORDS

misinformation; infodemic; social media; natural disaster; preparedness

Introduction

Social media platforms have increasingly become essential tools during natural disasters, enabling real-time communication, rapid information dissemination, and enhanced public awareness and safety [1]. However, these platforms also facilitate the rapid spread of misinformation, significantly complicating disaster management and public health responses [1-3]. The dissemination of misinformation during disasters can lead to

XSL•FO
RenderX

heightened public anxiety, confusion, resource misallocation, reduced effectiveness of emergency responses, and diminished trust in official communications [4-7]. For example, during Hurricane Sandy in 2012, false tweets about the New York Stock Exchange flooding and fabricated images of the storm spread widely, causing public panic and confusion [4]. Similarly, misinformation during Hurricanes Harvey and Irma in 2017 led to widespread fears about mandatory ID checks at shelters, discouraging undocumented immigrants from seeking safety [5]. Furthermore, the spread of misinformation can also delay emergency response efforts, as observed in the 2018 Kerala floods, where misleading social media posts hindered rescue operations [6].

A key challenge in addressing misinformation during natural disasters lies in the timing and immediacy of such events. They often occur with little warning, leaving minimal time for preparation. For example, while a hurricane may be predicted a few days in advance, its impact is typically sudden and severe. This rapid onset and escalation of natural disasters can amplify the spread of misinformation on social media platforms in unique ways. For this reason, examining misinformation specifically within the context of natural disasters offers valuable insights into the role of social media in crisis communication.

Additionally, the sheer volume of misinformation circulating online poses substantial challenges for government agencies, humanitarian organizations, and health systems, entities that are paradoxically central to disaster response. To counter these challenges, coordinated efforts between governments, social media platforms, and the public are essential to ensure the integrity of information during disasters. For example, automated misinformation detection systems implemented by platforms like Twitter (subsequently rebranded X) have been proven beneficial during hurricanes by flagging misleading content related to emergency shelters and procedures [5]. In parallel, public digital literacy initiatives have demonstrated potential in reducing misinformation by the public's ability to critically assess online content [8]. Nevertheless, research suggests that misinformation spreads more rapidly when official communications are unclear or delayed, highlighting the need for timely, transparent, and proactive crisis communication strategies [9].

There are several terms to describe inaccurate information, such as fake news, rumors, propaganda, infodemic, disinformation, and misinformation. This review will focus on misinformation, while adopting a broader lens to also encompass false information intentionally shared (disinformation), as it provides a more holistic account of the impact on public health and disaster response, whether intentional or not.

While numerous accounts document specific incidents of misinformation during events such as hurricanes, floods, and earthquakes [4-6,9], much of the literature focuses on isolated cases. This narrow focus limits our understanding of broader patterns and systematic impacts across different types of disasters and geographic regions. The absence of an integrated synthesis constrains our ability to comprehend how misinformation influences public perceptions, behavioral responses, and the overall effectiveness of emergency response [2,10,11].

Therefore, this narrative review aims to address this gap by synthesizing and consolidating available evidence on misinformation on social media during natural disasters, identifying key patterns, dynamics, and impacts. Additionally, a conceptual model will be developed as part of the review to clarify the relationships between misinformation origins, public impact, and disruptions in emergency responses, informed by sourced evidence [4-6,8,9,12-15]. By synthesizing available research, this review may support policy makers, emergency responders, and public health officials in designing targeted strategies to mitigate the spread of misinformation. Ultimately, strengthening the resilience of disaster response systems and safeguarding public health outcomes.

Methods

Study Design

The methodology for this narrative review is grounded in established principles from relevant literature to guide the review process effectively. These encompass a structured process and have been adapted from the guidelines recommended by the Joanna Briggs Institute [16] as follows.

- The topic and research question were defined, and the criteria that determine whether literature will be included or excluded were created.
- A search strategy was developed and executed, specifying keywords, subject headings, and Boolean operators.
- An initial screening of titles and abstracts from the search results was conducted to assess relevance.
- A detailed full-text review of articles was performed, based on predefined inclusion and exclusion criteria. The screening process results were documented using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for transparency.
- Relevant information was extracted from the included studies to support and inform the findings of the review using a standardized data extraction tool.

Also, the Cochrane Handbook for Systematic Reviews of Interventions was used to develop the protocol for this review (PROSPERO [International Prospective Register of Systematic Reviews] CRD42024542111).

The initial search was done by SH. The screening and the full text review were done by SH and EP. Data extraction and synthesis were done by all authors. Any disagreement was solved in consensus meetings, which led to an agreement. The research team consisted of members with different backgrounds relevant to this study.

SH's experience supporting disaster management, digital transformation programs, and national COVID-19 pandemic response initiatives has deeply influenced her approach toward organizational efficiency, resilience, and strategic foresight.

MG, an associate professor of the organization, management, and digital, applies sociotechnical systems thinking to align social and technical dimensions in health care and business,

XSL•FO
RenderX

advancing theory by linking STS principles to digital transformation challenges while also integrating broader cultural, behavioral, and political perspectives on social phenomena.

EP is a public-health specialist and senior lecturer (University of Crete, University of West Attica, and Maastricht University) whose 2-decade career spans R&D (Research and Development), evidence generation and synthesis, HTA, digital health, and artificial intelligence (AI). Since returning to academia, she combines academic research with policymaking and advocacy work, currently serving as president for Global Health and vice president for HTA (Health Technology Assessment) at The European Public Health Association (EUPHA) and chair for RWE & AI at HTAi.

SE, Professor of Public Health Technology Assessment and Scientific Director of the Care and Public Health Research Institute, specializes in health-economic review methods, scoping, narrative, systematic, and meta-analyses and created the widely adopted Consensus Health Economic Criteria list and accompanying guidelines for systematic reviews of economic evaluations.

All research team members are aware that experiences derived from their different roles and positions have shaped their own perspectives. However, the diversity in backgrounds helped to broadly reflect on the findings of this study.

Search Strategy

The search was conducted in April 2024. The primary source to retrieve information was the PubMed database, because of its primacy as a reputable source of public health research. The paucity of results from that source led to expanding the search, adding free searching on Google Scholar, for works published in the English language.

The decision to limit the selection to English-language papers was based on 3 factors. First, the research team possesses proficiency in English, which ensures accurate interpretation and analysis of the included studies. Second, the databases used for literature searches (eg, PubMed) primarily index journals in the English language. Finally, conducting a multilingual review would require significant resources, including translation and verification by native speakers of each language.

There was no date restriction imposed to ensure the capture of all relevant publications related to misinformation on social media in the context of a natural disaster, specifically related to natural hazards. Also, no restriction was imposed on the country or on the age of the target populations.

The search process began with the identification of relevant keywords or phrases that would deliver the desired review results.

A simple search was then done using free text terms which included a search on the topic, looking at words in subject headings, titles, abstracts, and authors' keywords, also scanning for synonyms, alternative spelling variants, acronyms, abbreviations, encompassing (1) "exploded" subject heading, include narrower subject headings found in the hierarchy as free text terms; (2) text mining (subject terms, index terms, descriptors, and MeSH [Medical Subject Headings]); (3) truncation was used where a search for a term that begins with a word was needed; and (4) once all free-text terms and controlled vocabulary terms had been identified, the next step was to use the correct Boolean operators to combine the terms using "or" OR "and."

The three categories of interest were (1) "misinformation" (also "infodemic"), (2) "social media," and (3) "natural disaster" (expanded to include 4 specific types: earthquake, fire, flooding, and tsunami). Combinations of these 3 categories of terms were applied to search titles, abstracts, subject headings, and author keywords. The search strategy also included synonyms, alternative spellings, acronyms, and abbreviations to ensure inclusive coverage of the topic. The review then used the snowballing technique to identify additional studies cited from the studies retrieved.

Peer-reviewed research reports published in journals and conference proceedings were eligible for inclusion. There was no restriction on the date published or the type of original research studies. This included research regarding natural field experiments, observational analysis, surveys, samples, and public information gathered through web scraping and social mining techniques. Systematic reviews meeting the eligibility criteria were examined for additional relevant references.

Inclusion Criteria

Studies were included if the authors only addressed misinformation on social media in the context of a specific instance of a natural disaster.

There was no restriction on the date published or the type of original research studies. This included research regarding natural field experiments, observational analysis, surveys, samples, and public information gathered through web scraping and social mining techniques.

Exclusion Criteria

The studies that did not investigate misinformation on social media in the context of a natural disaster or that were not in English, and duplicates of another paper, were also removed after screening. Also, studies that solely focused on how to design and build tools to detect misinformation were also excluded.

In addition, the exclusion criteria comprised editorials, letters to the editor, systematic reviews, abstracts, protocols, workshop summaries, perspectives, opinions, diagnosis methods, books, and book chapters, as well as summaries of other reviews, which were excluded before further screening. A documented record was kept of the search findings and translated into a PRISMA flow diagram [17] to provide transparent and complete reporting.

Data Extraction (Selection)

We used Microsoft Excel software to capture and synthesize the data. For the included studies, the elements in Tables S2 and S3 found in Multimedia Appendix 1 were analyzed in the full text. The characteristics of the included studies were assessed using a predefined criterion.

XSL•FO
RenderX

Study Characteristics

Each study was analyzed for the demographics, context, geographical location, study methods, objectives, study setting, data analysis techniques, the topic of interest, the period of time, data sources, data size, the collecting data method, misinformation effects, and lessons learned.

Evidence Synthesis

Evidence synthesis includes the type of social media platform, type of natural disaster, type of misinformation, public impact, and who is impacted.

The aim of the study is to (1) critically analyze the available evidence; (2) explore the dynamics of misinformation on social media in the context of natural disasters, shedding light on the challenges, implications, and potential solutions in this critical domain; and (3) the development of a conceptual model linking misinformation, public impact, and natural disasters based on available evidence from the papers sourced.

For the purpose of this review, the term "natural disaster" is used to explicitly refer to natural hazards, including but not limited to hurricanes, earthquakes, floods, tornadoes, and wildfires. The focus on natural hazards allows for a targeted exploration of how misinformation spreads and impacts populations during these specific types of crises.

Also, the conceptual model is obtained from the 9 studies to gain a greater understanding when examining the relationship between misinformation on social media in the context of natural disasters.

Definition of Terms

A disaster, according to the United Nations Office for Disaster Risk Reduction, is: "A serious disruption of the functioning of a community or a society at any scale due to hazardous events interacting with conditions of exposure, vulnerability, and capacity, leading to one or more of the following: human, material, economic, and environmental losses and impacts" [7]. Disasters can be caused by various kinds of hazards [18,19] and can have devastating impacts on people and communities. Disasters linked to natural hazards, including widespread fires, floods, storms, earthquakes, and droughts, may result in significant damage and loss of lives.

Disinformation is commonly defined as false information intentionally shared to deceive [20,21].

Infodemic, which was originally coined by David J Rothkopf in 2003, describes the overabundance of information, including misinformation, associated with significant events such as a federal election, pandemics, or natural disasters [22,23].

Misinformation is defined as the dissemination of inaccurate information without the intention to deceive [20,21].

Public health is the science and practice of preventing disease, extending life expectancy, and promoting overall health through organized societal efforts [24].

Public impact is defined as the influence or effect that actions and events have on the public [25].

Social media is defined as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, allowing the creation and exchange of user-generated content" [26].

Results

Overview

Initially, 173 studies were screened for eligibility based on their titles and abstracts, reducing the number of potential studies to 33. Following a full-text review, 9 studies met the inclusion criteria and were included in this narrative review (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the search findings.

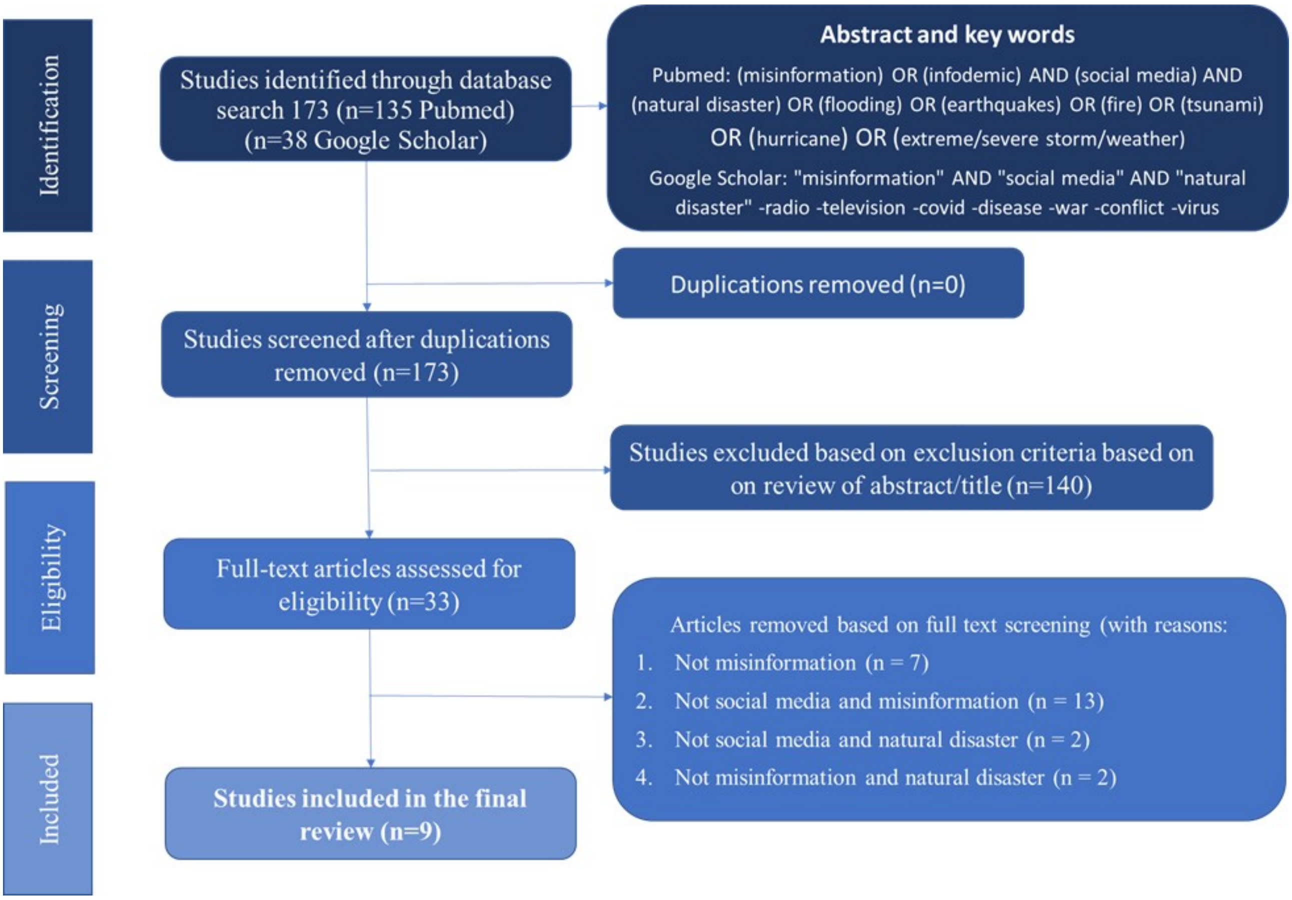


Table 1 summarizes the key characteristics of the 9 studies selected for the review. Of these, 4 studies used a case study design, 4 studies used content analysis, and 1 study used survey methods. All studies analyzed data extracted from social media platforms; 8 studies focused on Twitter, while 1 study examined several social media platforms.

XSL•FO
RenderX

Table . Summary of findings of characteristics of the 9 studies.

Author	Country of origin	Type of study and source	Social media platform	Type of natural hazard (disaster)	Type of misinformation	Study size
Gupta et al (2013) [4]	United States	• Content analysis, • peer-reviewed conference paper	Twitter (subsequently rebranded X)	Hurricane Sandy	Fake images of sharks swimming in the streets and manipulated photos of storm damage.	1.8 million tweets from 1.2 million unique users, with 10,350 tweets containing fake images and 5767 tweets containing real images.
Hunt et al (2020) [5]	United States	Case study, peer-reviewed	Twitter (X)	Hurricanes Harvey and Irma.	False claims that immigration status checks were being conducted at evacuation sites and shelters.	Hurricane Harvey: 2032 unique tweets (1440 debunking tweets). Hurricane Irma: 601 unique tweets (259 debunking tweets).
Vasudevan and Alathur (2022) [6]	India	• Case study, peer-reviewed	WhatsApp and Facebook	Heavy flooding	Misleading information on emergency instructions.	561 completed a survey who were affected by the flood in Kerala.
Rajdev and Lee (2015) [8]	United States	• Thesis: case study, • peer-reviewed conference paper	Twitter (X)	Moore Tornado and Hurricane Sandy	Spam and fake messages.	Collected 1% sample tweets posted during a period of each of the 2 natural disaster events.
Zhai et al (2023) [9]	United States	• Case study, peer-reviewed	Twitter (X)	Hurricane Sandy	False information about the disaster's impact and situation, such as exaggerated damages or incorrect emergency instructions.	691 tweets.
King and Wang (2023) [12]	United States	• Content analysis, peer-reviewed	Twitter (X)	Hurricane Harvey	Misleading information leading to changes in perception.	42 million tweets with 3589 original verified real or false tweets cross-checked with fact-checking websites and relevant federal agencies.
Dallo et al (2023) [13]	Global	• Content analysis peer-reviewed	Twitter (X)	Earthquake	About the ability to predict earthquakes.	82,129 tweets.
Oh et al (2010) [14]	Haiti	• Content analysis, • peer-reviewed conference paper	Twitter (X)	Haiti Earthquake	False claims about aid offers.	962 tweets.
Abdullah et al (2015) [15]	Japan	• Survey peer-reviewed	Twitter (X)	Generically focused	Unverified information and rumors.	133 participants (students from Iwate Prefectural University, Japan).

The narrative review evaluated the available evidence on misinformation via social media in the context of natural disasters and developed a conceptual model to enhance understanding of the intersection of misinformation origin, public impact, and emergency response disruption during natural disasters. The term "natural disaster" is used explicitly to refer to natural hazards, including but not limited to hurricanes, earthquakes, floods, tornadoes, and wildfires. Focusing on

XSL•FO
RenderX

natural hazards enabled a targeted examination of how misinformation spreads and affects populations during these specific types of crises.

Critical Analysis of Available Evidence

The 9 studies reviewed provide powerful insights into how misinformation spreads and the impacts to communities during natural disasters. Synthesizing evidence across key criteria, such as misinformation type and the social media platforms used, reveals distinct challenges posed by different disaster contexts (Multimedia Appendix 1) [4-6,8,9,12,14,15].

Type of Misinformation and Natural Disasters

Recurring themes in disaster-related misinformation include exaggerated reports, misleading emergency instructions, false predictions, and inaccurate information concerning public safety measures. Such misinformation not only disrupts disaster response efforts but also disproportionately affects vulnerable populations. Multimedia Appendices 2 and 3 illustrate the range of disasters examined and the corresponding misinformation trends identified in each case, highlighting the pervasive and far-reaching consequences of misinformation during crises.

Four studies focused specifically on hurricanes [4,5,9,12] and one on Hurricane Sandy and the 2013 Moore Tornado [8]. Two studies focused on earthquakes [13,14], 1 study was generically focused on “disasters” [15], and 1 study focused on the Kerala floods [6].

Social Media Platforms and the Spread of Misinformation

The reviewed studies indicate that misinformation during disasters spreads primarily through popular social media platforms, particularly Twitter (X), Facebook, and WhatsApp.

Twitter was identified by 8 studies as a social media platform for the dissemination of misinformation [4,5,8,9,12-15]. These studies emphasized Twitter’s role in spreading rumors, false predictions, and misleading information during disasters such as Hurricane Sandy, Hurricanes Harvey and Irma, the 2013 Moore Tornado, and the Haiti and Tohoku earthquakes.

Social media platforms like Facebook, WhatsApp, and Twitter are essential for rapid disaster communication, enabling institutions to disseminate critical updates, supporting evacuation efforts, and facilitating resource allocation such as food and medical supplies [6]. For instance, during Hurricane Sandy, social media was crucial for disseminating safety updates [9]; however, these platforms also acted as conduits for misinformation.

The dual role of social media, as both a valuable communication tool and a channel for misinformation, highlights its complex influence on public perception and response efforts [15]. An example of this was seen during the 2018 floods in Kerala, India, where WhatsApp served as the main communication channel for identifying places of safety and communicating the needs of affected individuals. However, this same communication channel was exploited for spreading misinformation, causing confusion and disruption to relief operations [6]. The rapid and unverified dissemination of false information among affected communities contributed to the chaos in relief operations.

The duality inherent in social media highlights the importance of actively monitoring and managing these platforms to ensure they aid rather than hinder disaster response efforts [5].

Impact of Exposure to Misinformation in the Context of a Natural Disaster

Misinformation during natural disasters can profoundly affect the public and disrupt disaster response systems. Vulnerable populations, along with the general public, face heightened risks as misinformation spreads through various sources, including individuals, influencers, bots, and fake profiles, leading to widespread confusion and panic [14,15].

The 9 studies reviewed highlight the rapid and far-reaching consequences of misinformation. Several studies reported associations between misinformation circulating on social media and heightened public panic, perceived misallocation of resources, and a decline in trust in emergency response systems [14]. Public confusion and fear are particularly acute when misinformation distorts the perceived severity of a disaster or spreads false emergency instructions [4,6,8,9,12,14,15].

For example, during Hurricane Sandy, exaggerated reports of damage and incorrect emergency instructions were circulated on Twitter, leading to public confusion and resource misallocation [8]. Similarly, false predictions about earthquakes shared on Twitter triggered panic and disrupted effective response strategies [13-15]. Furthermore, during the Kerala floods, WhatsApp emerged as a major source of misinformation, where false reports about aid and evacuation points delayed relief efforts [6].

In 1 notable instance, misinformation was specifically targeted at immigrant communities. During Hurricanes Harvey and Irma, false claims about immigration enforcement at evacuation sites generated fear and deterred individuals from seeking help [5].

Exposure to conflicting reports and unverified claims makes it challenging for the public to discern credible information, leading to skepticism and diminished trust in authorities and media sources [4].

Multimedia Appendix 4 presents the 3 key themes identified from the 9 studies illustrating the public impact from misinformation during a natural disaster, emphasizing its disruptive effects on trust, resource allocation, and public safety during natural disasters.

The evidence highlights the critical need for robust strategies to combat the spread of misinformation during natural disasters. Suggested approaches include digital literacy campaigns, timely debunking efforts, and coordinated action between governments, social media platforms, and public health organizations. Addressing these challenges is essential for strengthening emergency response systems and restoring public trust in official information.

XSL•FO
RenderX

Dynamics of Misinformation on Social Media in the Context of Natural Disasters

The dynamics of misinformation on social media during natural disasters represent a complex phenomenon that poses substantial challenges while also offering opportunities for targeted interventions. The rapid and widespread dissemination of misinformation endangers public safety and disrupts response efforts, yet it simultaneously highlights the potential for developing tailored mitigation strategies.

In this context, dynamics refers to the evolving interplay between social behaviors, technological mechanisms, and emotional responses that influence how misinformation originates, spreads, and influences public understanding and disaster response. An understanding of these dynamics, including the associated challenges, implications, and potential responses, is essential for designing effective strategies to mitigate the impact of misinformation and improve disaster resilience.

Challenges

The role of AI in misinformation management is inherently bidirectional. On one hand, algorithmic tools can detect and counter false information, while on the other hand, social media content algorithms often prioritize engagement over accuracy, curating personalized news feeds that amplify sensational content. As a result, misinformation, particularly that which evokes strong emotional responses, often receives greater visibility than verified information [5]. Platforms such as Twitter, Facebook, and WhatsApp facilitate the rapid sharing of sensational or emotionally charged, easier-to-read content and tend to resonate more with users experiencing stress during disasters. The emotional resonance contributes to the accelerated spread of misinformation [4]. The speed and volume of such content can overwhelm emergency communication channels, confuse the public, and erode trust in official sources [5,8,9,14].

Furthermore, the decentralized nature of social media allows anyone to act as a source of information, regardless of credibility. This absence of gatekeeping makes it difficult for the public to distinguish between trustworthy sources and those spreading misinformation. The resulting confusion can lead to panic, as evidenced during Hurricane Sandy and the Kerala floods, where false reports led to delayed relief efforts [4,6,8].

The influencers, users with a substantial social media following, exacerbate the problem. During the 2013 Moore Tornado, high-profile accounts were responsible for the spread of misinformation, which misdirected resources and heightened public confusion [8].

Compounding these structural issues is the psychological environment of disaster contexts. Zhai et al [9] observed that emotionally charged misinformation is more likely to resonate with fearful and uncertain audiences. This resonance increases the likelihood of misinformation being shared, creating a feedback loop in which it spreads rapidly and becomes increasingly more difficult to debunk.

These dynamics highlight the complexity of misinformation during disasters and the critical importance of regulating influential content sources, strengthening digital literacy, and enhancing platform accountability.

Implications

The consequences of misinformation during natural disasters are immediate and far-reaching. It not only causes immediate confusion and panic but also delays emergency responses and undermines long-term trust in official information sources. These effects persist well beyond the disaster itself. When the public is exposed to conflicting or inaccurate information during a disaster, their trust in future communications is diminished, making it more difficult for authorities to manage subsequent emergencies effectively [13,27].

As noted by Dallo et al [13], this erosion of trust weakens community cohesion and reduces collective disaster resilience. When public trust deteriorates, communities are less able to respond in a unified and effective manner, potentially leading to greater societal fragmentation and reduced capacity to withstand future disasters [28]. These outcomes highlight the importance of sustained efforts to build and maintain public trust in official communications during disasters.

Additionally, misinformation released during disasters can lead to misallocation of resources, inappropriate public behavior, and delayed emergency response. When individuals act on misinformation, such as evacuating in response to nonexistent threats or requests for aid in areas not requiring immediate need, emergency services can become overwhelmed [4,6,12]. For instance, Rajdev and Lee [8] reported how misinformation during Hurricane Sandy misled the public into taking misguided actions, which strained emergency resources and complicated the overall response effort. In such cases, disaster management teams must contend not only with the crisis but also with the secondary challenges created by misinformation.

Potential Solutions

Addressing the challenges posed by misinformation during natural disasters requires a multipronged approach. One key solution involves the use of advanced technological tools, such as AI algorithms, to detect and flag misinformation on social media platforms. These systems can identify patterns of misinformation and alert both users, official agencies, and platform administrators, enabling timely corrective action [5,9,14]. Additionally, partnerships with fact-checking organizations allow social media platforms, social media platforms to implement real-time verification mechanisms, which can help curb the rapid spread of unverified content [14].

Public education also plays a crucial role in countering misinformation. Digital literacy programs that teach individuals how to assess the credibility of online content can reduce susceptibility to misinformation during disasters. Such programs should emphasize critical thinking and promote awareness of the risks associated with sharing unverified information [8,14,29]. Furthermore, targeted educational campaigns aimed at specific demographics, such as older adults, who are more likely to unknowingly spread misinformation, can be particularly effective [13].

XSL•FO
RenderX

Moreover, collaboration between governments, social media platforms, and public health institutions is essential. Through strategic partnerships, these entities can coordinate the dissemination of timely, verified information during disasters [4,6]. Such collaborations should prioritize amplifying credible sources and adjusting platform algorithms to promote accurate information over sensationalized or misleading content. Furthermore, proactive engagement such as preemptive debunking and real-time updates during a disaster can help to limit the spread and influence of misinformation [5,9,14,30].

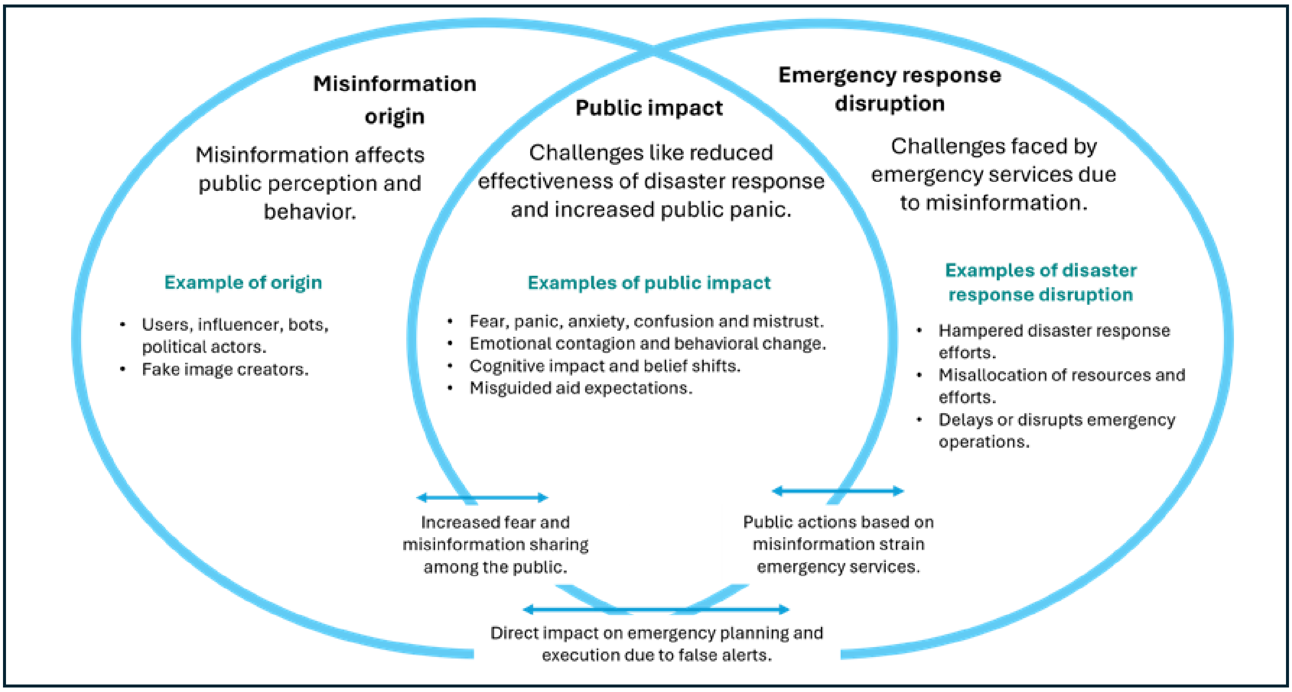
Communication strategies must also account for the emotional responses that misinformation can trigger during disasters. Understanding the emotional dynamics within affected communities enables the development of messages that are empathetic, clear, and trustworthy. This approach can help reduce panic and confusion while reinforcing public trust in official sources. For example, during the Kerala floods, misinformation shared via WhatsApp contributed to widespread anxiety and disrupted relief efforts [6]. Emotionally sensitive communication may alleviate distress and enhance the overall effectiveness of disaster response.

Finally, predisaster preparedness initiatives should include public education about misinformation and promote critical thinking before a natural disaster occurs. Programs such as community disaster preparedness events can inform the public about the dangers of misinformation and encourage reliance on official communication channels [13]. The proactive awareness campaigns are vital for ensuring that individuals are better equipped to navigate the complex information landscape during disasters, thereby strengthening societal resilience and reducing the harmful impact of misinformation.

Conceptual Model Linking Misinformation, Natural Disasters, and Public Impact

The research findings underscore the multifaceted impact of misinformation on natural disasters and the public. The conceptual model presented in Figure 2 [4-6,8,9,12,14,15], structured as a Venn diagram, offers a systematic framework that has been largely absent in existing literature. It elucidates the complex interrelationships between the origins of misinformation, its effects on public perception, and the resulting disruptions to emergency response systems during disasters.

Figure 2. Conceptual model extrapolated from the 9 studies, illustrating how misinformation on social media impacts the public and disrupts disaster recovery.



The conceptual model synthesizes findings from the 9 included studies, organized into 3 interrelated domains: misinformation origin, public impact, and emergency response disruption. Each domain was examined through extraction and thematic analysis of study-level data, enabling the identification of key patterns related to misinformation actors, societal reactions, and institutional challenges (Figure 2).

By illustrating the intersections and overlaps between the domains, the model highlights the compounding effects of misinformation in disaster contexts, demonstrating how these dynamics escalate challenges and increase the burden on response systems [1].

The interconnected nature of these concepts suggests that efforts to identify and address misinformation at its source may enhance public health outcomes and improve disaster response effectiveness.

While misinformation is typically unintentional, in the context of natural disasters, it may be amplified by coordinated disinformation campaigns, automated bots, or actors with strategic intent. This convergence increasingly blurs the distinction between misinformation and disinformation.

XSL•FO
RenderX

Misinformation Origin

Misinformation during emergencies often originates from a range of sources, including ordinary users, social media influencers, bots, and fake accounts [4,5,13-15]. Ordinary users may unknowingly share unverified or manipulated content, often motivated by a sense of urgency or a desire to inform others [31]. Influencers, driven by the pursuit of attention and engagement, may amplify misinformation due to its sensational or emotionally resonant nature. In contrast, bots and fake accounts are typically programmed to deliberately spread false information with the intent to create confusion and panic [21,27]. These actors contribute to a feedback loop, whereby misinformation gains visibility and traction, becoming increasingly difficult to correct once it has spread widely [32].

Additionally, the design of social media platforms, which prioritizes immediacy and engagement, plays a pivotal role in the amplification of misinformation. Content such as exaggerated reports of disaster severity, false safety procedures, or fabricated health advisories often spreads more rapidly than verified information from official sources [28,32,33].

Algorithms that prioritize engaging content over factual accuracy contribute significantly to this phenomenon [5]. Zhai et al [9] demonstrated how social network dynamics and sentiment contagion fuel the spread of misinformation, particularly in emotionally charged disaster contexts. Rapid content sharing without verification contributes to what is referred to as a misinformation cascade, making it challenging for the public to distinguish between accurate and false information.

Users often place trust in their personal networks for disaster-related updates, assuming credibility based on social proximity rather than content accuracy. The misplaced trust exacerbates the problem. During Hurricane Sandy, for example, exaggerated reports of damage and doctored images circulated widely on social media, causing unnecessary fear and panic [4,8].

As misinformation spreads, it creates confusion and panic, directly undermining the effectiveness of official communication channels during disasters [4,6,8,9,12-15]. It misguides public behavior and erodes trust in official communications, thereby complicating disaster response efforts [6,13]. The convergence between misinformation origin and its public impact reveals the challenge of maintaining public trust and ensuring effective communication during disasters [8].

Public Impact

The societal and psychological effects of misinformation during disasters are profound. The misinformation distorts public perceptions, leading to confusion and inappropriate actions, such as unnecessary evacuations or ignoring official guidance [9]. This confusion, compounded by the public's increasing reliance on social media for real-time updates, erodes trust in credible information and reduces compliance with official instructions [5,13].

The resulting behavior shifts, driven by misinformation, can delay coordinated disaster responses, misallocate resources, and undermine the overall effectiveness of emergency operations [6]. Dallo et al [13] similarly observed, stating that misinformation generates widespread confusion and fear among the public, significantly impeding disaster management efforts.

As reliance on social media increases, distinguishing between verified and false information becomes increasingly difficult, further diminishing public trust and adherence to official directives [5]. The erosion of public trust weakens effective communication and coordination during disaster response, contributing to fragmented and delayed decision-making [4-6,8,12-15].

Heightened emotional states during crises amplify the damaging effects of misinformation, destabilizing already fragile situations. For instance, during Hurricane Sandy and the 2018 floods in Kerala, misinformation fueled public confusion and anxiety, disrupting emergency responses and delaying critical relief efforts [6,8].

Moreover, beyond immediate operational challenges, the long-term mental health consequences of exposure to misinformation are significant. Anxiety, depression, and posttraumatic stress disorder may arise when receiving misinformation about the safety of affected areas or the availability of essential resources, placing further strain on mental health services [6,34,35].

The intersection between public impact and emergency response disruption becomes evident in the actions of misinformed individuals. Fueled by misinformation, public behaviors such as confusion, panic, and inappropriate actions directly undermine the efficacy of official communications and disaster response strategies [4,9,13]. These actions not only hinder response coordination but also lead to the misallocation of resources, ultimately reducing the effectiveness of emergency operations [6].

Emergency Response Disruption

While social media serves as a vital tool for crisis communication, it simultaneously poses significant challenges to emergency response operations. The rapid spread of misinformation can saturate communication channels, misdirect resources, and create false alerts, all of which complicate coordination and operational effectiveness [9,12,14].

Emergency responders are frequently forced to simultaneously manage both the actual disaster and counteract the spread of misinformation. This dual burden diverts attention and resources from where they are most urgently needed, thereby weakening the overall emergency response effort [6,9,12,14,15].

Disruption in resource deployment further complicates crisis management, particularly when misinformation about aid availability, safe zones, or evacuation procedures leads to confusion and misallocation. Events such as earthquakes and hurricanes have demonstrated how such misinformation delays relief efforts and overwhelms emergency services [6,9]. As responders are required to counter misinformation while managing real-time emergencies, operational delays become inevitable [6,8,9,12]. Compounding the problem is the speed at which misinformation spreads, often outpacing official corrections, making debunking efforts an ongoing challenge

XSL•FO
RenderX

[4,5]. Hunt et al [5] observed that the effectiveness of debunking strategies is heavily dependent on both the timing and platform used. When corrective messaging lags, public safety is undermined and trust in emergency services erodes [9,14,15].

The interplay between misinformation origin, public impact, and emergency response disruption highlights the compounded challenges misinformation creates during disasters. As illustrated in Figure 2, the domain overlap reinforces a feedback loop showing how misinformation alters public behavior, amplifies panic, and disrupts emergency response operations, collectively degrading the effectiveness of crisis management and compromising public safety. Addressing these challenges requires robust mechanisms to combat them. Governments, social media platforms, and relevant organizations must collaborate to enhance public awareness, refine communication strategies, and build public resilience against misinformation. By understanding how misinformation spreads and impacts both the public and emergency services, more effective interventions can be developed to strengthen disaster resilience and improve public health outcomes.

Discussion

Principal Findings

The narrative review has critically evaluated the spread of misinformation on social media in the context of natural disasters, shedding light on key challenges, implications, and potential solutions.

A conceptual model was developed to illustrate the interconnected relationships between misinformation origin, public impact, and disaster response, grounded in evidence from the 9 reviewed studies.

The main findings reveal the dual role of social media during disasters; while facilitating the rapid dissemination of vital information, they also serve as a vector for misinformation. The speed and scale at which the misinformation spreads can undermine official communication efforts, create public confusion, disrupt disaster resilience, hinder public health efforts, and divert critical emergency resources [6,35]. Several studies reported cases in which misinformation regarding aid availability and safe locations led to unnecessary chaos, overwhelming emergency services, and delayed assistance to those in need [6,8,12,36]. When the public perceives official sources as unreliable or slow, they increasingly turn to unofficial and less credible alternatives, heightening the risk of acting on misinformation [28].

Additionally, the erosion of trust in official communications weakens community resilience [14] and has psychological implications that extend beyond the immediate disaster response. Exposure to misinformation has been associated with long-term mental health impacts, including anxiety, stress, and posttraumatic symptoms, especially when the public is misinformed about safety conditions or resource availability [4,37].

Addressing this duality requires a deep understanding of the origins of misinformation, its public impact, and its capacity to disrupt disaster response systems.

The 3 interconnected concepts—misinformation origin, public impact, and emergency response disruption—emerged across the studies and informed the development of the conceptual model (Figure 2). By illustrating the overlap between these domains, the model provides a structured framework for understanding how misinformation disrupts public health, emergency response effectiveness, and disaster resilience.

Misinformation during disasters originates from a variety of sources, including ordinary users, influencers, and bots [4,6,13-15]. While some share content with good intentions, others, particularly influencers, can unintentionally amplify false narratives due to their large followings [6,9,12,13,38]. Automated bots further exacerbate the problem by generating and spreading misinformation that appears legitimate through tactics such as hashtag hijacking and the use of official links, as observed during the 2013 Moore Tornado [8].

The rapid spread of misinformation is driven in part by the algorithmic architecture of social media platforms, which prioritizes engagement-based content (shares, likes, and comments) over accuracy [5]. This environment facilitates the viral dissemination of misinformation, contributing to heightened stress, confusion, anxiety, and public panic [5,35].

Numerous studies confirm the detrimental effects of misinformation on public well-being during disasters. When individuals receive conflicting reports, their trust in official sources diminishes, leading to heightened emotional responses and behavior that undermines response efforts, such as unnecessary evacuations or disregard for safety protocols [4,5,14,35]. Events like Hurricane Sandy and the Kerala floods exemplify how misinformation and increased public anxiety eroded trust in official communications and complicated relief operations [4,6,8].

The misallocation of resources is another consequence of misinformation. False alerts have repeatedly caused emergency services to divert attention to noncritical areas, delaying vital assistance elsewhere.

For instance, during Hurricane Sandy and the 2013 Moore Tornado, misinformation about resource availability misled responders and complicated coordination, while during the Kerala floods, false data led to misdirected relief operations [4,6,8,9,36]. These disruptions illustrate how misinformation not only undermines immediate disaster response efforts but also impairs longer-term coordination and recovery. The psychological burden faced by affected populations is further intensified when community recovery is delayed due to the diversion of resources and persistent misinformation [36].

From the evidence, several lessons emerge. First, the speed of the misinformation spread necessitates proactive and timely countermeasures. Once misinformation gains traction, it often outpaces correction efforts, causing widespread harm. Real-time debunking campaigns led by trusted authorities must be prioritized and well-resourced to counteract this effect [8,9,14].

XSL•FO
RenderX

Communication strategies must also emphasize accuracy, clarity, and empathy to rebuild public trust [6,14,39].

Second, public education is vital. Preemptive educational initiatives that promote digital literacy and critical thinking can empower individuals to evaluate online information and reject misinformation [8,14,40]. These campaigns should target all demographics, including vulnerable groups such as older adults who are more susceptible to false narratives [13]. By fostering media literacy, these campaigns can reduce the spread of misinformation and enhance public resilience [14,41].

Third, technological advances should be leveraged. AI detection tools for real-time verification can identify patterns and flag or remove misinformation before it spreads widely [12,42]. These technologies must be integrated into disaster communication systems to enhance responsiveness and limit harm.

Early intervention is critical. Rapid deployment of countermeasures can contain the spread of misinformation and reduce its impact. Delayed responses, by contrast, allow false narratives to proliferate, compounding the damage [8,9].

To conclude, future efforts must focus on a combination of proactive strategies, technological innovation, and public education to effectively combat the challenges posed by misinformation during disasters. Disaster preparedness plans must formally integrate management of misinformation as a key component, ensuring that emergency responders are equipped to swiftly identify and address false narratives.

The discussion should be interpreted in light of the narrative review's strengths and limitations.

Strengths

A notable strength of this narrative review lies in its robust and transparent methodological framework, which adheres to the established guidelines from the Joanna Briggs Institute and the Cochrane Handbook for Systematic Reviews. The rigorous observance of PRISMA guidelines further ensures methodological transparency, enhancing the reliability and reproducibility of the findings.

Additionally, this review synthesizes evidence from a diverse array of studies, incorporating data from different geographic regions and varied natural disaster contexts. By not imposing date restrictions, this paper captures a broad temporal perspective, allowing for an inclusive assessment of the evolving dynamics of misinformation on social media during natural disasters.

Furthermore, the conceptual model developed in this narrative review advances the literature by illustrating the complex interrelationships between misinformation origins, public impacts, and disruptions to emergency response systems.

Limitations

There may be several limitations of the narrative review.

- The retrieved results included only the studies that were indexed in PubMed or found in a free search on Google Scholar. Thus, any studies not indexed on these sites were excluded from the review.
- The selection of search terms might not have been sufficiently comprehensive to capture all existing literature on misinformation on social media during natural disasters.
- The review did not examine narrative, social, and other theories posed as reasons for the spread of the misinformation.
- Language restrictions to English-only studies may limit the generalizability of the findings.

Conclusions

This narrative review critically examined the role of misinformation on social media during disasters, specifically natural hazards, highlighting its challenges, implications, and potential solutions. The review highlights the pervasive and damaging impact of misinformation, including its disruption of disaster response, erosion of public trust, and amplification of psychological distress. The conceptual model developed (Figure 2) provides a structured framework to deepen understanding of these interconnected dynamics.

The findings also demonstrate the complex, dual role that social media plays during disasters. While it facilitates the rapid dissemination of vital information, it simultaneously acts as a breeding ground for misinformation. The speed and scale at which misinformation spreads can undermine official communications, erode public trust, and create confusion and anxiety. These disruptions not only compromise disaster resilience but also hinder the effectiveness of emergency response systems and degrade public health outcomes, as illustrated by the conceptual model.

Moving forward, proactive strategies, technological innovation, and public education must be prioritized. This includes integrating misinformation management into disaster preparedness plans, enhancing public awareness, deploying advanced verification tools, and fostering trust in credible sources. By combining these approaches, individuals and communities will be better equipped to navigate the challenges posed by misinformation, thereby strengthening disaster resilience and improving public health outcomes.

In addition, future research, currently limited in scope, should focus on the following areas.

- A clearer understanding of the current state of misinformation during natural disasters.
- The experiences and perspectives of public health and disaster response professionals, particularly how misinformation affects roles, decision-making, and the effectiveness of response strategy.
- Emotional and psychological dimensions of misinformation in the context of natural hazards.

Ongoing research and policy efforts are critical for refining disaster preparedness and ensuring that future disaster responses are not compromised by the harmful effects of misinformation.

XSL•FO
RenderX

Acknowledgments

This study was supported by the Department of Health Services Research, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands.

The study, “The Impact of Misinformation on Social Media in the Context of a Natural Disaster,” was presented at the European Public Health Conference Lisbon 2024 as an oral presentation in the Realm of the Information Ecosystem and its Public Health Impact Session. The manuscript benefited from discussions and input received.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables used for data extraction.
[DOCX File, 25 KB - infodemiology_v5i1e70413_app1.docx]

Multimedia Appendix 2
Type of natural disasters.
[PNG File, 44 KB - infodemiology_v5i1e70413_app2.png]

Multimedia Appendix 3
Type of misinformation spread.
[PNG File, 91 KB - infodemiology_v5i1e70413_app3.png]

Multimedia Appendix 4
Public impact because of exposure to misinformation extrapolated from the 9 papers.
[PNG File, 34 KB - infodemiology_v5i1e70413_app4.png]

References

1. Muniz-Rodriguez K, Ofori SK, Bayliss LC, et al. Social media use in emergency response to natural disasters: a systematic review with a public health perspective. Disaster Med Public Health Prep 2020 Feb;14(1):139-149. [doi: 10.1017/dmp.2020.3] [Medline: 32148219]
2. Chary MA, Overbeek DL, Papadimoulis A, Sheroff A, Burns MM. Geospatial correlation between COVID-19 health misinformation and poisoning with household cleaners in the Greater Boston Area. Clin Toxicol (Phila) 2021 Apr;59(4):320-325. [doi: 10.1080/15563650.2020.1811297] [Medline: 32901533]
3. Islam A, Laato S, Talukder S, Sutinen E. Misinformation sharing and social media fatigue during COVID-19: an affordance and cognitive load perspective. Technol Forecast Soc Change 2020 Oct;159:120201. [doi: 10.1016/j.techfore.2020.120201] [Medline: 32834137]
4. Gupta A, Lamba H, Kumaraguru P, Joshi A. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. Presented at: Proceedings of the 22nd International Conference on World Wide Web (WWW ’13 Companion); May 13-17, 2013; Rio de Janeiro Brazil p. 729-736. [doi: 10.1145/2487788.2488033]
5. Hunt K, Wang B, Zhuang J. Misinformation debunking and cross-platform information sharing through Twitter during Hurricanes Harvey and Irma: a case study on shelters and ID checks. Nat Hazards 2020 Aug;103(1):861-883. [doi: 10.1007/s11069-020-04016-6]
6. Vasudevan J, Alathur S. Misinformation in social media during disasters. Int J Inf Syst Soc Change 2022 Jan;13(1):1-15 [FREE Full text] [doi: 10.4018/IJISSC.303596]
7. The Sendai framework terminology on disaster risk reduction “disaster”. United Nations Office for Disaster Risk Reduction (UNDRR). 2017. URL: https://www.undrr.org/terminology/disaster [accessed 2025-03-20]
8. Rajdev M, Lee K. Fake and spam messages: detecting misinformation during natural disasters on social media. Presented at: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT); Dec 6-9, 2015; Singapore. [doi: 10.1109/WI-IAT.2015.102]
9. Zhai W, Yu H, Song CY. Disaster misinformation and its corrections on social media: spatiotemporal proximity, social network, and sentiment contagion. Ann Am Assoc Geogr 2024 Feb 7;114(2):408-435. [doi: 10.1080/24694452.2023.2271549]
10. Chou WYS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
11. Chen L, Zhang G, Qiu X, Hu Z. Social media for intelligent public information and warning in disasters: an interdisciplinary review. Int J Inf Manage 2018;43:102-107. [doi: 10.1016/j.ijinfomgt.2019.04.004]

XSL•FO
RenderX

12. King KK, Wang B. Diffusion of real versus misinformation during a crisis event: a big data-driven approach. Int J Inf Manage 2023 Aug;71:102390. [doi: 10.1016/j.ijinfomgt.2021.102390]
13. Dallo I, Elroy O, Fallou L, Komendantova N, Yosipof A. Dynamics and characteristics of misinformation related to earthquake predictions on Twitter. Sci Rep 2023 Aug 17;13(1):13391. [doi: 10.1038/s41598-023-40399-9] [Medline: 37592002]
14. Oh O, Kwon KH, Rao HR. An exploration of social media in extreme events: rumor theory and Twitter during the Haiti earthquake 2010. 2010 Presented at: ICIS 2010 Proceedings p. 231 URL: https://aisel.aisnet.org/icis2010_submissions/231 [accessed 2024-04-14]
15. Abdullah NA, Nishioka D, Tanaka Y, Murayama Y. User's action and decision making of retweet messages towards reducing misinformation spread during disaster. J Inf Process 2015;23(1):31-40. [doi: 10.2197/ipsjjip.23.31]
16. JBI manual for evidence synthesis. Joanna Briggs Institute (JBI). 2020. URL: https://jbi-global-wiki.refined.site/space/MANUAL [accessed 2024-03-14]
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
18. Health emergency and disaster risk management framework. World Health Organization. 2019. URL: https://iris.who.int/bitstream/handle/10665/326106/9789241516181-eng.pdf?sequence=1 [accessed 2025-03-20]
19. World Health Organization. A strategic framework for emergency preparedness. 2016. URL: https://iris.who.int/bitstream/handle/10665/254883/9789241511827-%20eng.pdf,jsessionid=7FC0AC38E601560FF076384A459C41A7?sequence=1 [accessed 2025-03-20]
20. Joseph AM, Fernandez V, Kritzman S, et al. COVID-19 misinformation on social media: a scoping review. Cureus 2022 Apr;14(4):e24601. [doi: 10.7759/cureus.24601] [Medline: 35664409]
21. Wang Y, Bye J, Bales K, et al. Understanding and neutralising COVID-19 misinformation and disinformation. BMJ 2022 Nov 22;379:e070331. [doi: 10.1136/bmj-2022-070331] [Medline: 36414251]
22. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb 29;395(10225):676. [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
23. Simon FM, Camargo CQ. Autopsy of a metaphor: the origins, use and blind spots of the ‘infodemic’. New Media & Society 2023 Aug;25(8):2219-2240. [doi: 10.1177/14614448211031908]
24. Public impact fundamentals. Centre for Public Impact. 2024. URL: https://centreforpublicimpact.org/resource-hub/public-impact-fundamentals [accessed 2025-03-19]
25. Introduction to public health: public health 101 series. Centers for Disease Control. 2024. URL: https://www.cdc.gov/training-publichealth101/media/pdfs/introduction-to-public-health.pdf [accessed 2025-03-19]
26. Zhang S, Pian W, Ma F, Ni Z, Liu Y. Characterizing the COVID-19 infodemic on Chinese social media: exploratory study. JMIR Public Health Surveill 2021 Feb 5;7(2):e26090. [doi: 10.2196/26090] [Medline: 33460391]
27. Muhammed TS, Mathew SK. Enhancing disaster communication: strategies for effective social media use. J Commun Healthc 2022;15(1):64-72. [doi: 10.1007/s41060-022-00311-6]
28. Starbird K. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on Twitter. ICWSM 2017;11(1):230-239. [doi: 10.1609/icwsm.v11i1.14878]
29. Pennycook G, McPhetres J, Zhang Y, Lu JG, Rand DG. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. Psychol Sci 2020 Jul;31(7):770-780. [doi: 10.1177/0956797620939054] [Medline: 32603243]
30. Su Z, Agyingi E. Modeling the impact of misinformation on the transmission dynamics of COVID-19. AppliedMath 2024;4(2):544-560. [doi: 10.3390/appliedmath4020029]
31. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 9;359(6380):1146-1151 [FREE Full text] [doi: 10.1126/science.aap9559] [Medline: 29590045]
32. Southwick L, Guntuku SC, Klinger EV, Seltzer E, McCalpin HJ, Merchant RM. Characterizing COVID-19 content posted to TikTok: public sentiment and response during the first phase of the COVID-19 pandemic. J Adolesc Health 2021 Aug;69(2):234-241. [doi: 10.1016/j.jadohealth.2021.05.010] [Medline: 34167883]
33. Erku DA, Belachew SA, Abrha S, et al. When fear and misinformation go viral: pharmacists’ role in deterring medication misinformation during the “infodemic” surrounding COVID-19. Res Social Adm Pharm 2021 Jan;17(1):1954-1963. [doi: 10.1016/j.sapharm.2020.04.032] [Medline: 32387230]
34. Garfin DR, Silver RC, Holman EA. The novel coronavirus (COVID-2019) outbreak: amplification of public health consequences by media exposure. Health Psychol 2020 May;39(5):355-357. [doi: 10.1037/hea0000875] [Medline: 32202824]
35. Why misinformation is dangerous, especially during disasters. Red Cross Canada. 2023. URL: https://www.redcross.ca/blog/2023/5/why-misinformation-is-dangerous-especially-during-disasters [accessed 2024-04-15]
36. Shams AB, Apu EH, Rahman A, et al. Real-time credible online health information inquiring: a novel search engine misinformation notifier extension (seminext) during COVID-19-like disease outbreak. Research Square. Preprint posted online on 2020. [doi: 10.21203/rs.3.rs-60301/v2]
37. Purohit H, Castillo C, Diaz F, Sheth A, Meier P. Emergency-relief coordination on social media: Automatically matching resource requests and offers. FM 2013;19(1). [doi: 10.5210/fm.v19i1.4848]

XSL•FO
RenderX

38. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 2;41:433-451. [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
39. Houston JB, Hawthorne J, Perreault MF, et al. Social media and disasters: a functional framework for social media use in disaster planning, response, and research. Disasters 2015 Jan;39(1):1-22. [doi: 10.1111/disa.12092] [Medline: 25243593]
40. Pennycook G, Bear A, Collins ET, Rand DG. The implied truth effect: attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. Manage Sci 2020 Nov;66(11):4944-4957. [doi: 10.1287/mnsc.2019.3478]
41. Lewandowsky S, Ecker UKH, Cook J. Beyond misinformation: Understanding and coping with the “post-truth” era. J Appl Res Mem Cogn 2017;6(4):353-369. [doi: 10.1016/j.jarmac.2017.07.008]
42. Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. Sci Commun 2017 Oct;39(5):621-645. [doi: 10.1177/1075547017731776]

Abbreviations

AI: artificial intelligence
EUPHA: The European Public Health Association
HTA: Health Technology Assessment
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
R&D: Research and Development

Edited by I Brooks; submitted 20.12.24; peer-reviewed by D Harsono, N Calleja; revised version received 24.05.25; accepted 03.06.25; published 31.07.25.

Please cite as:
Hilberts S, Govers M, Petelos E, Evers S
The Impact of Misinformation on Social Media in the Context of Natural Disasters: Narrative Review
JMIR Infodemiology 2025;5:e70413
URL: https://infodemiology.jmir.org/2025/1/e70413
doi:10.2196/70413

© Sonya Hilberts, Mark Govers, Elena Petelos, Silvia Evers. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 31.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Availability and Use of Digital Technology Among Women With Polycystic Ovary Syndrome: Scoping Review

Pamela J Wright, PhD, MS, MEd; Charlotte Burts, BSN; Carolyn Harmon, PhD, DNP; Cynthia F Corbett, PhD

University of South Carolina, 1601 Greene Street, Columbia, SC, United States

Corresponding Author:
Pamela J Wright, PhD, MS, MEd
University of South Carolina, 1601 Greene Street, Columbia, SC, United States

Abstract

Background: Polycystic ovary syndrome (PCOS) is a common endocrinopathy among women that requires self-management to improve mental and physical health outcomes and reduce risk of comorbidity. Digital technology has rapidly emerged as a valuable self-management tool for people with chronic health conditions. However, little is known about the digital technology available for and used by women with PCOS.

Objective: The purpose of this scoping review was to identify what is known about digital technology currently available and used by women with PCOS for PCOS-specific knowledge, self-management, or social support.

Methods: The databases PubMed, Embase, CINAHL, and Compendex were searched using Medical Subject Headings terms for PCOS, digital technology, health knowledge, self-management, and social support. Inclusion criteria were full-text, peer-reviewed publications of primary research from 2010 to 2025 in English about digital technology used for PCOS-specific knowledge, self-management, or social support by women aged 18 years and older with PCOS. Exclusion criteria were articles about pediatric populations and digital technology used for intervention recruitment or by health care providers to diagnose or treat patients.

Results: In total, 34 full-text articles met the inclusion criteria. Given the scope of digital technology, eligible studies were grouped into 7 domains: mobile apps (n=14), internet-based programs (eg, Google; n=6), social media (n=6), SMS text message (n=2), machine learning (n=2), artificial intelligence (eg, ChatGPT [OpenAI]; n=3), and web-based intervention platforms (n=1). Findings highlighted participants' varied perceptions of technology usefulness based on reliability of health care information, application features, accuracy of PCOS or fertility prediction, social group engagement, user-friendly interfaces, cultural sensitivity, and accessibility.

Conclusions: There is potential for digital technology to transform PCOS self-management, but further design and development are needed to optimize the technologies for women with PCOS. Future research should focus on including end users during the design phase of digital technology, refining predictive models, improving app inclusivity, conducting frequent reliability testing, and enhancing user engagement and support via additional features to promote more comprehensive self-management of PCOS.

(JMIR Infodemiology 2025;5:e68469) doi:10.2196/68469

KEYWORDS

polycystic ovary syndrome; digital technology; mobile apps; digital health; self-management; mobile phone

Introduction

Polycystic ovary syndrome (PCOS) is a complex, heterogenous collection of symptoms due to hormonal dysregulation [1], and it is the most common endocrinopathy among women across all races and ethnicities [2]. Characterized by hyperandrogenism, ovulatory dysfunction, or polycystic ovaries, PCOS often manifests with a range of symptoms such as irregular menstrual cycles, hirsutism, and metabolic disturbances [1]. Obesity, insulin resistance, and dyslipidemia are common clinical features in women with PCOS and increase the risk for cardiometabolic diseases and reproductive cancers by ≥50% [3] while negatively impacting women's health-related quality-of-life and psychological morbidity [4]. Treatment guidelines for PCOS are complex and multifaceted, as the Endocrine Society Practice Guidelines recommend medical adherence, cognitive strategies, and lifestyle changes to self-manage day-to-day PCOS symptoms [5] and reduce affected women's risk of comorbid conditions (eg, cardiometabolic) [6]. Thus, there is an urgent need for innovative approaches to help women with PCOS effectively manage their condition.

Digital technology is defined as tools, systems, or devices that can generate, store, or process data, such as smartphone apps, wearable trackers, online platforms, and social media communities [7]. Over the past 3 decades, a range of digital technologies to support health have rapidly developed and realized widespread adoption. Studies have revealed that women

XSL•FO
RenderX

are more frequent users of online health information than men [8]. However, the gendered dimensions of using digital technologies to self-manage chronic health conditions has received little attention. Despite the prevalence of chronic health conditions among women across the lifespan, the little research available is primarily about pregnancy-related conditions or the perinatal period [9,10].

As a chronic condition, PCOS transcends the reproductive years. Study findings indicate that women with PCOS report inadequate information and disconcerting interactions with health care professionals [11,12], such that many women with PCOS pursue health care information and social support via the internet [13,14]. Approximately 98% of women with PCOS searched for PCOS-specific information on Google and 19% joined PCOS support groups or forums found during web-based searches [10]. However, little is known about the breadth of digital technology available for or used by women with PCOS or the features offered to help promote health for women with PCOS. Thus, the purpose of this scoping review was to identify available digital technology used by women with PCOS for PCOS-specific knowledge, self-management, or social support.

Methods

Overview

A scoping review was chosen due to the broad nature of the research question in a relatively new area of interest, that being both digital technology and its purposeful use among women with PCOS. A scoping review is ideal when attempting to categorize the volume, nature, and features of emerging research [15]. This scoping review was conducted using the updated Joanna Briggs Institute (JBI) guidance for scoping reviews [16].

The research question guiding this review was: "What is known about the digital technology currently available and used by women with PCOS for PCOS knowledge, self-management, or social support?" Inclusion criteria were full-length, peer reviewed articles reporting primary research about digital technology used for PCOS-specific knowledge, self-management, or social support by adult (≥18 y) women with PCOS, available in English, and published from the years 2010 to 2025. The year 2010 was chosen to be inclusive of all digital technologies, but with awareness of the rapid nature of technological development in the digital application and website realm [17]. Exclusion criteria were articles that were not primary research (eg, editorials and reviews), research about the pediatric population with PCOS, and studies that used digital technology for intervention recruitment or medical treatment. The pediatric population (≤18 y) was excluded because, although a PCOS diagnosis can be considered in pubescent girls, this age poses diagnostic problems due to characteristics of normal puberty that often overlap with signs and symptoms of PCOS [18]. In addition, the health literacy, design preferences, and access to digital technology of the pediatric population differ from those of adults. We excluded articles about digital technology used for recruitment and by health care providers for diagnosis and medical treatment because we were interested in PCOS-specific digital technology used by women with PCOS for PCOS-specific knowledge, self-management, or social support.

Thus, the participants included women aged 18 years and older with PCOS. The concept for this scoping review was available digital technology used by women with PCOS for knowledge, self-management, or social support. The context is open in terms of geographic regions but specific to digital technology used by women with PCOS and not digital technology used for research recruitment nor in health care settings for diagnosis or procedures.

Search Strategy

The search strategy was created using Medical Subject Headings terms for PCOS, digital technology, health knowledge, self-management, and social support. Social support was defined as empathy, encouragement, information, and feedback provided by others with shared experiences [19]. Adapting the strategy based on the database PubMed, the databases Embase, CINAHL, and Compendex were also searched for relevant articles. The search strategies used can be found in Multimedia Appendix 1.

Evidence Screening and Selection Process

The search was completed by 2 reviewers (CB and PJW) using Covidence systematic review software (Veritas Health Innovation). After removing duplicates, the articles were screened by title and abstract. Considering the research question and inclusion and exclusion criteria, copies of the full articles were obtained for studies that appeared relevant. If relevance was unclear from the abstract, the full article was kept for further review. The interrater reliability was strong between the 2 reviewers of the initial screening with a Cohen κ of 0.88. The remaining articles were divided into three-fourth of the research team (CB, PJW, and CH) for individual full-text review. The 3 reviewers collaboratively discussed the inclusion of each article. The selected articles were then read by 2 of the reviewers (PJW and CB) to assign a quality rating. The final selected articles were read again by 3 reviewers (PJW, CB, and CH) to collect data.

Quality and Risk of Bias Assessment

The JBI, an international research organization and global leader in evidence-based health care, was chosen because it offers checklists and rigorous processes for diverse forms of evidence. Using the JBI checklists, 2 reviewers (CB and PJW) critically appraised each article for methodological quality and risk of bias to increase confidence in the findings. The third reviewer (CH) was available to discuss concerns, of which there were none.

Data Extraction and Analysis

A data chart was created to compile basic study characteristics (eg, authors, publication year, country of origin, and research design), study purpose, the technology used, outcomes, and conclusions. Furthermore, 3 reviewers (CB, PJW, and CH) extracted data and reviewed the data chart to reduce the chance for errors and bias. The research team met several times to code data into categories and summarize the findings. Findings were viewed through the lens of the Unified Theory of Acceptance and Use of Technology (UTAUT). UTAUT was developed by integrating various behavioral theories to help explain the acceptance and use of digital technology. UTAUT emphasizes four key determinants of intention and usage: (1) performance

XSL•FO
RenderX

expectancy (perception that a system helps improve performance or productivity), (2) effort expectancy (perception of ease associated with using the system), (3) social influence (the impact of external social factors such as social norms on use), and (4) facilitating conditions (adequate resources and support to use the system) [20].

Rigor

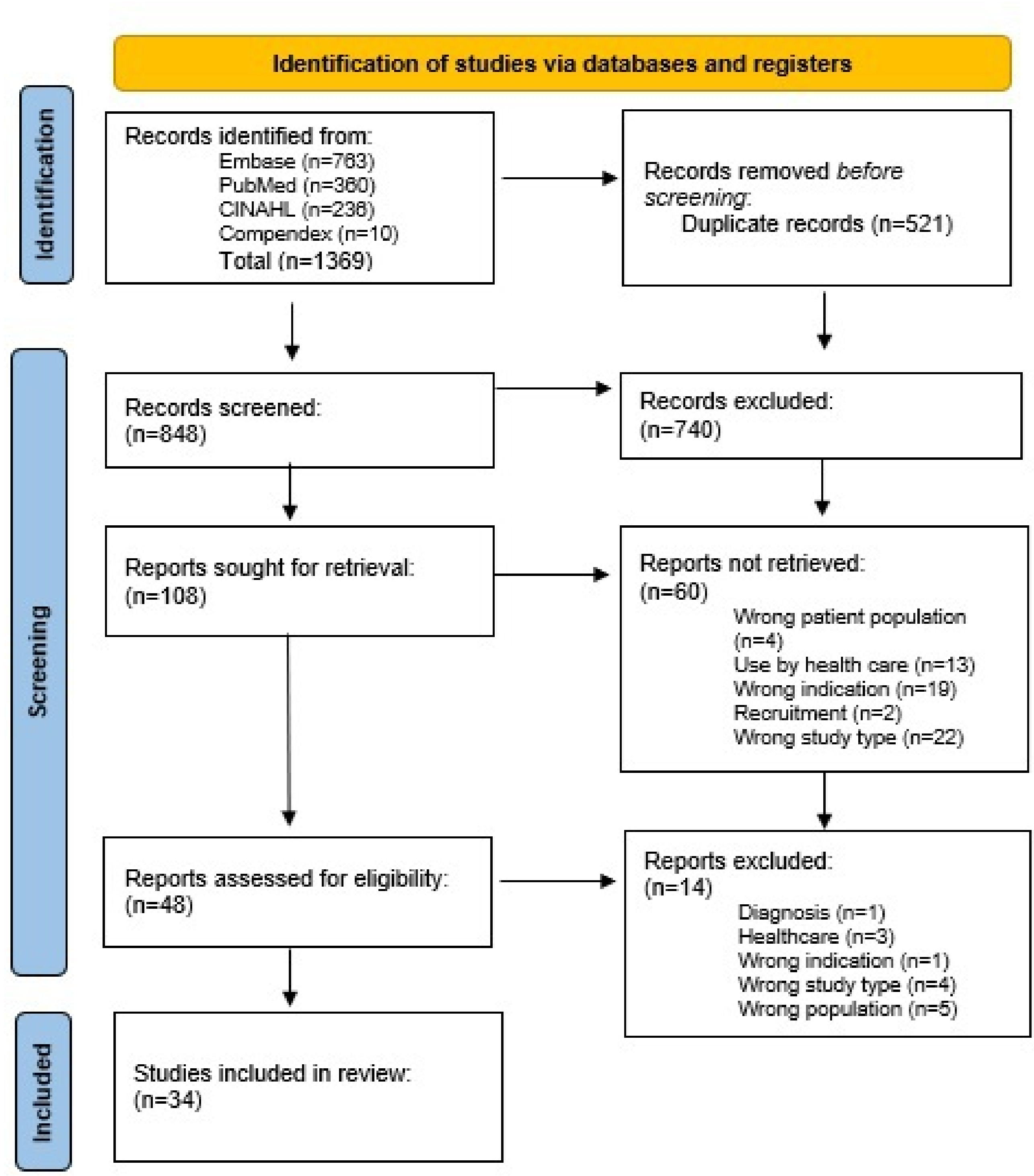
In addition to iterative review, study rigor was increased by using a scoping review protocol and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) checklist (Checklist 1).

Results

Overview

A scoping review of the literature revealed 34 full-text articles that met inclusion and exclusion criteria (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flow chart.



XSL•FO
RenderX

All selected articles met the criteria on the JBI checklists, and were, therefore, included in the review. Data that were systematically extracted and included in the final summary table were author, publication year, research design, study location, sample size, purpose of the technology, and key findings (Table 1). The studies used various research designs, including cross-sectional and randomized controlled trials. Sample sizes varied widely, ranging from 1 [21] to 416,712 [22]. Given the scope of digital technology, eligible studies were grouped into 7 domains: mobile apps (n=14), internet-based programs (eg, Google; n=6), social media (n=6), SMS text message (n=2), machine learning (n=2), artificial intelligence (AI; eg, ChatGPT [OpenAI]; n=3), and web-based intervention platforms (n=1).

XSL•FO
RenderX

Table . Characteristics of included studies (N=34).

Author	Year	Design	Location	Sample size, n	Purpose of technology	Findings
Mobile apps						
Alotaibi and Shaman [23]	2017	RCT[a]	Saudi Arabia	50	To connect women with PCOS[b] with health care providers and help diagnose PCOS	Increased PCOS awareness
Boyle et al [24]	2018	Cross-sectional	Australia	264	To identify pre-existing mobile apps relating to women's health or PCOS that may support self-management	16 apps identified: 8 diet only, 3 fertility or menstruation alone or combined with diet, 5 general information
Rodriguez et al [25]	2020	Multiple methods	United States	9	To assess risk factors and diagnose PCOS based on symptoms	Overall, good estimation of risk, but occasional false positives due to overestimation of certain risk factors
Choi et al [26]	2021	Multiple methods	Korea	30	To promote lifestyle modification and social support for women with PCOS	An integrated app is feasible to develop, but requires further research
Jain et al [22]	2021	Cross-sectional	United States, United Kingdom, India, Philippines, and Australia	416,712	To identify the most prevalent symptoms of PCOS	Common symptoms identified: bloating, high cholesterol, and elevated blood glucose
Sang et al [27]	2022	RCT	China	100	To provide PCOS education, promote lifestyle modification, and provide emotional social support via WeChat (Tencent Holdings Limited)	Statistically greater level of change BMI (87/100, 87.2% lost weight compared with only 68/100, 67.5% in the control) and higher live pregnancy and birth rates
Wang et al [28]	2022	RCT	China	100	To provide theory-based lifestyle modification via an app	Theory-based apps can support weight loss and mental health among women with PCOS
Lee and Lee [29]	2023	RCT	Korea	28	To provide PCOS education and a lifestyle modification program featuring exercise and diet	Apps can support weight loss and reductions in depressive symptoms among overweight women with PCOS
Ou et al [30]	2023	RCT	China	100	To provide health information and patient-provider communication via WeChat and a diet or physical activity tracker via a weight management app	Apps are effective means to provide health information and enable patient-provider communication to help promote health outcomes

XSL•FO
RenderX

Author	Year	Design	Location	Sample size, n	Purpose of technology	Findings
Peven et al [31]	2023	Multiple methods	United Kingdom	24	To track symptoms	88% (21/88) exact match between physician and symptom checker classifications, with 100% sensitivity and 75% specificity. Positive predictive value of 80% and negative predictive value of 100%.
Stujenske et al [32]	2023	Cross-sectional	United States	386	To track menstrual cycles and predict fertility	3 main tracking devices: smartphone apps, temperature tracking, and at-home urine hormone tests. 40% (154/386) used for determining fertility, 27% for tracking symptoms (104/386), and 18% (70/386) for receiving reproductive health education
Trépanier et al [33]	2023	Descriptive	Global	119 apps	To provide menstrual and pain tracking	Free mobile apps with varied approaches; overall, good functionality but poor design and lack of evidence-based approaches to pain
Dilimulati et al [34]	2024	RCT	China	80	To provide videos about healthy lifestyle behaviors	WeChat more engaging and produced more healthy lifestyle changes versus metformin alone; however, metformin significantly lowered HOMA-IR[c]
Hamzehgardeshi et al [35]	2024	RCT	Iran	60	To provide health information and counseling via WhatsApp	WhatsApp (Meta) effective means to deliver intervention that promotes healthy lifestyle behaviors
Internet						
Mallappa Saroja and Hanji Chandrashekar [36]	2010	Qualitative	Global	15	To provide information about PCOS-specific symptom management	Quality of information poor; no website with editorial review
Holbrey and Coulson [37]	2013	Qualitative	United Kingdom	50	To provide social support via online support group discussion forums	Positives: emotional, informational social support; negative: source of worry and anxiety

XSL•FO
RenderX

Author	Year	Design	Location	Sample size, n	Purpose of technology	Findings
Authier et al [38]	2020	Qualitative	France	211	To provide knowledge and social support to infertile women with PCOS	Positive source of experiential sharing; inaccurate and incomplete health information
Hoyos et al [14]	2020	Cross-sectional	United States	759	To find information about PCOS symptoms, support groups, and other PCOS-specific topics via Google	98.2% (745/759) of women used Google to search their symptoms because >50% women were dissatisfied with care; 18.8% (143/759) sought social support.
Malhotra and Kempegowda [39]	2023	Cross-sectional	Global	—[d]	To increase PCOS awareness and provide PCOS-specific information via Google	Global increase of PCOS awareness especially during PCOS Awareness Month, and increased volume of PCOS searches in recent years
Gomula et al [40]	2024	Qualitative	United States and United Kingdom	15	To find health information and social support and compare self to "normal" women via Google	Overwhelming amount of information that lacked reliability and cultural sensitivity and created conflict between phenotypes
Social media						
Gour et al [41]	2021	Quasi-experimental	Turkey	41	To provide PCOS-specific knowledge via video-based, structured, educational module posted on preferred social media site	PCOS-specific knowledge significantly increased among those with at least minimum level of health literacy and familiarity with the Internet
Elhariry et al [42]	2022	Qualitative	Global	8	To understand the impact of PCOS influencers on social media platforms	PCOS influencers report motive as PCOS awareness but all with marketing interest
Clarke et al [43]	2023	Cross-sectional	Jamaica	80 YouTube videos	To provide PCOS content on YouTube	About 30% (24/30) of videos created by physicians. Greater global quality score and video power index among videos posted by health care professionals.
Horvath et al [44]	2024	Qualitative	Global	238 videos	To provide PCOS-specific health information via Tik Tok	TikTok attracts considerable engagement; however, most videos provided low-quality information

XSL•FO
RenderX

Author	Year	Design	Location	Sample size, n	Purpose of technology	Findings
Naroji et al [45]	2024	Cross-sectional	Global	34,308	To provide PCOS-related information across the social media platforms of TikTok, Instagram, and Redditt	About 1.8million views on top PCOS content on TikTok. Weight and diet most common topics. Interactions with medical providers noted in 30% of content. Reddit most engagement under self-management topics.
Afaq et al [46]	2025	Cross-sectional	Global	12,200 posts	To provide PCOS-specific information and social support via X	Higher-quality content with physicians contributing 30% of the discourse; varied engagement
SMS text messaging						
Jiskoot et al [47]	2020	RCT	Netherlands	183	To promote weight loss among women with PCOS and obesity	Group with SMS text message lost more weight than other groups
Dietz de Loos et al [48]	2021	RCT	Netherlands	183	To provide reminders	SMS text message helpful to remind, encourage, and motivate for lifestyle behaviors
Machine learning						
Zigarelli et al [49]	2022	Cross-sectional	India	541	To predict risk factors and diagnosis of PCOS based on health records for patients and providers	Good prediction accuracy: 80% first run and 82.5% second run.
Karia et al [50]	2023	Correlational	India	—	To track complete menstrual health and predict likelihood of PCOS	72.1% - 90.4% accuracy
Artificial intelligence						
Devranoglu et al [51]	2024	Cross-sectional	Turkey	460 queries	To provide answers about PCOS and fertility treatments via ChatGPT	Although fluctuations in performance, accuracy rated high
Shamanna et al [21]	2025	Case study	India	1	To deliver personalized nutrition by predicting postprandial glucose response	Decreased bodyweight, waist circumference, blood pressure, and fasting insulin after 360 days
Ulug et al [52]	2025	Cross-sectional	Turkey	63 questions	To provide answers to PCOS nutrition-related queries via ChatGPT	ChatGPT provides high quality information; readability difficult for average user
Web-based intervention platform						

XSL•FO
RenderX

Author	Year	Design	Location	Sample size, n	Purpose of technology	Findings
Percy et al [53]	2024	Mixed methods	United Kingdom	13 interviews and 58 surveys	To provide theory-based comprehensive PCOS self-management program	Effective for self-management, specifically modest weight loss and reduced depressive symptoms; need for tailoring across lifespan

[a]RCT: randomized controlled trial.

[b]PCOS: polycystic ovary syndrome.

[c]HOMA-IR: Homeostatic Model Assessment of Insulin Resistance.

[d]Not applicable.

Mobile Apps

Mobile apps, software applications designed to run on a small mobile device such as a phone, tablet, or wearable, were the most featured digital technology among the selected articles (14/34; 41%). Mobile apps were used for the following reasons: (1) menstrual tracking to assess for the likelihood of PCOS or predict fertility, (2) symptom tracking to assess for risk factors of PCOS, (3) lifestyle modification, (4) PCOS information, (5) social support, and (6) connection with health care professionals. In one article, the authors identified 16 existing PCOS-specific mobile apps that focused on menstrual tracking, nutrition, or general PCOS education [24], while another article identified 119 free mobile apps for menstrual tracking for women with or without PCOS (eg, Flo [Flo Health, Inc] and Clue [BioWink GmbH]) [32]. Findings by researchers in 2 studies about the use of apps featuring lifestyle modification revealed positive outcomes such as decreased waist circumference [28], modest weight loss, and decreased depressive symptoms [29], and positive chemical and clinical pregnancy rates [26] compared with a standard care control group. Apps can be a feasible means to deliver a PCOS-specific lifestyle intervention [29,30,34,35].

Internet

The category of internet included any type of content electronically transmitted and stored. In 15% of articles (5/34), the researchers studied the Google search browser as used by organizations to promote PCOS awareness [33] and by women with PCOS to search for PCOS health information [30,33,39], nutrition and exercise [14], and social support [14,37,38,40]. In an analytical cross-sectional survey (N=759), researchers reported 98.2% (745/759) of women with PCOS used Google to search their symptoms and 18.8% (143/759) sought support from others who understand PCOS by joining online groups [14]. The Google search engine was reported to be effective at promoting PCOS awareness [39]; however, women with PCOS felt overwhelmed by the amount of information and doubted the information's reliability [38,40].

Social Media

Social media refers to interactive online platforms. Researchers who authored 4 of the articles (12%) included information about online platforms: YouTube (Google) [43] and TikTok (ByteDance), Instagram (Meta), Reddit [45], and X, formerly known as Twitter [42,46]. Authors of the article about YouTube described the analysis of PCOS-specific videos (n=80) posted to YouTube. Of these videos, 37% (30/80) were uploaded by women with PCOS and 29% (23/80) were uploaded by health care professionals. Videos posted by women with PCOS had greater popularity as evidenced by total number of likes; however, they had the lowest quality indicator scores compared with those posted by health care professionals [43]. Contrarily, Twitter or X had higher-quality content with physicians contributing most of the discourse [46]. Engagement was higher on TikTok (approximately 1.8 million viewers) than Instagram and Reddit. The most searched topics on TikTok and Instagram were "weight" and "diet" whereas health information and discussion about health care encounters were the most searched topics on Reddit [45]. A conflict of interest among PCOS-specific content on social media platforms was the presence of promotional advertising: 29% (23/80) of all YouTube videos [42], 45% (23/50) of all Tik Tok videos, and 89% (45/50) of all Instagram posts [45]. As such, social media platforms often involve people, called influencers, who affect the purchasing decisions of others because of their authority, knowledge, or position. Elhariry and colleagues [42] identified the top 100 PCOS influencers and interviewed 8 of them. Influencers self-identified as PCOS advocates and strived to post on multiple media platforms. Connections with marketing companies enabled them to spread their message through brand sponsorships.

SMS Text Messaging

SMS text messaging was featured in 9% of articles (2/34) that reported different analyses from the same study. In the original article, the authors reported the effectiveness of a group-based 3-component lifestyle intervention (nutritional advice, exercise, and cognitive behavioral therapy) to promote weight loss among women with PCOS and obesity. Furthermore, 1 of the 3 study arms received the intervention with an additional 9 months of feedback through SMS text message via their mobile phone [47]. In the second article, they reported on the effectiveness of the intervention on PCOS phenotypical features, such as hyperandrogenism [48]. Overall, the 3-component lifestyle intervention promoted weight loss as compared with the control group, and more so in the group receiving SMS text messages. The authors concluded that SMS text message feedback was useful to remind, encourage, and motivate the women with

XSL•FO
RenderX

PCOS to continue the lifestyle behavior and maintain weight loss.

Machine Learning

In 9% of studies (2/34), apps that included machine learning were developed and tested to track menstrual characteristics [49] and predict the likelihood of a PCOS diagnosis [35]. Machine learning using health care records, patients, and health care providers resulted in 80% - 82.5% accuracy [49], whereas machine learning using various machine learning algorithms ranged from 72.1% to 90.4% [50]. Both study authors noted that currently the accuracy of trackers and PCOS predictors vary widely, especially among free apps, further increasing health disparity in rural areas and among those with lower socioeconomic status.

Artificial Intelligence

AI is technology that enables computers and machines to simulate human intelligence. Thus, AI systems are made to understand and respond to human language, recognize objects, and learn from new information [54]. Furthermore, 6% (2/34) of articles were about the use of ChatGPT to provide PCOS information; 1 about fertility treatments [51] and 1 about nutrition [52]. In both articles, the accuracy of the information was rated high; however, the readability was a challenge for the average user. Shamanna and colleagues [21] completed a case study examining the use of AI to predict postprandial glycemic response to meals and suggesting alternative meal options to avoid glucose spikes. After 360 days of use, the 38-year-old woman with PCOS decreased body weight, waist circumference, blood pressure, and fasting insulin level [21].

Web-Based Intervention Platform

Of the 34 articles, the authors of 1 article (4%) discussed the adaptation of a web-based self-management program based on the identified barriers to self-management and psychological well-being reported from women with PCOS [53]. The existing HOPE program was chosen given its research with multiple other patient groups, such as those with diabetes and cancer, the easier accessibility of web-based intervention platform, and the potential to reach many women with PCOS. Using theory and input from end users, 6 modules were created to include health education, self-care, mindfulness, nutrition, communication skills with health care providers, and goal setting. Now called HOPE PCOS, the program offered features to address multiple components of PCOS self-management and was deemed ready for feasibility testing [53]. However, it was not complete as it omitted physical activity or health-specific needs of older women with PCOS.

Key Findings With a UTAUT Lens

Performance Expectancy

The primary reasons women with PCOS used digital technology was to find health information about PCOS treatment and self-management, track menstrual characteristics to determine fertility, and assess the likelihood of a PCOS diagnosis. As expected, women quickly and conveniently found health information with a Google search, yet they reported the volume of information as overwhelming and the content contradictory [40]. Many of the free apps for menstrual tracking lacked reliability [33,50]. Women with PCOS also expected a confidential, secure platform for social support with others who understand PCOS. Overall, women with PCOS, the end users of PCOS-specific digital technology, expected an interactive, comprehensive, evidence-based, and user-friendly platform that assisted with self-management, health behavior tracking, interaction with health care providers, and social support [14,24,26,33,40,53]. Digital technology incorporating these features was found to help promote weight loss and decrease depressive symptoms [28,29,45] and improve biochemical markers such as hyperandrogenism and ovulatory dysfunction [48].

Effort Expectancy

Women with PCOS report the expectation of an easily accessible and user-friendly platform [23,25,27,28]. Most apps were found to be easily accessible and user-friendly, although engagement was often low, especially in interventions relying on self-motivation [20,33]. However, non–PCOS-specific apps (eg, Flo and Clue) were criticized for insufficient customization, making the information more difficult to use [22,24]. Only one form of digital technology, the website HOPE PCOS, incorporated end user input during the design phase [53].

Social Influence

PCOS awareness was increased through information posted on social media. Some studies noted low involvement in public social support groups [14]. Apps with integrated social functions, such as private networks, demonstrated higher engagement [23,27], especially if the end users shared more in common (eg, life phase) with peers than the illness experience [23,39,40].

Facilitating Conditions

Several facilitating conditions were found to be deficient and negatively affected digital technology use among women with PCOS, including lack of accessibility (eg, cost of Wi-Fi or platform) [33,53], content that required high health literacy [46,51,52], and concerns about the reliability of health information [14,24,40,44]. Women with PCOS indicated that the currently available digital technology is helpful but does not address all their needs [24]. Thus, the lack of a comprehensive set of features deterred regular use of any one type of digital technology [24,47,48].

Discussion

Principal Findings

Of the selected 34 articles for this scoping review, the researchers' main goals across the studies included (1) assessing usability, engagement, and performance of different types of digital technology to meet needs of women with PCOS; (2) enhancing knowledge, symptom tracking, and lifestyle modification via digital technology; and (3) predicting the risk of PCOS through machine learning and symptom-tracking apps. A clear expectation of any digital technology was PCOS education. In a large international cross-sectional study of women with PCOS aged 18 - 45 years (N=1256), approximately

XSL•FO
RenderX

65% (n=816) reported dissatisfaction with information received from health care professionals and the need for self-directed research [55]. Findings from this scoping review indicate concern about the quality and reliability of online health information. Most online information was posted by other women with PCOS based on anecdotal evidence. Higher-quality content was found on Twitter or X and ChatGPT. The primary users of both platforms have college degrees and represent technical sectors, education, research, or business [56,57]. Unfortunately, as information quality increases, readability for the average user decreased.

A significant amount of PCOS content is associated with commercial interests or promoted by influencers with market interest [40,42,58]. The impact of social media influencers on health outcomes is a growing area of research. In a cross-sectional survey of participants, 59% (137/232) stated that they prefer to follow social medial influencers on social media platforms compared with nonsponsored posts [58]. Most research about influencers have focused on their marketing of diets or the impact of their imagery on body image dissatisfaction [59]. While not entirely negative, this type marketing can be misleading and harmful for individuals with vulnerabilities such as body image concerns [60], which are prevalent among women with PCOS [61].

Consistent with previous studies, most women with PCOS preferred apps for PCOS self-management and personal data tracking given their ease of use in terms of time and space [62]. Many apps are currently available for menstrual tracking; however, most are not equipped to interpret irregular cycles [63], a clinical feature of many women with PCOS [1]. Machine learning advances hold potential to improve the accuracy of menstrual trackers and PCOS predictive models [49,50]. Women with PCOS across the lifespan have also expressed need for reliable PCOS-specific apps that offer an evidence-based, comprehensive approach to self-management, are interactive with both others with PCOS and health care providers, and target multiple outcomes [64,65]. In this scoping review, the app AskPCOS (Monash Centre) and the web-based program HOPE PCOS, were the only digitally based platforms to approximate a comprehensive approach [24,53].

Social media platforms were a common source of social support among women with PCOS because of the asynchronous access to a larger group of individuals who share similar life experiences or interests. Similar to findings from a systematic review about social connectedness among individuals with chronic health conditions [66], engagement diminished over time as women with PCOS sought more common ground than the illness experience, such as life stage (eg, student, mother, and retiree) and, therefore, overall health goal (eg, fertility vs risk reduction of comorbidities) [14,23,26,31,40]. Thus, many social media platforms could benefit from tailoring to reach specific subgroups of women with PCOS, including those with different phenotypes and sociodemographic characteristics (ie, race and residential location).

While digital technology confers several benefits, other facilitating conditions, such as accessibility due to cost or geographic location, remain issues among many women with PCOS. These digital determinants of health encompass technological factors that can influence the access and use of information for self-management, which, in some cases, can potentiate existing sociodemographic inequities in health care [67]. Digital health literacy, or the individual's ability to find, understand, appraise, and apply health information from electronic sources [68] is an emerging priority as people regularly interface with digital technology for health information. The design and development teams of digital technologies could benefit from consideration of digital determinants of health at the onset and the inclusion of a diverse sample of end users for input on use, useability, and aesthetics. In addition, the development of a system to regularly evaluate, score, and post the reliability of health care information on digital technology could empower women with PCOS as they choose digital technology to use and apply electronic health care information.

Assessing the findings of the scoping review through the lens of UTAUT helps more clearly identify future directions for digital technology used by women with PCOS. For performance expectancy, women with PCOS prefer an interactive PCOS-specific app with comprehensive features, including health care education (eg, diagnosis and treatment), lifestyle modification (eg, physical activity, nutrition, and stress management), personal data tracking, interaction with health care providers, and peer support. They expect reliable, evidence-based information that involves accurate predictions of diagnosis and fertility, promotes positive health outcomes and is unencumbered by commercial interests. For effort expectancy, women with PCOS want an easily accessible, user-friendly, and pleasing interface. For social influence, women with PCOS readily use more popular modes of digital technology to connect with others with PCOS; however, motivation to maintain usage is low due to exclusion of subgroups of women with PCOS and lack of sociodemographic context sensitivity. For facilitating conditions, digital technology with more features and higher reliability may cost more, presenting an economic barrier for some women with PCOS. While mobile phone use is rapidly increasing, including across low- and middle-income earning countries, Wi-Fi access may be unavailable [69].

Limitations

By using a systematic, comprehensive search strategy, the reach and rigor of the scoping review was increased. However, only records available in English were examined, which may skew the distribution and recognition of studies. The search of only 4 databases was a limitation, as it may have resulted in the omission of relevant literature. However, we included Embase and PubMed, which are considered among the most comprehensive research databases. Digital technology is a rapidly evolving sector, and advancements are made constantly. Additional PCOS digital technologies may have been developed since the time the scoping review was conducted. While wearables, such as fitness trackers, are most likely used by women with PCOS and have potential to aid self-management, no articles were found about wearables specifically for or preferred by women with PCOS.

XSL•FO
RenderX

Conclusion

The findings of this scoping review revealed that digital health technologies, particularly mobile apps, hold promise for supporting women with PCOS. Future research should focus on advancing user-centered design for apps aimed at managing PCOS, with an emphasis on usability and performance expectancy while accounting for sociodemographic differences and disparities. Health care professionals should also be included to corroborate health care information and increase the likelihood of their endorsement to patients with PCOS. There is potential for digital technology to transform PCOS self-management, but further design and development are needed to optimize these technologies for women with PCOS.

Acknowledgments

The authors would like to thank the ACORN Center in the University of South Carolina College of Nursing for its support during this study. The authors appreciate Dr Sara Donevant for her review and support and Amy Edwards for her assistance with the search strategies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File, 15 KB - infodemiology_v5i1e68469_app1.docx]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[DOCX File, 27 KB - infodemiology_v5i1e68469_app2.docx]

References

1. Azziz R. Epidemiology and pathogenesis of polycystic ovary syndrome in adults. : UpToDate; 2020. URL: https://www.uptodateonline.ir/contents/UTD.htm?39/39/40568?source=related_link [accessed 2024-06-05]
2. Engmann L, Jin S, Sun F, et al. Racial and ethnic differences in the polycystic ovary syndrome metabolic phenotype. Am J Obstet Gynecol 2017 May;216(5):493. [doi: 10.1016/j.ajog.2017.01.003] [Medline: 28104402]
3. Zhu T, Cui J, Goodarzi MO. Polycystic ovary syndrome and risk of type 2 diabetes, coronary heart disease, and stroke. Diabetes 2021 Feb;70(2):627-637. [doi: 10.2337/db20-0800] [Medline: 33158931]
4. Behboodi Moghadam Z, Fereidooni B, Saffari M, Montazeri A. Measures of health-related quality of life in PCOS women: a systematic review. Int J Womens Health 2018;10:397-408. [doi: 10.2147/IJWH.S165794] [Medline: 30123008]
5. Teede H, Tay CT, Laven J, et al. International evidence-based guidelines for the assessment and management of polycystic ovary syndrome. 2023 URL: https://www.monash.edu/__data/assets/pdf_file/0003/3379521/Evidence-Based-Guidelines-2023.pdf [accessed 2025-06-09]
6. Kang E, Kim S, Rhee YE, Lee J, Yun YH. Self-management strategies and comorbidities in chronic disease patients: associations with quality of life and depression. Psychol Health Med 2021 Sep;26(8):1031-1043. [doi: 10.1080/13548506.2020.1838585] [Medline: 33095059]
7. Kraus S, Jones P, Kailer N, Weinmann A, Chaparro-Banegas N, Roig-Tierno N. Digital transformation: an overview of the current state of the art of research. Sage Open 2021 Jul;11(3):1-22. [doi: 10.1177/21582440211047576]
8. Bidmon S, Terlutter R. Gender differences in searching for health information on the internet and the virtual patient-physician relationship in Germany: exploratory results on how men and women differ and why. J Med Internet Res 2015 Jun 22;17(6):e156. [doi: 10.2196/jmir.4127] [Medline: 26099325]
9. Ajayi KV, Wachira E, Onyeaka HK, Montour T, Olowolaju S, Garney W. The use of digital health tools for health promotion among women with and without chronic diseases: insights from the 2017-2020 health information national trends survey. JMIR Mhealth Uhealth 2022 Aug 19;10(8):e39520. [doi: 10.2196/39520] [Medline: 35984680]
10. Chou WYS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009 Nov 27;11(4):e48. [doi: 10.2196/jmir.1249] [Medline: 19945947]
11. Ismayilova M, Yaya S. "I'm usually being my own doctor": women's experiences of managing polycystic ovary syndrome in Canada. Int Health 2023 Jan 3;15(1):56-66. [doi: 10.1093/inthealth/ihac028] [Medline: 35567790]
12. Aljumah R, Phillips S, Harper JC. An online survey of postmenopausal women to determine their attitudes and knowledge of the menopause. Post Reprod Health 2023 Jun;29(2):67-84. [doi: 10.1177/20533691231166543] [Medline: 36994487]
13. Huang DR, Goodship A, Webber I, et al. Experience and severity of menopause symptoms and effects on health-seeking behaviours: a cross-sectional online survey of community dwelling adults in the United Kingdom. BMC Womens Health 2023 Jul 14;23(1):373. [doi: 10.1186/s12905-023-02506-w] [Medline: 37452317]

XSL•FO
RenderX

14. Hoyos LR, Putra M, Armstrong AA, et al. Measures of patient dissatisfaction with health care in polycystic ovary syndrome: retrospective analysis. J Med Internet Res 2020 Apr 21;22(4):e16541. [doi: 10.2196/16541] [Medline: 32314967]
15. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
16. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
17. Digital transformation: statistics and facts. Statista. 2024. URL: https://www.statista.com/topics/6778/digital-transformation/ [accessed 2025-03-12]
18. Roe AH, Dokras A. The diagnosis of polycystic ovary syndrome in adolescents. Rev Obstet Gynecol 2011;4(2):45-51. [doi: 10.1542/peds.2015-1430] [Medline: 22102927]
19. Feeney BC, Collins NL. A new look at social support: a theoretical perspective on thriving through relationships. Pers Soc Psychol Rev 2015 May;19(2):113-147. [doi: 10.1177/1088868314544222] [Medline: 25125368]
20. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
21. Shamanna P, Maheshwari A, Keshavamurthy A, et al. Digital twin technology in resolving polycystic ovary syndrome and improving metabolic health: a comprehensive case study. AACE Clin Case Rep 2025;11(1):70-74. [doi: 10.1016/j.aace.2024.11.004] [Medline: 39896947]
22. Jain T, Negris O, Brown D, Galic I, Salimgaraev R, Zhaunova L. Characterization of polycystic ovary syndrome among Flo app users around the world. Reprod Biol Endocrinol 2021 Mar 3;19(1):36. [doi: 10.1186/s12958-021-00719-y] [Medline: 33658043]
23. Alotaibi M, Shaman AA. Enhancing polycystic ovarian syndrome awareness using private social network. Mhealth 2020;6:33. [doi: 10.21037/mhealth.2019.12.02] [Medline: 33209914]
24. Boyle JA, Xu R, Gilbert E, et al. Ask PCOS: identifying need to inform evidence-based app development for polycystic ovary syndrome. Semin Reprod Med 2018 Jan;36(1):59-65. [doi: 10.1055/s-0038-1667187] [Medline: 30189452]
25. Rodriguez EM, Thomas D, Druet A, Vlajic-Wheeler M, Lane KJ, Mahalingaiah S. Identifying women at risk for polycystic ovary syndrome using a mobile health app: virtual tool functionality assessment. JMIR Form Res 2020 May 14;4(5):e15094. [doi: 10.2196/15094] [Medline: 32406861]
26. Choi HK, Lee SH, Yang SY. Development of an integrated mobile application for lifestyle modification in women with polycystic ovarian syndrome. J Clin Nurs 2023 Jan;32(1-2):49-57. [doi: 10.1111/jocn.16253] [Medline: 35150026]
27. Sang M, Wu Q, Tao Y, et al. Usage of mobile health interventions among overweight/obese PCOS patients undergoing assisted reproductive technology treatment during the COVID-19 pandemic. Gynecol Endocrinol 2022 Sep 2;38(9):776-780. [doi: 10.1080/09513590.2022.2112170]
28. Wang L, Liu Y, Tan H, Huang S. Transtheoretical model-based mobile health application for PCOS. Reprod Health 2022 May 12;19(1):117. [doi: 10.1186/s12978-022-01422-w] [Medline: 35549736]
29. Lee H, Lee SH. Effectiveness of an integrated mobile application for lifestyle modifications in overweight women with polycystic ovarian syndrome: a randomized controlled trial. Life (Basel) 2023 Jul 10;13(7):1533. [doi: 10.3390/life13071533] [Medline: 37511908]
30. Ou Y, Wu L, Liu L, Ye X, Chen Y, Yuan H. Effects of an “internet +” multidisciplinary management model oriented by nurse specialists on obesity combined with polycystic ovary syndrome. Clin Exp Obstet Gynecol 2023;50(6):125. [doi: 10.31083/j.ceog5006125]
31. Peven K, Wickham AP, Wilks O, et al. Assessment of a digital symptom checker tool’s accuracy in suggesting reproductive health conditions: clinical vignettes study. JMIR Mhealth Uhealth 2023 Dec 5;11:e46718. [doi: 10.2196/46718] [Medline: 38051574]
32. Stujenske TM, Mu Q, Pérez Capotosto M, Bouchard TP. Survey analysis of quantitative and qualitative menstrual cycle tracking technologies. Medicina (Kaunas) 2023 Aug 22;59(9):1-10. [doi: 10.3390/medicina59091509] [Medline: 37763628]
33. Trépanier LCM, Lamoureux É, Bjornson SE, Mackie C, Alberts NM, Gagnon MM. Smartphone apps for menstrual pain and symptom management: a scoping review. Internet Interv 2023 Mar;31:100605. [doi: 10.1016/j.invent.2023.100605] [Medline: 36761398]
34. Dilimulati D, Shao X, Wang L, et al. Efficacy of WeChat-based digital intervention versus metformin in women with polycystic ovary syndrome: randomized controlled trial. J Med Internet Res 2024 Oct 2;26(26):e55883. [doi: 10.2196/55883] [Medline: 39357046]
35. Hamzehgardeshi Z, Elyasi F, Moosazadeh M, et al. The effect of social network based motivational interviewing on health behaviors among infertile women with polycystic ovary syndrome: a randomized controlled trial. Sci Rep 2024 May 27;14(1):12049. [doi: 10.1038/s41598-024-61161-9] [Medline: 38802409]
36. Mallappa Saroja CS, Hanji Chandrashekar S. Polycystic ovaries: review of medical information on the internet for patients. Arch Gynecol Obstet 2010 May;281(5):839-843. [doi: 10.1007/s00404-010-1378-4] [Medline: 20143080]

XSL•FO
RenderX

37. Holbrey S, Coulson NS. A qualitative investigation of the impact of peer to peer online support for women living with polycystic ovary syndrome. BMC Womens Health 2013 Dec 17;13(51):51. [doi: 10.1186/1472-6874-13-51] [Medline: 24341398]
38. Authier M, Normand C, Jego M, Gaborit B, Boubli L, Courbiere B. Qualitative study of self-reported experiences of infertile women with polycystic ovary syndrome through on-line discussion forums. Ann Endocrinol (Paris) 2020 Oct;81(5):487-492. [doi: 10.1016/j.ando.2020.07.1110] [Medline: 32827451]
39. Malhotra K, Kempegowda P. Increasing trends in online information-seeking and decreasing trends of funded research for polycystic ovary syndrome: an 18-year joinpoint analysis of Google Trends data from 2004 to 2022. Digit Health 2023;9:20552076231220787. [doi: 10.1177/20552076231220787] [Medline: 38152443]
40. Gomula J, Warner M, Blandford A. Women’s use of online health and social media resources to make sense of their polycystic ovary syndrome (PCOS) diagnosis: a qualitative study. BMC Womens Health 2024 Mar 5;24(1):157. [doi: 10.1186/s12905-024-02993-5] [Medline: 38443902]
41. Gour A, Dubey P, Goel A, Halder A. Remote assessment and reinforcement of patient awareness of role of lifestyle modification and treatment adherence in polycystic ovary syndrome using an online video based educational module. J Turk Ger Gynecol Assoc 2022 Mar 8;23(1):1-7. [doi: 10.4274/jtgga.galenos.2021.2021-9-29] [Medline: 35263832]
42. Elhariry M, Malhotra K, Solomon M, Goyal K, Kempegowda P. Top 100 #PCOS influencers: understanding who, why and how online content for PCOS is influenced. Front Endocrinol (Lausanne) 2022;13:1084047. [doi: 10.3389/fendo.2022.1084047] [Medline: 36568090]
43. Clarke S, Jangid G, Nasr S, Atchade A, Moody BL, Narayan G. Polycystic ovary syndrome (PCOS): a cross-sectional observational study analyzing the quality of content on YouTube. Cureus 2023 Sep;15(9):e45354. [doi: 10.7759/cureus.45354] [Medline: 37849574]
44. Horvath A, Chaffin K, Ahmad S, et al. The discourse surrounding polycystic ovary syndrome on TikTok: a social media analysis. Healthcare (Basel) 2024 Nov 12;12(22):2253. [doi: 10.3390/healthcare12222253] [Medline: 39595451]
45. Naroji S, John J, Gomez-Lobo V. Understanding PCOS-related content across social media platforms-a cross-sectional analysis. J Pediatr Adolesc Gynecol 2024 Apr;37(2):142-148. [doi: 10.1016/j.jpag.2023.10.007] [Medline: 37992985]
46. Afaq M, Vashisht K, Al-Dhoon A, et al. Impact of #PCOSweightloss: a global X hashtag analysis study of weight loss narratives in the PCOS community. Digit Health 2025;11:20552076251314100. [doi: 10.1177/20552076251314100] [Medline: 39886094]
47. Jiskoot G, Timman R, Beerthuizen A, Dietz de Loos A, Busschbach J, Laven J. Weight reduction through a cognitive behavioral therapy lifestyle intervention in PCOS: the primary outcome of a randomized controlled trial. Obesity (Silver Spring) 2020 Nov;28(11):2134-2141. [doi: 10.1002/oby.22980] [Medline: 32969197]
48. Dietz de Loos ALP, Jiskoot G, Timman R, Beerthuizen A, Busschbach JJV, Laven JSE. Improvements in PCOS characteristics and phenotype severity during a randomized controlled lifestyle intervention. Reprod Biomed Online 2021 Aug;43(2):298-309. [doi: 10.1016/j.rbmo.2021.05.008] [Medline: 34238659]
49. Zigarelli A, Jia Z, Lee H. Machine-aided self-diagnostic prediction models for polycystic ovary syndrome: observational study. JMIR Form Res 2022 Mar 15;6(3):e29967. [doi: 10.2196/29967] [Medline: 35289757]
50. Karia A, Poojary A, Tiwari A, Sequeira L, Sokhi MK. BeRedy (period tracker & PCOS diagnosis). Presented at: 2023 International Conference on Communication System, Computing and IT Applications (CSCITA); Mar 31 to Apr 1, 2023. [doi: 10.1109/CSCITA55725.2023.10104755]
51. Devranoglu B, Gurbuz T, Gokmen O. ChatGPT’s efficacy in queries regarding polycystic ovary syndrome and treatment strategies for women experiencing infertility. Diagnostics (Basel) 2024 May 22;14(11):1082. [doi: 10.3390/diagnostics14111082] [Medline: 38893609]
52. Ulug E, Gunesli I, Acıkgoz Pinar A, Yildiz BO. Evaluating reliability, quality, and readability of ChatGPT’s nutritional recommendations for women with polycystic ovary syndrome. Nutr Res 2025 Jan;133:46-53. [doi: 10.1016/j.nutres.2024.11.005] [Medline: 39673813]
53. Percy C, Turner A, Orr C. Developing a novel web-based self-management support intervention for polycystic ovary syndrome: mixed methods study with patients and health care professionals. JMIR Form Res 2024 Mar 7;8:e52427. [doi: 10.2196/52427] [Medline: 38451567]
54. Secinaro S, Calandra D, Secinaro A, Muthurangu V, Biancone P. The role of artificial intelligence in healthcare: a structured literature review. BMC Med Inform Decis Mak 2021 Apr 10;21(1):125. [doi: 10.1186/s12911-021-01488-9] [Medline: 33836752]
55. Gibson-Helm M, Teede H, Dunaif A, Dokras A. Delayed diagnosis and a lack of information associated with dissatisfaction in women with polycystic ovary syndrome. J Clin Endocrinol Metab 2017 Feb 1;102(2):604-612. [doi: 10.1210/jc.2016-2963] [Medline: 27906550]
56. X (formerly Twitter): statistics and facts. Statista. URL: https://www.statista.com/topics/737/twitter/#statisticChapter [accessed 2025-03-13]
57. Singh S. ChatGPT statistics (March 2025): number of users & queries. Demandsage. URL: https://www.demandsage.com/chatgpt-statistics/ [accessed 2025-03-13]

XSL•FO
RenderX

58. Byrne E, Kearney J, MacEvilly C. The role of influencer marketing and social influencers in public health. Proc Nutr Soc 2017;76(OCE3). [doi: 10.1017/S0029665117001768]
59. Powell J, Pring T. The impact of social media influencers on health outcomes: systematic review. Soc Sci Med 2024 Jan;340:116472. [doi: 10.1016/j.socscimed.2023.116472] [Medline: 38070305]
60. Bocci Benucci S, Fioravanti G, Silvestro V, et al. The impact of following Instagram influencers on women's body dissatisfaction and eating disorder symptoms. Nutrients 2024 Aug 16;16(16):2730. [doi: 10.3390/nu16162730] [Medline: 39203866]
61. Davitadze M, Malhotra K, Khalil H, et al. Body image concerns in women with polycystic ovary syndrome: a systematic review and meta-analysis. Eur J Endocrinol 2023 Aug 2;189(2):R1-R9. [doi: 10.1093/ejendo/lvad110] [Medline: 37619990]
62. Graca S, Alloh F, Lagojda L, et al. Polycystic ovary syndrome and the internet of things: a scoping review. Healthcare (Basel) 2024 Aug 21;12(16):1671. [doi: 10.3390/healthcare12161671] [Medline: 39201229]
63. Epstein DA, Lee NB, Kang JH, et al. Examining menstrual tracking to inform the design of personal informatics tools. Proc SIGCHI Conf Hum Factor Comput Syst 2017 May 2;2017:6876-6888. [doi: 10.1145/3025453.3025635] [Medline: 28516176]
64. Brennan L, Teede H, Skouteris H, Linardon J, Hill B, Moran L. Lifestyle and behavioral management of polycystic ovary syndrome. J Womens Health (Larchmt) 2017 Aug;26(8):836-848. [doi: 10.1089/jwh.2016.5792] [Medline: 28570835]
65. Saslow LR, Aikens JE. Lifestyle interventions for polycystic cvary syndrome: cross-sectional survey to assess women's treatment and outcome preferences. JMIR Form Res 2020 Sep 2;4(9):e17126. [doi: 10.2196/17126] [Medline: 32876573]
66. Wright PJ, Raynor PA, Bowers D, et al. Leveraging digital technology for social connectedness among adults with chronic conditions: a systematic review. Digit Health 2023;9(1-25):20552076231204746. [doi: 10.1177/20552076231204746] [Medline: 37799504]
67. Chidambaram S, Jain B, Jain U, et al. An introduction to digital determinants of health. PLOS Digit Health 2024 Jan;3(1):e0000346. [doi: 10.1371/journal.pdig.0000346] [Medline: 38175828]
68. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
69. Elkasabi M, Khan A. The evolution of mobile phone surveys in low- and middle-income countries: a study of coverage structure. Int J Public Opin Res 2023 Dec 1;35(4):edad031. [doi: 10.1093/ijpor/edad031]

Abbreviations

AI: artificial intelligence
JBI: Joanna Briggs Institute
PCOS: polycystic ovary syndrome
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by T Mackey; submitted 06.11.24; peer-reviewed by A Delaforce, L Perrier, R Carron, SF Albasri, T Wu; revised version received 23.03.25; accepted 21.04.25; published 12.06.25.

Please cite as:
Wright PJ, Burts C, Harmon C, Corbett CF
Availability and Use of Digital Technology Among Women With Polycystic Ovary Syndrome: Scoping Review
JMIR Infodemiology 2025;5:e68469
URL: https://infodemiology.jmir.org/2025/1/e68469
doi:10.2196/68469

© Pamela J Wright, Charlotte Burts, Carolyn Harmon, Cynthia F Corbett. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 12.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Harnessing Facebook to Investigate Real-World Mentions of Adverse Events of Glucagon-Like Peptide-1 Receptor Agonist (GLP-1 RA) Medications: Observational Study of Facebook Posts From 2022 to 2024

Amrutha S Alibilli[1*]; Vidur Jain[1*]; Heran Mane[1*], BS; Xiaohe Yue[1], MS; Alexandria Ratzki-Leewing[2,3*], MSc, PhD; Junaid S Merchant[1*], MS, PhD; Shaniece Criss[4*], MPH, MPA, SCD; Quynh C Nguyen[5*], MSPH, PhD; Rozalina G McCoy[2,3,6*], MS, MD; Thu T Nguyen[1,3*], MSPH, SCD

[1]Department of Epidemiology and Biostatistics, University of Maryland, College Park, 4254 Stadium Drive, College Park, United States

[2]Division of Gerontology, Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, United States

[3]University of Maryland Institute for Health Computing, North Bethesda, MD, United States

[4]Department of Health Sciences, Furman University, Greenville, SC, United States

[5]National Institute of Nursing Research, National Institutes of Health, Bethesda, MD, United States

[6]Division of Endocrinology, Diabetes, and Nutrition, Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Thu T Nguyen, MSPH, SCD
Department of Epidemiology and Biostatistics, University of Maryland, College Park, 4254 Stadium Drive, College Park, United States

Abstract

Background: In recent years, there has been a dramatic increase in the popularity and use of glucagon-like peptide-1 receptor agonists (GLP-1 RAs) for weight loss. As such, it is essential to understand users' real-world discussions of short-term, long-term, and co-occurrent adverse events associated with currently used GLP-1 RA medications.

Objective: This study aims to quantitatively analyze temporal and co-occurrent GLP-1 RA adverse event trends through discussions of GLP-1 RA weight loss medications on Facebook from 2022 to 2024.

Methods: We collected 64,202 Facebook posts (59,293 posts after removing duplicate posts) from January 1, 2022, to May 31, 2024, through CrowdTangle, a public insights tool from Meta. Using English language social media posts from the United States, we examined discussions of adverse event mentions for posts referencing 7 GLP-1 RA weight loss product categories (ie, semaglutide, Ozempic, Wegovy, tirzepatide, Mounjaro, Zepbound, and GLP-1 RA as a class). All analyses were conducted using Python (version 3; Python Software Foundation) in a Google Colab environment.

Results: Temporal time series analysis revealed that the GLP-1 RAs' adverse event mentions on social media aligned with several key events: the Food and Drug Administration's approval of Wegovy for pediatric weight management in December 2022, increased media coverage in August 2023, celebrity endorsement in December 2023, and Medicare Part D coverage expansion for weight loss medications in March 2024. Gastrointestinal (GI)–related adverse events (general term) were most prevalent for posts mentioning the GLP-1 RA class (210/4885, 4.30%) and Mounjaro (241/4031, 5.98%). In contrast, the most prevalent adverse event mentions noted for tirzepatide were headache (78/4202, 1.86%) and joint pain (71/4202, 1.69%). Hypertension (13/1769, 0.73%) was frequently mentioned in Zepbound posts, while pancreatitis was commonly associated with Mounjaro posts (44/4031, 1.08%), and 2.85% (139/4885) of posts broadly referring to the GLP-1 RA class. Furthermore, an integrated node network analysis revealed 3 distinct GLP-1 RA adverse events–mentioned clusters: cluster 1 (purple) contained allergies, anxiety, depression, chronic obstructive pulmonary disease, fatigue, fever, hypertension, indigestion, insomnia, gastroesophageal reflux disease, hives, swelling, restlessness, and seizures. Cluster 2 (pink) contained constipation, dehydration, headache, diarrhea, dizziness, hypoglycemia, sweating, and jaundice. Cluster 3 (brown) contained GI symptoms, such as nausea, pancreatitis, rash, and vomiting. The GI symptoms, such as nausea, vomiting, pancreatitis, diarrhea, and indigestion, were strongly associated together (≥100 co-occurrence mentions), while the mentioned neurological symptoms, such as anxiety, depression, and insomnia, were highly correlated with each other (50 - 100 co-occurrence mentions).

Conclusions: This social media study highlights the adverse event mention patterns for posts referencing GLP-1 RA medications. While further research is needed to rigorously examine and validate these findings, this study demonstrates the importance of

XSL•FO
RenderX

monitoring social media discussions to predict novel, underreported, or rare drug adverse events, thereby improving patient care, clinical research, and health policy interventions.

(JMIR Infodemiology 2025;5:e73619) doi:10.2196/73619

KEYWORDS

social media; semaglutide; tirzepatide; Ozempic; Wegovy; Mounjaro; Zepbound; GLP-1 RAs; adverse events; obesity management; weight loss medication

Introduction

In recent years, there has been a dramatic increase in the popularity and use of glucagon-like peptide-1 receptor agonists (GLP-1 RAs) for the treatment of type 2 diabetes and obesity [1]. The Food and Drug Administration (FDA) first approved Ozempic (semaglutide) and Mounjaro (tirzepatide) for the management of type 2 diabetes in December 2017 and May 2022, respectively. Wegovy (semaglutide) and Zepbound (tirzepatide) were then approved for the management of obesity and overweight in June 2021 and November 2023, respectively. Given that nearly 38 million US adults are living with diabetes and more than 100 million have obesity, the increased demand for these drugs has contributed to an alarming global shortage of medications for people in need [2,3]. Furthermore, the low availability, high costs, and inconsistent and inadequate insurance coverage of GLP-1 RAs have contributed to the rise of compounded versions of GLP-1 RAs. However, while often lower in cost and more readily available (including without a prescription), compounded GLP-1 RAs have unknown dosage, safety, and efficacy profiles [4,5].

An important consideration for GLP-1 RAs' use is their adverse event profile, including significant GI discomfort and other potential physical and psychological symptoms [6]. GLP-1 RAs' functions include stimulating glucose-dependent insulin secretion, inhibiting glucagon secretion, and delaying gastric emptying [7]. These mechanisms of action contribute to reducing blood glucose levels and facilitating weight loss. However, by reducing patients' appetite and nutritional intake, GLP-1 RAs may lead to malnutrition, dehydration, and rapid loss of both fat and lean (ie, muscle and skeletal) body mass [8]. Previous clinical studies have discovered prevalent adverse events associated with GLP-1 RAs, including GI symptoms, such as nausea, diarrhea, vomiting, and reduced appetite [9]. Others revealed notable mental health adverse events, including insomnia, anxiety, and depression [6,10]. Importantly, the individual experience with medications—concerning effectiveness, tolerability, and safety—may differ among those taking medications in the real world compared with those enrolled and studied in clinical trials, and some events may be too infrequent to be detected in clinical trials. Moreover, rare adverse events may not be ascertained in clinical trial protocols. Therefore, further research is necessary to understand the complex adverse events of GLP-1 RAs on physical and mental health conditions experienced and reported by individuals consuming these medications in contemporary practice.

Since the most potent GLP-1 RAs (including the dual glucose-dependent insulinotropic polypeptide and GLP-1 RA, tirzepatide) were introduced into practice recently, it is important to monitor for emerging adverse events of these medications [11]. Clinical studies identify many of GLP-1 RAs' adverse events, but they cannot capture the totality of adverse events experienced by and important to patients. For instance, clinical studies' sample size and duration of follow-up are often limited, reducing the likelihood of detecting rare or long-term adverse events. Individuals taking GLP-1 RAs also differ from those included in clinical trials as they likely have different clinical profiles, lifestyles, environmental contexts, and dietary and physical activity cointerventions. Furthermore, the robustness of education and clinical monitoring available before and while taking GLP-1 RA medications contributes to individuals' perceptions of the medications. It is also essential to monitor adverse events of compounded GLP-1 RA products, whose safety profiles are expected to differ from the FDA-approved products and are largely unknown at this time.

Social media platforms provide a relatively untapped, extensive source of public discussion from which health professionals can develop targeted interventions. A meta-analysis study analyzed social media pharmacovigilance studies and FDA reports and found that more adverse event discussions on social media corresponded to a higher hazard and faster drug product recall [12]. Therefore, social media may reveal invaluable insights regarding mentioned adverse events that may not be presented in other clinical or observational studies.

Building upon prior findings, this social media–based study aims to use advanced data extraction techniques and conduct quantitative analyses to investigate mentioned GLP-1 RA adverse event phenotypes. Our approach harnesses social media data as a complement to traditional clinical and pharmacovigilance data sources to identify user-reported mentions of adverse events for posts referencing GLP-1 RAs. Additionally, we examine temporal trends of social media GLP-1 RA discussions to identify factors that drive their popularity from Oprah Winfrey's endorsement of GLP-1 RAs [13] to Medicare's coverage expansion [14]. Finally, we identify clusters of co-mentioned adverse events, introducing a novel way to characterize adverse event interconnections and risk profiles. Quantitative temporal analyses are vital to highlight possible unreported GLP-1 RA adverse events, providing timely information to inform community-driven interventions to promote public health awareness about GLP-1 RA use.

Methods

Data Source

We collected Facebook posts through CrowdTangle, a public insights tool from Meta [15] and engagement metrics (eg, views and post likes). CrowdTangle has been discontinued. Currently

XSL•FO
RenderX

available tools for researchers include Meta Content Library and Content Library API [16]. CrowdTangle provides data from public accounts, including 7 million Facebook pages, groups, and verified profiles. Data collection can be refined by specifying various parameters such as time frame, specific pages, groups or profiles, keywords or tags, language, geolocation, and presence of memes. The returned data include posts, image URLs, image caption text, date of the post, and engagement metrics (eg, views and post likes).

Data Preparation (Sample and Inclusion Criteria)

Using CrowdTangle, we obtained and processed Facebook English language posts by US-based users between January 1, 2022, and May 31, 2024, that mentioned weight loss medications, specifically "Semaglutide/semaglutide," "Ozempic/Ozempic/ozempy," "Wegovy/wegovy," "Tirzepatide/tirzepatide," "Mounjaro/mounjaro," "Zepbound/zepbound," and "GLP-1/GLP-1s/GLP-1 agonist/GLP-1 RAs." To limit our collection to US-based discussions on weight loss medications, we leveraged CrowdTangle's "Page Admin Country" filtering parameter. This field identifies the country of the primary administrators of Facebook Pages. Because the filter was not available for Facebook Groups, we collected posts derived solely from Facebook Pages where the United States was listed as the "Page Admin Country. Among these posts, we identified those that mentioned any potential short-term or long-term adverse events, defined as any undesirable experience associated with a drug or treatment [17] (Table 1). The list of adverse events was determined based on past literature, clinical expertise, and the SIDER database version 4.1, which cataloged 5868 documented adverse events for 1430 commercially available drugs [18]. To examine the magnitude of adverse events mentioned for weight loss medications, duplicate posts with the same message were filtered out.

XSL•FO
RenderX

Table . List of adverse events associated with different weight loss medications.

Adverse event category	Adverse event
Gastrointestinal	• Gastrointestinal (general term) • Nausea • Vomiting • Diarrhea • Constipation • Abdominal pain • Loss of appetite • Indigestion • Jaundice • Pancreatitis • Gallbladder issues • Gastroesophageal reflux disease • Liver damage
Neuromuscular	• Headache • Dizziness • Seizures • Muscle cramps • Tingling sensation • Joint pain • Muscle weakness • Back pain
Cardiovascular and respiratory	• Persistent cough • Swelling • Increased heart rate • Shortness of breath • Hypertension • Chronic obstructive pulmonary disease • Blood clots • Heart palpitations
Dermatologic and immunologic	• Rash • Sweating • Hives • Sore throat • Fever • Dry mouth • Hair loss • Allergies
Endocrine and metabolic	• Elevated blood sugar levels • Irregular menstrual cycle • Thyroid tumor • Erectile dysfunction • Kidney damage • Hypoglycemia
Mental health and behavioral	• Depression • Anxiety • Mood swings • Insomnia • Restlessness
General symptoms	• Vision changes • Fatigue • Dehydration

Statistical Analysis

Descriptive analyses were conducted using Python (version 3) in a Google Colab environment. For each weight loss medication, we calculated the frequency and percentage of adverse events mentioned—calculated as the total mention of a certain adverse events over the total number of posts for a particular weight loss medication. To examine temporal trends, we recorded the total frequency of adverse events by calendar month. We plotted frequency and percentage of adverse events

XSL•FO
RenderX

mentions from January 2022 to May 2024 to assess for changes in the mentions of adverse events for four specific events: (1) FDA approval of Wegovy for weight management in children aged 12 - 17 years in December 2022 [19], (2) increased media coverage of GLP-1 RA medications on TikTok and other platforms in August of 2023 [20], (3) Oprah Winfrey's endorsement of weight loss medications in December 2023 [14], and (4) expanded FDA indication for Wegovy for the prevention of cardiovascular events and mortality in adults with obesity and cardiovascular disease in March 2024, paving the way for Medicare coverage of weight management medications for reduction of cardiovascular events [21].

To investigate the co-occurrence of adverse events mentioned, a cluster analysis was conducted. Clusters were formed based on the Louvain community detection algorithm using the Python library (python-louvain) [22], which can detect communities (via modularity optimization) and subsequently improve the quality of community structures [23]. A spring layout algorithm for positioning with k=7.8 and the number of iterations=100 was applied to arrange the nodes to effectively repel them to minimize overlap and enhance the clarity of the node-network structure [24]. Using this method, we constructed a network of adverse events mentioned where 2 nodes mentioned in the same post are connected by an edge. A cluster was based on the co-occurrence of mentioned adverse events. The size of each node reflects the frequency of the mentions, while the width of each edge (thickness of the lines) corresponds to its weight, indicating the strength of co-occurrence between comentioned adverse events.

Ethical Considerations

This study was determined to be not human participant research by the University of Maryland College Park Institutional Review Board (2072551-1). In addition, the social media posts were anonymized, upholding user privacy.

Results

Descriptive Patterns of Adverse Events

Table S1 in Multimedia Appendix 1 shows the FDA-reported adverse events of the drugs examined in this study with their brand name, generic name, and approved indication for the drug. We collected 64,202 (59,293 posts after removing duplicates) posts that included a discussion of the 7 weight loss products. Of these posts, 28.9% (17,146/59,293) on semaglutide, 31.6% (18,733/59,293) on Ozempic, 14.4% (8527/59,293) on Wegovy, 7.1% (4202/59,293) on tirzepatide, 6.8% (4031/59,293) on Mounjaro, 3.0% (1769/59,293) on Zepbound, and 8.5% (4885/59,293) referenced the GLP-1 RA class without mentioning a specific drug. Among all the posts, 13.8% (8171/59,293) mentioned adverse events (Table 2). Table 2 shows the counts and percentages of all posts about each medication that mentions specific adverse events. For example, nausea is mentioned in 3.23% (185/4885) of posts referencing GLP-1. Within each column of the different medications, adverse event mentions are depicted in decreasing frequency within broad categories of adverse events such as "Gastrointestinal" or "Mental Health and Behavior," where the highest percentage within the category is near the top, and the lowest percentage is near the bottom. The percentages in boldface represent the highest frequency within each column.

XSL•FO
RenderX

Table . Frequency and percentage of adverse event mentions by weight loss medications[a].

	Adverse events	GLP-1[b]	Semaglutide	Ozempic	Wegovy	Mounjaro	Tirzepatide	Zepbound
	Total posts	4885	17,146	18,733	8527	4031	4202	1769
Category	Total number of adverse event mentions within drug posts	1138	1486	2579	1495	868	436	169
Gastrointestinal	Gastrointestinal (general term), n (%)	**210 (4.30)**[c]	95 (0.55)	**488 (2.61)**	**296 (3.47)**	**241 (5.98)**	19 (0.45)	16 (0.90)
	Nausea, n (%)	158 (3.23)	**197 (1.15)**	317 (1.69)	175 (2.05)	86 (2.13)	48 (1.14)	**23 (1.30)**
	Vomiting, n (%)	113 (2.31)	99 (0.58)	308 (1.64)	228 (2.67)	153 (3.80)	17 (0.40)	11 (0.62)
	Pancreatitis, n (%)	139 (2.85)	52 (0.30)	165 (0.88)	91 (1.07)	44 (1.09)	12 (0.29)	7 (0.40)
	Constipation, n (%)	59 (1.21)	64 (0.37)	143 (0.76)	82 (0.96%)	38 (0.94)	26 (0.62)	16 (0.90)
	Diarrhea, n (%)	64 (1.31)	59 (0.34)	155 (0.83)	83 (0.97)	36 (0.89)	19 (0.45)	12 (0.68)
	Abdominal pain, n (%)	51 (1.04)	22 (0.13)	112 (0.60)	34 (0.40)	19 (0.47)	6 (0.14)	2 (0.11)
	Gallbladder issues, n (%)	13 (0.27)	4 (0.02)	61 (0.33)	36 (0.42)	3 (0.07)	1 (0.02)	1 (0.06)
	Indigestion, n (%)	2 (0.04)	5 (0.03)	6 (0.03)	0 (0.00)	5 (0.12)	0 (0.00)	0 (0.00)
	Loss of appetite, n (%)	4 (0.08)	2 (0.01)	5 (0.03)	4 (0.05)	0 (0.00)	0 (0.00)	0 (0.00)
	GERD[d], n (%)	5 (0.10)	2 (0.01)	4 (0.02)	1 (0.01)	1 (0.02)	0 (0.00)	0 (0.00)
	Liver damage, n (%)	0 (0.00)	1 (0.01)	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
	Jaundice, n (%)	0 (0.00)	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Mental health and behavior	Depression, n (%)	42 (0.86)	151 (0.88)	165 (0.88)	134 (1.57)	19 (0.47)	13 (0.40)	9 (0.62)
	Anxiety, n (%)	39 (0.80)	59 (0.34)	70 (0.37)	36 (0.42)	19 (0.47)	13 (0.31)	9 (0.51)
	Mood swings, n (%)	3 (0.06)	24 (0.14)	15 (0.08)	14 (0.16)	2 (0.05)	3 (0.07)	0 (0.00)
	Insomnia, n (%)	4 (0.08)	7 (0.04)	6 (0.03)	1 (0.01)	5 (0.12)	2 (0.05)	0 (0.00)
	Restlessness, n (%)	1 (0.02)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)

XSL•FO
RenderX

	Adverse events	GLP-1[b]	Semaglutide	Ozempic	Wegovy	Mounjaro	Tirzepatide	Zepbound
Dermatologic and immunologic	Hair loss, n (%)	48 (0.98)	50 (0.29)	153 (0.82)	77 (0.90)	80 (1.98)	15 (0.36)	17 (0.96)
	Rash, n (%)	8 (0.16)	51 (0.30)	25 (0.13)	5 (0.06)	6 (0.15)	4 (0.10)	1 (0.06)
	Fever, n (%)	2 (0.04)	5 (0.03)	7 (0.04)	5 (0.06)	3 (0.07)	1 (0.02)	2 (0.11)
	Sweating, n (%)	3 (0.06	9 (0.05)	13 (0.07)	4 (0.05)	2 (0.05)	1 (0.02)	0 (0.00)
	Allergies, n (%)	1 (0.02)	14 (0.08)	7 (0.04)	3 (0.04)	2 (0.05)	1 (0.02)	0 (0.00)
	Hives, n (%)	6 (0.12)	1 (0.01)	7 (0.04)	1 (0.01)	1 (0.02)	1 (0.02)	0 (0.00)
	Sore throat, n (%)	2 (0.04)	0 (0.00)	3 (0.02)	3 (0.04)	0 (0.00)	0 (0.00)	0 (0.00)
	Dry mouth, n (%)	0 (0.00)	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	1 (0.02)	0 (0.00)
Neuromuscular	Headache, n (%)	21 (0.43)	111 (0.65)	22 (0.12)	13 (0.15)	**7 (0.17)**	**78 (1.86)**	3 (0.17)
	Joint pain, n (%)	10 (0.20)	97 (0.57)	4 (0.02)	2 (0.02)	4 (0.10)	71 (1.69)	1 (0.06)
	Dizziness, n (%)	8 (0.16)	15 (0.09)	19 (0.10)	16 (0.19)	2 (0.05)	0 (0.00)	1 (0.06)
	Muscle weakness, n (%)	2 (0.04)	2 (0.01)	2 (0.01)	2 (0.02)	2 (0.05)	1 (0.02)	0 (0.00)
	Back pain, n (%)	0 (0.00)	7 (0.04)	1 (0.01)	0 (0.00)	2 (0.05)	0 (0.00)	0 (0.00)
	Seizures, n (%)	0 (0.00)	4 (0.02)	6 (0.03)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
	Muscle cramps, n (%)	0 (0.00)	3 (0.02)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
	Tingling sensation, n (%)	0 (0.00)	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
General symptom	Fatigue, n (%)	29 (0.59)	111 (0.65)	65 (0.35)	47 (0.55)	21 (0.52)	33 (0.79)	18 (1.02)
	Vision changes, n (%)	4 (0.08)	3 (0.02)	4 (0.02)	3 (0.04)	3 (0.07)	2 (0.05)	2 (0.11)
	Dehydration, n (%)	4 (0.08)	12 (0.07)	8 (0.04)	3 (0.04)	2 (0.05)	0 (0.00)	0 (0.00)
Cardiovascular and respiratory	Hypertension, n (%)	32 (0.66)	53 (0.31)	24 (0.13)	25 (0.29)	11 (0.27)	21 (0.50)	13 (0.73)
	Persistent cough, n (%)	0 (0.00)	0 (0.00)	22 (0.12)	0 (0.00)	22 (0.55)	0 (0.00)	0 (0.00)
	Swelling, n (%)	1 (0.02)	24 (0.14)	10 (0.05)	7 (0.08)	2 (0.05)	4 (0.10)	0 (0.00)
	COPD[e], n (%)	4 (0.08)	0 (0.00)	7 (0.04)	0 (0.00)	1 (0.02)	0 (0.00)	0 (0.00)
	Blood clots, n (%)	4 (0.08)	2 (0.01)	2 (0.01)	1 (0.01)	1 (0.02)	1 (0.02)	0 (0.00)
	Increased heart rate, n (%)	0 (0.00)	2 (0.01)	2 (0.01)	3 (0.04)	1 (0.02)	1 (0.02)	0 (0.00)
	Heart palpitations, n (%)	0 (0.00)	3 (0.02)	2 (0.01)	0 (0.00)	2 (0.05)	1 (0.02)	0 (0.00)
	Shortness of breath, n (%)	0 (0.00)	1 (0.01)	0 (0.00)	4 (0.05)	0 (0.00)	0 (0.00)	0 (0.00)

XSL•FO
RenderX

	Adverse events	GLP-1[b]	Semaglutide	Ozempic	Wegovy	Mounjaro	Tirzepatide	Zepbound
Endocrine and metabolic	Hypoglycemia, n (%)	29 (0.59)	31 (0.18)	80 (0.43)	33 (0.39)	9 (0.22)	7 (0.17)	2 (0.11)
	Thyroid tumor, n (%)	8 (0.16)	15 (0.09)	54 (0.29)	20 (0.23)	9 (0.22)	3 (0.07)	1 (0.06)
	Erectile dysfunction, n (%)	1 (0.02)	10 (0.06)	2 (0.01)	2 (0.02)	0 (0.00)	5 (0.12)	0 (0.00)
	Kidney damage, n (%)	4 (0.08)	4 (0.02)	5 (0.03)	1 (0.01)	2 (0.05)	1 (0.02)	0 (0.00)
	Elevated blood sugar levels, n (%)	0 (0.00)	0 (0.00)	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
	Irregular menstrual cycles, n (%)	0 (0.00)	0 (0.00)	1 (0.01)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)

[a]Percentage = (Number of posts mentioning the adverse event for a medication/Total number of posts about the medication) × 100. The values in boldface indicate the most commonly mentioned adverse event for the particular GLP-1 RA category.

[b]GLP-1: glucagon-like peptide-1.

[c]The values in boldface indicate the most commonly mentioned adverse event for the particular GLP-1 RA category.

[d]GERD: gastroesophageal reflux disease.

[e]COPD: chronic obstructive pulmonary disease.

Distinct GLP-RA adverse events mention patterns emerged. General (nonspecified) GI concerns were among the most frequently reported, representing 4.3% (210/4885) of general GLP-1 RA posts and 6.0% (241/4031) and 3.5% (296/8527) of posts referencing Mounjaro and Wegovy, respectively (Table 2). A lower percentage of GI issues were reported in posts mentioning Ozempic (488/18,733, 2.6%), semaglutide (95/17,146, 0.5%), tirzepatide (19/4202, 0.5%), and Zepbound (16/1769, 0.9%). Nausea was among the most commonly reported specific symptoms, appearing in 3.2% (158/4885) of posts about general GLP-1 RA use, 2.1% (86/4031) of posts about Mounjaro and Wegovy each, 1.7% (317/18,733) about Ozempic, and 1.3% (23/1769) of posts about Zepbound; it was mentioned in approximately 1% (197/17,146) of semaglutide and tirzepatide posts. Vomiting was also mentioned in 3.8% (153/4031) of Mounjaro posts, 2.7% (228/8527) of Wegovy posts, 2.3% (113/4885) of general GLP-1 RA posts, 1.6% of Ozempic posts (308/18,733), and 0.6% of semaglutide (99/17,146) and Zepbound (11/1769) posts each. Pancreatitis was a concern that appeared in 2.9% (139/4885) of general GLP-1 RAs posts.

Several other adverse events were referenced in posts. Depression was noted across various GLP-1 RAs, although it was less prevalent than general GI concerns or nausea and vomiting (Table 2). Wegovy posts mentioned depression 1.6% (134/8527) of the time, while general GLP-1 RAs (42/4885), Ozempic (165/18,733), and semaglutide (151/17,146) posts discussed depression 0.9% of the time. Depression was less often noted in Zepbound (9/1769, 0.6%), Mounjaro (19/4031, 0.5%), and tirzepatide (13/4202, 0.4%)-related posts. Hair loss was more prominent in Mounjaro posts (80/4031, 2.0%) than in general GLP-1 RA (48/4885) or Zepbound (17/1769) posts, which were discussed in 1.0% of posts each. Headaches (78/4202) and joint pain (71/4202) appeared in 2.0% of tirzepatide posts, compared with 0.6% (111/17,146) and 0.7% (97/17,146) of semaglutide posts. Fatigue was also reported across all weight loss medications. Table S2 in Multimedia Appendix 1 shows the percentage of posts that mention each adverse event by medication. For example, 1.69% (317/18,733) of posts mentioning Ozempic adverse events include "nausea." Figure S1 in Multimedia Appendix 1 displays the prominent adverse events frequencies (total≥75) for the weight loss medications.

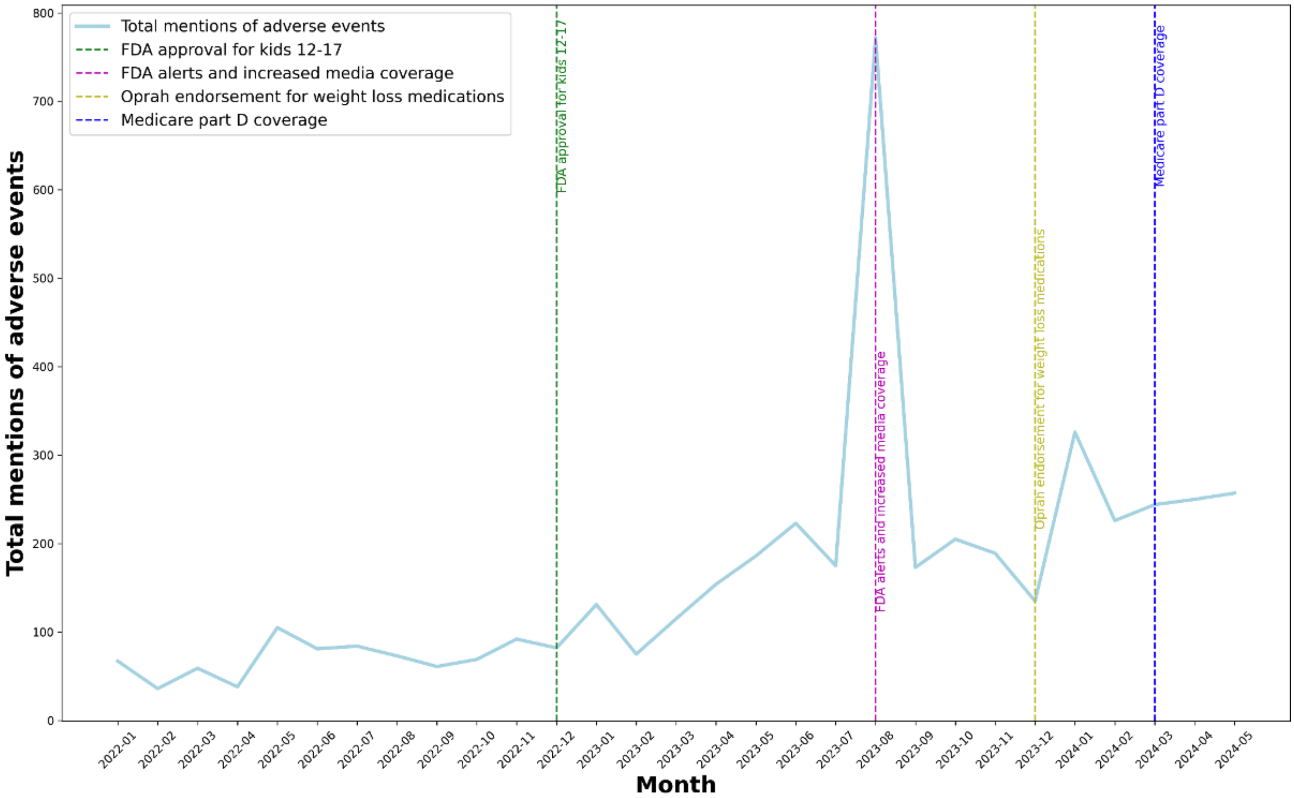
Temporal Trends With Key Events

We examined the potential impact of several key events occurring during the temporal analysis period on observed side effect mention trends (Figure 1). In December 2022, the FDA approved Wegovy for the treatment of obesity in children aged 12 - 17 years [19]. This event was followed by increased media coverage in August 2023 that highlighted potential adverse event mentions associated with these drugs and interactions with other medications, coinciding with a surge in media coverage across social media platforms including Reddit, TikTok, Instagram, Facebook, and Twitter [19]. At the end of 2023, Oprah Winfrey publicly endorsed the use of weight loss medications (December 14, 2023) [14], while Medicare Part D expanded its coverage for weight loss medications starting in March 2024 [21]. Figure S2 and Table S3 in Multimedia Appendix 1 show the percentage of adverse events mentioned to total drug mentions over time. Of note, although the number of adverse events mentioned has increased over time, the percentage of posts that mention adverse events has decreased over time (Figure S2 in Multimedia Appendix 1), indicating

XSL•FO
RenderX

changes in the composition of these discussions from adverse events to other topics related to these medications.

Figure 1. Temporal time series analysis of mention of all adverse events from January 2022 to May 2024. FDA: Food and Drug Administration.

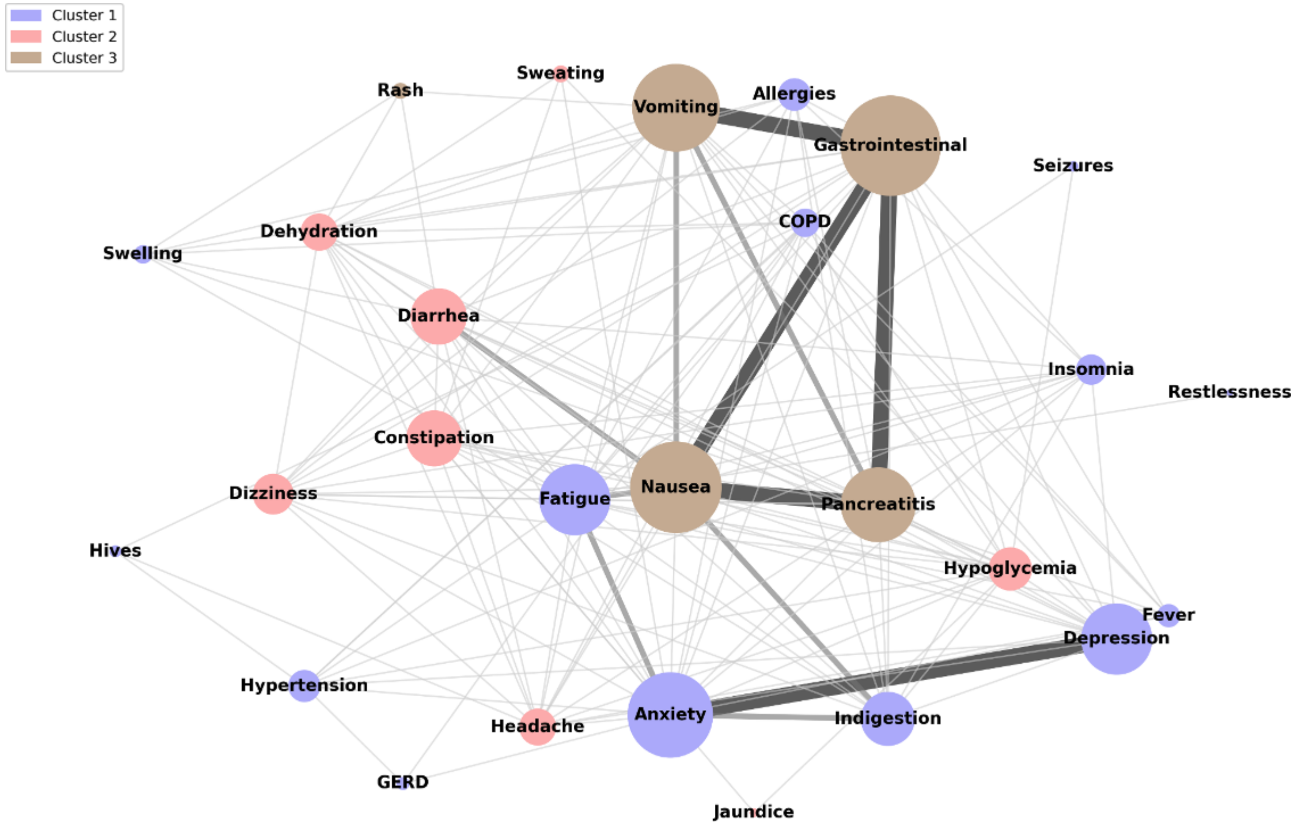


Co-Occurrence Network Trends

In a network analysis of noted adverse events for posts referencing these medications, a total of 2618 adverse event co-occurrence interactions and 27 nodes (circles representing adverse events) were identified from the dataset. As shown in Figure 2, 3 distinct communities were formed, each representing a grouping of adverse events mentioned with similar characteristics based on their high-frequency co-occurrence patterns. Cluster 1 (purple) contains allergies, anxiety, depression, chronic obstructive pulmonary disease, fatigue, fever, hypertension, indigestion, insomnia, gastroesophageal reflux disease, hives, swelling, restlessness, and seizures. Cluster 2 (pink) contains constipation, dehydration, headache, diarrhea, dizziness, hypoglycemia, sweating, and jaundice. Cluster 3 (brown) contains GI symptoms, such as nausea, pancreatitis, rash, and vomiting. Interactions with a co-occurrence count of ≥100 were visualized with thick lines, while those with counts between 50 and 100 were shown with a moderately thick line, and interactions with counts <50 were represented by the thinnest lines. As depicted in Figure 2, nausea, vomiting, pancreatitis, and GI distress are connected by thick lines, reflecting their strong co-occurrence. Anxiety and depression are also highly correlated. Next, the intermediately thick lines reveal moderately strong associations between fatigue, anxiety, indigestion, nausea, diarrhea, vomiting, and pancreatitis. The most common co-occurrences are GI and vomiting (298), GI and pancreatitis (152), GI and nausea (132), anxiety and depression (210), and nausea and pancreatitis (124). Moderate co-occurrences include anxiety and indigestion (96), anxiety and fatigue (76), and fatigue and nausea (74).

XSL•FO
RenderX

Figure 2. Network analysis of adverse events co-occurrences and community clusters. COPD: chronic obstructive pulmonary disease; GERD: gastroesophageal reflux disease.



Discussion

Principal Findings

This study used Facebook data to examine the adverse events discussed on the social media platform with respect to different GLP-1 RAs commonly used for weight management from January 2022 to June 2024. Semaglutide and Ozempic (the brand of semaglutide FDA-approved for the treatment of type 2 diabetes) were the most mentioned products, while Wegovy (the brand of semaglutide FDA-approved for the treatment of obesity), tirzepatide, Mounjaro (the brand of tirzepatide FDA-approved for the treatment of type 2 diabetes), Zepbound (the brand of tirzepatide FDA-approved for the treatment of obesity), and the overall GLP-1 RA class were discussed less often, reflecting the public's attention to Ozempic as a prominent GLP-1 RA agent and the first medication receiving market approval. The most commonly described adverse events were GI symptoms for most GLP-1 RA categories, while for tirzepatide social media posts, the most prevalent ones were headache and joint pain. The popularity of these medications and mentions of their adverse events surged in alignment with novel indications, approvals (particularly for children), and celebrity endorsement. We also uncovered 3 distinct clusters of referenced co-occurring adverse events, characterized by (1) somatic or metabolic symptoms (purple), (2) neurological or inflammatory discomfort (pink), and (3) gastrointestinal distress (brown). Together, these findings offer a glimpse into the public's social media discussions of adverse events related to GLP-1 RA medications. Posts discussing GLP-1 RAs as a class as well as specific posts discussing Mounjaro were most likely to mention adverse events (1138/4885, 23.3% and 868/4031, 21.5%, respectively), while posts discussing the other products (ie, semaglutide, tirzepatide, and Wegovy) included adverse events <10% of the time. However, it is important to note that Mounjaro and Zepbound (both tirzepatide products) were approved during the study period in May 2022 and November 2023, respectively, while Ozempic and Wegovy (both semaglutide products) were approved in 2017 and 2021, respectively. Thus, temporal differences in market access may have impacted the public's discussion of adverse events for these medications. More recently approved medications may garner more discussion about adverse events, while discussions of medications that have been available and used longer may have shifted toward other topics, such as health insurance coverage or novel indications [25,26]. GLP-1 RA adverse events may be dose dependent and diminish with time. Nevertheless, in our study, we did not collect dose information or include it in our analysis as doses were not consistently posted or available, and it was not possible to assess what dosages of the medications the adverse events were experienced on and attributed to.

GI adverse events, including nausea and vomiting, were the most frequently mentioned for all the drugs, consistent with both clinical trial and social media analysis data [27,28]. Additionally, our study found that Ozempic and all semaglutide medications were more frequently associated with reports of depression and anxiety than other weight loss medications. While our analyses cannot establish a causal relationship or provide a mechanistic pathway, the clustering of depressive symptoms among individuals discussing semaglutide is

XSL•FO
RenderX

concerning and builds upon prior literature, raising concerns about an increased risk of suicidal ideation with semaglutide therapy [29,30]. Therefore, it may be important to monitor individuals treated with semaglutide—and other GLP-1 RAs—for mood concerns until more conclusive clinical data are available. Nevertheless, it remains to be determined whether semaglutide impacts mood through its actions on the central nervous system and reward pathways, or whether individuals at high risk for depression are more likely to take semaglutide and share their concerns on social media.

Not surprisingly, we found significant increases in social media posts—both overall and those discussing medication adverse events—corresponding to the broader context during our study period between January 2022 and June 2024. After the FDA approved Wegovy for weight management in children aged 12 - 17 years in December 2022 [19], there was a corresponding increase in adverse event mentions in the following months. This trend aligns with a previous study that highlights potential adverse event risks, including GI issues, mental health concerns, and metabolic changes in younger populations [31]. During August 2023, there was an increase in media coverage of GLP-1 RA—particularly of compounded GLP-1 RA products—across social media platforms, concurrent with an FDA warning against use of unapproved GLP-1 RA drugs. This time period also corresponded to a second spike in adverse event mentions as illustrated in Figure 2 [32]. Our findings are consistent with an earlier report that revealed increased patient concerns and discussions about GLP-1 RA adverse events following increased media attention [33]. The next spike in social media discussions was observed after Oprah Winfrey shared that she had used weight loss medication as part of her health and fitness routine in December 2023, which contributed to a 400% increase in semaglutide prescriptions [34]. This surge coincided with a rise in public discourse about associated GI adverse events as well as more severe complications such as pancreatitis and kidney damage [35,36]. Finally, there was an increase in mentions of adverse event concerns after the March 2024 FDA expansion of Wegovy's approved indication to include the prevention of cardiovascular events in patients with obesity and cardiovascular disease, thereby paving the way for Medicare coverage of Wegovy for patients with obesity but not type 2 diabetes. A subsequent public poll by the Kaiser Family Foundation similarly noted an increase in public discussions of GLP-1 RA adverse events, including GI issues and fatigue [37]. Unlike previous studies, our study uniquely analyzed real-time social media discussions on GLP-1 RA medications, capturing shifts in adverse events mentioned in response to key regulatory and media events.

Previous clinical studies have also examined the adverse event profiles of semaglutide and tirzepatide GLP-1 RAs receptor agonists. For instance, in SURPASS-2, a 40-week phase 3 clinical trial comparing tirzepatide and semaglutide in 1879 participants with type 2 diabetes [10], the most commonly reported adverse events were GI: nausea, diarrhea, vomiting, and reduced appetite. Social media results not only corroborated the high prevalence of these GI adverse events but also exposed some more severe and rare adverse events, such as pancreatitis, which appeared more prominently in social media discussions than in clinical trials. For example, pancreatitis, which was experienced by 0.6% of patients in SUSTAIN-2, was discussed in 2.85% (139/4885) of GLP-RA's social media posts. Joint pain was an additional adverse event that appeared on social media but is not yet attributed to these medications in clinical trials [38]. Similarly, headache was the most frequently mentioned adverse events mentioned in 1.86% (78/4202) of tirzepatide social media posts. In contrast, a clinical study revealed reduced headache risk in patients with hypertension treated with tirzepatide [39]. These findings warrant further clinical investigation to examine the complex adverse events of tirzepatide on mitigating and exacerbating headaches and other neurological symptoms. Additionally, our study identified 3 unique clusters of adverse event profiles attributed to GLP-1 RA agents. We found co-occurrences between GI adverse events, such as nausea, vomiting, pancreatitis, and GI distress. Anxiety and depression were also highly correlated with each other. These findings build upon an earlier mixed methods study that analyzed Reddit, YouTube, and TikTok to assess the impact of GLP-1 RAs on mental health [7], finding 3 overarching mental health–related themes: insomnia (n=620 matches), anxiety (n=353), and depression (n=204); these mental health conditions were also detected in our analysis. Another study analyzed Twitter, Reddit, PubMed, SIDER, and manufacturer reports to identify clusters of adverse events frequently co-mentioned in the same posts [23]. The most prevalent cluster of co-mentioned adverse events was GI adverse events, corroborating our analytical findings. Next, mental health adverse events were a frequent cluster, aligning with the neurological adverse events cluster in our integrated node network analysis. This study not only leveraged a novel platform, Facebook, to reveal additional insights into relationships between comentioned adverse events but also corroborated the results of previous studies using similar methodologies. Our broader co-occurrence trends reveal broad adverse event categories that can be further explored to analyze individual differences between patients. Further clinical research is necessary to understand the complex relationships of co-occurring adverse events.

Strengths and Limitations

This quantitative social media analysis study has several strengths. Primarily, it evaluated adverse event profiles attributed by individuals to different GLP-1 RA agents over a 2-year period and across 63,023 posts. Social media posts represent a wide range of individuals, including those from diverse backgrounds who are underrepresented in clinical trials (racial and ethnic minority patients, women of childbearing age, rural residents, etc), in different drug markets, and from all regions of the United States [40]. Facebook is a popular social media platform with 68% of US adults using this digital platform in 2023 [41]. On social media, people have the opportunity to share their real-world experiences, beliefs, concerns, and outcomes and may be more likely to report some experiences they may not report in an in-person study.

However, the study's limitations must also be taken into consideration. While in general, social media users represent a diverse sample, we do not know the demographic characteristics of the social media users in our specific dataset. Symptoms

XSL•FO
RenderX

mentioned in these posts are self-reported, and we cannot discern whether they were personally experienced, represent concerns about potentially experiencing those adverse events, or were posted in response to hearing about adverse events from others. Additionally, social media users may differ from the general US population treated with GLP-1 RA medications. However, more than 7 in 10 Americans use social media to engage with others, receive news, and share information, and the social media user base is increasingly becoming more representative of the broader US population [42]. As such, social media posts about GLP-1 RA medications may represent a combination of the experiences of the broader population taking these medications, those interested in taking these medications, or expressions of concern about these adverse events. They do not solely represent the adverse events reported by the subset of patients with type 2 diabetes or obesity who are prescribed these medications in clinical practice. Furthermore, for those taking medications, we do not know the dosage or the length of time that they were taking these medications. Thus, while social media posts cannot replace clinical trials or large cohort studies, they provide a meaningful, complementary data source to detect adverse events that may warrant further study. Furthermore, this study includes a subset of weight loss medications as identified from Facebook posts. Future studies can expand on this work to examine other drugs used for weight loss and investigate these questions using additional social media platforms. While the analyzed social media posts were in the English language and based in the United States, social media posts in other languages and from other countries can also be analyzed in the future to expand the scope and generalizability of the study's findings. To construct a list of potential adverse events, we used past literature, clinical expertise, and the SIDER database version 4.1 and did not use the subscription-based product, Medical Dictionary for Regulatory Activities [43]. It would be valuable for future studies to investigate the extent to which findings differ depending on approaches used to identify adverse events. While compounded versions of GLP-1 RAs are not FDA-approved and have been associated with significant adverse drug reactions, this study did not distinguish adverse events from FDA-approved GLP-1 RAs compared with compounded versions of GLP-1 RAs [31]. It would be valuable for future studies to examine differences in reported adverse events of FDA-approved and compounded versions of these drugs.

Conclusions

Social media provides a space for users to share personal experiences, challenges, and outcomes with weight loss medications in a way that is infeasible to ascertain from other data sources and study designs. Through our comprehensive analysis of Facebook posts during 2022 and 2024, we were able to systematically identify and characterize adverse event profiles attributed to GLP-1 RAs. Using advanced data filtering and extraction techniques, we were able to identify patterns and trends within these data to effectively highlight the unique role of social media in examining the public narrative about increasingly popular weight loss medications. As the rates of GLP-1 RA use continue to grow, it is imperative to investigate their short-term and long-term adverse events and user perceptions of adverse events associated with each drug, along with the complex relationships between concurrent adverse events. Overall, this study enhances our understanding of the nuanced perspectives of the therapeutic landscape, helping facilitate further discovery in this area.

Acknowledgments

The authors thank Raj Bhansali and Linna Kuang for assistance with earlier versions of this manuscript. Research reported in this publication was supported by the National Institute on Minority Health and Health Disparities (R01MD015716 [TTN]). The research was supported in part by the Intramural Research Program of the NIH (QCN). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Drs. McCoy, Ratzki-Leewing, and Nguyen are investigators at the University of Maryland-Institute for Health Computing, which is supported by funding from Montgomery County, Maryland, and The University of Maryland Strategic Partnership: MPowering the State, a formal collaboration between the University of Maryland, College Park, and the University of Maryland, Baltimore.

Data Availability

Publicly available Facebook data were collected using CrowdTangle's user interface dashboard, a public insights tool from Meta now discontinued. More information can be found in the study by Shiffman and Silverman [15].

Authors' Contributions

TTN and RGM contributed to the conceptualization of the study. TTN contributed to the resources, funding, and supervision of the study. AA, VJ, XY, and HM were involved in the methodology, validation, data curation, and formal analysis. TTN, AA, and VJ wrote the original draft of the manuscript. AA, VJ, XY, HM, ARL, JSM, SC, QCN, and RGM were involved in reviewing and editing the manuscript.

Conflicts of Interest

TTN is a member of the *JMIR Infodemiology* Editorial Board. All other authors declare no conflicts of interest.

Multimedia Appendix 1

XSL•FO
RenderX

Supplementary tables and figures displaying additional details regarding glucagon-like peptide-1 receptor agonists and adverse events.
[DOCX File, 851 KB - infodemiology_v5i1e73619_app1.docx]

References

1. Singh G, Krauthamer M, Bjalme-Evans M. Wegovy (semaglutide): a new weight loss drug for chronic weight management. J Investig Med 2022 Jan;70(1):5-13. [doi: 10.1136/jim-2021-001952] [Medline: 34706925]
2. Whitley HP, Trujillo JM, Neumiller JJ. Special report: potential strategies for addressing GLP-1 and dual GLP-1/GIP receptor agonist shortages. Clin Diabetes 2023;41(3):467-473. [doi: 10.2337/cd23-0023] [Medline: 37456085]
3. National diabetes statistics report. Centers for Disease Control and Prevention (Diabetes). 2024. URL: https://www.cdc.gov/diabetes/php/data-research/index.html [accessed 2025-03-04]
4. Neumiller JJ, Bajaj M, Bannuru RR, et al. Compounded GLP-1 and dual GIP/GLP-1 receptor agonists: a statement from the American Diabetes Association. Diabetes Care 2025 Feb 1;48(2):177-181. [doi: 10.2337/dci24-0091] [Medline: 39620926]
5. Panic G, Yao X, Gregory P, Austin Z. How do community pharmacies in Ontario manage drug shortage problems? Results of an exploratory qualitative study. Can Pharm J (Ott) 2020;153(6):371-377. [doi: 10.1177/1715163520958023] [Medline: 33282028]
6. Arillotta D, Floresta G, Guirguis A, et al. GLP-1 receptor agonists and related mental health issues; insights from a range of social media platforms using a mixed-methods approach. Brain Sci 2023 Oct 24;13(11):1503. [doi: 10.3390/brainsci13111503] [Medline: 38002464]
7. Zheng Z, Zong Y, Ma Y, et al. Glucagon-like peptide-1 receptor: mechanisms and advances in therapy. Signal Transduct Target Ther 2024 Sep 18;9(1):234. [doi: 10.1038/s41392-024-01931-z] [Medline: 39289339]
8. Sargeant JA, Henson J, King JA, Yates T, Khunti K, Davies MJ. A review of the effects of glucagon-like peptide-1 receptor agonists and sodium-glucose cotransporter 2 inhibitors on lean body mass in humans. Endocrinol Metab (Seoul) 2019 Sep;34(3):247-262. [doi: 10.3803/EnM.2019.34.3.247] [Medline: 31565876]
9. Frías JP, Davies MJ, Rosenstock J, et al. Tirzepatide versus semaglutide once weekly in patients with type 2 diabetes. N Engl J Med 2021 Aug 5;385(6):503-515. [doi: 10.1056/NEJMoa2107519] [Medline: 34170647]
10. Chen W, Cai P, Zou W, Fu Z. Psychiatric adverse events associated with GLP-1 receptor agonists: a real-world pharmacovigilance study based on the FDA Adverse Event Reporting System database. Front Endocrinol 2024;15:38390214. [doi: 10.3389/fendo.2024.1330936]
11. FDA approves new medication for chronic weight management. FDA. 2024. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-new-medication-chronic-weight-management [accessed 2025-03-04]
12. Gao Y, Duan W, Rui H. Does social media accelerate product recalls? Evidence from the pharmaceutical industry. Inf Syst Res 2022 Sep;33(3):954-977. [doi: 10.1287/isre.2021.1092]
13. Leonard E. Oprah Winfrey reveals she uses weight-loss medication as a “maintenance tool”: “I’m absolutely done with the shaming” (exclusive). 2023. URL: https://people.com/oprah-winfrey-reveals-weight-loss-medication-exclusive-8414552 [accessed 2025-03-04]
14. Wreschnig LA, Congressional Research Service. Medicare coverage of GLP-1 Drugs. 2024. URL: https://crsreports.congress.gov/product/details?prodcode=IF12758 [accessed 2025-03-06]
15. Shiffman N, Silverman B. CrowdTangle opens public application for academics. Meta Research. URL: https://research.facebook.com/blog/2020/7/crowdtangle-opens-public-application-for-academics/ [accessed 2025-03-04]
16. Transparency Center. CrowdTangle. 2024 Aug 16. URL: https://transparency.meta.com/researchtools/other-datasets/crowdtangle [accessed 2025-05-08]
17. What is a serious adverse event? US Food and Drug Administration. 2019. URL: https://www.fda.gov/safety/reporting-serious-problems-fda/what-serious-adverse-event [accessed 2025-05-20]
18. Kuhn M, Letunic I, Jensen LJ, Bork P. The SIDER database of drugs and side effects. Nucleic Acids Res 2016 Jan 4;44(D1):D1075-D1079. [doi: 10.1093/nar/gkv1075] [Medline: 26481350]
19. News details. Novo Nordisk. 2022. URL: https://www.novonordisk-us.com/content/nncorp/us/en_us/media/news-archive/news-details.html [accessed 2025-03-04]
20. Ring S, Court E, Kresge N. Ozempic’s surging popularity has Novo Nordisk (NVO) struggling to meet demand. Bloomberg.com. 2023. URL: https://www.bloomberg.com/news/articles/2025-02-21/novo-s-ozempic-wegovy-shortage-ends-threatening-copycats [accessed 2025-03-04]
21. Noguchi Y. Medicare plans can now cover Wegovy for patients at risk of heart disease. NPR. 2024. URL: https://www.npr.org/sections/health-shots/2024/03/22/1240170094/wegovy-medicare-part-d-weight-loss-drugs [accessed 2022-03-04]
22. Aynaud T. taynaud/python-louvain. 2025. URL: https://github.com/taynaud/python-louvain [accessed 2025-04-07]
23. Bartal A, Jagodnik KM, Pliskin N, Seidmann A. Utilizing AI and social media analytics to discover adverse side effects of GLP-1 receptor agonists. arXiv. Preprint posted online on 2024. [doi: 10.48550/ARXIV.2404.01358]
24. Shen L, Tai Z, Shen E, Wang J. Graph exploration with embedding-guided layouts. arXiv. Preprint posted online on 2023. [doi: 10.48550/arXiv.2208.13699]

XSL•FO
RenderX

25. Sidik SM. Ozempic and Mounjaro aren’t the same. here’s how weight-loss drugs compare. Scientific American. 2024. URL: https://www.scientificamerican.com/article/mounjaro-and-ozempic-arent-the-same-heres-how-weight-loss-drugs-compare/ [accessed 2025-03-04]
26. Logan P. On the Increase in Use of GLP-1s blogs. 2024. URL: https://medicine.iu.edu/blogs/bioethics/on-the-increase-in-use-of-glp-1s [accessed 2025-03-04]
27. Yu J, Lee J, Lee SH, Cho JH, Kim HS. A study on weight loss cause as per the side effect of liraglutide. Cardiovasc Ther 2022;2022:5201684. [doi: 10.1155/2022/5201684] [Medline: 36540096]
28. Yao H, Zhang A, Li D, et al. Comparative effectiveness of GLP-1 receptor agonists on glycaemic control, body weight, and lipid profile for type 2 diabetes: systematic review and network meta-analysis. BMJ 2024 Jan 29;384:e076410. [doi: 10.1136/bmj-2023-076410] [Medline: 38286487]
29. Heaney C. Study links semaglutide and suicidal ideation risk. newsGP. 2024. URL: https://www1.racgp.org.au/newsgp/clinical/study-links-semaglutide-and-suicidal-ideation-risk [accessed 2025-03-04]
30. Schoretsanitis G, Weiler S, Barbui C, Raschi E, Gastaldon C. Disproportionality analysis from World Health Organization data on semaglutide, liraglutide, and suicidality. JAMA Netw Open 2024 Aug 1;7(8):e2423385. [doi: 10.1001/jamanetworkopen.2024.23385] [Medline: 39163046]
31. Nowogrodzki J. Should young kids take the new anti-obesity drugs? What the research says. Nature New Biol 2024 Sep;633(8031):750-751. [doi: 10.1038/d41586-024-02983-5] [Medline: 39289578]
32. FDA’s concerns with unapproved GLP-1 drugs used for weight loss. FDA. 2024. URL: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/fdas-concerns-unapproved-glp-1-drugs-used-weight-loss [accessed 2025-03-04]
33. Ovalle D, McGinley L. Patients grapple with side effects of popular weight-loss drugs. The Washington Post. 2023. URL: https://www.washingtonpost.com/health/2023/08/08/weight-loss-drugs-side-effects-wegovy-ozempic/ [accessed 2025-03-04]
34. Hopper L. Ozempic popularity among privately insured may boost disparities. USC Today. 2024. URL: https://today.usc.edu/weight-loss-drugs-popularity-ozempic-wegovy-may-worsen-disparities/ [accessed 2025-03-04]
35. Thompson D. Cost keeps many who need them from getting new weight-loss meds. 2024. URL: https://www.healthday.com/health-news/weight-loss/cost-keeps-many-who-need-them-from-getting-new-weight-loss-meds [accessed 2025-03-04]
36. Hopper L. USC Study: exploding popularity of Ozempic, Wegovy among privately insured patients may worsen disparities. USC Alfred E Mann School of Pharmacy and Pharmaceutical Sciences. 2024. URL: https://mann.usc.edu/news/usc-study-exploding-popularity-of-ozempic-wegovy-among-privately-insured-patients-may-worsen-disparities/ [accessed 2025-03-04]
37. Montero A, Sparks G, Presiado M, Hamel L. KFF health tracking poll may 2024: the public’s use and views of GLP-1 drugs. KFF. 2024. URL: https://www.kff.org/health-costs/poll-finding/kff-health-tracking-poll-may-2024-the-publics-use-and-views-of-glp-1-drugs/ [accessed 2025-03-04]
38. Bliddal H, Bays H, Czernichow S, et al. Once-weekly semaglutide in persons with obesity and knee osteoarthritis. N Engl J Med 2024 Oct 31;391(17):1573-1583. [doi: 10.1056/NEJMoa2403664] [Medline: 39476339]
39. Azzam AY, Essibayi MA, Farkas N, et al. Efficacy of tirzepatide dual GIP/GLP-1 receptor agonist in patients with idiopathic intracranial hypertension. a real-world propensity score-matched study. medRxiv. Preprint posted online on Nov 13, 2024. [doi: 10.1101/2024.11.12.24317193]
40. Stewart J, Krows ML, Schaafsma TT, et al. Comparison of racial, ethnic, and geographic location diversity of participants enrolled in clinic-based vs 2 remote COVID-19 clinical trials. JAMA Netw Open 2022 Feb 1;5(2):e2148325. [doi: 10.1001/jamanetworkopen.2021.48325] [Medline: 35157053]
41. Gottfried J. Americans’ social media use. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/ [accessed 2024-08-30]
42. Atske S. Social media use in 2021. Pew Research Center: Internet, Science & Tech. 2021. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2022-11-28]
43. MedDRA. Meddra.org. 2019. URL: https://www.meddra.org/ [accessed 2025-03-20]

Abbreviations:

FDA: Food and Drug Administration (United States)
GPP-1 RA: glucagon-like peptide-1 receptor agonist

XSL•FO
RenderX

Edited by T Mackey; submitted 08.03.25; peer-reviewed by A Breitschaft, O Allela; revised version received 22.05.25; accepted 27.05.25; published 24.07.25.

Please cite as:
Alibilli AS, Jain V, Mane H, Yue X, Ratzki-Leewing A, Merchant JS, Criss S, Nguyen QC, McCoy RG, Nguyen TT
Harnessing Facebook to Investigate Real-World Mentions of Adverse Events of Glucagon-Like Peptide-1 Receptor Agonist (GLP-1 RA) Medications: Observational Study of Facebook Posts From 2022 to 2024
JMIR Infodemiology 2025;5:e73619
URL: https://infodemiology.jmir.org/2025/1/e73619
doi:10.2196/73619

© Amrutha S Alibilli, Vidur Jain, Heran Mane, Xiaohe Yue, Alexandria Ratzki-Leewing, Junaid S Merchant, Shaniece Criss, Quynh C Nguyen, Rozalina G McCoy, Thu T Nguyen. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 24.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Messaging and Information in Mental Health Communication on Social Media: Computational and Quantitative Analysis

Rebecca K Ivic[1], PhD; Amy Ritchart[1], MFA; Shaheen Kanthawala[2], PhD; Heather J Carmack[3], PhD

[1]College of Communication and Information Sciences, University of Alabama, 901 University Ln, Tuscaloosa, AL, United States

[2]Department of Journalism & Creative Media, University of Alabama, Tuscaloosa, AL, United States

[3]Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, United States

Corresponding Author:
Rebecca K Ivic, PhD
College of Communication and Information Sciences, University of Alabama, 901 University Ln, Tuscaloosa, AL, United States

Abstract

Background: Mental health organizations have the vital and difficult task of shaping public discourse and providing important information. Social media platforms such as X (formerly known as Twitter) serve as such communication channels, and analyzing organizational health information offers valuable insights into their guidance and linguistic patterns, which can enhance communication strategies for health campaigns and interventions. The findings inform strategies to enhance public engagement, trust, and the effectiveness of mental health messaging.

Objective: This study examines the predominant themes and linguistic characteristics of messages from mental health organizations, focusing on how these messages' structure information, engage audiences, and contribute to public information and discourse on mental health.

Methods: A computational content analysis was conducted to identify thematic clusters within messages from 17 unique mental health organizations, totaling 326,967 tweets and approximately 7.2 million words. In addition, Linguistic Inquiry and Word Count (LIWC) was used to analyze affective, social, and cognitive processes in messages with positive versus negative sentiment. Differences in sentiment were assessed using a Mann-Whitney *U* test.

Results: The analysis revealed that organizations predominantly emphasize themes related to community, well-being, and workplace mental health. Sentiment analysis indicated significant differences in affect (P<.001), social processes (P<.001), and cognitive processing (P<.001) between positive and negative messages, with effect sizes that were small to medium. Notably, while messages frequently conveyed positive sentiment and social engagement, there was a lower emphasis on cognitive processing, suggesting that more complex discussions about mental health challenges may be underrepresented.

Conclusions: Organizations use social media to promote engagement and support, often through positively valanced messages. Yet the limited emphasis on cognitive processing may indicate a gap in how organizations address more nuanced or complex mental health issues. Findings demonstrate the need for communication strategies that balance information with depth and clarity, ensuring that messages are trustworthy, actionable, and responsive to multiple mental health needs. By refining digital messaging strategies, organizations can enhance the effectiveness of health communication and improve engagement with mental health resources.

(JMIR Infodemiology 2025;5:e48230) doi:10.2196/48230

KEYWORDS

health communication; mental health; mental health organizations; computational analysis; LIWC; Linguistic Inquiry and Word Count; large scale data

Introduction

Background

Mental health conditions represent an urgent and escalating global public health crisis. The burden of mental illness is global, yet structural barriers, including underfunded resources and lack of access to care, persist across regions and economic contexts. The World Health Organization's (WHO) Comprehensive Mental Health Action Plan 2013 - 2030 emphasizes the need for coordinated, scalable interventions, but a widening treatment gap remains, particularly in low-resource settings [1]. In this landscape, digital platforms serve as important places for mental health information, offering opportunities for intervention, advocacy, and public engagement at a scale typically unattainable through traditional health communication channels. Organizational structures play a vital role in shaping public discourse and mobilizing resources through social media [2-5]. They have the potential to facilitate large-scale mental health campaigns using health communication

XSL•FO
RenderX

strategies. As the global mental health crisis intensifies, it is critical to understand how organizations use digital platforms inform and build trust and ensure that mental health messaging is clear, accessible, and actionable.

Mental Health Organizations on Social Media

Health communication campaigns through advocacy strategies have been identified as a possibility for positive change for people experiencing these challenges [6]. The WHO called for key organizations to assist with mental health issue resolutions [7]. Similarly, the Center for Disease Control and Prevention put forth guidelines on health communication through text-based social media platforms like Facebook (Meta Platforms) or X (formerly known as Twitter) [8]. The importance of messages from such organizations has been examined in past research. Smith-Frigerio [6] examined 2 mental health advocacy groups' social media strategy. She found their messaging included awareness-raising strategies, support of various policy initiatives, and the promotion of diversity and inclusivity. Another study followed 3 mental health organizations' messages and engagement for a year and noted that content focused on mental health received more engagement as compared to relationship building or event promotion [9]. When investigating visual social media posts, Jia et al [10] noted that posts depicting knowledge about mental health disorders, their treatments, and antistigma frames were more heavily engaged content.

To better understand how mental health organizations communicate effectively, this study applies the Crisis and Emergency Risk Communication (CERC) framework to analyze the affective, social, and cognitive dimensions of their messaging. CERC helps organizations craft communication that enhances sensemaking and efficacy, ensuring messages are clear, actionable, and trustworthy. Traditionally used in public health crises [11,12], CERC is relevant to mental health communication, where strategic messaging can reduce stigma, build trust, and improve access to care [1,7]. This study extends CERC's application to social media, analyzing how mental health organizations engage with these principles in digital discourse.

Health Communication and Computational Text Analysis

A body of research has investigated the use of computational qualitative analysis in a broad range of disciplines, including health communication [12], psychology [13], tourism [14,15], physician education [5,16], and other domains. Importantly, we recognize there are advantages and disadvantages when using such approaches. Leximancer, for instance, was used [17] to strengthen the validity of their study of Indigenous voices in a public health curriculum. Gibson et al [18] examined a large volume of social media data regarding e-cigarette and tobacco coverage, using automated coding to content analyze the data. To generate reliability in the content analysis, they calculated the consistency of reliability for weekly estimates with a threshold set to >.70. Based on this, they found that X and YouTube (Alphabet Inc) contained fewer themes compared to other data sources. Vos and Buckner [11] examined an emerging crisis about the H7N9 virus and messages, drawing upon the CERC framework to address the ongoing communication that happened at various stages of the crisis. To analyze 25,598 unique messages, automated coding was used to identify sensemaking and efficacy messages in messages.

Social media sites can have a profound effect on communication regarding mental health, but there remains a dearth of knowledge on the detection of concepts and themes in a larger corpus of mental health discourse, especially from the perspective of official organizations, though this technique has been used in other fields, such as digital entrepreneurship [19,20]. Likewise, the application of Leximancer to health issues research is recent. There is a compelling opportunity for scholars to use automated text analysis to examine health-related data, especially when other approaches cannot process the volume of data. Considering this, automated data analysis to quickly understand the corpus of large amounts of words and text in data can be useful in health research, although it is not a panacea for understanding phenomena [21].

In prior research, these approaches have been used to explore larger data sets and mental health but have yet to fully emerge in health communication-related research, though they have been explained using interpretivist approaches from survey data with larger response rates (n=934; [22]). Other studies [23] investigated the environmental effects of COVID-19 using Leximancer. They first used a segmented regression technique and then applied automated content analysis on environment-related subreddits to gain insight into the evolution of risk amplification and ripple effects. In this study, Leximancer was used to demonstrate how large data sets may be used in qualitative research in the context of health communication, specifically in nuanced communication through mental health. The capabilities regarding automated textual analysis and interactive, visual depictions of the relationships among the data have long been highlighted as important for extracting the meaning of themes. Examining the predominant themes in mental health communication can yield important insights into the issues and challenges organizations face in disseminating and ultimately reaching the public.

The following research question is offered:

Research question 1 (RQ1): What are the most prevalent themes present in messages about mental health led by mental health organizations?

Message Sentiment

Linguistic features are integral to shaping the themes expressed in communication, which is an inherently complex process. As a fundamental component of communication, language influences how messages are interpreted and significantly impacts psychological health [24]. Linguistic Inquiry and Word Count (LIWC) has long been used in research to understand cognitive complexities in language use and texts, including social media data, and in relation to mental health. Research has indicated that depressed patients have been more likely to use negative emotion words and express more self-focused thoughts [25], while the use of types of words, such as causation words and self-discrepancies, have been linked to positive health outcomes [26]. Therefore, to analyze language use in the messages from mental health organizations in terms of their

XSL•FO
RenderX

affect, social, and cognitive processes using LIWC, this research adds deeper value into the study of the relationship between mental health and communication. A total of 3 fundamental processes are analyzed: affect (emotional contents of language), social (process of interactions with others), and cognitive processes (how information is processed and language regarding insights, causation, and discrepancies).

Social media platforms serve as key sources of mental health information, offering vital insights into how organizations communicate and engage the public. Researchers have increasingly used LIWC to analyze mental health–related content on X, examining linguistic patterns and their role in shaping information delivery. This study explores how mental health organizations structure their messaging within a large text corpus. Based on this, we propose the following hypothesis (H1):messages from mental health organizations will vary such that those with positive sentiments will have significant differences in affect, social, and cognitive processes compared to those with negative sentiments.

Methods

Data Collection

Tweets from mental health care organizations were collected using the Twitter API using March 2023 as a deliberate sampling frame due to the data size and scope. The organizations were identified through a snowball sampling approach, which began with accounts that were reshared or followed by a population of relevant accounts. The goal was to compile a list of accounts explicitly focused on mental health promotion or services or for which mental health was a core organizational value. Queries were conducted using organization names and their common abbreviations with specific parameters (eg, NIMHgov for the National Institutes of Mental Health), following methods of previous studies [27]. The initial sample included accounts reshared or followed by relevant organizations, and the iterative nature of the snowball sampling allowed the sample to expand gradually. This process resulted in a final dataset of 17 unique accounts queried. The dataset includes 326,967 tweets from 17 unique mental health organizations (see Multimedia Appendix 1). This dataset consists solely of tweets originally posted by these organizations and does not include mentions, replies, or retweets from other users. Account-level metadata, such as follower count and creation date, was not collected as part of the extraction process and is not available for analysis. Sampling procedures and account inclusion criteria were consistent with established scholarship [27]. Multimedia Appendix 1 provides a list of the accounts, including their X handles, full names, and the number of messages they posted at the time of data collection.

Data were extracted from JSON files and saved as a CSV file to facilitate analysis and ensure alignment with the study's focus on mental health communication. The dataset included details such as the text of each post, post type (eg, social media engagement metrics such as likes, quote counts, replies, and retweets), timestamp, and number of impressions. To assess the extent of irrelevant content, a review was conducted on a randomly selected 10% (32,900) subset of tweets. Following this review, thematic extraction was refined using Leximancer's automated procedures. To minimize data noise, the thematic relevance threshold was adjusted to retain only central, frequently co-occurring concepts in the final analysis. Concepts with weak connectivity, such as isolated fundraising messages, promotional event announcements, or generic engagement phrases ("join us" or "support this"), were ranked lower and excluded from core thematic groupings. Leximancer identified thematic clusters through semantic co-occurrence analysis, assigning prominence scores using a color-coded red, green, and blue (RGB) scale. Concepts most frequently appearing in central discourse were highlighted in hotter colors (red, orange, and yellow), while less prominent or potentially irrelevant content was represented in cooler colors (green, blue, and violet).

Ethical Considerations

We were mindful in our use of publicly available social media data, particularly given the ethical complexities surrounding mental health content. Ethical standards in this area are still evolving, and while many users may have awareness that their data can be accessed under terms and conditions, the extent of that understanding likely varies considerably [28,29]. We removed any potentially identifying information from the dataset, and usernames and IDs were not included in the queried API. As a result, the dataset only consisted of the aforementioned information, and no personally identifying information was included in the data analysis. We recognized that some groups may be either under- or overrepresented when analyzing the themes and sentiment in a larger scale analysis, and in addressing a stigmatized health issue [30]. We were cognizant of the potential harms and risks and sought to be mindful of these in the analysis by analyzing only text contents and excluding any user IDs or potentially identifying information.

Data Analysis

RQ1 was analyzed using Leximancer, an automated machine learning software that identifies thematic clusters without requiring predefined coding categories. Leximancer detects co-occurring terms within a dataset, using semantic relationships rather than word frequency alone to reveal conceptually connected themes [31]. This approach enables researchers to analyze patterns in large text corpora by grouping related ideas based on their contextual associations. While Leximancer generates data visualizations that organize key concepts into thematic clusters, researcher interpretation remains essential [3,13]. In these visualizations, themes are named based on the most prominent concept within each cluster, and researchers can refine these labels by reviewing associated terms. The strength of each concept is visually represented through a color gradient, ranging from red (most prominent) to violet (least prominent), reflecting its frequency and relevance within the data set.

Concept maps were generated using Leximancer to identify key topics emerging from the messages. Standard semantic and relational extraction techniques were applied. In the semantic extraction phase, the software identified fundamental concepts, terms that frequently appear together, by analyzing their

XSL•FO
RenderX

frequency, occurrence, and co-occurrence. These concepts were then weighted and placed in a co-occurrence matrix, with a thesaurus of highly relevant words and phrases generated to refine semantic meaning [32]. In the relational extraction phase, the discovered concepts were analyzed to determine their associations. After computing concept count, co-occurrence frequency, and relative importance, a data visualization was created. This visualization prioritizes highly co-occurring concepts, clustering them into themes based on their strongest connections.

To address H1, LIWC was used to measure the emotional tone of messages made, comprising roughly 7.2 million words and the full 326,967 tweets. We computed the mean score, SD, skewness, and kurtosis for each dimension of affect (mean 0.111, SD 0.079; skewness=1.516, kurtosis=5.171), social (mean 0.1752, SD 0.1055; skewness=0.973, kurtosis=1.838), and cognitive (mean 0.094, SD 0.077; skewness=0.914, kurtosis=1.374). Out of each of the dimensions, the distribution of affect scores was found to be significantly left-skewed ($z=-250.589$; $P<.001$), indicating that most mental health organizational messages with negative sentiment held higher affect scores compared to those with positive sentiment. Likewise, because messages were made by distinct mental health organizations and presumably different users (thus treating cases as independent observations), a Mann-Whitney *U* test was conducted.

Results

Overview

The findings are presented by first addressing the results of RQ1, followed by H1. RQ1 focused on identifying the most prevalent themes in messages about mental health shared by mental health organizations. The concept map output, which considers the dataset regardless of the date or time of messages, revealed a y-shaped structure with 2 distinct paths, where the left path is slightly more prominent than the right. Figure 1 shows the concept map. Following recommendations for conducting research with Leximancer [32], we identified the dominant concepts within these clusters as themes, including example messages were relevant. Leximancer identified thematic clusters based on semantic co-occurrence patterns in the dataset. The most prominent themes, community, mental health, and well-being, emerged without predefined coding categories. These themes were later interpreted through the lens of CERC, which emphasizes trust-building, efficacy, and social support. While Leximancer does not apply theoretical labels, the resulting themes align with CERC principles, reinforcing its relevance to digital mental health messaging and are later included in the discussion.

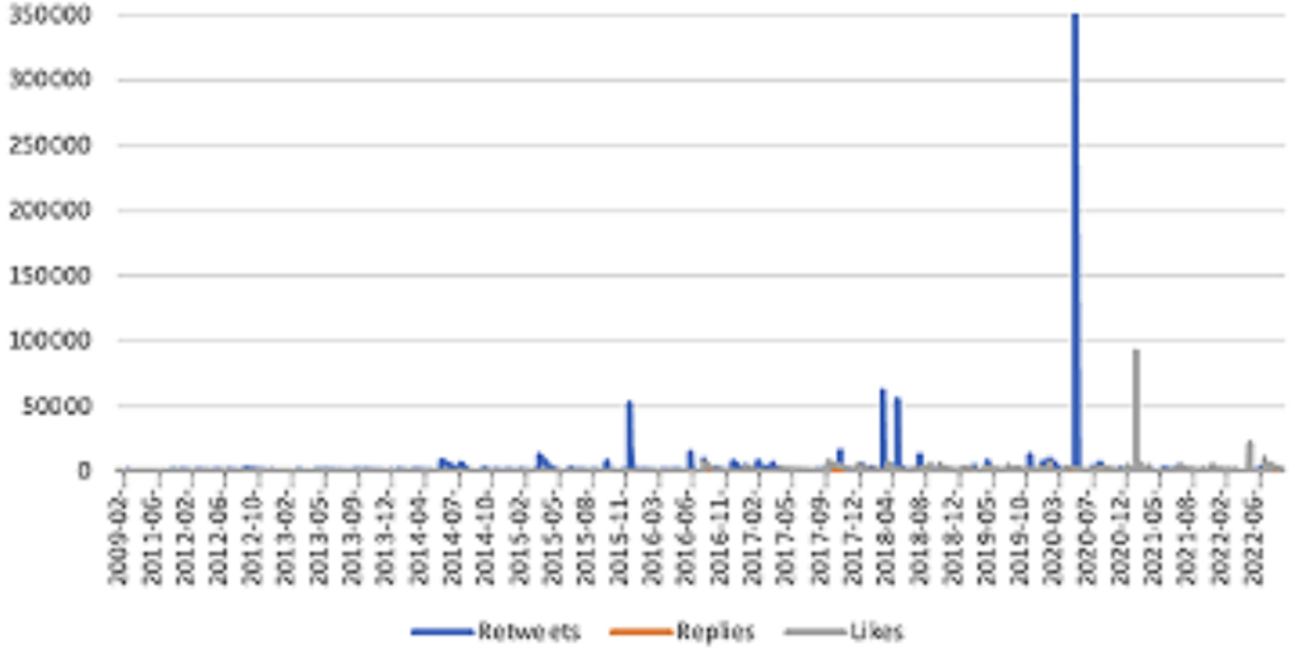
XSL•FO
RenderX

Figure 1. Concept map of the data.

XSL•FO
RenderX

A time series analysis further highlights the evolution of engagement with the message data (Figure 2). The y-axis represents the total number of engagements, including likes, retweets, and replies, while the x-axis denotes the date and time of posts. This analysis illustrates how engagement fluctuated over time, with periods of heightened activity aligning with significant mental health-related events. For example, the largest spike in activity occurred in March 2020, marked by a surge in retweets, coinciding with the WHO's declaration of COVID-19 as a global pandemic.

Figure 2. Temporal engagement over time with mental health messages.



RQ1 Findings

The most prominent thematic cluster identified was community, followed by mental health and well-being. The first 2 clusters demonstrated a linear relationship, where topics associated with community were directly linked to mental health, without further connections to other themes in the concept map. In contrast, the discourse surrounding mental health extended to less prominent, peripheral themes such as depression. Although these themes were less centralized, they held notable importance within the overall communication landscape. For example, some messages highlighted calls to action, such as: "What are you doing for your mental health this holiday season?" Posts such as these were often accompanied by links to resources promoting self-care and mental health practices, encouraging individual engagement with mental health management.

While well-being was not the most dominant theme, it served as a crucial connecting point in the y-shaped conceptual structure, bridging otherwise distinct clusters. The theme of well-being acted as a connecting point, linking concepts such as feelings, social support, and Twitter users. This indicates that mental health organizations strategically promoted prosocial messaging aimed at fostering shared communication and engagement. Many messages sought to directly interact with audiences through participatory activities, such as fundraising initiatives or encouraging retweets. In some cases, these messages attempted to challenge public perceptions of mental illness, as illustrated by posts like: "What does mental illness look like to you? [...] Send yours to [fundraising, redacted]." In addition, organizations shared practical resources designed to assist individuals in need, including helplines and support services. For instance, messages advertised emotional support systems: "Mentalhealth helpline offering specialist emotional support, guidance, and information" and "Textcare provides emotional support and connection when you need it most."

In the concept map's right-side path, less prominent themes emerged, including workplace and employees. These themes illustrated the role of work environments in shaping mental health and well-being, with organizations sharing messages emphasizing the importance of supportive workplace cultures. For example:

> *Organisational culture is critical to #wellbeing in the #workplace" linked to reports outlining strategies for fostering mental health in professional settings. Similarly, some messages underscored the value of positive work environments with messages such as, "Creating a great work environment matters. #healthyworkplace.*

Although themes on the periphery of the concept map, such as employee were ranked as less prominent, they provided meaningful insights into specific facets of the discussion. For example, the employee cluster included terms like "GPTW" (Great Place to Work), which may strengthen messaging about workplace mental health. These findings suggest that peripheral themes, while less frequent, often enrich the central discourse by adding nuanced perspectives and addressing less commonly explored areas.

These thematic clusters reflect a comprehensive strategy by mental health organizations to engage diverse audiences. Dominant themes serve as anchors, while secondary and peripheral themes act as bridges, connecting audiences to

XSL•FO
RenderX

lesser-explored topics. These layered insights not only enhance the understanding of mental health communication but also highlight opportunities for expanding outreach and improving messaging strategies. The implications of these findings are explored further in the discussion.

H1 Results

H1 examined whether there would be a significant difference in emotional processes in mental health organizational messages with positive sentiment compared to negative sentiment. LIWC was used to measure sentiment-related linguistic features, including affect, social, and cognitive processes. This analysis was conducted independently of Leximancer and was not used to categorize themes. A Mann Whitney-*U* test was performed, which showed significant differences in affect (U=6.582E+9, z=−250.589; P<.001), social processes (U=7.582E+9, z=−213.367; P<.001), and cognitive processes (U=1.275E+10, z=−21.660; P<.001) between mental health messages with positive sentiment compared to those with negative sentiment. The effect sizes for these differences were small (cognitive: r=−.205) to medium (affect: r=−.556, social: r=−.469).

These results suggest that the mean rank of messages containing negative tone were altogether lower than mean rankings of messages that were positive in tone. Thus, although the negative effect sizes suggest that each of affect, social, and cognitive dimensions held lower means, the values were highly statistically significant. Therefore, we suggest the results may support H1, which predicted differences in each of these processes in the emotional tone in messages.

Discussion

Principal Findings

Effective health communication goes beyond simply raising awareness; it requires transparency, credibility, and a commitment to delivering high-quality health information that resonates across different populations. This study underscores the essential role of health communication in shaping public perceptions of mental health and ensuring that information shared by organizations is accessible, trustworthy, and actionable. These findings highlight how mental health organizations use social media to frame information around well-being, but they also reveal gaps in addressing structural barriers to care and the need for more nuanced engagement strategies.

The analysis reveals that mental health organizations primarily emphasize community, well-being, and support networks in their messaging. While these themes foster positive engagement, there is a pressing need for clearer, more actionable communication strategies that go beyond broad encouragement to provide concrete guidance. Effective messaging must acknowledge the complexities of mental health struggles, offer practical pathways to care, and address the nuances of different lived experiences. Trust remains central to mental health communication, and organizations must ensure their messaging is both scientifically accurate and culturally responsive to reach audiences effectively.

Regarding H1, the linguistic analysis further underscores the role of emotion and social connection in digital mental health messaging. Findings indicate that while positive sentiment is a dominant strategy, it may not always align with the lived experiences of individuals facing mental health challenges. Overly optimistic framing, while engaging, could inadvertently dismiss the complexity of mental health struggles, or discourage open conversations about distress. A balance is needed: one that promotes hope while validating the realities of mental health needs. Health communication efforts must consider how organizations can refine their language to be impactful, ensuring that health information is framed in ways that build trust, encourage help-seeking, and reduce stigma.

These findings align with the CERC framework, which emphasizes that effective health communication should engage and provide clear, actionable information that enables individuals to respond to their circumstances [11,12]. While organizations successfully foster social connection, many messages lack the informational clarity and efficacy-building components that CERC underscores as essential in public health communication. Without concrete guidance, audiences may struggle to translate engagement into action. Applying informed strategies, such as reinforcing practical coping strategies, clear pathways to care, and guidance on recognizing and addressing mental health concerns, can help organizations move beyond awareness-building toward delivering information that empower individuals to take meaningful steps in managing their health.

Mental health communication takes place globally on social media platforms, with differing social and cultural contexts that shape how messages are received and interpreted [33]. As mental health organizations engage with global audiences, understanding how language functions across affective, social, and cognitive dimensions [34] are essential for researchers and practitioners. However, while fostering a positive emotional tone can enhance engagement, it is equally important to recognize that cultural variations in emotional expression and information shared may lead to unintended consequences. Overemphasizing positivity may inadvertently invalidate negative emotions, reinforcing stigma or limiting dialogue about health struggles, particularly in areas where stigmatization is an issue. To maximize impact, organizations must ensure that their messaging is culturally responsive, contextually appropriate, and globally inclusive of lived experiences.

By identifying key themes and emotional tones that effectively engage audiences, researchers can guide the development of evidence-based communication strategies for mental health organizations [10,35]. The study's use of computational analysis alongside LIWC demonstrates a scalable approach to analyzing social media data, offering insights that extend beyond mental health to broader health communication. These findings provide organizations with actionable strategies to craft messages that are informative and impactful, ensuring they resonate with audiences encourage engagement.

These results highlight the important role of language in shaping mental health communication through information shared on social media, particularly the impact of positive sentiment on affective and social processes. The lower cognitive processing

XSL•FO
RenderX

scores observed in mental health-related messages suggest that individuals facing mental health challenges may find it difficult to engage in complicated reasoning or articulate their experiences fully. This has important implications for organizations aiming to craft messages that are both accessible and solutions oriented. Rather than relying solely on positivity, messaging should balance encouragement with clarity, offering tangible strategies that individuals can apply to maintain or restore wellness. Findings from H1 reinforce the need for a deeper understanding of how language influences engagement with mental health communication on social media. A more nuanced approach to message design, grounded in both scientific accuracy and sensitivity to lived experiences, can enhance the effectiveness of mental health discourse, ensuring it supports those who need it most.

Limitations and Future Directions

This study examined overarching thematic patterns in mental health communication across organizations, rather than stratifying by account type. While differences in messaging strategies may exist between government agencies, advocacy groups, and service providers, our analysis was designed to capture broad trends rather than organizational distinctions. Future research could explore these variations by applying stratified thematic analysis. This study builds on previous scholarship while recognizing its potential limitations [36].

A limitation of this study is that themes were analyzed at the aggregate level without stratifying by organization type. Future research could examine whether certain themes are more prominent among different categories of mental health organizations, providing additional insights into institutional messaging strategies. Using X as a data source means that the sample reflects conversations from a public stream, which may introduce biases based on the users who actively participate on the platform [36]. Automated sentiment analysis, while valuable, has inherent limitations in capturing the full nuance and complexity of human emotions. Although emotional tone can be quantitatively measured, it may not fully account for the range of interpretations and contextual meanings in messages. Furthermore, snowball sampling was used to identify relevant accounts by tracking those frequently retweeted or followed. This method ensured consistency in the dataset but may have led to overrepresentation of highly visible accounts, potentially overlooking lesser-known organizations and users. Social media research also faces broader challenges, as the nature of online information and messages varies across platforms. Differences in audience demographics, communication styles, and modes of expression highlight the importance of comparing discussions across multiple platforms to better understand the specific contexts in which these conversations take place.

Analyzing large-scale social media data presents several challenges. Researchers must navigate high data volumes, the evolving nature of digital texts, user behaviors, data legitimacy, and ethical considerations when handling sensitive content [28,29]. Future research can build on these findings by exploring how emotional tone and language use in mental health organizations' messaging influence health literacy, attitudes, beliefs, and behaviors.

Conclusions

This study investigates how mental health organizations leverage social media to shape public discourse, focusing on the themes and linguistic strategies that define their messaging and information shared. The findings reveal that these organizations primarily emphasize community, positive sentiment, and workplace-related discussions on well-being. Health communication plays a crucial role in shaping public understanding and engagement, the clarity, accessibility, and trustworthiness of these messages are essential. Effective communication not only informs but also empowers individuals by providing actionable, evidence-based information. By analyzing these digital messages and the information shared, this study contributes to a deeper understanding of how mental health is framed on social media, offering insights into how messaging strategies can be refined to enhance public trust, reduce stigma, and promote meaningful engagement with mental health information.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

RKI served as the lead author and contributed to the study design, data collection, data analysis, and leading stages of manuscript development, including drafting, writing, and revisions. AR contributed to the development of the literature review, supported insights on the data analysis, and provided critical feedback on manuscript drafts. SK assisted with the literature review, contributed to the insights on the methodological framework, and offered insights into the study design. HJC provided input on the introduction, helped reshape the literature review and conceptual framing of the arguments, and provided substantive input on the Discussion and Conclusion sections.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of accounts queried from X (formerly known as Twitter) API.

XSL•FO
RenderX

[DOCX File, 15 KB - infodemiology_v5i1e48230_app1.docx]

References

1. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. ICWSM 2021;7(1):128-137. [doi: 10.1609/icwsm.v7i1.14432]
2. Kostygina G, Tran H, Binns S, et al. Boosting health campaign reach and engagement through use of social media influencers and memes. Social Media + Society 2020 Apr;6(2):205630512091247. [doi: 10.1177/2056305120912475]
3. Angus D, Rintel S, Wiles J. Making sense of big text: a visual-first approach for analysing text data using Leximancer and Discursis. Int J Soc Res Methodol 2013 May;16(3):261-267. [doi: 10.1080/13645579.2013.774186]
4. Krishen AS, Petrescu M. The world of analytics: interdisciplinary, inclusive, insightful, and influential. J Market Anal 2017 Mar;5(1):1-4. [doi: 10.1057/s41270-017-0016-4]
5. Pill S, Harvey S, Hyndman B. Novel research approaches to gauge global teacher familiarity with game-based teaching in physical education: an exploratory #Twitter analysis. Asia Pac J Health Sport Phys Educ 2017 May 4;8(2):161-178. [doi: 10.1080/18377122.2017.1315953]
6. Smith-Frigerio S. Grassroots mental health groups' use of advocacy strategies in social media messaging. Qual Health Res 2020 Dec;30(14):2205-2216. [doi: 10.1177/1049732320951532] [Medline: 32842878]
7. Mental health action plan 2013 - 2020. World Health Organization. 2015. URL: https://doi.org/entity/mental_health/publications/action_plan/en/index.html [accessed 2024-12-01]
8. Coronavirus disease 2019 (COVID-19). Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/coronavirus/2019-ncov/communication/guidance-list.html [accessed 2023-04-16]
9. Ju R, Jia M, Cheng J. Promoting mental health on social media: a content analysis of organizational tweets. Health Commun 2023 Jul;38(8):1540-1549. [doi: 10.1080/10410236.2021.2018834] [Medline: 34955059]
10. Jia M, Ju R, Zhu J. Understanding mental health organizations' Instagram through visuals: a content analysis. Health Commun 2024 Apr;39(4):767-777. [doi: 10.1080/10410236.2023.2185350] [Medline: 36856059]
11. Vos SC, Buckner MM. Social media messages in an emerging health crisis: tweeting bird flu. J Health Commun 2016;21(3):301-308. [doi: 10.1080/10810730.2015.1064495] [Medline: 26192209]
12. Britt BC, Britt RK, Hayes JL, Panek ET, Maddox J, Musaev A. Oral healthcare implications of dedicated online communities: a computational content analysis of the r/Dentistry subreddit. Health Commun 2021 May;36(5):572-584. [doi: 10.1080/10410236.2020.1731937] [Medline: 32091259]
13. Cretchley J, Rooney D, Gallois C. Mapping a 40-Year history with Leximancer: themes and concepts in the Journal of Cross-Cultural Psychology. J Cross Cult Psychol 2010 May;41(3):318-328. [doi: 10.1177/0022022110366105]
14. Scott N, Smith AE. Use of automated content analysis techniques for event image assessment. Tour Recreat Res 2005 Jan;30(2):87-91. [doi: 10.1080/02508281.2005.11081477]
15. Tkaczynski A, Rundle-Thiele SR, Cretchley J. A vacationer-driven approach to understand destination image. J Vacat Mark 2015 Apr;21(2):151-162. [doi: 10.1177/1356766714567796]
16. Hyndman B, Pill S. What's in a concept? A Leximancer text mining analysis of physical literacy across the international literature. Eur Phy Educ Rev 2018 Aug;24(3):292-313. [doi: 10.1177/1356336X17690312]
17. Lee V, Coombe L, Mahoney R, Allen C, Robinson P. Incorporating indigenous and non-indigenous worldviews through innovative text analysis: an evaluation of indigenous public health curricula. Int J Qual Methods 2018 Dec 1;17(1):160940691881937. [doi: 10.1177/1609406918819377]
18. Gibson LA, Siegel L, Kranzler E, et al. Combining crowd-sourcing and automated content methods to improve estimates of overall media coverage: theme mentions in e-cigarette and other tobacco coverage. J Health Commun 2019;24(12):889-899. [doi: 10.1080/10810730.2019.1682724] [Medline: 31718524]
19. Lai SL, To WM. Content-based analysis of social media and official websites: a grounded approach. JECR. 2015. URL: http://www.jecr.org/sites/default/files/16_2_p05.pdf [accessed 2024-12-01]
20. Nambisan S. Digital entrepreneurship: toward a digital technology perspective of entrepreneurship. Entrepreneurship Theory and Practice 2017 Nov;41(6):1029-1055. [doi: 10.1111/etap.12254]
21. Lemon L, Hayes J. Enhancing trustworthiness of qualitative findings: using Leximancer for qualitative data analysis triangulation. TQR 2020;25(3). [doi: 10.46743/2160-3715/2020.4222]
22. Usher W. Living in quiet desperation: the mental health epidemic in Australia's higher education. Health Educ J 2020 Mar;79(2):138-151. [doi: 10.1177/0017896919867438]
23. VanDyke MS, Britt BC, Britt RK, Franco CL. How environment-focused communities discuss COVID-19 online: an analysis of social (risk) amplification and ripple effects on Reddit. Environmental Communication 2023 Apr 3;17(3):322-338. [doi: 10.1080/17524032.2022.2056222]
24. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language. use: our words, our selves. Annu Rev Psychol 2003;54(1):547-577. [doi: 10.1146/annurev.psych.54.101601.145041] [Medline: 12185209]
25. Rude S, Gortner EM, Pennebaker J. Language use of depressed and depression-vulnerable college students. Cogn Emot 2004 Dec;18(8):1121-1133. [doi: 10.1080/02699930441000030]

XSL•FO
RenderX

26. Petrie KJ, Booth RJ, Pennebaker JW. The immunological effects of thought suppression. J Pers Soc Psychol 1998;75(5):1264-1272. [doi: 10.1037/0022-3514.75.5.1264]
27. Sanders WS, Wang YJ, Zheng Q. Brand's social media presence as networks: the role of interactivity and network centrality on engagement. Commun Res Rep 2019 Mar 15;36(2):179-189. [doi: 10.1080/08824096.2019.1590192]
28. Nicholas J, Onie S, Larsen ME. Ethics and privacy in social media research for mental health. Curr Psychiatry Rep 2020 Nov 23;22(12):84. [doi: 10.1007/s11920-020-01205-9] [Medline: 33225404]
29. Kanthawala S, Cotter K, Foyle K, Decook JR. It's the methodology for me: a systematic review of early approaches to studying tiktok. Presented at: Hawaii International Conference on System Sciences; Jan 4-7, 2022; Maui, Hawaii. [doi: 10.24251/HICSS.2022.382]
30. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453. [doi: 10.1126/science.aax2342] [Medline: 31649194]
31. Sotiriadou P, Brouwers J, Le TA. Choosing a qualitative data analysis tool: a comparison of NVivo and Leximancer. Ann Leis Res 2014 Apr 3;17(2):218-234. [doi: 10.1080/11745398.2014.902292]
32. Smith AE, Humphreys MS. Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. Behav Res Methods 2006 May;38(2):262-279. [doi: 10.3758/bf03192778] [Medline: 16956103]
33. Spitzberg B. Methods of interpersonal skill assessment. In: Green J, Burleson B, editors. Handbook of Communication and Social Interaction Skills: Lawrence Erlbaum; 2023:111-152.
34. Stupinski AM, Alshaabi T, Arnold MV, et al. Quantifying language changes surrounding mental health on Twitter. arXiv. Preprint posted online on Jun 2, 2021. [doi: 10.48550/arXiv.2106.01481]
35. Gohil S, Vuik S, Darzi A. Sentiment analysis of health care tweets: review of the methods used. JMIR Public Health Surveill 2018 Apr 23;4(2):e43. [doi: 10.2196/publichealth.5789] [Medline: 29685871]
36. Wilk V, Cripps H, Capatina A, Micu A, Micu AE. The state of #digitalentrepreneurship: a big data Leximancer analysis of social media activity. Int Entrep Manag J 2021 Dec;17(4):1899-1916. [doi: 10.1007/s11365-020-00729-z]

Abbreviations

CERC: Crisis and Emergency Risk Communication
LIWC: Linguistic Inquiry and Word Count
RGB: red, green, and blue
RQ: research question
WHO: World Health Organization

Edited by T Mackey; submitted 16.04.23; peer-reviewed by M Peters, K Mikołaj; revised version received 20.03.25; accepted 21.04.25; published 03.07.25.

Please cite as:
Ivic RK, Ritchart A, Kanthawala S, Carmack HJ
Messaging and Information in Mental Health Communication on Social Media: Computational and Quantitative Analysis
JMIR Infodemiology 2025;5:e48230
URL: https://infodemiology.jmir.org/2025/1/e48230
doi:10.2196/48230

© Rebecca K Ivic, Amy Ritchart, Shaheen Kanthawala, Heather Carmack. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 3.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Postpartum Depression and Maternal-Infant Bonding Experiences in Social Media Videos: Qualitative Content Analysis

Kunmi Sobowale[1], MD; Jamie Sarah Castleman[2], MS; Sophia Yingruo Zhao[3], BSc

[1]Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, 760 Westwood Plaza, Los Angeles, CA, United States

[2]Ponce Health Sciences University School of Medicine, St. Louis, MO, United States

[3]University of California Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Kunmi Sobowale, MD
Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, 760 Westwood Plaza, Los Angeles, CA, United States

Abstract

Background: While the negative effects of postpartum depression on maternal-infant bonding are well-documented, our understanding of how it exerts these effects remains incomplete. A better understanding of how maternal postpartum depression affects bonding can enable clinicians to better identify and support mothers with difficulties bonding with their children.

Objective: This study aims to describe the bonding experiences of mothers with postpartum depression through an analysis of short-form videos and user engagement.

Methods: We collected publicly available highly-viewed TikTok videos using hashtags associated with postpartum depression and associated engagement metrics in May 2023. After manual screening, we extracted 533 videos related to the mother-infant bond, from which we analyzed a random subset of 159 videos. We abstracted categories from videos using a hybrid deductive and inductive approach. Negative binomial regression models of video likes, views, shares, and comment count were used with content categories and the creator's numbers of followers as independent variables.

Results: Abstraction of content from mother-infant bond videos resulted in six categories: (1) navigating anxiety and anger, (2) creating physical and emotional boundaries, (3) overwhelmed by demands of caregiving, (4) subverted expectations, (5) enduring and finding strength through the challenge of postpartum depression, and (6) can't remember early life. Subverted expectations and navigating anxiety and anger categories were associated with increased views (rate ratio [RR] 1.72, 95% CI 1.22 - 2.43; RR 1.61, 95% CI 1.09 - 2.38, respectively), likes (RR 3.61, 95% CI 2.55 - 5.11; RR 3.96, 95% CI 2.69 - 5.85, respectively), shares (RR 2.95, 95%CI 2.09 - 4.18; RR 2.45, 95% CI 1.66 - 3.61, respectively), and comments (RR 2.78, 95% CI 1.97 - 3.94; RR 1.89, 95% CI 1.28 - 2.79, respectively). Sensitivity analysis with creators with fewer followers mostly aligned with these results.

Conclusions: This qualitative content analysis of short-form videos identified specific ways postpartum depression impacts the mother-infant bond, highlighting strategies for clinicians to support bonding. Analysis of engagement metrics further demonstrated the types of experiences that most resonate with viewers. Our findings demonstrate the potential of this qualitative method to augment understanding of lived experiences.

(JMIR Infodemiology 2025;5:e59125) doi:10.2196/59125

KEYWORDS

postpartum depression; perinatal depression; postpartum anxiety; depression; TikTok; social media; video; content; postpartum; anxiety; maternal-infant; qualitative; content analysis; bonding; experience; deductive; inductive; regression model

Introduction

A mother's bond to her child in the first few years of life greatly influences the child's development and future mental health. The bond is a caregiver's emotions, cognitions (ie, thoughts and beliefs), and behaviors toward the infant [1,2]. Bonding promotes child development, including physical, cognitive, and socioemotional domains [3-5], with more potent effects for the latter postnatally [3-7]. Longitudinally, stronger postpartum bonding is associated with less child internalizing and externalizing problems [6-8].

While the negative effects of poor maternal-infant bonding (MIB) are well-documented, our understanding of bonding, particularly in the context of depression, is lacking. Postpartum depression afflicts at least 13% of all birthing mothers annually in the United States and can harm child development and mental health in part by impairing mother-child bonding [7,9]. Yet, recent studies on the operationalization of the MIB find the construct, content, and structural validity of MIB measures,

XSL•FO
RenderX

including measures validated in women with postpartum depression, to be insufficient [10,11]. A lack of diverse samples in measurement development may contribute to this issue [11]. Most qualitative studies focus on maternal depression more broadly rather than the MIB in mothers with depression. Such studies could provide valuable insights about the MIB in the context of postpartum depression [12]. A better understanding of how postpartum depression exerts this effect can help clinicians identify and support mothers with difficulties bonding with their children.

One way to increase understanding of MIB is to examine the experiences of women with postpartum depression, as described on social media. With the rise of social media, mothers have taken to the internet to discuss mental health difficulties using a variety of platform types (blogs, microblogs, forums, photo sharing, etc). Women share experiences of their own accord, allowing for the concerns of those with lived experiences of postpartum depression to come to light. Understanding these experiences is crucial as social approval, fear of judgment, or fear of having their child removed may make women less likely to disclose their concerns to clinicians [13]. Existing studies primarily use text-based social media such as forums [14-18]. To our knowledge, no studies to date have used video-based social media to examine MIB in the context of postpartum depression. Video-based social media data can provide a detailed understanding of the experiences of women with postpartum depression as compared to text alone, capturing nonverbal communication and additional contextual factors. TikTok, a social media platform featuring short-form videos, has significantly higher use (over 1 billion monthly active users) [19] than text-based platforms used in prior studies, potentially allowing for more diverse perspectives and broader reach.

Therefore, this study aims to categorize common ways mothers endorsing postpartum depression describe their bond with their child using content analysis of highly accessed videos on TikTok. In addition, this research examines user engagement (ie, views, likes, comments, and shares) across these categories.

Methods

Overview

We conducted a qualitative content analysis of videos from TikTok to document the variety of experiences of bonding in mothers with postpartum depression. Our approach was hybrid in nature, combining data-driven coding with literature-based coding [20]. The study followed the Standards for Reporting Qualitative Research reporting guideline [21].

Ethical Considerations

The University of California, Los Angeles Institutional Review Board (#23-001157) deemed this nonhuman participant research exempt. Although we used publicly available data, given reidentification risks, we deidentified data throughout the study by not using identifiable information. Furthermore, we do not provide hashtags used for this study.

Data Collection and Preprocessing

We identified videos using hashtags associated with postpartum depression, a common method in social media studies [22]. Content creators' use of these hashtags suggests that they may self-identify with this diagnosis or its symptoms. Specific hashtags were identified using a variety of sources. First, the term "postpartum depression" was searched using the TikTok search feature. Common hashtags used in the initial 20 videos were collected. In addition, hashtags for postpartum depression that were used in other social media studies were evaluated [15]. We chose the 3 hashtags with the most views. Using these hashtags, we used the app Apify in May 2023 to collect publicly available popular videos and associated engagement metrics. These hashtags had a combined 1.1 billion views when videos were collected in May 2023.

We manually reviewed videos to exclude videos without content related to postpartum depression, and subsequently, not related to the mother-infant bond. We also excluded videos that were not in English, although the speakers were not required to be native English speakers. Given the large number of relevant videos (n=533) and resource constraints, we decided to use a random sample of approximately a third of the total number of videos collected (159/533, 30%) for analysis. This number of videos is in line with other content analysis studies of TikTok, which examine between 25 and 150 videos [23-26].

Engagement Metrics

To explore user engagement with videos, we examined video views, likes, shares, and comments. Views indicate how many people were exposed to the content. The number of likes shows how many people responded positively to a video by endorsing it. Shares capture the number of times users spread the video to their network, indicating how much it spread beyond the original audience. A high share count suggests that a video garnered enough interest to be actively passed on. The number of comments reflects how much discussion the video generates by prompting reactions and opinions from viewers.

Content Analysis

The research team consisted of a male psychiatrist with experience working with mothers with postpartum depression and infant mental health, along with a female medical student and a female undergraduate student who are interested in obstetrics and child development and psychology, respectively (KS, JSC, and SYZ).

A coding rubric was developed in several stages using an inductive and deductive (ie, hybrid) approach [27,28]. Videos were the unit of analysis with visual, audio, and text (both in the video and its description) components. First, deductively, one author (KS) identified an initial set of codes based on existing literature to design a preliminary codebook. Subsequently, over 5 rounds, coders independently watched 45 videos (5 - 10 per round) from the eligible videos and then reviewed them as a group. In initial rounds, coders gained familiarity with videos and determined the validity of the initial set of codes. Discrepancies in coding were discussed and resolved as a group. Furthermore, inductively, we added a new preliminary code whenever a video featured an experience of

XSL•FO
RenderX

the MIB that could not be suitably coded into any existing categories. Bonding experience codes that frequently recurred were made into permanent codes. We reached data saturation finding repetition of codes during our coding of our final random sample of 159 videos. Code definitions are provided in Table S1 in Multimedia Appendix 1 with codes derived from the literature cited. Videos were coded under multiple codes when applicable. Interrater reliability (IRR) was acceptable (Krippendorff α=0.67 - 0.71) [29] in pairwise ratings of coders for coding 30 randomly selected videos. We used Excel (Microsoft Corp) to facilitate coding.

Our data abstraction from codes to categories was guided by the nature of the data and our research objectives [30]. Given the short-form nature of many TikTok videos with limited depth and our desire to closely represent the postpartum depressed mothers' bonding experiences, we abstracted from codes to a higher-order category grouping without much interpretation. Thus, we chose not to use a thematic framework to avoid imposing external structures on the data. Instead, we used Miro boards to independently sort and group codes into categories based on similarities and relationships between codes. Categories had to be internally homogeneous and distinguishable from other categories. We then compared independently created categories among coders and aligned them to finalize the categories.

We reviewed creators' profiles and videos to determine their country of residence and the approximate age of their child when the mothers experienced postpartum depression. To estimate the child's age, we primarily used the age of the child when the video was created. However, if the mothers specifically mentioned experiencing depression during a particular period, such as the "first few months," we used that age instead.

Statistical Analysis

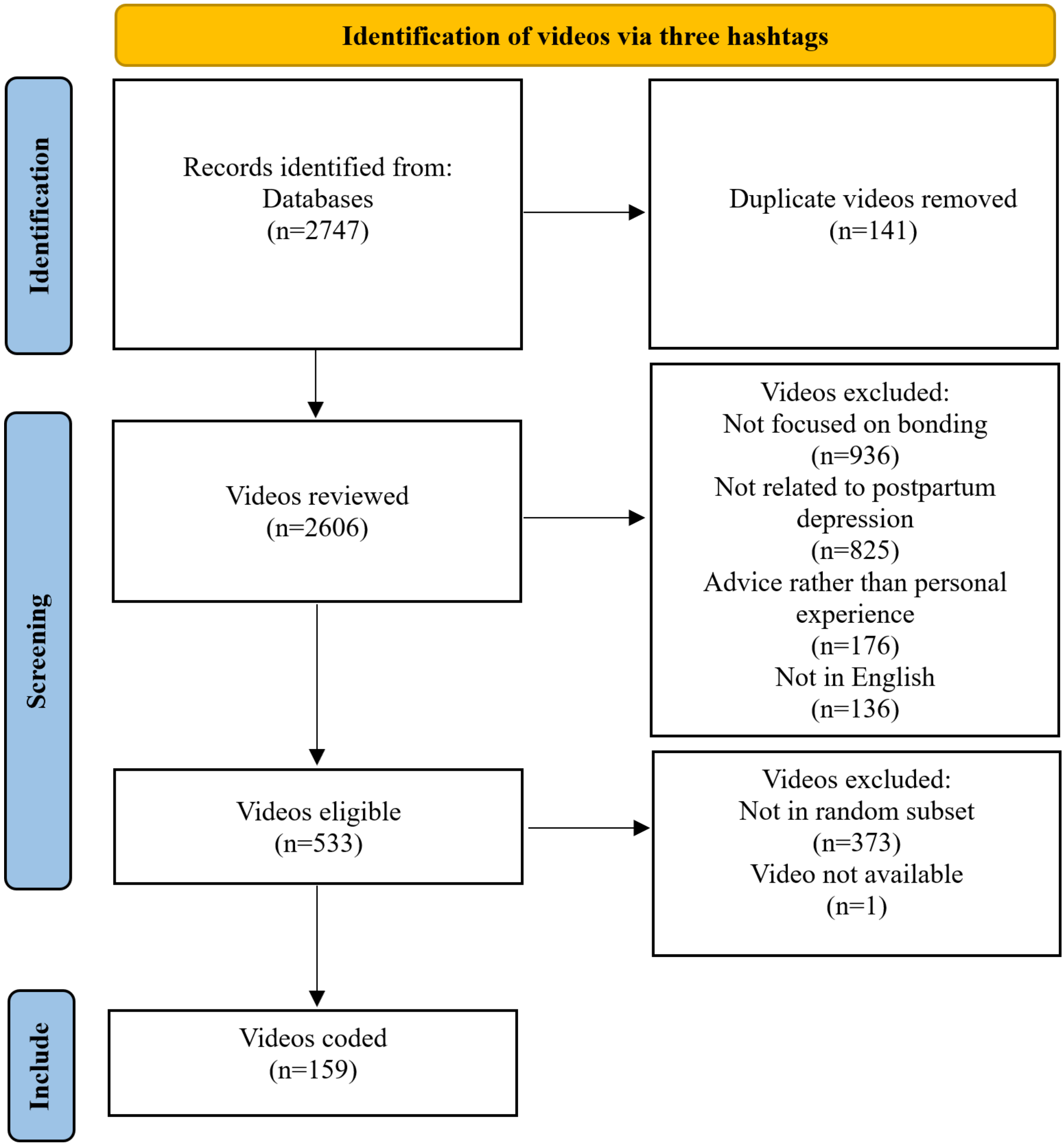
Because engagement metrics were skewed, we fit a negative binomial regression. Our dependent variables included likes, views, shares, and comment counts. The independent variables were the categories of video content. We also included the content creator's number of followers (log-scaled to account for overdispersion) as an explanatory variable because it is generally associated with video engagement. We included all independent variables simultaneously (ie, no stepwise regression) in the model. We used Python (version 3.8.8; Python Software Foundation) for data analysis. Statistical tests were 2-sided with significance set at $P<.05$.

Results

Figure 1 outlines the review process that resulted in 533 relevant videos.

XSL•FO
RenderX

Figure 1. Flow diagram of included videos.



Characteristics of the Included Short-Form Videos

The 159 videos from 147 content creators in our analytic sample had a median of 42,200 (IQR 10,800 - 15,6750) views, 4347 (IQR 970-12,100) likes, 35 (IQR 5-196) shares, and 80 (IQR 26-365) comments. Altogether, these videos have been viewed over 35 million times. Video creators (n=147) had a median of 31,800 (IQR 5265 - 137,900) followers ranging from 46 to over 1,700,000. Creators were primarily based in the United States (141/159, 88%). Other countries represented included the United Kingdom (7/159, 4%), Canada (5/159, 3%), Australia (2/159, 1%), South Africa (1/159, 1%), the Philippines (1/159, 1%), and New Zealand (1/159, 1%). The country of one creator could not be determined (1/159, 1%). Most mothers experienced postpartum depression during the first year of their child's life (147/159, 82%). Mothers of toddlers, most of whom had children younger than 2 years of age, were also represented (17/147, 12%). In some cases, the age of the children could not be determined (9/147, 6%).

Content Categories

Overview

We abstracted 6 categories from the video codes, which are described below. Table 1 lists categories with their respective codes and examples.

XSL•FO
RenderX

Table . Categories with code examples of the content of depressed mothers' bond with their child.

Category, code	Example video description
Navigating anxiety and anger	
Obsessive thoughts or worries about the child's health or safety and compulsive behaviors to cope with it	The video shows a mother enacting an internal argument with her brain regarding her child's safety when a family member holds the baby. Her brain worries that another family member will drop her baby and wants to take the baby back. Despite the creator trying to rationalize that the baby is safe in an effort to quell her intrusive thoughts, she feels a panic attack beginning.
Obsessive thoughts or worries of harming the baby and compulsive behaviors to cope with it	The creator was visibly upset as she shared ego-dystonic intrusive thoughts of harming her baby in the first few months postpartum. These included thoughts of dropping her baby. She could not control these thoughts, resulting in her having panic attacks and needing another person to stay with her. She reports being afraid to tell people, including her husband, but eventually, she did open up to her obstetrician.
Feelings of anger or aggression toward the child	The video shows a mother holding her head as she seems to reminisce. Text overlaid on the video describes a memory of wanting to throw her baby across the room when rocking the baby to sleep. But instead, she threw a piece of exercise equipment. Mother mouths "crazy" along with the lyrics from a song playing.
Involuntary emotional or mental responses	Video of mother verbally describing her postpartum depression relapse. She reports crying whenever she sees her child for no apparent reason. She starts tearing up after changing the camera angle from herself to show her baby.
Creating physical and emotional boundaries	
Mothers put up physical or emotional boundaries for fear of harming the baby	The creator stares at the camera as it zooms in with text overlaid on the video. Text relays an instance when she was in distress, waiting in the closet for a relative to come so she could avoid hitting her child. Music plays in the background.
Mothers put up physical or emotional boundaries for relief	The creator states that it is okay for moms to want to hide and get a break. Music in the background has lyrics about needing "me time" and "free time."
Mothers try to minimize their child's exposure to their distress or hope it does not or did not cause distress	The video shows a mother who appears to be putting the final touches on her appearance with text that she puts on a happy face for her baby despite going through postpartum depression and wanting to "curl in a ball and sleep."
Overstimulated by too much sensory input from the child	The mother mentions that hearing her child cry causes her physical pain and agony.
Escaping from the caregiving role	Mother is tearful with overlaid text that she could never understand how people could leave their children behind until she had postpartum depression.
Overwhelmed by demands of caregiving	
Overwhelmed by caregiving responsibilities	The video shows a mother initially smiling but then gazing into the distance, thinking about something deeply. As she thinks, text bubbles appear and pile up on one another. These thoughts state things like: "What if she falls?"; "laundry"; "My house is a mess"; and "I shouldn't have yelled." Among the many thoughts piling up are: "I can't keep this up" and "Am I a good mom?"
Depression makes it difficult for the mother to take care of their child	The video shows a mother's internal struggle with her feelings of depression and lack of motivation to get off the couch to care for her child, as well as feelings of guilt about being an incompetent mother. Mother shares the thoughts racing in her mind, such as: "You're a bad mom for not playing with him."
Breastfeeding as a burdensome and distressing experience	The video shows a mother with a can of formula with the text: "I'm not okay. I bought a can of formula last night because I'm so burnt out over nursing." Mother covers her face with her hands near the end of the video as if exhausted or ashamed with text that she "feel[s] like a failure for quitting."

XSL•FO
RenderX

Category, code	Example video description
Not feeling the expected emotional connection or enjoyment with the child	The mother uses sound bite, saying, "This is how it feels?!" to express her surprise after finally feeling a strong connection to her child after struggling with postpartum depression (written in text).
Subverted expectations	
Feelings of guilt or regret regarding their care or relationship with their child	The video shows a montage of videos of the creator's child as a newborn. In the video text, she expresses regret that she could not cherish her child's newborn phase because she was overtired and overwhelmed.
Feel incompetent as a mother	The video shows a mother on one side of a door embracing her child with text that she is trying to be a good mom. On the other side of the door, she personifies postpartum depression, trying to break through the door to disrupt the mother's efforts.
Enduring and finding strength through the challenge of postpartum depression	
Mother finds hope or resilience through their relationship with their child	The video shows a collage of photos of a mother and baby appearing happy. The text says: "Postpartum was not good to me, but she is the sole reason I'm still here and always will be."
Responsibility or duty to the child	The video has narration stating difficulty getting out of bed but having to do so "to keep them [children] alive." The video goes on to say that although the mother feels like she is "hanging on by a thread," she continues to get up each day for her children.
Positive feelings or thoughts toward the child or a positive relationship with the child despite postpartum depression	The mother poses a question to herself, asking if she wants to see life with her baby even at the cost of struggling with postpartum depression and thoughts of suicide. After she answers yes, a montage of photos and videos of happy moments of the mother and baby together begins.
Can't remember early life	
Postpartum depression makes it hard to remember child's early life	The video shows a mother pretending to pull a knife out of her chest with text that she could not remember the newborn stage of her child's life because she was "so depressed." In the video description, the creator writes that she wishes she had taken more videos and pictures during that time.

Navigating Anxiety and Anger

Mothers with depression shared feelings of anxiety and anger when interacting with their children. Many mothers experienced perseverative intrusive thoughts of something terrible happening to their child. These thoughts frequently related to the child passing away while asleep, though scenarios varied from the child choking while eating to others making the child sick. One creator mouthed a sound bite to recreate a conversation with her obstetrician, who instructed her that pulling over her car every time her baby fell asleep to check if her baby was breathing was a sign of postpartum anxiety. Some mothers had ego-dystonic intrusive thoughts of harming their baby, but more often, anger toward the child was a genuine feeling. Triggers for anger included difficulties with child sleep, crying, screaming, and sensory overstimulation. Still, more broadly, many mothers noted that their thoughts and feelings arose without warning. Only a few mothers explicitly mentioned anxiety or anger as a manifestation of postpartum depression. For example, one mother said, "It can come in so many different forms: anger, rage, depression, sadness."

In reaction to anxiety and anger, mothers engaged in behaviors to protect their children. For example, mothers compulsively checked if their child was safe while sleeping or vigilantly watched them, often resulting in poor sleep; some used baby monitoring devices to try to alleviate their fears. Mothers sometimes displaced their anger by hitting objects or themselves. In addition, they physically separated themselves from their child if they had these thoughts of hurting them. Several mothers brought up reluctance to share these feelings with their doctors and family members for fear of negative judgment.

Creating Physical and Emotional Boundaries

Mothers exhibited avoidant behaviors or attitudes toward their children. Some mothers expressed the need for a break from caregiving for personal relief due to childcare-related stress. Sensory overload, including loud noises and frequent touch, was a common stressor. One mother described feeling physically in pain due to hearing their child's cries. Other mothers needed to physically relocate in order to calm their fluctuating emotions (sadness, anger, and frustration). For example, some women who felt frustrated would remove themselves from the vicinity of their child to avert lashing out at the child. In addition to personal relief, women expressed fear that their emotional state may negatively impact their child. For example, some creators disclosed that they would conceal their own distress by leaving the room when crying or putting on a happy façade in hopes of minimizing the impact their observable distress may have on their child. Other women described having frequent fantasies where they are forced out of a caregiving role (eg, sickness, suicidal thoughts, and running away).

Overwhelmed by Demands of Caregiving

Mothers struggling with postpartum depression were often overwhelmed by daily caregiving tasks, causing them to feel

XSL•FO
RenderX

depleted or on the verge of a breakdown. As the tasks piled up on each other, many shared a feeling of inadequacy, doubting their competence as mothers. These tasks include feeding, diapering, meal preparation, toileting, and managing sleep routines. Challenges with breastfeeding were pronounced, leaving mothers feeling incapable, which further worsened their mental health. Consequently, some moms switched to formula feeding, though they felt guilty about whether this was best for their child. Caregiving was often described as all-consuming, limiting mothers' time to care for themselves, as evidenced by several videos showing mothers alone in a secluded area (eg, bathroom) away from their children. Relatedly, many women noted a lack of caregiving support, and several videos featured women caring for multiple children. Depressive symptoms such as low energy made it hard to care for their child as desired. Some women were able to get support from family at times when depression left them unable to take care of their children.

Subverted Expectations

The gap between the idealized image of motherhood and experiencing motherhood with postpartum depression left women with negative images of themselves as a mother. Due to depression, mothers expressed being unable to experience the joy and bond with their children that they anticipated. Consequently, mothers expressed feeling incompetent, with many creators asking, "Am I a good mother?" Mothers revealed feeling shame and guilt over their perceived nonconventional emotional state. One mother recounted expecting an instantaneous bond with her child and described feeling shame and guilt when she did not feel the connection early on. Depression also caused tiredness and sensitive emotional states, making it difficult for mothers to fulfill their own image of an ideal parent. Subverted expectations affected mothers both in the present and past. Some mothers expressed regret over missing out on the regular interactions they believed a parent and child should experience in early life due to depression or related tiredness, feeling overwhelmed, or delayed bonding.

Enduring and Finding Strength Through the Challenge of Postpartum Depression

Mothers with postpartum depression shared how they found the strength to persevere through their relationship with their children. Despite struggling to have the energy to get up each day and fulfill the demanding tasks of motherhood due to their depression, these mothers identified that their children were their primary motivating factor to carry on. Often, mothers stated that the smiles on their children's faces were uplifting. Music used in these videos often had resilience-themed lyrics such as "Rise up," "You are my sunshine," and "Don't give up." Some mothers shared a neutral sentiment, feeling that they must take care of their child—and would not let depression hinder their caregiving. Additionally, some mothers expressed their children sparked their desire to seek treatment.

Can't Remember Early Life

Some mothers noticed they had little memory of their child's early life months. These missing memories include important milestone moments, as well as day-to-day caregiving. With this realization came deep sadness that postpartum depression had robbed these special moments from them. This realization often came when their child was older or as their depression began to improve. Frequently, mothers were only reminded of these memories when presented with a picture of their child. Some women used photo montages in their videos to symbolize missed memories.

User Engagement

User engagement statistics for these categories across all videos are provided in Table 2. Overall, videos featuring content on navigating anxiety and anger, as well as subverted expectations, received the most views, likes, and comments. Users most frequently shared videos on subverted expectations and creating physical and emotional boundaries. Sensitivity analyses using normalized (log-scaled) follower counts found similar trends (Table S2 in Multimedia Appendix 1).

Table . User engagement of videos by content category.

Category	Videos, n (%)[a]	Views, median (IQR)	Likes, median (IQR)	Shares, median (IQR)	Comments, median (IQR)
Overwhelmed by demands of caregiving	64 (40)	39,700 (10,586 - 142,300)	3966 (1047 - 14,900)	52 (7 - 282)	128 (42 - 550)
Subverted expectations	56 (35)	50,050 (17,575 - 171,275)	4419 (1191 - 17,225)	80 (6 - 292)	147 (31 - 375)
Navigating anxiety and anger	53 (33)	67,600 (18,800-155,800)	6123 (1654 - 13,600)	49 (5 - 290)	132 (40 - 532)
Enduring and finding strength through the challenge of postpartum depression	26 (16)	27,600 (11,025-86,925)	1856 (553 - 10,734)	10 (1 - 110)	40 (16 - 285)
Creating physical and emotional boundaries	25 (16)	42,200 (10,800-91,800)	2774 (867 - 9932)	69 (4 - 223)	77 (25 - 315)
Can't remember early life	10 (6)	32,450 (25,125-211,075)	4180 (1970 - 6808)	24 (10 - 89)	36 (17 - 97)

[a]Data are expressed as n (%) of 159 videos because a single video can have multiple categories.

XSL•FO
RenderX

In regression analyses, videos on navigating anxiety and anger and subverted expectations increased the views, likes, shares, and comments by 61% to nearly 400% (Table 3). Content featuring the overwhelmed by demands of caregiving category also increased likes, shares, and comments, but not views. Content on resilience in the face of postpartum depression was more highly liked (rate ratio 1.60, 95%CI 1.00 - 2.57) but less likely to receive comments (rate ratio 0.48, 95%CI 0.30 - 0.78). Across all metrics, users engaged less with videos on creating physical and emotional boundaries. We found a small negative association between the log number of followers and all engagement types. Sensitivity analysis of creators with less than the median number of followers mostly corroborated these main findings with the following exceptions: the associations of enduring and finding strength through the challenge of postpartum depression, overwhelmed by demands of caregiving categories, and number of followers with engagement were consistently negative, nonsignificant, and positive, respectively (Table S3 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Association between video content categories and user engagement.

Variable	Risk ratio (95% CI)	*P* value
Views		
Overwhelmed by demands of caregiving	1.12 (0.78 - 1.61)	.53
Subverted expectations	1.72 (1.22 - 2.43)	.002[a]
Navigating anxiety and anger	1.61 (1.09 - 2.38)	.02[a]
Enduring and finding strength through the challenge of postpartum depression	0.75 (0.46 - 1.20)	.23
Creating physical and emotional boundaries	0.40 (0.26 - 0.61)	<.001[a]
Can't remember early life	0.95 (0.47 - 1.92)	.88
Creator's number of followers[b]	0.92 (0.85 - 0.98)	.02[a]
Likes		
Overwhelmed by demands of caregiving	1.61 (1.13 - 2.31)	.009[a]
Subverted expectations	3.61 (2.55 - 5.11)	<.001[a]
Navigating anxiety and anger	3.96 (2.69 - 5.85)	<.001[a]
Enduring and finding strength through the challenge of postpartum depression	1.60 (1.00 - 2.57)	.052
Creating physical and emotional boundaries	0.30 (0.19 - 0.46)	<.001[a]
Can't remember early life	1.45 (0.72 - 2.94)	.30
Creator's number of followers[b]	0.89 (0.83 - 0.96)	.002[a]
Shares		
Overwhelmed by demands of caregiving	1.78 (1.24 - 2.55)	.002[a]
Subverted expectations	2.95 (2.09 - 4.18)	<.001[a]
Navigating anxiety and anger	2.45 (1.66 - 3.61)	<.001[a]
Enduring and finding strength through the challenge of postpartum depression	1.09 (0.68 - 1.76)	.72
Creating physical and emotional boundaries	0.63 (0.41 - 0.98)	.04[a]
Can't remember early life	0.87 (0.43 - 1.77)	.70
Creator's number of followers[b]	0.89 (0.83 - 0.95)	.001[a]
Comments		
Overwhelmed by demands of caregiving	2.67 (1.87 - 3.83)	<.001[a]
Subverted expectations	2.78 (1.97 - 3.94)	<.001[a]
Navigating anxiety and anger	1.89 (1.28 - 2.79)	.001[a]
Enduring and finding strength through the challenge of postpartum depression	0.48 (0.30 - 0.78)	.003[a]
Creating physical and emotional boundaries	0.44 (0.29 - 0.68)	<.001[a]
Can't remember early life	0.51 (0.25 - 1.05)	.07
Creator's number of followers[b]	1.00 (0.93 - 1.07)	.99

[a]Values are significant.

[b]Number of followers is in log scale.

XSL•FO
RenderX

Discussion

Principal Findings

This qualitative content analysis is the first study to examine the MIB in the context of postpartum depression using short-form videos. Nearly 150 mothers shared their experiences bonding with their children, resulting in 6 categories that advance understanding of this topic. Analysis of engagement metrics across categories added information on topics that resonated most with users. These findings support the use of short-form videos to enhance our understanding of lived experiences.

Video Content

Many of these categories reflect themes in the literature. The subverted expectations category aligns with prior studies discussing the clash between mothers' personal or societal expectations and the lived reality of postpartum depression [31,32]. Regarding bonding specifically, we found women worried that they were not the "good mother" they expected and felt associated guilt and shame. Unrealistic expectations of having an instantaneous emotional connection with their child contributed frequently to their mothering assessment. Regret was another prominent feeling, with mothers reporting remorse for missing out on moments with their children. As in Beck's [12] paper, depression was frequently described as "robbing" moms of precious time with their young children.

Several studies report that mothers with depression feel overwhelmed by caregiving responsibilities. Our results similarly find that the symptoms of depression, such as lack of energy, low mood, and lack of joy, compound the everyday challenges of raising a young child [12,33]. We found that poor maternal sleep and breastfeeding often contributed to mothers feeling like they could not meet these demands. Mothers could not refuel because they often could not take a break, or as one mother described, get "me" time.

Although anxiety and anger have long been recognized as manifestations of postpartum depression [34,35], few women in this study directly associated their anxiety or anger with their depression. Previous qualitative and quantitative studies of postpartum depression [35,36] describe symptoms such as intense intrusive thoughts and outbursts of anger, both provoked and unprovoked, which align with the experiences categorized as navigating anxiety and anger in this study. The low attribution of these symptoms to depression in this study might result from greater awareness and screening for postpartum depression compared to anxiety or anger. Therefore, even if anxiety or anger are sequelae or independent conditions, mothers might not have perceived them as symptoms of depression.

Despite this, mothers commonly described anxiety and anger as negatively affecting their bond with their children. Experiences of obsessions and compulsions, as highlighted in other qualitative studies [12] and our research, demonstrate the significant role that postpartum-specific anxieties have on the MIB. Indeed, postpartum anxiety impairs bonding, leading to pathological anger and infant-focused anxiety [37]. Prior research [12,36] has also shown that depressed mothers experience anger but find ways to avoid taking it out on their children, akin to how mothers on TikTok coped.

The creating physical and emotional boundaries category, characterized by behaviors or wishes to be separated from the infant, shared some similarities with a study of posts on a postpartum depression internet forum [17]. That study found that desires to be separated from the child manifested with thoughts or feelings of hurting the child. However, mothers in this study were more avoidant rather than aggressive. In addition, though some mothers noted thoughts of suicide, descriptions were less explicitly expressed in videos than in internet forum posts. These differences may be due to reluctance to share these thoughts by video, which is less identity-preserving.

In contrast, the enduring and finding strength through the challenge of postpartum depression and can't remember early life categories are less represented in the literature. Prior studies similarly found that mothers with postpartum depression expressed positive feelings toward their child or felt an obligation to meet their child's needs despite postpartum depression. Our results go a step further: for some mothers, their bond with their child offered comfort and reassurance. We know of a similar finding in one study of West African mothers residing in the United Kingdom who described their baby as a source of pride or distraction from distress [38]. Some mothers had difficulty remembering their child's early life because they were trying to "survive." Without these memories that connect and strengthen the bond between mother and child, these mothers felt deep sadness and regret. Traumatic events, such as childbirth trauma, can cloud memory, but only 2 included videos in this category mentioned trauma. Given the lack of prior literature on this phenomenon, it should be further explored in future work.

Determining mothers' motivations for sharing these personal experiences on social media is challenging without direct inquiry. Studies have shown that self-expression is a primary motivation for people with mental health conditions to share content on social media [39-41]. Creators, particularly on platforms like TikTok, are often willing to share personal experiences because they perceive a safe environment free from negative judgment [39,40]. This perceived safety encourages those with stigmatized conditions or identities to share their experiences [39,42]. Other motivations, such as the desire to receive social support or gain social status, are possible and could influence the type of content shared. The congruence of our findings with previous research using interviews, including in women diagnosed with clinical depression [12], lends credibility to our approach of analyzing short-form videos. Future research should interview creators to better understand their motivations and how these motivations influence the content they produce.

The technological features of social media platforms likely influence how women express their lived experiences. Women used the various multimedia affordances provided by the video-sharing platform to share their experiences. In addition to text, they often added sound bites or music to convey meaning. Visual effects such as photo montages and

XSL•FO
RenderX

re-enactments of how postpartum depression affected bonding were another way mothers used the medium. Although most videos were less than 30 seconds, this multimodal content added to the richness of the content.

User Engagement

Videos featuring 3 categories well-represented in the literature—subverted expectations, navigating anxiety and anger, and overwhelmed by caregiving—received more views, likes, shares, and comments. According to the uses and gratifications theory [43], people engage with content that meets their desires and needs. Users "like" content to express appreciation for the material they find relevant, interesting, or emotionally engaging [39,44]. Users are more likely to engage with mental health content that shares personal experiences and experiential knowledge [45,46]. This trend is noticeable on TikTok, where users frequently share personal stories and information, often presented in creative or humorous ways. In addition, users are drawn to experiences that resonate with their own [39,47]. The challenges of motherhood, many of which are unexpected, are relatable experiences represented by these categories. However, in postpartum depression, these challenging experiences pervade the mother's thoughts and actions, impairing the bond with their child.

Decreased likes, shares, and comments for the creating physical and emotional boundaries category indicate that users were less inclined to engage with videos about taking a break or fantasizing about separation from their child, possibly due to the stigma associated with appearing to want to abandon their responsibilities as a mother. However, content featuring anger generated more engagement, suggesting that anger without the expectation of abandoning the mother role may be the key difference. Content on resilience was well-received, nearing significance for increased likes. However, these videos are less likely to be commented on. This may reflect that users wanted to show support but may not have personally felt resilient or that it was actionable (eg, sharing advice). A study of depression on the Chinese version of TikTok found parallel results where positive portrayals of general depression had less engagement, which was attributed to positive content garnering less sympathy [48].

Our findings are further strengthened by the observation that videos from creators with small or modest followings were well-represented and received significant engagement. This suggests that diverse perspectives on postpartum depression, not just those from influencers with large followings, are valued. This could be attributed to TikTok's algorithm, which promotes videos that people consistently watch and interact with to a broader audience, resulting in increased engagement. The potential of content creators with small followings to create a viral video is one of the draws for people to create content on TikTok [49].

Clinical Implications

Our findings provide practical ways for clinicians to support bonding in depressed mothers. Across all these categories, mothers with postpartum depression tried various ways to cope with the challenges of bonding with their children. However, internalization of their bonding difficulties caused self-blame with associated shame and guilt, as well as isolation. Normalizing mothers' experiences, given their commonality and the resulting shame, is therapeutic. For instance, clinicians can reassure mothers that it often takes time to develop an emotional bond with their child. Based on the high engagement, these videos may already be beneficial by normalizing these experiences. Self-compassion approaches that reduce shame and guilt effectively enhance bonding and reduce depression [50,51]. Sharing ways to improve sleep for moms and infants and strengthen social support would be helpful for overwhelmed mothers, especially those with multiple young children. Health care professionals can change the narrative by emphasizing that not every mother can breastfeed.

Further, our findings indicate that assessment of bonding in the context of postpartum depression should consider anxiety, anger, and parenting stress. Asking only about depressive symptoms will likely miss many women with bonding difficulties. Although clinicians may be reluctant to ask about anger toward the child or the mother's escape fantasies, our results suggest that they should. It is important to acknowledge that mothers' reluctance to share their struggles because of fears about child protective services removing their children is not without basis. Informing mothers on the difference between what is considered maltreatment and what is not may provide reassurance. We believe giving mothers information and allowing them to share as they so choose is a beneficial intervention within itself and opens the door to future conversations.

Limitations

Although our findings are well supported by prior work, only publicly available videos with many views were reviewed. This enabled us to examine videos that resonate with many users, but further work is needed to determine if these categories are truly representative of bonding for women struggling with depression. In addition, while many social media researchers use hashtags to identify and learn about specific phenomena [22], it is unknown whether mothers in the videos have depressive symptoms or clinical depression. Nonetheless, subclinical depressive symptoms alone can negatively affect bonding [52]. In addition, demographic information such as maternal age, race, or sex could not be determined. We did not see videos featuring men with postpartum depression in our video corpus, highlighting an area for future investigation. The vast majority of women were from the United States followed by other Western countries. This is likely because we required videos to be in English. The concurrent findings in mothers from non-English speaking cultures are reassuring, but our approach limits generalizability.

Another limitation is the relatively low but acceptable IRR for the codes. The low IRR reflects the challenges of applying 21 nonmutually exclusive potential codes to short-form videos. Moreover, these codes were applied to 30 overlapping videos used for calculating IRR between each pair of coders, resulting in some of the codes being rarely applied. Consequently, even a single discrepancy between 2 coders when applying a code could lead to a moderate Krippendorff α reliability score. Two of the 3 codes with the lowest IRR of 0 ("Mother perceives their

XSL•FO
RenderX

child has negative feelings about them" and "Mothers put up physical or emotional boundaries for relief") were derived from the prior literature review. This suggests that these 2 codes may have been less relevant to the current dataset. As a result, one of these codes was excluded when codes were abstracted into categories ("Mother perceives their child has negative feelings about them"). While coders conducted several rounds of review to determine if initial deductive codes were present, further rounds may have helped validate these codes and improve IRR."

Conclusions

Overall, our research demonstrates that short-form videos provide a meaningful lens through which to explore mothers' experiences with postpartum depression, particularly its impact on the MIB. Knowledge gleaned helps us understand mothers' specific challenges and thereby develop ways to support them.

Acknowledgments

This work is supported by the UCLA Depression Grand Challenges Judith Young Fellowship awarded to KS. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

Datasets generated or analyzed during this study cannot be made publicly available due to potentially identifying information (eg, biometric) that may compromise privacy.

Authors' Contributions

KS conceived and designed the study. All authors reviewed and coded videos. All authors contributed to the interpretation of data. KS conducted the statistical analysis. All authors contributed to manuscript drafting and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Codes and definitions of content of depressed mothers' bond with their child; normalized user engagement of videos by content category; association between video content categories and user engagement for creators with below the median number of followers; and references.

[DOCX File, 37 KB - infodemiology_v5i1e59125_app1.docx]

References

1. Radoš SN, Hairston I, Handelzalts JE. The concept analysis of parent-infant bonding during pregnancy and infancy: a systematic review and meta-synthesis. J Reprod Infant Psychol 2024 Mar;42(2):142-165. [doi: 10.1080/02646838.2022.2162487] [Medline: 36588501]
2. Suzuki D, Ohashi Y, Shinohara E, et al. The current concept of paternal bonding: a systematic scoping review. Healthcare (Basel) 2022 Nov 11;10(11):2265. [doi: 10.3390/healthcare10112265] [Medline: 36421589]
3. Le Bas G, Youssef G, Macdonald JA, et al. The role of antenatal and postnatal maternal bonding in infant development. J Am Acad Child Adolesc Psychiatry 2022 Jun;61(6):820-829. [doi: 10.1016/j.jaac.2021.08.024] [Medline: 34555489]
4. Le Bas GA, Youssef GJ, Macdonald JA, et al. Maternal bonding, negative affect, and infant social-emotional development: a prospective cohort study. J Affect Disord 2021 Feb 15;281:926-934. [doi: 10.1016/j.jad.2020.11.031] [Medline: 33229017]
5. Branjerdporn G, Meredith P, Strong J, Garcia J. Associations between maternal-foetal attachment and infant developmental outcomes: a systematic review. Matern Child Health J 2017 Mar;21(3):540-553. [doi: 10.1007/s10995-016-2138-2] [Medline: 27456309]
6. Sierau S, Dähne V, Brand T, Kurtz V, von Klitzing K, Jungmann T. Effects of home visitation on maternal competencies, family environment, and child development: a randomized controlled trial. Prev Sci 2016 Jan;17(1):40-51. [doi: 10.1007/s11121-015-0573-8] [Medline: 26103919]
7. Mason ZS, Briggs RD, Silver EJ. Maternal attachment feelings mediate between maternal reports of depression, infant social–emotional development, and parenting stress. J Reprod Infant Psychol 2011 Sep;29(4):382-394. [doi: 10.1080/02646838.2011.629994]
8. Behrendt HF, Scharke W, Herpertz-Dahlmann B, Konrad K, Firk C. Like mother, like child? Maternal determinants of children's early social-emotional development. Infant Ment Health J 2019 Mar;40(2):234-247. [doi: 10.1002/imhj.21765] [Medline: 30731022]

XSL•FO
RenderX

9. Fransson E, Sörensen F, Kunovac Kallak T, et al. Maternal perinatal depressive symptoms trajectories and impact on toddler behavior—the importance of symptom duration and maternal bonding. J Affect Disord 2020 Aug 1;273:542-551. [doi: 10.1016/j.jad.2020.04.003] [Medline: 32560952]
10. Wittkowski A, Vatter S, Muhinyi A, Garrett C, Henderson M. Measuring bonding or attachment in the parent-infant-relationship: a systematic review of parent-report assessment measures, their psychometric properties and clinical utility. Clin Psychol Rev 2020 Dec;82:101906. [doi: 10.1016/j.cpr.2020.101906] [Medline: 32977111]
11. Mathews TL, Emerson MR, Moore TA, Fial A, Hanna KM. Systematic review: feasibility, reliability, and validity of maternal/caregiver attachment and bonding screening tools for clinical use. J Pediatr Health Care 2019 Nov;33(6):663-674. [doi: 10.1016/j.pedhc.2019.04.018]
12. Beck CT. Postpartum depressed mothers' experiences interacting with their children. Nurs Res 1996;45(2):98-104. [doi: 10.1097/00006199-199603000-00008] [Medline: 8604372]
13. Doyle FL, Dickson SJ, Eapen V, et al. Towards preventative psychiatry: concurrent and longitudinal predictors of postnatal maternal-infant bonding. Child Psychiatry Hum Dev 2023 Dec;54(6):1723-1736. [doi: 10.1007/s10578-022-01365-0] [Medline: 35616764]
14. Progga FT, Rubya S. "Just like therapy!": investigating the potential of storytelling in online postpartum depression communities. Presented at: GROUP '23; Jan 8-11, 2023; Hilton Head, United States. [doi: 10.1145/3565967.3570977]
15. Yao X, Mikhelson M, Micheletti M, et al. Understanding postpartum parents' experiences via two digital platforms. Proc ACM Hum-Comput Interact 2023 Apr 14;7(CSCW1):1-23. [doi: 10.1145/3579540]
16. Moore D, Ayers S, Drey N. A thematic analysis of stigma and disclosure for perinatal depression on an online forum. JMIR Ment Health 2016 May 19;3(2):e18. [doi: 10.2196/mental.5611] [Medline: 27197516]
17. Kantrowitz-Gordon I. Internet confessions of postpartum depression. Issues Ment Health Nurs 2013 Dec;34(12):874-882. [doi: 10.3109/01612840.2013.806618] [Medline: 24274243]
18. Evans M, Donelle L, Hume-Loveland L. Social support and online postpartum depression discussion groups: a content analysis. Patient Educ Couns 2012 Jun;87(3):405-410. [doi: 10.1016/j.pec.2011.09.011] [Medline: 22019021]
19. Thanks a billion!. TikTok Newsroom. URL: https://newsroom.tiktok.com/en-us/1-billion-people-on-tiktok [accessed 2023-12-26]
20. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2006 Mar;5(1):80-92. [doi: 10.1177/160940690600500107]
21. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
22. Chancellor S, De Choudhury M. Methods in predictive techniques for mental health status on social media: a critical review. NPJ Digital Med 2020;3(1):43. [doi: 10.1038/s41746-020-0233-7] [Medline: 32219184]
23. McCashin D, Murphy CM. Using TikTok for public and youth mental health—a systematic review and content analysis. Clin Child Psychol Psychiatry 2023 Jan;28(1):279-306. [doi: 10.1177/13591045221106608] [Medline: 35689365]
24. Basch CH, Donelle L, Fera J, Jaime C. Deconstructing TikTok videos on mental health: cross-sectional, descriptive content analysis. JMIR Form Res 2022 May 19;6(5):e38340. [doi: 10.2196/38340] [Medline: 35588057]
25. Schaadhardt A, Fu Y, Pratt CG, Pratt W. "Laughing so I don't cry": how TikTok users employ humor and compassion to connect around psychiatric hospitalization. Presented at: 2023 CHI Conference on Human Factors in Computing Systems; Apr 23-28, 2023; Hamburg, Germany. [doi: 10.1145/3544548.3581559]
26. Lookingbill V, Le K. "There's always a way to get around the guidelines": nonsuicidal self-injury and content moderation on TikTok. Social Media + Soc 2024 Apr;10(2):20563051241254371. [doi: 10.1177/20563051241254371]
27. Krippendorff K. Content Analysis: An Introduction to Its Methodology, 4th edition: Thousand Oaks, CA: SAGE Publications, Inc; 2019.
28. Armat MR, Assarroudi A, Rad M, Sharifi H, Heydari A. Inductive and deductive: ambiguous labels in qualitative content analysis. TQR 2018. [doi: 10.46743/2160-3715/2018.2872]
29. Krippendorff K. Reliability in content analysis: some common misconceptions and recommendations. Hum Comm Res 2004 Jul;30(3):411-433. [doi: 10.1111/j.1468-2958.2004.tb00738.x]
30. Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. Nurse Educ Today 2017 Sep;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
31. Mollard EK. A qualitative meta-synthesis and theory of postpartum depression. Issues Ment Health Nurs 2014 Sep;35(9):656-663. [doi: 10.3109/01612840.2014.893044] [Medline: 25162187]
32. Mauthner NS. "Feeling low and feeling really bad about feeling low": women's experiences of motherhood and postpartum depression. Can Psychol 1999;40(2):143-161. [doi: 10.1037/h0086833]
33. Abrams LS, Curran L. "And you're telling me not to stress?" A grounded theory study of postpartum depression symptoms among low-income mothers. Psychol Women Q 2009 Sep;33(3):351-362. [doi: 10.1177/036168430903300309]
34. Ou CH, Hall WA. Anger in the context of postnatal depression: an integrative review. Birth 2018 Dec;45(4):336-346. [doi: 10.1111/birt.12356] [Medline: 29781142]
35. Beck CT, Indman P. The many faces of postpartum depression. J Obstet Gynecol Neonatal Nurs 2005 Sep;34(5):569-576. [doi: 10.1177/0884217505279995]

XSL•FO
RenderX

36. Beck CT. Postpartum depression: a metasynthesis. Qual Health Res 2002 Apr;12(4):453-472. [doi: 10.1177/104973202129120016] [Medline: 11939248]
37. Fallon V, Silverio SA, Halford JCG, Bennett KM, Harrold JA. Postpartum-specific anxiety and maternal bonding: further evidence to support the use of childbearing specific mood tools. J Reprod Infant Psychol 2021 Apr;39(2):114-124. [doi: 10.1080/02646838.2019.1680960] [Medline: 31642692]
38. Gardner PL, Bunton P, Edge D, Wittkowski A. The experience of postnatal depression in West African mothers living in the United Kingdom: a qualitative study. Midwifery 2014 Jun;30(6):756-763. [doi: 10.1016/j.midw.2013.08.001] [Medline: 24016554]
39. Schluchter T. Investigating user perceptions of mental health content on TikTok: a comprehensive exploration. 2024 May 11 Presented at: CHI '24; May 11-16, 2024; Honolulu, United States p. 1-6. [doi: 10.1145/3613905.3647964]
40. Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S. #WhyWeTweetMH: understanding why people use Twitter to discuss mental health problems. J Med Internet Res 2017 Apr 5;19(4):e107. [doi: 10.2196/jmir.6173] [Medline: 28381392]
41. Oeldorf-Hirsch A, Sundar SS. Social and technological motivations for online photo sharing. J Broadcast Electron Media 2016 Oct;60(4):624-642. [doi: 10.1080/08838151.2016.1234478]
42. Messina N. It's there for you, but is it #FYP?: a methodological reflection of researching content vs. community on TikTok. SPIR 2021. [doi: 10.5210/spir.v2021i0.11982]
43. Ruggiero TE. Uses and gratifications theory in the 21st century. Mass Commun Soc 2000 Feb;3(1):3-37. [doi: 10.1207/S15327825MCS0301_02]
44. Lowe-Calverley E, Grieve R. Thumbs up: a thematic analysis of image-based posting and liking behaviour on social media. Telematics Inf 2018 Oct;35(7):1900-1913. [doi: 10.1016/j.tele.2018.06.003]
45. Martini T, Czepielewski LS, Baldez DP, et al. Mental health information online: what we have learned from social media metrics in BuzzFeed's Mental Health Week. Trends Psychiatry Psychother 2018;40(4):326-336. [doi: 10.1590/2237-6089-2018-0023] [Medline: 30156651]
46. Choi B, Kim H, Huh-Yoo J. Seeking mental health support among college students in video-based social media: content and statistical analysis of YouTube videos. JMIR Form Res 2021 Nov 11;5(11):e31944. [doi: 10.2196/31944] [Medline: 34762060]
47. Milton A, Ajmani L, DeVito MA, Chancellor S. "I see me here": mental health content, community, and algorithmic curation on TikTok. Presented at: CHI '23; Apr 23-28, 2023; Hamburg, Germany. [doi: 10.1145/3544548.3581489]
48. Li J, Tang L, Pu Y. My story of depression: a content analysis of autobiographic videos on Douyin. Health Commun 2024 May;39(5):906-914. [doi: 10.1080/10410236.2023.2191887] [Medline: 36941529]
49. Le Compte D, Klug D. "It's viral!"—a study of the behaviors, practices, and motivations of TikTok users and social activism. Presented at: CSCW '21; Oct 23-27, 2021. [doi: 10.1145/3462204.3481741]
50. Fernandes DV, Canavarro MC, Moreira H. The role of mothers' self-compassion on mother-infant bonding during the COVID-19 pandemic: a longitudinal study exploring the mediating role of mindful parenting and parenting stress in the postpartum period. Infant Ment Health J 2021 Sep;42(5):621-635. [doi: 10.1002/imhj.21942] [Medline: 34407224]
51. Jefferson FA, Shires A, McAloon J. Parenting self-compassion: a systematic review and meta-analysis. Mindfulness (N Y) 2020 Sep;11(9):2067-2088. [doi: 10.1007/s12671-020-01401-x]
52. Moehler E, Brunner R, Wiebel A, Reck C, Resch F. Maternal depressive symptoms in the postnatal period are associated with long-term impairment of mother-child bonding. Arch Womens Ment Health 2006 Sep;9(5):273-278. [doi: 10.1007/s00737-006-0149-5] [Medline: 16937313]

Abbreviations

IRR: interrater reliability
MIB: maternal-infant bonding
RR: rate ratio

Edited by T Mackey; submitted 06.04.24; peer-reviewed by A Lavallée, A Christoforou; revised version received 14.01.25; accepted 04.02.25; published 15.05.25.

Please cite as:
Sobowale K, Castleman JS, Zhao SY
Postpartum Depression and Maternal-Infant Bonding Experiences in Social Media Videos: Qualitative Content Analysis
JMIR Infodemiology 2025;5:e59125
URL: https://infodemiology.jmir.org/2025/1/e59125
doi:10.2196/59125

XSL•FO
RenderX

© Kunmi Sobowale, Jamie Sarah Castleman, Sophia Yingruo Zhao. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 15.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Global Surveillance of Public Interest in Cosmetic Tourism for Aesthetic Eyelid Surgery Abroad: Cross-Sectional Infodemiology Investigation of Internet Search Trends and Social Media Content

Daniel B Azzam, MBA, MD; Yi Ling Dai, MD; Victoria S North, MD; Alison B Callahan, MD; Katrinka L Heher, MD; Mitesh K Kapadia, MD; M Reza Vagefi, MD

Division of Oculofacial Plastic & Orbital Surgery, New England Eye Center, Tufts Medicine, Biewend Building 11th Fl., 260 Tremont St, Boston, MA, United States

Corresponding Author:
Daniel B Azzam, MBA, MD
Division of Oculofacial Plastic & Orbital Surgery, New England Eye Center, Tufts Medicine, Biewend Building 11th Fl., 260 Tremont St, Boston, MA, United States

Abstract

Background: Global medical tourism for aesthetic surgery has become a popular phenomenon through ease of access in the digital era, though such services are not without potential risks. The application of infodemiology for global health surveillance may provide unique insights into unknown patient travel patterns and surgeon workforce dynamics abroad.

Objective: This study aimed to evaluate American cosmetic tourism trends in oculofacial plastic surgery, including demand profile and qualifications of the most sought-after international eyelid surgeons on social media.

Methods: This cross-sectional infodemiology study queried Google Trends to assess US interests in aesthetic eyelid surgery abroad in 25 destination countries from 2013 to 2023. The highest-rated content posted by 55 eyelid surgeons (US: n=11; international: n=44) on a social media platform (Instagram; Meta Platforms) was evaluated. The main outcomes included Google search volumes for aesthetic eyelid surgery for each destination country, as well as specialty training and professional medical society affiliations of popular eyelid surgeons on social media in each of these countries.

Results: The top 5 destinations Americans sought for aesthetic eyelid surgery abroad were South Korea, Mexico, Canada, Turkey, and China. Interest in eyelid surgery abroad remained stable over the last decade despite 118% growth in blepharoplasty searches. Social media indicated eyelid surgeons abroad were more often general plastic surgeons than in the United States (30/44, 68% vs 2/11, 18%; *P*=.003). US surgeons more frequently completed oculofacial plastics, facial plastics, or aesthetic plastics fellowships compared with international surgeons (9/11, 82% vs 10/44, 23%; *P*<.001) and had membership in professional medical societies (11/11, 100% vs 22/44, 50%; *P*=.002).

Conclusions: American demand for international eyelid surgery remained stable over the past decade despite a 2-fold increase in the US interest for blepharoplasty. Digital epidemiology data reveal a shortage of international surgeons with specialized aesthetic eyelid fellowship training or professional society affiliations on social media among the preferred destinations for Americans seeking aesthetic eyelid surgery. These findings may provide beneficial insights for patients interested in traveling abroad for eyelid surgery, as well as for surgeons or academic societies seeking to increase social media presence or patient-directed educational content via social media engagement.

(JMIR Infodemiology 2025;5:e64639) doi:10.2196/64639

KEYWORDS

cosmetic tourism; oculofacial plastic surgery; aesthetic eyelid surgery; Google Trends; social media; travel medicine; global health; infodemiology; digital epidemiology; Instagram; eyelids; aesthetic; medical tourism; eyelid surgery; plastic surgery; blepharoplasty

Introduction

The digital transformation driven by the internet and social media has reduced international boundaries in medical care, enabling aesthetic surgeons worldwide to market their services to patients abroad for cosmetic tourism [1].

Travel abroad for discounted cosmetic surgery may offer some putative benefits, including anonymity, access to procedures unavailable locally, and all-inclusive medical tourism packages. However, such services are not without potential risks or untoward consequences. Considerations include variable preoperative counseling, inadequate follow-up, inconsistent quality of care, language barrier, postoperative care burden on

XSL•FO
RenderX

the home health care system, and financial and legal challenges in the event of complications [2,3]. Furthermore, centers abroad may not hold surgeons to the same training standards or safety regulations as the United States [4].

Eyelid surgery ranks third in aesthetic surgeries worldwide [5]. Google Trends (Alphabet Inc) is a valuable tool for health information–seeking behavior and has been employed to analyze aesthetic surgery demand within oculofacial plastic surgery (OPS), facial cosmetic surgery, facial feminization surgery, and cosmetic tourism for plastic surgery [3,6-8].

Social media is a dominant platform for medical advice, patient education, and business marketing in this field [9]. A study of nearly 400 survey participants showed that 49% (n=193) of patients found their oculofacial plastic surgeon on social media, with more than two-thirds choosing Instagram (Meta Platforms) as the preferred social media platform to find an oculofacial plastic surgeon [10].

Previous studies on social listening within OPS have largely focused on marketing strategies through content category analysis [11]. In this regard, Cheng et al [9] showed that the OPS content category amassing the most views was "live procedure or surgeries" followed by "educational" and "patient experience." Park et al [12] demonstrated that OPS photographs were more successful than videos, carousel presentation was better than collage, and featuring the posting doctor, especially when smiling and wearing a white coat, increased public engagement. Similarly, other social media listing research in aesthetic plastic surgery and ophthalmology have demonstrated the success of various hashtag utilization in social media marketing and the prevalence of inaccurate medical information among patient-generated Reddit discussions [13,14].

This study fills an important gap in the social listening literature for eyelid surgery from the perspective of global health and, to the authors' knowledge, is the first paper examining social media as a space for cosmetic tourism patients to find international OPS providers. The authors aimed to use a digital epidemiology approach to analyze current global health trends in OPS cosmetic tourism sought by Americans over the last decade. Furthermore, this study undertakes an analysis of social media data to compare training backgrounds and professional academic society affiliations of social media's most popular international eyelid surgeons. These findings may provide beneficial insights for patients interested in traveling abroad for eyelid surgery, as well as for surgeons or academic societies seeking to increase social media presence or patient-directed educational content via social media engagement.

Methods

Ethical Considerations

This cross-sectional digital epidemiology study followed the World Medical Association's ethical principles for medical research involving human participants outlined in the Declaration of Helsinki as amended in 2013. Institutional Review Board approval was not required as publicly available data were used. No identifiable patient information was involved in this study, and no compensation to participants took place. These methods were adapted from previous publications [6,13,15,16].

Google Trends

In order to ensure that search terms accurately represent the intent of Americans searching the topic of aesthetic eyelid surgery abroad in an international destination, the Google Trends query was designed in the following way. Google Trends relative search volume (RSV) for US interest in aesthetic eyelid surgery abroad from April 1, 2013, to March 31, 2023, was collected on April 1, 2024. The query output is search volume as arbitrary values ranging from 0 to 100 for each query (maximum 5 search terms at a time) and referenced to the highest peak search popularity (set at 100) for the given terms, region, and time. As such, a standard RSV reference range was implemented across all 25 search terms by using a bracket type elimination where all queries were tested until the maximum was identified (Korea), and therefore Korea was always included as the most popular peak term (100) within each set of 5-term queries to ensure that all RSV values were relative to the same reference point. The location of origin of the Google searches was limited to the US geographic filter, and the "Cosmetic Surgery" category filter was applied to limit the output to aesthetic-surgery–related searches. A variety of search term combinations were generated and tested through trial and error until the term combinations with the greatest output were identified in order to minimize cells with low data counts. As a result, "eyelid nation" was found to produce the most data for temporal country analysis, while "blepharoplasty" produced the most data for geographic analysis in the United States. Quotation marks were not included in any queries.

A total of 25 destination countries were included according to the International Society of Aesthetic Plastic Surgery (ISAPS) Global Survey 2021 most popular cosmetic surgery destinations [5]. As such, the following twenty-five search terms were applied: (1) eyelid Argentina, (2) eyelid Australia, (3) eyelid Brazil, (4) eyelid Canada, (5) eyelid China, (6) eyelid Colombia, (7) eyelid France, (8) eyelid Germany, (9) eyelid Greece, (10) eyelid India, (11) eyelid Italy, (12) eyelid Japan, (13) eyelid Korea, (14) eyelid Mexico, (15) eyelid Netherlands, (16) eyelid Portugal, (17) eyelid Romania, (18) eyelid Russia, (19) eyelid Saudi Arabia, (20) eyelid Spain, (21) eyelid Taiwan, (22) eyelid Thailand, (23) eyelid Turkey, (24) eyelid United Kingdom, and (25) eyelid Venezuela. Control terms as a proxy for levels of general internet traffic in the same time period included "weather," "sports," "google," and "news." RSV data for aesthetic eyelid surgery overall in each US state during this period were extracted using the search topic "blepharoplasty."

To evaluate the most popular destination countries based on average search interest, the average search volume from 2013 - 2023 for each country were compared. To analyze the changes in search interest over time, the following calculations were performed to normalize aggregate RSV data on a standard scale of 0 to 100. First, the search volume of all international countries was summed together for each month from April 2013 to March 2023 to be termed eyelid surgery abroad. The same was done for the control terms. The monthly aggregate data for each category were then normalized to be set on a scale of 0 to

XSL•FO
RenderX

100 in the same fashion performed by Google Trends, which sets the maximum value of each category to 100. Normalization is achieved by dividing each monthly data point by the maximum value of the category, and then multiplying by 100. Next, the normalized monthly data for eyelid surgery abroad, blepharoplasty overall, and control were averaged into seasonal data based on the following seasons: Winter (December, January, and February), Spring (March, April, and May), Summer (June, July, and August), and Fall (September, October, and November).

Social Media

After identifying America's most desired eyelid cosmetic tourism destinations as above, the most popular international aesthetic eyelid surgeons on social media within those countries were identified in the following manner. The social media platform, Instagram, was chosen as it has been previously demonstrated that a majority of patients prefer Instagram for finding an oculofacial plastic surgeon [10].

Instagram was manually queried on May 1, 2024, for top eyelid surgery–related posts using the following search strategy based on previous methodologies [13]. Both hashtags and topic searches, including a variety of medical and layperson terminology, were used for the United States and each of the top 5 countries of highest demand extracted from the Google Trends analysis above (South Korea, Mexico, Canada, Turkey, and China). Through trial and error testing, additional similar hashtags or topic searches were added as needed to generate a minimum of 10 top posts for each country of interest. Instagram uses algorithm-based sorting to show a grid of the 9 top posts for a query based on factors like engagement (likes, comments, and shares), relevance to the query terms, and recency [13,17]. Search terms included the following: (1) #blepharoplastykorea, (2) #eyelidsurgerykorea, (3) #koreaeyelidsurgery, (4) eyelid surgery Korea, (5) #blepharoplastymexico, (6) #eyelidsurgerymexico, (7) #blepharoplastytijuana, (8) #eyelidsurgerytijuana, (9) eyelid surgery Mexico, (10) #eyelidsurgerycanada, (11) eyelid surgery Canada, (12) #blepharoplastyturkey, (13) #eyelidsurgeryturkey, (14) #eyelidsurgeryistanbul, (15) #blepharoplastychina, (16) #eyelidsurgerychina, (17) eyelid surgery China, (18) #blepharoplasty (representing United States), (19) #eyelidsurgery (representing United States), and (20) #eyelidlift (representing United States).

Content analysis was performed via a human annotator evaluating each top post for topic outputs to ensure they aligned with the theme identified for that topic. Inclusion criteria for content included posts relevant to eyelid surgery posted by an eyelid surgeon (ie, promoting services or medical tourism, photographs or videos in the operating room, before and after photos, anatomical diagrams, and patient-directed educational material). Duplicate posts, posts not relevant to eyelid surgery, or posts with an inability to access the appropriate information (blocked access or insufficient surgeon details) were excluded from the analysis. After removing duplicates, 55 unique posts were analyzed. The Instagram profile for each post was then mined for the surgeon's website, or if not listed then the surgeon's practice website was identified using search engine informatics querying the surgeon's name on Google. Data extracted from each surgeon's professional website included procedures, residency, fellowship, country of practice, professional medical society affiliations, warranties, and medical tourism packages. Data collected for membership to professional organizations were verified on the corresponding academic society website member lists and included American Society of Ophthalmic Plastic & Reconstructive Surgery (ASOPRS); Canadian Society of Oculoplastic Surgeons (CSOPS); American Academy of Facial Plastic & Reconstructive Surgery (AAFPRS); Fellow of the American College of Surgeons; American Society of Plastic Surgeons; European Board of Plastic Reconstructive & Aesthetic Surgery; Fellow of the Royal College of Surgeons of Canada; ISAPS; Oriental Society of Aesthetic Plastic Surgery; International Confederation for Plastic, Reconstructive & Aesthetic Surgery; Korean Academy of Facial Plastic and Reconstructive Surgery; Korean Society of Plastic and Reconstructive Surgeons; Asociación Mexicana de Cirugía Plástica Estética y Reconstructiva; and Canadian Society of Plastic Surgeons.

Statistical Analysis

Statistical analysis used Microsoft Excel Version 16.66.1 (Microsoft Corporation) and GraphPad Prism QuickCalcs (GraphPad Software). Descriptive statistics were used to analyze the search interest volume over time, as well as the demographics of the international eyelid surgeons. Chi-square tests compared proportions for categorical variables between international and US eyelid surgeons.

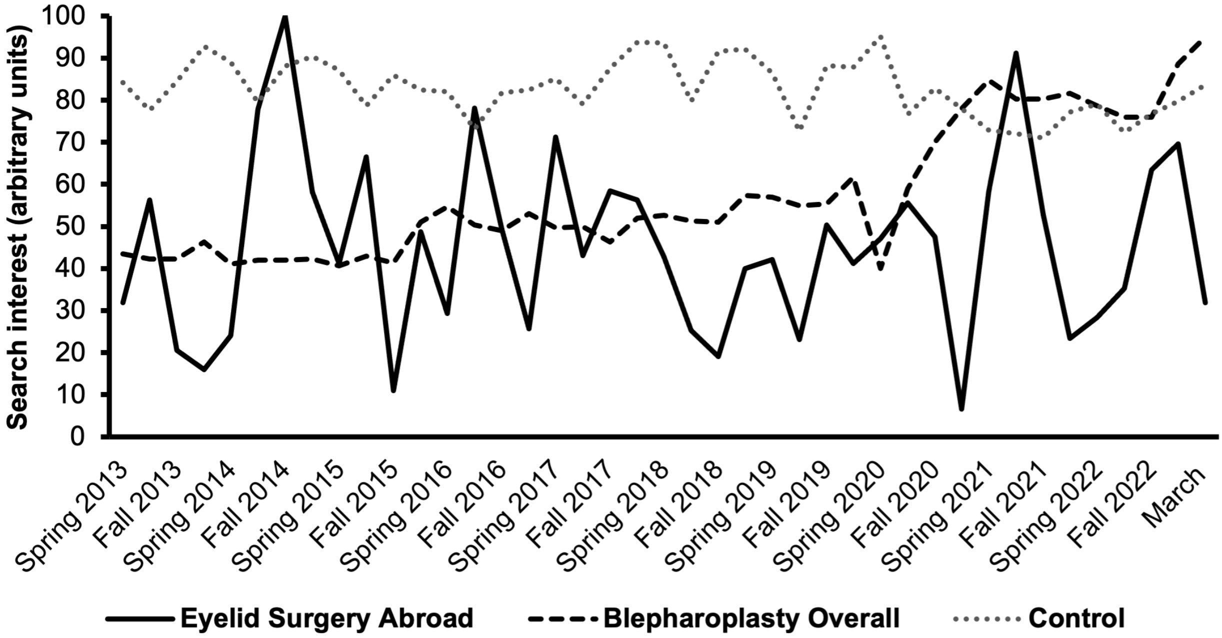
Results

Google Trends

Between 2013 and 2023, the top 5 destinations for Americans seeking eyelid surgery abroad were South Korea, Mexico, Canada, Turkey, and China (Multimedia Appendix 1). Despite 118% growth in blepharoplasty searches, interest in eyelid surgery abroad remained steady (Figure 1). The notable growth occurred recently, averaging 30.61 (SD 2.49) increased RSV from June 2020 to March 2023 compared with previous years from April 2013 to May 2020 (95% CI 25.73-35.49 RSV; P<.001). US states with the highest blepharoplasty RSV were Florida, California, Hawaii, Nevada, and New York. While complete state-level data for interest in eyelid surgery in all the destination countries were unavailable, California led the United States in searches for eyelid surgery in South Korea.

XSL•FO
RenderX

Figure 1. Google Trends search interest in cosmetic tourism for eyelid surgery**.** Despite periodic fluctuations, American Google Trends search interest in cosmetic tourism for eyelid surgery remained stagnant (0% growth) from 2013 to 2023, while overall interest in blepharoplasty rose 118.4%.



Social Media

In total, 55 top Instagram posts were included from aesthetic eyelid surgeons across ophthalmology, Otolaryngology-Head and Neck Surgery (OHNS), and plastic surgery. Within each discipline, qualifying fellowship training specializing in aesthetic eyelid surgery was examined. For ophthalmology, this included ASOPRS fellowship, CSOPS fellowship, or an unspecified OPS fellowship; for OHNS, AAFPRS facial plastic and reconstructive surgery fellowship or other unspecified or international facial plastic and reconstructive surgery fellowships; for plastic surgery, aesthetic plastic surgery fellowship.

When combining both the United States and international eyelid surgeons within each discipline, ophthalmology (7/10, 70%) and OHNS (9/13, 69%) were far more likely to have had fellowship training that included aesthetic eyelid surgery than plastic surgeons (1/32, 3%; *P*<.001) (Table 1). Of the plastic surgeons performing aesthetic eyelid surgery, 19% (4 abroad and 2 in the United States) were initially trained as general surgeons and subsequently obtained further training in general plastic surgery.

XSL•FO
RenderX

Table . Summary of training and professional society affiliations of most popular aesthetic eyelid surgeons on Instagram.

Specialty	Korea, n	Mexico, n	Canada, n	Turkey, n	China, n	Intl[a], n (%)	United States, n (%)	*P* value
Ophthalmology	0	0	3	3	0	6 (14)	4 (36)	.08
ASOPRS[b] fellowship	0	0	0	0	0	0 (0)	3 (27)	<.001
CSOPS[c] fellowship	0	0	1	0	0	1 (2)	0 (0)	.61
Other OPS[d] fellowship	0	0	2	0	0	2 (5)	1 (2)	.55
Total OPS fellowship	0	0	3	0	0	3 (7)	4 (36)	.009
OHNS[e]	1	1	2	4	0	8 (18)	5 (46)	.06
AAFPRS[f] fellowship	0	0	1	0	0	1 (2)	5 (46)	<.001
Other FPS[g] fellowship	0	0	1	2	0	3 (7)	0 (0)	.37
Total FPS fellowship	0	0	2	2	0	4 (9)	5 (46)	.004
Plastic surgery	9	10	5	5	1	30 (68)	2 (18)	.003
Aesthetic fellowship	0	0	0	1	0	1 (2)	0 (0)	.61
Total fellowship overall	0	0	5	3	0	8 (18)	9 (82)	<.001
Professional society[h]	5	8	7	2	0	22 (50)	11 (100)	.002
Total	10	11	10	12	1	44	11	—[i]

[a]Intl: international.

[b]ASOPRS: American Society of Ophthalmic Plastic & Reconstructive Surgery.

[c]CSOPS: Canadian Society of Oculoplastic Surgeons.

[d]OPS: oculofacial plastic surgery.

[e]OHNS: Otolaryngology–Head and Neck Surgery.

[f]AAFPRS: American Academy of Facial Plastic and Reconstructive Surgery.

[g]FPS: Facial Plastics.

[h]Professional Organizations included American Academy of Facial Plastic & Reconstructive Surgery (AAFPRS); American Society of Ophthalmic Plastic & Reconstructive Surgery (ASOPRS); Canadian Society of Oculoplastic Surgeons (CSOPS); Fellow of the American College of Surgeons (FACS); American Society of Plastic Surgeons (ASAPS); European Board of Plastic Reconstructive & Aesthetic Surgery (EBOPRAS); Fellow of the Royal College of Surgeons of Canada (FRCSC); International Society of Aesthetic Plastic Surgery (ISAPS); Oriental Society of Aesthetic Plastic Surgery (OSAPS); International Confederation for Plastic, Reconstructive & Aesthetic Surgery (IPRAS), Korean Academy of Facial Plastic and Reconstructive Surgery (KAFPRS), Korean Society of Plastic and Reconstructive Surgeons (KSPRS), Asociación Mexicana de Cirugía Plástica Estética y Reconstructiva (AMCPER), and Canadian Society of Plastic Surgeons (CSPS).

[i]Not applicable.

Looking specifically at the United States versus international surgeons, the US surgeons more frequently completed fellowships that encompassed aesthetic eyelid surgery–specific training compared with international surgeons (9/11, 82% vs 10/44, 23%; *P*<.001) (Table 1). Within ophthalmology, US aesthetic eyelid surgeons more often had OPS training compared with international surgeons (4/11, 36% vs 3/44, 7%; *P*=.009). Within OHNS, US surgeons more often had facial plastic surgery (FPS) fellowship training compared with international surgeons (5/11, 46% vs 4/44, 9%; *P*=.004), and especially had more AAFPRS fellowship training (5/11, 46% vs 1/44, 2%; *P*<.001). Interestingly, 100% (9/9) of the US ophthalmologists and US OHNS had done an OPS or FPS fellowship, respectively, compared with only 50% (3/6) of the international ophthalmologists and 50% (4/8) of the international OHNS, although this was not statistically significant (*P*=.06 and *P*=.09, respectively). General plastic surgery accounted for the majority of training for the most popular international eyelid surgeons, contrasting with a low proportion in the United States (30/44, 68% vs 2/11, 18%; *P*=.003). These plastic surgeons rarely had

XSL•FO
RenderX

aesthetic plastic surgery fellowship training both abroad and in the US (1/30, 2% vs 0/2, 0%; *P*=.61).

Aesthetic eyelid surgeons in the United States were more likely to have active membership in recognized professional medical societies than their international counterparts (11/11, 100% vs 22/44, 50%; *P*=.002). A summary of these professional societies is listed in Table 1.

Among aesthetic eyelid surgeons abroad, 27% (12/44) offered medical tourism travel packages, while only 5% (2/44) mentioned warranty systems for the financial handling of revisional surgery on their websites.

Discussion

Principal Findings

This infodemiology study indicates Americans primarily seek aesthetic eyelid surgery abroad in Asia and Europe (South Korea, Turkey, and China) as well as the US neighboring countries (Mexico and Canada). These findings align with existing plastic surgery literature. ISAPS 2021 ranked Turkey, Colombia, Mexico, Thailand, and Spain as top aesthetic surgery destinations [5]. Previous studies suggest these trends may reflect proximity (eg, Mexico and Canada in this United States–based study, or Spain in the largely European-based ISAPS study) and niche surgical offerings (eg, Thailand for gender reassignment or Colombia for breast, body, and extremity) [5,18]. South Korea's popularity for eyelid surgery among Americans may be linked to the nation's commonly sought-after idealized appearance referred to as the Korean look achieved through a distinct set of facial cosmetic procedures defined in South Korea and popularized by global exportation of South Korean popular culture [19]. These aesthetic surgeries produce a desired appearance for the East Asian face, focused on widening the eyes, narrowing the cheekbones and jawbones, and augmenting the nose tip. In this study, most interest in South Korean eyelid surgery originated from California, home to half a million of the United States' 1.7 million Korean Americans, suggesting a shared cultural desire for the Korean look [20].

This study demonstrates overall American interest in blepharoplasty doubled in the last decade with most growth occurring after the United States lifted COVID-19 restrictions in May 2020. However, this has not translated to a greater preference for international destinations for aesthetic eyelid surgery, possibly attributed to reopening patterns after COVID-19 travel restrictions. These findings expand on previous studies that found declines in domestic interest in oculofacial plastic, facial plastic, or general plastic surgery and a temporary rise in interest abroad during the pandemic [3,6]. By 2021, the United States rebounded to 85.7% of prepandemic cosmetic surgery volume, making the US surgeons more accessible than the surgeons abroad postrestrictions [5].

The US aesthetic eyelid surgeons popular on social media more often had fellowship training specializing in blepharoplasty techniques, including OPS fellowship or FPS fellowship. They were also more likely to hold membership in a professional medical society than their international counterparts. A recent study on problematic Instagram medical marketing demonstrated a significant number of physicians advertising as cosmetic surgeons without appropriate plastic surgery credentials [13]. Ideally, consumers should be aware of differences in specialty training and qualifications when selecting an aesthetic eyelid surgeon.

A quarter of international eyelid surgeons offered medical tourism vacation packages, involving substantial upfront financial commitments that may pressure patients to proceed with surgery. Cosmetic tourists may also encounter a lack of language concordance, psychological stress from traveling and recovering alone, inadequate follow-up, limited legal recourse, and surgical complications including perioperative fatalities [1-3]. Of note, few international aesthetic eyelid surgeons were observed to offer a warranty. It is customary in the United States for eyelid surgery practices to have a revisions policy addressing the possible event of additional corrective procedures; while this may not necessarily be displayed on the surgeon's website, it is an important aspect of the surgical discussion and consent process. For international surgeons, on the other hand, displaying the warranty policy on the surgeon's website may be more essential as reassurance to the patient. Of note, the Centers for Disease Control and Prevention recommends that cosmetic tourism patients should be informed of their rights and legal recourse before agreeing to travel outside the United States for medical care because of the high costs of complications and limited ability for legal action [21].

Aesthetic surgery remains a key contributor in the US $27.8 billion global medical tourism industry, earning high satisfaction ratings in some populations [22]. Future investigations should explore cost, desire for local expertise, and perceptions of international experience to better appreciate the American interest in aesthetic eyelid surgery abroad. In addition, an examination of a particular destination's aesthetic OPS niche may shed further light on cosmetic preferences and tourism patterns. Finally, an understanding of how social media's relative lack of international aesthetic eyelid surgeons with fellowship training that includes blepharoplasty techniques correlates to surgical outcomes and patient satisfaction would provide further insight for Americans seeking cosmetic tourism abroad. Such information may also help guide professional academic society resource allocation and social media engagement.

Limitations

Google Trends lacks patient demographics and outcomes, can be influenced by media exposure or user manipulation, and may encompass users beyond those seeking surgery. Social media may portend representation gaps for less tech-savvy surgeons. Instagram restrictions limit data from China, precluding conclusions about the country. Surgeons' websites may not disclose all details. While the Google and Instagram platforms are popular United States–based channels for health information–seeking, these may not represent the only relevant media platforms from an international standpoint. Relevant data for each country may not have been captured by the specific hashtags used; however, the presence of several duplicate posts among variations of hashtags and topic queries supports the appropriateness of these terms in highlighting the top content.

XSL•FO
RenderX

Analysis of surgical qualifications is limited as different countries have inherently different training and licensing regulations, as well as cultural variations and legal requirements in how they may be advertised. The study is limited to internet search trends and social media trends, but the reasons behind these trends remain unclear. The search terms designed for geographic identification may not represent all queries for eyelid surgery abroad. Rate limiting on posts collected on Instagram may confound the output for all or a number of countries. In consistency with previous research, the sample of social media data was queried at one point in time to control for temporal variations in content engagement; however, this strategy may limit the generalizability of the content to other seasons of the year. Given these limitations, the study should be received as exploratory, laying a preliminary foundation of novel insight in the unexplored area of cosmetic tourism for eyelid surgery.

Conclusions

This study used Google Trends and social media to identify preferred international destinations among US travelers seeking cosmetic eyelid surgery and examined specialty training of the most sought-after aesthetic surgeons in those countries trending on social media. It highlights the shortage of international surgeons with aesthetic eyelid surgery–specific fellowship training and membership in a recognized professional medical society. Further research is necessary to evaluate how these trends correlate with demographics, surgical outcomes, and niche surgical offerings.

Acknowledgments

Funding for open access was provided by Tufts University Hirsh Health Sciences Library's Open Access Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Top 5 destinations Americans seek for eyelid surgery abroad. The most sought-after destinations for cosmetic eyelid surgery abroad were compared based on US consumers' Google searches for eyelid surgery internationally in 25 countries from 2013 to 2023.

[PNG File, 129 KB - infodemiology_v5i1e64639_app1.png]

References

1. Raggio BS, Brody-Camp SA, Jawad BA, Winters RD, Aslam R. Complications associated with medical tourism for facial rejuvenation: a systematic review. Aesthetic Plast Surg 2020 Jun;44(3):1058-1065. [doi: 10.1007/s00266-020-01638-w] [Medline: 32040602]
2. Pereira RT, Malone CM, Flaherty GT. Aesthetic journeys: a review of cosmetic surgery tourism. J Travel Med 2018 Jun 1;25(1). [doi: 10.1093/jtm/tay042] [Medline: 29924349]
3. Murphy D, Lane-O'Neill B, Dempsey MP. COVID-19 and cosmetic tourism: a Google Trends analysis of public interests and the experience from a tertiary plastic surgery centre. J Plast Reconstr Aesthet Surg 2022 Apr;75(4):1497-1520. [doi: 10.1016/j.bjps.2022.01.040] [Medline: 35148980]
4. Wilde C, Ross AR, Maharajan S. Health tourism and the need for occasional strong paternalism: complications and management of cosmetic anterior chamber iris implantation. Eye (Lond) 2018 Dec;32(12):1915-1916. [doi: 10.1038/s41433-018-0200-6] [Medline: 30171194]
5. ISAPS International Survey on Aesthetic/Cosmetic Procedures performed in 2021. : International Society of Aesthetic Plastic Surgeons; 2021 URL: https://www.isaps.org/media/vdpdanke/isaps-global-survey_2021.pdf
6. Azzam DB, Cypen SG, Tao JP. Oculofacial plastic surgery-related online search trends including the impact of the COVID-19 pandemic. Orbit 2021 Feb;40(1):44-50. [doi: 10.1080/01676830.2020.1852264] [Medline: 33317388]
7. Tijerina JD, Morrison SD, Nolan IT, Vail DG, Nazerali R, Lee GK. Google Trends as a tool for evaluating public interest in facial cosmetic procedures. Aesthet Surg J 2019 Jul 12;39(8):908-918. [doi: 10.1093/asj/sjy267] [Medline: 30304356]
8. Teixeira JC, Morrison SD, Brandstetter KA, Nuara MJ. Is there an increasing interest in facial feminization surgery? A search trends analysis. J Craniofac Surg 2020;31(3):606-607. [doi: 10.1097/SCS.0000000000006220] [Medline: 32011538]
9. Cheng T, Wang F, Barmettler A. #Oculoplastics: an analysis of TikTok's top oculoplastics content. Ophthalmic Plast Reconstr Surg 2022;38(5):452-457. [doi: 10.1097/IOP.0000000000002158] [Medline: 35323143]
10. Alshaalan HS, AlTamimi LA, Alshayie RA, Alsuhaibani AH. The impact of social media accounts on periocular cosmetic surgeries. Saudi J Ophthalmol 2021;35(3). [doi: 10.4103/SJOPT.SJOP_14_21]
11. Tingley J, Allen RC, Barmettler A. #OculoplasticsandSocialMedia: a review of social media in oculoplastics and relevant subspecialties. Orbit 2022 Apr;41(2):141-149. [doi: 10.1080/01676830.2021.1930064] [Medline: 34044737]
12. Park SSE, Akella SS, Moon JY, et al. Building your brand: analysis of successful oculoplastic surgeons on social media. Ophthalmic Plast Reconstr Surg 2020;36(6):582-589. [doi: 10.1097/IOP.0000000000001654] [Medline: 33156596]

XSL•FO
RenderX

13. Dorfman RG, Vaca EE, Mahmood E, Fine NA, Schierle CF. Plastic surgery-related hashtag utilization on Instagram: implications for education and marketing. Aesthet Surg J 2018 Feb 15;38(3):332-338. [doi: 10.1093/asj/sjx120] [Medline: 29040378]
14. Gunawardene AN, Suraneni S, Rohowetz LJ, Sridhar J. Characteristics and medical accuracy of online discussions of retinal conditions on a social media platform. J Vitreoretin Dis 2025 Feb 19:24741264251315139. [doi: 10.1177/24741264251315139] [Medline: 39991040]
15. Azzam DB, Nag N, Tran J, et al. A novel epidemiological approach to geographically mapping population dry eye disease in the United States through Google Trends. Cornea 2021 Mar 1;40(3):282-291. [doi: 10.1097/ICO.0000000000002579] [Medline: 33177410]
16. Nuti SV, Wayda B, Ranasinghe I, et al. The use of Google Trends in health care research: a systematic review. PLoS ONE 2014;9(10):e109583. [doi: 10.1371/journal.pone.0109583] [Medline: 25337815]
17. How posts appear on Instagram hashtag search results. Instagram. 2025. URL: https://help.instagram.com/777754038986618/?helpref=related_articles [accessed 2025-05-26]
18. Franzblau LE, Chung KC. Impact of medical tourism on cosmetic surgery in the United States. Plast Reconstr Surg Glob Open 2013 Oct;1(7):e63. [doi: 10.1097/GOX.0000000000000003] [Medline: 25289258]
19. Holliday R, Cheung O, Cho JH, Bell D. Trading faces: The ‘Korean Look’ and medical nationalism in South Korean cosmetic surgery tourism. Asia Pac Viewp 2017 Aug;58(2):190-202. [doi: 10.1111/apv.12154]
20. Ivey SL, Kim H, Yoo E, et al. Health and healthcare needs of Koreans in San Francisco Bay Area: The Korean Needs Assessment (KoNA) project. J Immigr Minor Health 2019 Oct;21(5):1043-1051. [doi: 10.1007/s10903-018-0823-5] [Medline: 30244329]
21. Medical tourism CDC yellow book 2024. US Centers For Disease Control and Prevention. 2023. URL: https://wwwnc.cdc.gov/travel/yellowbook/2024/health-care-abroad/medical-tourism [accessed 2025-05-26]
22. Campbell A, Restrepo C, Navas G. Patient satisfaction with medical tourism: a review of 460 international plastic surgery patients in Colombia. Plast Reconstr Surg Glob Open 2020 Dec;8(12):e3230. [doi: 10.1097/GOX.0000000000003230] [Medline: 33425575]

Abbreviations

AAFPRS: American Academy of Facial Plastic & Reconstructive Surgery
ASOPRS: American Society of Ophthalmic Plastic & Reconstructive Surgery
CSOPS: Canadian Society of Oculoplastic Surgeons
FPS: facial plastic surgery
ISAPS: International Society of Aesthetic Plastic Surgery
OHNS: Otolaryngology-Head and Neck Surgery
OPS: oculofacial plastic surgery
RSV: relative search volume

Edited by T Mackey; submitted 23.07.24; peer-reviewed by A Edalatpour, D Ashraf, MGC Iwunwa, S Mallick, S Grob; revised version received 28.03.25; accepted 13.04.25; published 02.06.25.

Please cite as:
Azzam DB, Dai YL, North VS, Callahan AB, Heher KL, Kapadia MK, Vagefi MR
Global Surveillance of Public Interest in Cosmetic Tourism for Aesthetic Eyelid Surgery Abroad: Cross-Sectional Infodemiology Investigation of Internet Search Trends and Social Media Content
JMIR Infodemiology 2025;5:e64639
URL: https://infodemiology.jmir.org/2025/1/e64639
doi:10.2196/64639

© Daniel B Azzam, Yi Ling Dai, Victoria S North, Alison B Callahan, Katrinka L Heher, Mitesh K Kapadia, M Reza Vagefi. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 2.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Discourse Toward Older Drivers in Japan Using Social Media Data From 2010 to 2022: Longitudinal Analysis

Akito Nakanishi[1], MPPS; Masao Ichikawa[2], PhD; Yukie Sano[3], PhD

[1]Graduate School of Science and Technology, University of Tsukuba, Ibaraki, Japan

[2]Institute of Medicine, University of Tsukuba, Ibaraki, Japan

[3]Institute of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

Corresponding Author:
Yukie Sano, PhD
Institute of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

Abstract

Background: As the global population ages, concerns about older drivers are intensifying. Although older drivers are not inherently more dangerous than other age groups, traditional surveys in Japan reveal persistent negative sentiments toward them. This discrepancy suggests the importance of analyzing discourse on social media, where public perceptions and societal attitudes toward older drivers are actively shaped.

Objective: This study aimed to quantify long-term public discourse on older drivers in Japan through Twitter (subsequently rebranded X), a leading social media platform. The specific objectives were to (1) examine the sentiments toward older drivers in tweets, (2) identify the textual contents and topics discussed in the tweets, and (3) analyze how sentiments correlate with various variables.

Methods: We collected Japanese tweets related to older drivers from 2010 to 2022. Each quarter, we (1) applied to the Japanese version of the Linguistic Inquiry and Word Count dictionary for sentiment analysis, (2) employed 2-layer nonnegative matrix factorization for dynamic topic modeling, and (3) applied correlation analyses to explore the relationships of sentiments with crash rates, data counts, and topics.

Results: We obtained 2,625,807 tweets from 1,052,976 unique users discussing older drivers. The number of tweets has steadily increased, with significant peaks in 2016, 2019, and 2021, coinciding with high-profile traffic crashes. Sentiment analysis revealed a predominance of negative emotions (n=383,520, 62.42%), anger (n=106,767, 17.38%), anxiety (n=114,234, 18.59%), and risk (n=357,311, 58.15%). Topic modeling identified 29 dynamic topics, including those related to driving licenses, crash events, self-driving technology, and traffic safety. The crash events topic, which increased by 0.28% per year, showed a strong correlation with negative emotion (r=0.76, P<.001) and risk (r=0.72, P<.001).

Conclusions: This 13-year study quantified public discourse on older drivers using Twitter data, revealing a paradoxical increase in negative sentiment and perceived risk, despite a decline in the actual crash rate among older drivers. These findings underscore the importance of reconsidering licensing policies, promoting self-driving systems, and fostering a more balanced understanding to mitigate undue prejudice and support continued safe mobility for older adults.

(JMIR Infodemiology 2025;5:e69321) doi:10.2196/69321

KEYWORDS

older driver; older people; ageism; social media; Twitter; sentiment analysis; topic modeling

Introduction

As the global population ages, concerns about older drivers are increasing. The share of the global population aged 70 years or older is projected to rise from 6.4% in 2023 to 11.7% in 2050 [1]. Japan is one of the most rapidly aging societies, with 23.6% of its population being 70 years or older in 2023. Concurrently, the proportion of driver's license holders among older Japanese people is also increasing, reaching 16.6% for those aged 70 years or older in 2023 [2]. In response to their decline in physical and cognitive functions, as well as tragic traffic crashes involving them, the National Police Agency of Japan has incrementally tightened licensing policies for them in 1998, 2002, 2009, 2017, and 2022 [3].

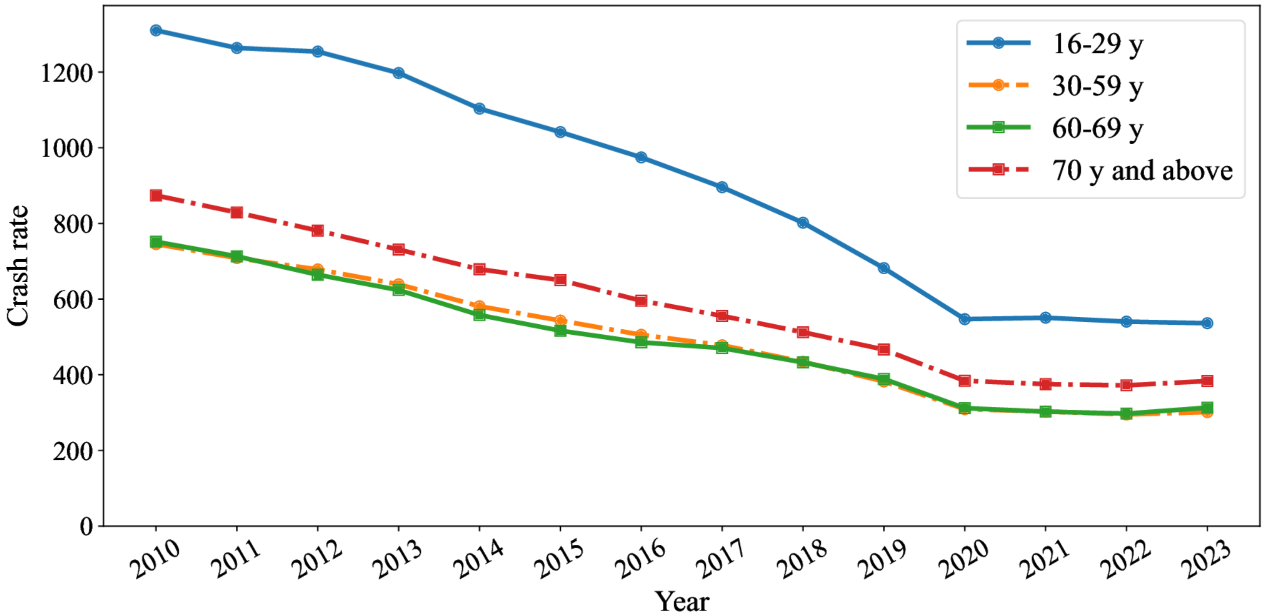
Research on older drivers' traffic safety frequently examines their risk of traffic crashes, neurological and physical impairments, and crash causes. A global meta-review [4] finds that older drivers have a higher risk of crashes and injuries even when controlling for travel time or distance, partly because they experience declines in vision, cognitive performance, reaction time, and physical ability, despite their strong adherence to traffic laws. This makes older drivers a higher-risk group in

XSL•FO
RenderX

traffic than younger drivers, who are more likely to be involved in crashes due to law violations or careless behavior. In addition to traffic crashes, studies have also discussed the effects of driving cessation, such as increased depressive symptoms [5,6] and reduced quality of life [7]. In Japan, research on older drivers is a relatively new field [8] and covers a range of topics, including dangerous driving associated with dementia [9], qualitative surveys of older drivers and their families [8,10,11], and discussions on licensing policies [12,13].

Although many studies highlight the risks associated with older drivers, they are not significantly more dangerous than other age groups in Japan when considering both their at-fault crash rate and the harm they cause to others. As shown in Figure 1, the at-fault crash rate (per 100,000 licensed drivers) for car and motorcycle drivers aged 70 years or older was 384 in 2023, which was higher than for those in middle age groups (30 - 59 y: 301 and 60 - 69 y: 313), but lower than for those in younger age groups (16 - 19 y: 1025 and 20 - 29 y: 497) [14]. Furthermore, the rate among those aged 70 years or older substantially decreased in the past decades from 874 in 2010 to 384 in 2023. It is also noted that their at-fault fatal crashes tend to result in fewer fatalities for occupants of other vehicles compared with fatal crashes caused by drivers in other age groups [15]. These statistics demonstrate that the risk of traffic crashes and resultant injuries imposed by older drivers is not as high as perceived, as driving failure frequencies do not significantly differ between age groups [4].

Figure 1. Traffic crash rates per 100,000 licensed drivers by year, categorized by age group (16 - 29 y, 30 - 59 y, 60 - 69 y, and 70 y and above).



Despite the decreasing crash rate, negative sentiments toward older drivers are still persistent in Japan. According to a survey by the National Police Agency, 85% (n=1698) of respondents (balanced by age and sex) perceive older drivers as dangerous, and 80% (n=1596) believe that licensing for older drivers should be revised [16]. Such sentiments are potentially influenced by media reports on older drivers and their crashes as well as stereotypes about older people. Recent studies suggested a negative tone toward older drivers in news reports [17], an increase in newspaper articles about them [18], and underreported of their crashes killing themselves [19]. However, discourse regarding older drivers in social media has not been explored despite the widespread use of various social media platforms.

Today, social media plays various roles in formal and informal communication at individual, organizational, and societal levels. The number of social media users escalates globally [20], with nearly 50% of Japan's population using Twitter (now X), a leading text-based social media platform [21]. Social media not only facilitates access to a broad range of information but also serves as a platform for public discussion, allowing individuals to express their views, opinions, and sentiments. As the reliance on online health information also intensifies [22], research into health-related topics on Twitter has expanded to include discussions on food security [23], type 1 diabetes [24], and various aspects of COVID-19, such as mask-wearing [25], vaccines [26-28], and the pandemic [29]. Consequently, applying text mining techniques, frequently employed to quantify the textual content and sentiment of tweets in these fields, offers valuable insights to facilitate a direct interpretation of societal discussions around health-related issues.

Given social media's impacts on public opinions and policy-making, gaining insights on older drivers and their crashes from social media—beyond conventional surveys [17,30]—is crucial for strategic planning of traffic safety. In this study, we aimed to (1) examine the sentiments toward older drivers in tweets, (2) identify the textual contents and topics discussed in the tweets, and (3) analyze how sentiments correlate with various variables. To achieve these goals, we conducted sentiment analysis, topic modeling, and correlation analysis, using over 2.6 million tweets on older drivers posted in the past 13 years.

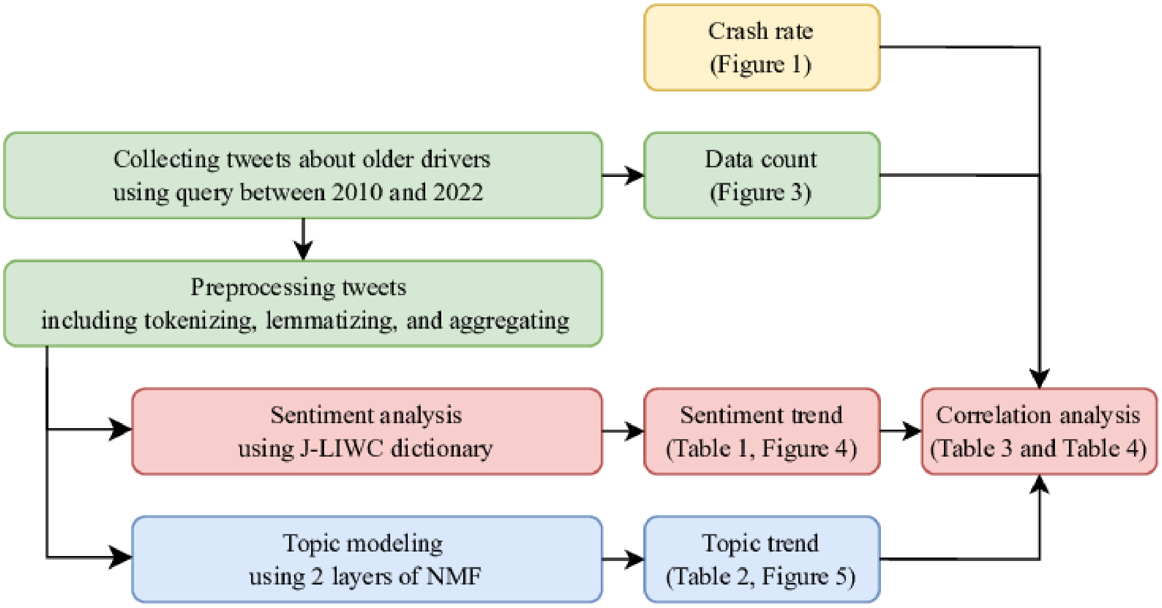
XSL•FO
RenderX

Methods

Study Workflow

Figure 2 illustrates the workflow of this study, covering data collection and processing, sentiment analysis, topic modeling, and correlation analysis.

Figure 2. Diagram of the study workflow. Tweets are collected and preprocessed, followed by sentiment analysis, topic modeling, and correlation analysis. Each figure and table number corresponds to the main text. J-LIWC: Japanese version of the Linguistic Inquiry and Word Count Dictionary; NMF: nonnegative matrix factorization.



Data Collection and Processing

We collected tweets about older drivers posted between January 1, 2010 and December 31, 2022, using the Twitter application programming interface v2 [31] with Japanese keywords "(高齢 OR 老人) AND (運転 OR ドライバー)", or literally, "(old age OR older people) AND (driving OR driver)." To ensure the relevance of collected tweets, we examined the contents in a random sample of 1000 tweets and determined that 967 tweets pertained to older drivers or their transportation issues.

As part of text preprocessing, we removed symbols, emojis [32], URLs, and hashtags (#...), and normalized characters by converting them to lowercase and half-width forms. Tweets were then tokenized into words using MeCab [33] with the mecab-ipadic dictionary, with stopwords removed (Multimedia Appendix 1) and verbs and adjectives lemmatized. In addition, we consolidated tweets from the same user within the same quarter by merging their tokenized word lists into single documents. This approach reduces individual biases, enables dynamic trend analysis by quarter, and improves topic modeling accuracy by processing aggregated and extended textual content [34]. After pooling tweets by user and quarter, we removed words with an occurrence rate of less than 0.1% per quarter, resulting in 9287 unique words. Hereafter, we refer to the collection of quarterly tweets aggregated by each user as a "document."

Sentiment Analysis

To clarify the Twitter users' feelings toward older drivers, we first conducted sentiment analysis using the Japanese version of the Linguistic Inquiry and Word Count Dictionary (J-LIWC) [35]. Sentiment analysis quantifies the sentiment expressed in documents, using techniques ranging from word-based and context-based methods to deep-learning approaches. We adopted a word-based method for its interpretability and consistency with topic modeling, specifically using J-LIWC, which is psychologically validated and reliable in both English and Japanese.

The J-LIWC includes 69 categories in 4 broad categories—"psychological processes" (37 categories), "linguistic processes" (14 categories), "punctuation" (12 categories), and "other grammar" (6 categories). For this study, we used the "psychological processes" of the J-LIWC, aligning with our research objectives. The "psychological processes" consists of 10 subcategories, namely "affective processes," "negative emotions," "social processes," "cognitive processes," "perceptual processes," "biological processes," "drives," "relativity," "personal concerns," and "informal language," and under these subcategories, there are several further subcategories, as illustrated in Table 1. Throughout this paper, we refer to the further subcategories of the J-LIWC as sentiments.

XSL•FO
RenderX

Table . Proportion of documents containing words corresponding to 10 sentiments in the Japanese version of the Linguistic Inquiry and Word Count Dictionary's "affective processes," "negative emotions," and "drives" within 614,429 documents discussing older drivers posted from 2010 to 2022. In addition to proportions, results of linear regression (the formula Yt=β0+β1Xt+ t) are represented by β0,β1, R^2, and *P* values.

Subcategory and sentiment		Tweets, n (%)	β0	β1	R^2	*P* value
Affective processes						
	Positive emotions	217,300 (35.37)	31.0	0.11	0.154	.004
	Negative emotions	383,520 (62.42)	46.1	0.36	0.462	<.001
Negative emotions						
	Anxiety	114,234 (18.59)	14.0	0.12	0.395	<.001
	Anger	106,767 (17.38)	11.0	0.17	0.686	<.001
	Sadness	23,609 (3.84)	1.70	0.05	0.230	<.001
Drives						
	Affiliation	105,984 (17.25)	15.6	0.04	0.049	.120
	Achievement	179,969 (29.29)	21.7	0.18	0.416	<.001
	Power	213,716 (34.78)	28.1	0.15	0.362	<.001
	Reward	105,067 (17.10)	13.1	0.10	0.412	<.001
	Risk	357,311 (58.15)	42.8	0.33	0.384	<.001

Topic Modeling

Next, we employed 2 layers of nonnegative matrix factorization (NMF) [36], a form of dynamic topic modeling. Topic modeling aims to extract latent textual data and generate topic distributions categorized as "static" (ignoring time) and "dynamic" (incorporating temporal variation) [37,38]. While many previous studies use latent Dirichlet allocation [39], it does not account for temporal trends, thus unsuitable for longitudinal studies. Several models have been developed to capture topic dynamics over time, such as dynamic topic model (DTM) [40], which is an extended version of latent Dirichlet allocation, 2 layers of NMF [35], and BERTopic [41]. DTM and 2 layers of NMF rely on a bag-of-words approach, whereas BERTopic uses word embeddings. This study opted for 2 layers of NMF as it provides a richer set of top words, higher coherence scores, faster performance than DTM, and, compared with BERTopic, offers objective evaluation metrics, enables topic distribution generation for each document, and eliminates the need for labor-intensive testing [36,42].

Two layers of NMF is a method for extracting dynamic topic distributions by applying a 2-stage NMF, an unsupervised dimensionality reduction technique that decomposes data into nonnegative factors. In the first stage, document-term matrix At for each time unit (quarter) *t* undergoes NMF, producing a "window topic," which represents topic distributions within each time unit. In the second stage, all "window topic"–term matrix B (a combination of all At) are further decomposed via NMF, resulting in a "dynamic topic," which incorporates temporal changes.

Following the previous study [36], we determined the optimal number of dynamic topics by testing a range between 25 and 90, selecting the one with the highest Topic Coherence using Word2Vec score [43]. TC-W2V score is defined as the mean pairwise cosine similarity between term vectors embedded using the word2vec model [44]. In this study, we employed the Japanese Social Media Corpus [45] as a large-scale word2vec model, trained on approximately 2 million Japanese words from social media and web sources (Multimedia Appendix 1).

Statistical Tests

In each section, we employ simple linear regression to analyze temporal trends in data counts, sentiments, and topics. The simple linear regression model is expressed as follows:

Yt= β0+β1Xt+ t

where Yt represents the dependent variable (eg, tweet count at quarter *t*) and Xt denotes the independent variable (eg, quarter *t*). The coefficient β1 indicates the rate of change, while the intercept β0 represents the baseline value when Xt is zero. The term t accounts for random fluctuations not explained by the model. For this analysis, we use the actual data count, the proportion of documents containing a given sentiment, and the mean topic proportion per time unit.

In addition, we examine the correlation between sentiments and traffic crash rates, data counts, and topics using Pearson correlation analysis. The correlation coefficient (*r*) is computed as follows:

r=∑(Xt−X−)(Yt−Y−)∑(Xt−X−)2∑(Yt−Y−)2

where Xt (eg, negative emotions at quarter *t*) and Yt (eg, tweet count at quarter *t*) represent the paired values of the independent and dependent variables, respectively, and X- and Y- denote their mean values. A positive *r* indicates a positive correlation, while a negative *r* suggests an inverse relationship.

To quantify the proportion of variance explained, we also compute the coefficient of determination (R^2) for regression models and the squared correlation coefficient (r^2) for correlation analysis. Significance levels are calculated under

XSL•FO
RenderX

the null hypotheses of β1=0 for linear regression and r=0 for correlation analysis.

Ethical Considerations

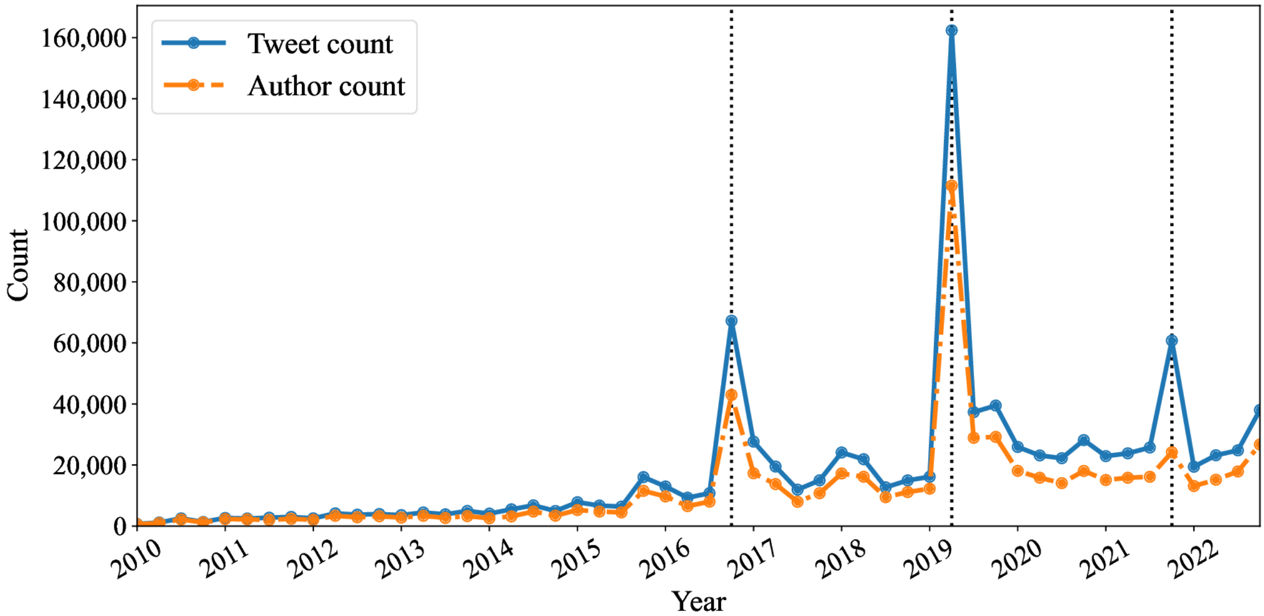
This study did not require institutional review board approval because it is an observational study that used only publicly accessible data and reported aggregate results with no personal identifiers. To maintain privacy and confidentiality, all personal identifiers were removed during data processing. Transparency was maintained throughout the study, with clear communication of its purpose, methods, and findings.

Results

Summary of Tweet Counts and Users

Our dataset contained 2,625,807 tweets from 1,052,976 unique users, comprising original tweets (n=767,419, 29.25%), retweets (n=1,678,133, 63.91%), replies (n=171,781, 6.54%), and quoted tweets (n=8474, 0.32%). To analyze primary opinions or original tweets, we excluded retweets and eliminated referenced text from quoted tweets and replies. Figure 3 illustrates the quarterly counts of original tweets and unique users discussing older drivers over a 13-year period. There has been a consistent presence of tweets and users, with a moderate increase over time. This trend is supported by linear regression (tweet count: β0=−3011.09, β1=832.75, R^2=0.256, P<.001; user count: β0=−1325.86, β1=534.90, R^2=0.238, P<.001).

Figure 3. Tweet and user counts related to older drivers by quarter from 2010 to 2022. Black vertical dotted lines indicate the 3 major tweet peaks.



Since 2016, the number of tweets per quarter has exceeded 10,000, rising to over 20,000 per quarter from 2019 onward. The highest peak occurred in the second quarter of 2019 (approximately 162,000 tweets), driven by a crash on April 19, 2019, in Higashi-Ikebukuro, Tokyo, where an 87-year-old man mistakenly stepped on the accelerator, killing a 3-year-old child and her mother and injuring 9 others [46]. The second peak, in the fourth quarter of 2016 (approximately 67,000 tweets), coincided with a series of crashes in Tochigi (November 10, 2016), Tokyo (November 12, 2016), and Miyazaki (November 13, 2016), where older drivers struck and killed pedestrians, sparking active public debate about older drivers [47]. Another significant peak in the fourth quarter of 2021 (approximately 60,000 tweets) followed a crash in Osaka, where an 89-year-old man mistakenly stepped on the accelerator, resulting in pedestrian casualties [48]. Hereinafter, we plot black dotted lines to indicate the 3 major tweet peaks in each figure.

Sentiments Toward Older Drivers

After processing, our final dataset comprised 614,429 documents posted by 404,689 unique users. Table 1 presents the proportion of documents containing words associated with sentiments in the primary emotional subcategories (“affective processes” and “negative emotions”) as well as “drives.”

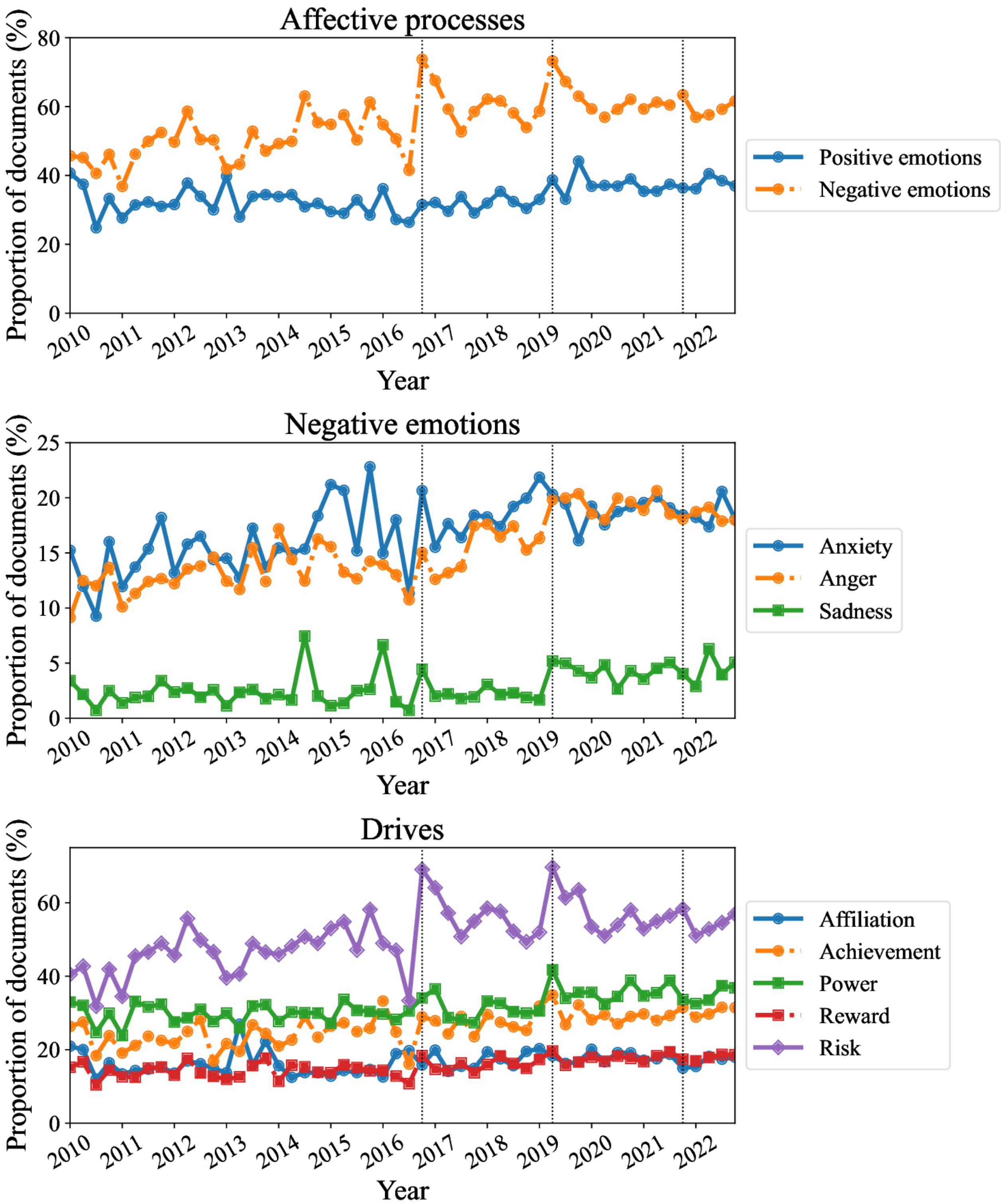
Across all quarters, negative emotions (n=383,520, 62.42%) were more prevalent than positive emotions (n=217,300, 35.37%) in “affective processes.” Within the subcategory of “negative emotions,” anxiety (n=114,234, 18.59%) and anger (n=106,767, 17.38%) were more frequent than sadness (n=23,609, 3.84%). In “drives,” risk (n=357,311, 58.15%) was the most frequently observed sentiment. Sentiments in other categories, none of which exceeded an average proportion of 50%, are presented in “S2. Sentiment analysis Details” in Multimedia Appendix 1. For instance, family (n=65,069, 10.59%) in “social processes,” insights (n=261,040, 42.48%) in “cognitive processes,” and health (n=80,967, 13.18%) in “biological processes” were the most predominant sentiments in their respective categories.

Temporal trends of sentiments are plotted in Figure 4, with the numeric results of linear regression presented in Table 1. Regarding “affective processes,” while positive emotions shows

XSL•FO
RenderX

a slight increase (β1=0.11, *P*=.004), negative emotions increased more prominently by 0.36% per quarter (*P*<.001), rising from approximately 40% to 60%, with spikes aligning with the 3 previously discussed data peaks, as indicated by black dotted lines. All sentiments in "negative emotions" display an increasing trend, with anxiety (n=114,234, 18.59%, β1=0.12%, *P*<.001) and anger (n=106,767, 17.38%, β1=0.17%, *P*<.001) showing relatively strong upward trends compared with sadness (n=23,609, 3.84%, β1=0.05%, *P*<.001). Regarding "drives," risk remained consistently above 40% and showed a noticeable increase over time (β1=0.33%, *P*<.001) relative to other sentiments. Additional analyses can be found in "S2. Sentiment Analysis Details" in Multimedia Appendix 1, indicating that insights (n=261,040, 42.48%, β1=0.25%, *P*<.001) and hear (n=87,281, 14.21%, β1=0.19%, *P*<.001) were relatively dominant and exhibited increasing trends.

Figure 4. The Japanese version of the Linguistic Inquiry and Word Count Dictionary quarterly trends in "affective processes," "negative emotions," and "drives." Black vertical dotted lines indicate the 3 major tweet peaks as illustrated in Figure 3.



XSL•FO
RenderX

Topics About Older Drivers

Topic modeling with 2 layers of NMF identified 29 dynamic topics that achieved the highest coherence score of 0.22 in an experiment varying the number of topics between 25 and 90 (Multimedia Appendix 1).

We further selected 16 topics that constituted at least 3.5% of all documents on average, naming them based on the top-weighted words and representative original tweets for each topic. Finally, we selected 9 topics where the R^2 value exceeded 0.1 in simple linear regression, represented by a red dashed line, as illustrated in Table 2. The details of all 16 topics, including typical tweets (in both Japanese and English), as well as the regression results for the 7 topics whose R^2 did not exceed 0.1, are provided in Multimedia Appendix 1.

Table . Description of 9 topics, including topic names, top 10 weighted words, proportions, and results of linear regression (the formula Yt=β0+β1Xt+ t).

Topic name	Top 10 weighted words	Tweets, n (%)	β0	β1	R^2	*P* value
Topic 1: License surrender	return, license, voluntary, revoke, system, forcefully, surrender, certificate, revocation, discount	43,270 (7.04)	4.37	0.08	0.278	<.001
Topic 2: Crash events	crash, cause, increase, traffic, report, death, decrease, frequent, occur, prevention	35,396 (5.76)	2.73	0.07	0.358	<.001
Topic 5: Traffic safety	traffic, safety, campaign, prevention, bicycle, drinking, nationwide, walking, public, child	25,139 (4.09)	6.13	−0.05	0.123	.010
Topic 6: Ikebukuro incident	runaway, Ikebukuro, Tokyo, NHK, defendant, crash, bereaved, family, memorial[a]	24,430 (3.98)	−0.10	0.09	0.343	<.001
Topic 7: Self-driving technology	self-driving, social, Japan, technology, necessary, safety, soon, experiment, popularization, development	24,235 (3.94)	1.95	0.05	0.151	.004
Topic 8: Social issues	issue, social, consider, life, rural, solution, necessary, difficult, traffic, local	23,935 (3.90)	1.57	0.05	0.294	<.001
Topic 9: License renewal	renewal, license, course, take, exam, test, center, excellent, Emperor, Kobe	23,525 (3.83)	5.83	−0.05	0.126	.010
Topic 10: Discussing senior driving	say, bad, hear, go, oneself, come, same, can, young, complain	23,511 (3.83)	1.30	0.07	0.402	<.001
Topic 14: Driving errors	brake, accelerator, step, mistake, wrong, pedal, crash, parking, MT[b], operation	21,741 (3.54)	2.23	0.04	0.199	<.001

[a]The frequently mentioned name of the individual involved in this crash has been masked to consider ethical concerns.

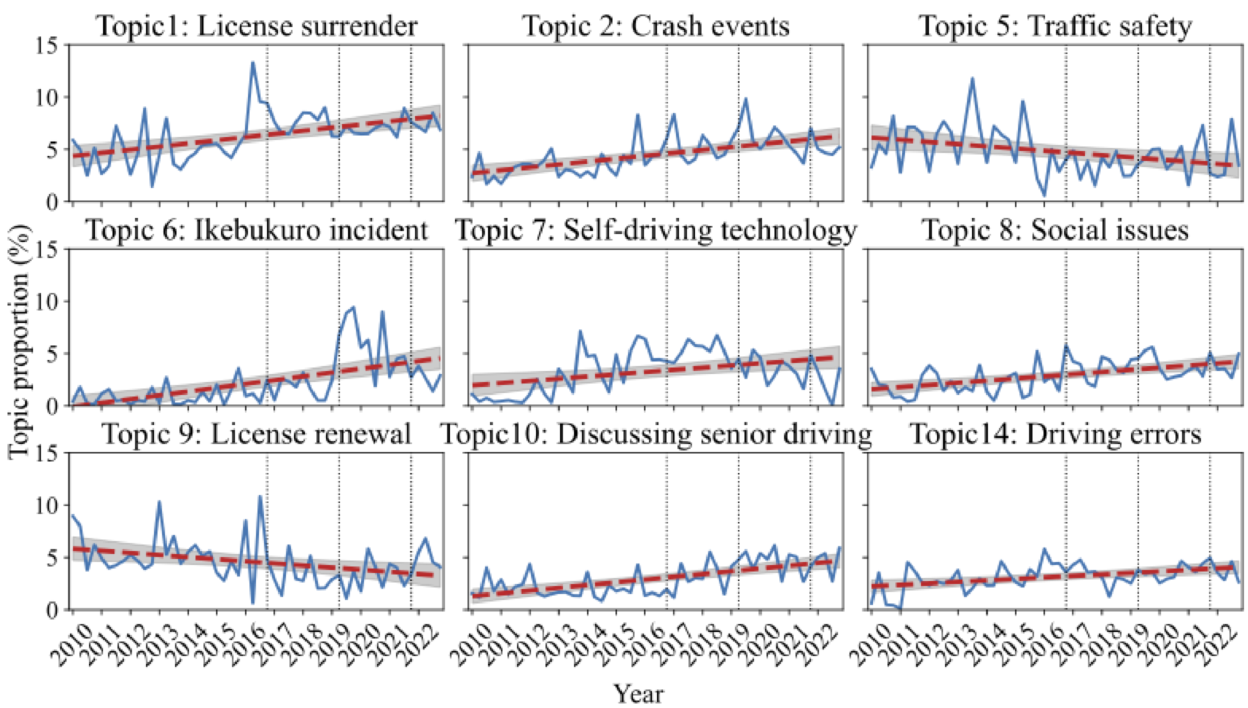
[b]MT indicates Manual Transmission, which is one type of vehicle transmission system.

XSL•FO
RenderX

The most prevalent topics were License surrender (n=43,270, 7.04%) and Crash events (n=35,396, 5.76%), followed by Traffic safety, Ikebukuro incident, Self-driving technology, Social issues, License renewal, Discussing senior driving, and Driving errors. Other topics include Thoughts on older drivers, Prevalent older drivers, and News media ("S3. Topic Modeling Details" in Multimedia Appendix 1).

Figure 5 illustrates the temporal trends of topic proportions for 9 topics with a red dashed line, representing the simple linear regression result for each topic. The proportions of topics 1 (License surrender, β1=0.08, *P*<.001), 2 (Crash event, β1=0.07, *P*<.001), 6 (Ikebukuro incident, β1=0.09, *P*<.001), 8 (Social issues, β1=0.05, *P*<.001), 10 (Discussing senior driving, β1=0.07, *P*<.001) and 14 (Driving errors, β1=0.04, *P*<.001) show consistent increases over time. Notably, topic 1 peaked in 2016, when measures to promote license surrender in rural areas gained attention. Topics 2 and 8 correspond to tweet count peaks, while topic 6 coincides specifically with the fourth quarter of 2019 subsequent periods. Although less pronounced, topic 7 (Self-driving technology, β1=0.05, *P*=.004) also exhibits an increasing trend, whereas topics 5 (Traffic safety, β1=−0.05, *P*=.01) and 9 (License renewal, β1=−0.05, *P*=.001) are declining. Additional analysis can be found in "S3. Topic Modeling Details" in Multimedia Appendix 1). Notably, News media exhibits topic peaks in the fourth quarter of 2016, reflecting unique characteristics not captured by simple linear regression.

Figure 5. Temporal trends of 9 topic mean proportions, shown as solid blue lines, with red dashed lines representing regression lines. Gray shaded areas represent 95% CIs for the regression lines. Black vertical dotted lines indicate the 3 major tweet peaks as illustrated in Figure 3.



Correlation Analysis on Sentiments

Table 3 and Table 4 present the results of the correlation analysis for positive emotions and negative emotions in "affective processes," risk in "drives" and anxiety, anger, and sadness in "negative emotions," across 4 crash rates, 2 data counts, and 9 topics. While the correlation analysis for crash rates among individuals aged more than 70 years and 80 years is based on yearly data (n=13) due to the limited availability of statistical data [14], the other analyses are based on quarterly data (n=52).

XSL•FO
RenderX

Table . Results of the correlation analysis for positive emotions, negative emotions, and risk.

Variables	Positive emotions			Negative emotions			Risk		
	R	*r* 2	*P* value	*R*	*r* 2	*P* value	*R*	*r* 2	*P* value
Crash rate (n=13)									
16 - 29	–0.747	0.559	.003	–0.750	0.563	.003	–0.702	0.493	.007
30 - 59	–0.714	0.510	.006	–0.784	0.615	.002	–0.739	0.546	.004
60 - 69	–0.687	0.473	.009	–0.799	0.638	.001	–0.754	0.568	.003
Over 70	–0.712	0.507	.006	–0.787	0.619	.001	–0.744	0.554	.004
Data count (n=52)									
Tweet count	0.317	0.101	.02	0.683	0.466	<.001	0.677	0.459	<.001
User count	0.311	0.097	.03	0.674	0.454	<.001	0.671	0.450	<.001
Topic (n=52)									
Topic 1	0.039	0.002	.78	0.396	0.157	.004	0.364	0.132	.008
Topic 2	0.173	0.030	.22	0.761	0.579	<.001	0.719	0.517	<.001
Topic 5	–0.037	0.001	.79	–0.173	0.030	.220	–0.111	0.012	.430
Topic 6	0.425	0.180	.002	0.546	0.298	<.001	0.554	0.306	<.001
Topic 7	0.003	0.000	.98	0.351	0.123	.010	0.381	0.146	.005
Topic 8	0.270	0.073	.05	0.714	0.510	<.001	0.693	0.480	<.001
Topic 9	0.217	0.047	.12	–0.496	0.246	<.001	–0.526	0.277	<.001
Topic 10	0.264	0.069	.06	0.425	0.181	.002	0.365	0.133	.008
Topic 14	0.177	0.031	.21	0.43	0.185	.001	0.41	0.168	.003

XSL•FO
RenderX

Table . Results of the correlation analysis for anxiety, anger, and sadness.

Variables	Anxiety			Anger			Sadness		
	R	r^2	*P* value	*R*	r^2	*P* value	*R*	r^2	*P* value
Crash rate (n=13)									
16 - 29	–0.754	0.568	.003	–0.938	0.881	<.001	–0.742	0.550	.004
30 - 59	–0.795	0.631	.001	–0.925	0.856	<.001	–0.744	0.554	.004
60 - 69	–0.818	0.669	<.001	–0.913	0.833	<.001	–0.743	0.551	.004
Over 70	–0.788	0.620	.001	–0.920	0.847	<.001	–0.735	0.540	.004
Data count (n=52)									
Tweet count	0.409	0.168	.003	0.519	0.269	<.001	0.442	0.196	.001
User count	0.407	0.166	.003	0.516	0.266	<.001	0.438	0.192	.001
Topic (n=52)									
Topic 1	0.419	0.175	.002	0.297	0.088	.030	0.158	0.025	.260
Topic 2	0.526	0.277	<.001	0.548	0.300	<.001	0.380	0.144	.005
Topic 5	–0.121	0.015	.390	–0.186	0.035	.190	–0.132	0.017	.350
Topic 6	0.375	0.141	.006	0.677	0.459	<.001	0.512	0.262	<.001
Topic 7	0.291	0.085	.040	0.235	0.055	.090	0.000	0.000	≥.99
Topic 8	0.497	0.247	<.001	0.485	0.235	<.001	0.396	0.157	.004
Topic 9	–0.514	0.264	<.001	–0.386	0.149	.005	–0.028	0.001	.840
Topic 10	0.385	0.148	.005	0.682	0.465	<.001	0.287	0.082	.040
Topic 14	0.246	0.061	.080	0.258	0.066	.070	0.366	0.134	.008

For the crash rate of individuals aged between 60 and 69 years, negative emotions (*r*=−0.80, *P*=.001) and anxiety (*r*=−0.82, *P*<.001) exhibit strong negative correlations, similar to those observed for individuals aged more than 70 years. Reversely, anger exhibits the strongest negative correlation with the youngest age group (16-29 y). Regarding the tweet count, negative emotions (*r*=0.68, *P*<.001), risk (*r*=0.68, *P*<.001), and anger (*r*=0.52, *P*<.001) exhibit moderate to strong positive correlations, similar to the user count.

For topics, topics 2 (Crash events, *r*=0.76, *P*<.001), 6 (Ikebukuro incident, *r*=0.55, *P*<.001), 8 (Social issues, *r*=0.71, *P*<.001), and 9 (License renewal, *r*=−0.50, *P*<.001) are significantly correlated with negative emotions, similar to risk. Furthermore, while positive emotions do not show significant correlations with any topic, other sentiments are correlated with multiple topics: anxiety with topics 2, 8, and 9; anger with topics 2, 6, 8, and 10 (Discussing senior driving); and sadness with topic 6.

Discussion

Principal Findings

In this study, we conducted a quantitative analysis of discourse regarding older drivers from 2010 to 2022 in Japan using Twitter. The number of tweets and users discussing older drivers has increased since 2016, with peaks observed in 2016, 2019, and 2021. Sentiment analysis revealed that negative emotions were more prevalent and increased over time compared with positive emotions. In addition, contexts related to anxiety, anger, and risk were prevalent and showed an upward trend. Topic modeling identified themes primarily related to driving licenses, crash events, personal perspectives, and traffic issues. Chronologically, despite decreasing trends in Traffic safety and License renewal, topics such as Crash events and License surrender are increasing. Finally, correlation analysis revealed that negative emotions were negatively correlated with crash rates among older drivers, positively correlated with tweet counts, and positively associated with topics such as Crash events and Ikebukuro incident.

This study quantified issues surrounding older drivers through Twitter, a leading social media platform, contributing to advancements in research on public health and ageism. The function of social media, such as posting and sharing tweets on any subject along with figures and URLs in Twitter, enables to prompt the swift dissemination of information [49] and derive population-level inferences. These have fostered new fields in public health sectors such as infoveillance [50], digital epidemiology [51], and digital disease detection [52]. They also led to an increase in health-related studies in Japan including hospitals [53], disease information [54], and eHealth literacy [55]. In addition to public health, the issue of older drivers has the aspects of ageism, which encompass our thoughts (stereotypes), emotions (prejudice), and behaviors (discrimination) toward others based on age [56]. Ageism toward

XSL•FO
RenderX

older people is especially high [57] and can adversely affect physical and mental health, such as reduced cognitive function [58], shorter life expectancy [59], deteriorated mental health [60], and increased isolation [61], prompting initiatives to counteract it [62]. Research on ageism is seen across multiple domains including health research [63], mental health services [64], long-term care facilities [65,66], workplaces [67,68], and media representations [69-71], with recent Twitter-based approaches employing thematic [72-76], descriptive [75,76], computational modeling [77,78], and quantitative text analyses [79-82]. This study, therefore, addresses the issues surrounding older drivers from both public health and ageism perspectives, contributing insights toward mitigating these challenges.

An abrupt increase in tweet counts sporadically observed over time is likely attributable to high-profile crashes caused by older drivers. The highest peak in the second quarter of 2019 corresponding to topic 6 (Ikebukuro incident) coincided with a crash that occurred in Higashi-Ikebukuro, Tokyo, drawing significant media and online attention for several months due to widespread sadness and anger toward the man, leading to the highest recorded number of driver's license surrenders in 2019 [2]. The peak in the fourth quarter of 2016, associated with topic 8 (Social issues) and topic 11 (News media) coincided with a series of crashes, and the peak in the fourth quarter of 2021, followed by a crash in Osaka, led to an increase in tweets related to topic 2 (Crash events) and topic 14 (Driving errors).

Regarding sentiments, all 3 peaks coincided with peaks in negative emotions and risk, with average proportions remaining somewhat higher after the peaks than before, while positive emotions were less relevant. Thus, the data peaks of tweets about older drivers are largely facilitated by contexts of negative emotions and risks, supported by the trends in their specific crashes aggregated by topics.

The detected topics help us understand how society forms opinions and perceptions toward older drivers in the long run. Among the 16 topics accounting for over 3.5%, we identified 7 increasing topics and 2 decreasing topics. Driving licenses are discussed in both topics 1 and 9, with public discourse increasingly focusing on license surrender (topic 1 is rising) rather than renewal (topic 9 is declining), suggesting a social shift toward discouraging older adults from driving and a potential impact on the number of license surrenders [2]. Topic 2 (Crash events), including tweets such as "It seems that crashes involving older drivers are occurring frequently" and "Let's prevent crashes involving older drivers," reflects general discussions about such events, often triggered by specific incidents, as seen in data peaks in 2016 and 2019. In contrast, topic 6 (Ikebukuro incident) includes a summary of the Ikebukuro incident and the perpetrator [46], showing a notable increase in 2019, and topic 14 (Driving errors) highlights that crashes caused by older drivers often occur due to physical or operational limitations, as shown in meta-review [4]. Over the data collection period, topic proportions shifted from topic 5 (Traffic safety; prominent until 2015) and topic 7 (Self-driving technology; high between 2013 and 2019) to topics 8 (Social issues) and 10 (Discussing senior driving), both of which have recently increased. These topics express concerns about older drivers from societal (topic 8) and family (topic 10) perspectives.

Other recurring topics include Thoughts on older drivers, Prevalent older drivers, and News media, which consistently appear in this discourse. Notably, topic 11 (News media), which includes references to major media outlets, showed a high proportion between 2016 and 2017, suggesting a potential media overemphasis on older driver–related issues. (Example tweets for these and other topics not discussed in the main text are provided in "S3. Topic Modeling Details" in Multimedia Appendix 1).

Our analysis clarifies the sentiments in public discourse and their correlations, particularly negative emotion, anxiety, anger, and risk, which have shown a persistent and increasing trend over time. The proportion of documents expressing negative emotions rose from approximately 40% to 60%, accounting for an increase of 1.4% per year. This trend is strongly positively correlated with data count and with topics 2 (Crash events), 6 (Ikebukuro incident), and 8 (Social issues). In contrast, negative emotions are most significantly and strongly negatively correlated with the crash rates of older adults (aged 60-69 y and 70+ y; P=.001), although similar trends are also observed in other age groups. These suggest that negative perceptions toward older drivers arise from heightened attention to their crash events and recognition of the issue as a social problem. Similar to negative emotions, the sentiment of risk increased from 40% to 60% (1.3% per year), with peaks around 70% in 2016 and 2019. This indicates that public perception includes not only negativity toward older drivers but also a sense of danger regarding their driving. Within "negative emotions," anxiety and anger, each accounting for approximately 20%, are more prevalent than sadness. Although they do not show sharp peaks, both have steadily increased over time. While their correlation patterns are similar to those of negative emotions, stronger correlations are observed; the negative correlation between anxiety and topic 9 (License renewal) suggests decreasing focus on continued driving, while the positive correlation between anger and topic 10 (Discussing senior driving) indicates emotional overrepresentation of older drivers in discourse. Positive emotions show no significant correlation with any of the variables. Sadness, however, is correlated with topic 6 (Ikebukuro incident), suggesting compassion for the mother and child who were victims in the incident. Overall, driving by older people is increasingly perceived as negative and risky, and this perception is influenced by strong negative emotions such as anxiety and anger, which amplify public discourse on the topic despite a declining trend in their actual crash rates.

Finally, we present our perspectives. First, negative stereotypes and prevailing public sentiment toward older drivers may influence licensing policies [3]. Our analysis reveals contrasting trends between topics related to license renewal and license surrender, suggesting a shift in public perception toward encouraging older individuals to stop driving. However, this shift tends to overlook the potential adverse effects of driving cessation, as noted in previous studies [5-7]. Although analysis of the relationship between public discourse and licensing policies remains limited, it is essential to reconsider current policies that may discourage older people from driving—such as overly strict licensing requirements [4,5] and campaigns

XSL•FO
RenderX

promoting license surrender [2]—in favor of approaches that both reduce traffic crashes and ensure their safety.

Second, the disproportionate public attention given to traffic crashes involving older drivers is a significant concern. Although the anger directed at tragic crashes involving older drivers is understandable, many older individuals follow traffic laws and drive responsibly [4], therefore it is unfair that the entire older population faces negative consequences as a result. To address this, it is necessary to implement risk-based licensing policies that focus on high-risk individuals, as well as to promote autonomous driving technologies and infrastructure that support continued safe mobility for older adults.

Third, public debate following crashes involving older drivers often centers on specific high-profile incidents rather than on the overall frequency or statistical context of such crashes. This may be partly due to media tendencies to sensationalize older driver involvement, as seen in the temporary increase in topic 11 (News Media). Media use influences the formation and stability of public opinion clusters [83] and can contribute to heightened public anxiety during periods of uncertainty, such as pandemics [84]. Providing the public with a more accurate and balanced understanding of older drivers—including countering media-driven bias—is crucial to fostering constructive discourse. This may involve encouraging more responsible media reporting and promoting public access to objective information on older driver safety.

In summary, this study offers a long-term analytical approach that integrates sentiment and textual content, contributing to research on public discourse related to aging issues and, more broadly, public health. Our findings enhance the understanding of societal perceptions and policy implications surrounding older drivers in Japan and may inform future discussions on aging societies around the world.

Limitations

Our study had several limitations. First, while Twitter is one of the most widely used social media platforms in Japan [85], Twitter users do not necessarily represent the general population [86]. Furthermore, Twitter provides access only to tweets from public accounts, deterring extreme or controversial textual content to be tweeted or we might have missed such tweets because past tweets can be deleted. In addition, display algorithms were modified around 2011 due to a rapid increase in Twitter users, resulting in nonstationarities in the dataset. To mitigate such user biases, several approaches have been proposed, including evaluating users based on the authenticity and bias of their engagements [87], and accounting for population demographics and word ambiguity [88].

Second, while the keyword (“(高齢 OR 老人) AND (運転 OR ドライバー)” in Japanese) effectively targets discussions about older drivers, there is still room for refinement. “運転” (driving) and “ドライバー” (driver) are the words related to driving, representing specifically older driving by combing words “高齢” (denoting old age) or “老人” (denoting older people). Although adding “事故” (crash) was considered, this word could collect data unrelated to crashes like falls in daily life, without words related to driving. Other age-related words like “80歳のドライバー” (80-year-old drivers) or “祖父” (grandfather) specify too detailed ages or familial relations, which are less general and more restrictive, so we did not use them aligning with our research objectives. In addition, we found that a small portion (33 out of 1000) of the sampled tweets were not related to older drivers but represented older people in nondriving contexts such as riding the bus or being crash victims. Therefore, while the data from this study appears to be sufficiently valid, there is a need to refine the keyword or explore other methods to collect more accurate data on older drivers.

Third, our methods have room for improvement. Word-based approaches for sentiment analysis and topic modeling offer valid and intuitive interpretations for emerging themes such as older driver–related discourse, but they may not fully capture contextual nuances such as negations or compound expressions. Although the J-LIWC dictionary covered 48% (4458 out of 9287) of the words used in this study—indicating substantial coverage with room for refinement—our findings remain persuasive, especially when compared with the Japanese version of the Moral Foundations Dictionary [89], which covers only 3% (275 words) (“S3. Sentiment Analysis Details” in Multimedia Appendix 1). Furthermore, the coherence score for the dynamic topic model was 0.22, lower than the commonly accepted baseline of 0.36 [36]. This may be influenced by the linguistic characteristics of Japanese and the use of a Japanese word2vec model. Finally, as correlation analysis does not imply causation, more rigorous statistical testing is needed. Therefore, although our word-based method using reliable dictionaries offers high interpretability, future work should explore alternative methods—such as deep-learning–based sentiment analysis or BERTopic for topic modeling—and incorporate techniques like supervised machine learning or community detection to gain deeper insights [24,41,90,91].

Conclusions

Despite the actual decline in crash rates among older drivers, previous surveys in Japan reveal persistent negative sentiments toward them, suggesting the need to understand discourse on social media, which serves as a platform for public debate. To quantitatively assess public awareness, we collected and analyzed tweets related to older drivers. The results revealed an increase in the number of tweets from 2010 to 2022, with certain peaks aligning with heightened attention to crashes involving older drivers. Negative emotions and risk consistently displayed high and rising levels, primarily correlated with the topic of crash events. Furthermore, there are diverse topics related to drivers’ licenses, crash events, and traffic safety. These imply unfair public recognition toward older drivers, suggesting the need for reconsidering current license policies, promoting self-driving systems, and facilitating accurate and balanced understanding, in order to provide continued safe mobility for older adults.

XSL•FO
RenderX

Acknowledgments

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI grants 23K21516 and 23K28192.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary details of the analysis.
[DOCX File, 1264 KB - infodemiology_v5i1e69321_app1.docx]

References

1. United Nations. Population division world population prospects 2024. 2024. URL: https://population.un.org [accessed 2025-04-23]
2. National Police Agency. Driver's license statistics: 2024. 2024. URL: https://www.npa.go.jp/publications/statistics/koutsuu/menkyo.html [accessed 2024-07-16]
3. National Police Agency. Report on measures against older drivers' traffic crashes. 2020. URL: https://www.npa.go.jp/koutsuu/kikaku/koureiunten/menkyoseido-bunkakai/prevention/final_report.pdf [accessed 2024-08-12]
4. Duc-nghiem N, Hoang-tung N, Kojima A, Kubota H. Traffic safety of elderly road user: the global trend and the Japanese case. J JSCE, Ser D3 2016;72(5):I. [doi: 10.2208/jscejipm.72.I_1249]
5. Marottoli RA, de Leon CFM, Glass TA, et al. Driving cessation and increased depressive symptoms: prospective evidence from the New Haven EPESE. J American Geriatrics Society 1997 Feb;45(2):202-206 [FREE Full text] [doi: 10.1111/j.1532-5415.1997.tb04508.x]
6. Chihuri S, Mielenz TJ, DiMaggio CJ, et al. Driving cessation and health outcomes in older adults. J Am Geriatr Soc 2016 Feb;64(2):332-341. [doi: 10.1111/jgs.13931] [Medline: 26780879]
7. Farquhar M. Elderly people's definitions of quality of life. Soc Sci Med 1995 Nov;41(10):1439-1446. [doi: 10.1016/0277-9536(95)00117-p] [Medline: 8560312]
8. Asahara K, Joichi Y, Tanimoto Y, Hikawa Y, Miyoshi Y. Literature review of research trends on older adult drivers in Japan from 2009 to 2019: analysis by text mining. J Yonago Med Assoc 2020;71:18-26 [FREE Full text]
9. Nakano T, Ogura Y, Yamasaki H, Yamada M. Assessment method of elderly cognitive function while driving. IEEJ TransEIS 2016;136(4):502-508. [doi: 10.1541/ieejeiss.136.502]
10. Nakahira Y, Okuda M, Torii J. Awareness in care managers and welfare volunteers regarding driving by elderly people. Bull Ehime Pref Univ Health Sci 2011;8(1):19-25 [FREE Full text]
11. Watanabe T, Fujikake K, Suzuki R. Survey on driving behavior of elderly drivers with dementia from the perspective of caregiving families. J Jpn Med J 2009;4443:89-93 [FREE Full text]
12. Ichikawa M, Inada H, Nakahara S. Effect of a cognitive test at license renewal for older drivers on their crash risk in Japan. Inj Prev 2020 Jun;26(3):234-239. [doi: 10.1136/injuryprev-2018-043117] [Medline: 30928916]
13. Ichikawa M, Inada H, Nakahara S. Increased traffic injuries among older unprotected road users following the introduction of an age-based cognitive test to the driver's license renewal procedure in Japan. Accid Anal Prev 2020 Mar;136:105440. [doi: 10.1016/j.aap.2020.105440] [Medline: 31962239]
14. National Police Agency. Traffic accident statistics: Annual report 2023. 2023. URL: https://www.npa.go.jp/publications/statistics/koutsuu/toukeihyo_e.html [accessed 2024-07-16]
15. Ichikawa M, Nakahara S, Taniguchi A. Older drivers' risks of at-fault motor vehicle collisions. Accid Anal Prev 2015 Aug;81:120-123. [doi: 10.1016/j.aap.2015.05.004] [Medline: 25980917]
16. National Police Agency of Japan. Materials for the prevention measures for older drivers. 2019. URL: https://www.npa.go.jp/koutsuu/kikaku/koureiunten/menkyoseido-bunkakai/prevention/final_report.pdf [accessed 2024-06-12]
17. Hoshi M. Is driving by the elderly dangerous. Economic Prism. 2015. URL: https://cir.nii.ac.jp/crid/1520010380785444992 [accessed 2024-07-17]
18. Nemoto M, Taniguchi A. Qualitative analysis of newspaper of traffic safety for older drivers. 2020 Presented at: Proceedings of the 62nd Conference on Infrastructure Planning and Management of the Japan Society of Civil Engineers URL: http://library.jsce.or.jp/jsce/open/00039/202011_no62/62-29-01.pdf
19. Ichikawa M, Tanaka R, Nakanishi A, Sano Y. News coverage of older drivers' fatal car crashes: Is it over-represented? J Epidemiol 2024;advpub:JE202402. [doi: 10.2188/jea.JE20240299]
20. Kemp S. Digital 2024: Global overview report. 2024. URL: https://datareportal.com/reports/digital-2024-global-overview-report [accessed 2024-08-15]
21. World Population Review. Twitter users by country / X users by country 2024. 2024. URL: https://worldpopulationreview.com/country-rankings/twitter-users-by-country [accessed 2024-08-15]

XSL•FO
RenderX

22. Song H, Omori K, Kim J, et al. Trusting social media as a source of health information: online surveys comparing the United States, Korea, and Hong Kong. J Med Internet Res 2016 Mar 14;18(3):e25. [doi: 10.2196/jmir.4193] [Medline: 26976273]
23. Molenaar A, Lukose D, Brennan L, Jenkins EL, McCaffrey TA. Using natural language processing to explore social media opinions on food security: sentiment analysis and topic modeling study. J Med Internet Res 2024 Mar 21;26:e47826. [doi: 10.2196/47826] [Medline: 38512326]
24. Bedford-Petersen C, Weston SJ. Mapping individual differences on the internet: case study of the type 1 diabetes community. JMIR Diabetes 2021 Oct 15;6(4):e30756. [doi: 10.2196/30756] [Medline: 34652277]
25. Xu WW, Tshimula JM, Dubé È, et al. Unmasking the Twitter discourses on masks during the COVID-19 pandemic: user cluster-based BERT topic modeling approach. JMIR Infodemiology 2022;2(2):e41198. [doi: 10.2196/41198] [Medline: 36536763]
26. Zhang J, Wang Y, Shi M, Wang X. Factors driving the popularity and virality of COVID-19 vaccine discourse on Twitter: text mining and data visualization study. JMIR Public Health Surveill 2021 Dec 3;7(12):e32814. [doi: 10.2196/32814] [Medline: 34665761]
27. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978. [doi: 10.2196/21978] [Medline: 33108310]
28. Kobayashi R, Takedomi Y, Nakayama Y, et al. Evolution of public opinion on COVID-19 vaccination in Japan: large-scale Twitter data analysis. J Med Internet Res 2022 Dec 22;24(12):e41928. [doi: 10.2196/41928] [Medline: 36343186]
29. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: using latent Dirichlet allocation for topic modeling on Twitter. PLOS ONE 2020;15(9):e0239441. [doi: 10.1371/journal.pone.0239441] [Medline: 32976519]
30. Show 59th year research report: study on the actual conditions and awareness of elderly drivers. Japan Safe Driving Center. 1985. URL: https://www.jsdc.or.jp/Portals/0/pdf/library/research/s59_1.pdf [accessed 2024-07-18]
31. Twitter API. 2023. URL: https://developer.x.com/ [accessed 2025-04-23]
32. kyokomi. Emoji. 2024. URL: https://github.com/kyokomi/emoji [accessed 2024-07-18]
33. Sato T, Hashimoto T, Okumura M. Implementation of a word segmentation dictionary called mecab-ipadic-NEologd and study on how to use it effectively for information retrieval (in Japanese). Proceedings of the Twenty-three Annual Meeting of the Association for Natural Language Processing 2017:NLP2017-NLPB6-1 [FREE Full text]
34. Hong L, Davison BD. Empirical study of topic modeling in twitter. Presented at: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Jul 25-28, 2010; Washington DC District of Columbia p. 80-88. [doi: 10.1145/1964858.1964870]
35. Igarashi T, Okuda S, Sasahara K. Development of the Japanese version of the linguistic inquiry and word count dictionary 2015. Front Psychol 2022;13:841534. [doi: 10.3389/fpsyg.2022.841534] [Medline: 35330723]
36. Greene D, Cross JP. Exploring the political agenda of the European parliament using a dynamic topic modeling approach. Polit Anal 2017 Jan;25(1):77-94. [doi: 10.1017/pan.2016.7]
37. Blei DM. Probabilistic topic models. Commun ACM 2012 Apr;55(4):77-84. [doi: 10.1145/2133806.2133826]
38. Vayansky I, Kumar SAP. A review of topic modeling methods. Inf Syst 2020 Dec;94:101582. [doi: 10.1016/j.is.2020.101582]
39. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text]
40. Blei DM, Lafferty JD. Dynamic topic models. Presented at: Proceedings of the 23rd International Conference on Machine Learning; Jun 25-28, 2010; Pittsburgh, Pennsylvania p. 113-120. [doi: 10.1145/1143844.1143859]
41. Grootendorst MR. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv. Preprint posted online on 2022. [doi: 10.48550/arXiv.2203.05794]
42. Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Front Sociol 2022;7:886498. [doi: 10.3389/fsoc.2022.886498] [Medline: 35602001]
43. O'Callaghan D, Greene D, Carthy J, Cunningham P. An analysis of the coherence of descriptors in topic modeling. Expert Syst Appl 2015 Aug;42(13):5645-5657. [doi: 10.1016/j.eswa.2015.02.055]
44. Tomas M, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv. Preprint posted online on 2013. [doi: 10.48550/arXiv.1301.3781]
45. Sakaki T, Mizuki S, Gunji N. BERT pre-trained model trained on large-scale Japanese social media corpus. 2019. URL: https://github.com/hottolink/hottoSNS-bert [accessed 2024-07-18]
46. Asahi Newspaper Digital. Older driver incident in Ikebukuro: 87-year-old driver hits 10 people, killing two including a young girl. 2019. URL: https://www.asahi.com/articles/ASM4M46VCM4MUTIL014.html [accessed 2024-08-13]
47. Transport Ministry of Land, Infrastructure and Japan Tourism. Fukuoka prefecture road traffic environment safety promotion liaison committee 1st meeting: "Ansuiren seminar.". 2016. URL: https://www.qsr.mlit.go.jp/fukkoku/site_files/file/iinkai_kaigi/dourokoutuu/161116shiryou.pdf [accessed 2024-08-13]
48. Asahi Newspaper Digital. 89-year-old suspect: "It's time to stop driving" - accidents involving elderly drivers continue. 2021. URL: https://www.asahi.com/articles/ASPCK6KS1PCKPTIL041.html [accessed 2024-08-13]

XSL•FO
RenderX

49. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? Presented at: Proceedings of the 19th International World Wide Web Conference; Apr 26-30, 2010; Raleigh North Carolina USA p. 591-600. [doi: 10.1145/1772690.1772751]
50. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet. J Med Internet Res 2009 Mar 27;11(1):e11. [doi: 10.2196/jmir.1157] [Medline: 19329408]
51. Salathé M, Bengtsson L, Bodnar TJ, et al. Digital epidemiology. PLOS Comput Biol 2012;8(7):e1002616. [doi: 10.1371/journal.pcbi.1002616] [Medline: 22844241]
52. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the web for public health surveillance. N Engl J Med 2009 May 21;360(21):2153-2155. [doi: 10.1056/NEJMp0900702] [Medline: 19423867]
53. Sugawara Y, Murakami M, Narimatsu H. Use of social media by hospitals and clinics in Japan: descriptive study. JMIR Med Inform 2020 Nov 27;8(11):e18666. [doi: 10.2196/18666] [Medline: 33245281]
54. Mitsutake S, Takahashi Y, Otsuki A, et al. Chronic diseases and sociodemographic characteristics associated with online health information seeking and using social networking sites: nationally representative cross-sectional survey in Japan. J Med Internet Res 2023 Mar 2;25:e44741. [doi: 10.2196/44741] [Medline: 36862482]
55. Mitsutake S, Oka K, Okan O, et al. eHealth literacy and web-based health information-seeking behaviors on COVID-19 in Japan: internet-based mixed methods study. J Med Internet Res 2024 Jul 11;26:e57842. [doi: 10.2196/57842] [Medline: 38990625]
56. World Health Organization. Global report on ageism. 2021. URL: https://www.who.int/publications/i/item/9789240016866 [accessed 2024-07-18]
57. Officer A, Thiyagarajan JA, Schneiders ML, Nash P, de la Fuente-Núñez V. Ageism, healthy life expectancy and population ageing: how are they related? Int J Environ Res Public Health 2020 May 1;17(9):3159. [doi: 10.3390/ijerph17093159] [Medline: 32370093]
58. Lamont RA, Swift HJ, Abrams D. A review and meta-analysis of age-based stereotype threat: negative stereotypes, not facts, do the damage. Psychol Aging 2015 Mar;30(1):180-193. [doi: 10.1037/a0038586] [Medline: 25621742]
59. Levy BR, Slade MD, Kunkel SR, Kasl SV. Longevity increased by positive self-perceptions of aging. J Pers Soc Psychol 2002;83(2):261-270. [doi: 10.1037//0022-3514.83.2.261]
60. Wurm S, Benyamini Y. Optimism buffers the detrimental effect of negative self-perceptions of ageing on physical and mental health. Psychol Health 2014;29(7):832-848. [doi: 10.1080/08870446.2014.891737] [Medline: 24527737]
61. Wethington E, Pillemer K, Principi A. Research in social gerontology: social exclusion of aging adults. In: Riva P, Eck J, editors. Social Exclusion: Springer; 2016:177-195. [doi: 10.1007/978-3-319-33033-4_9]
62. Officer A. A global campaign to combat ageism. Bull World Health Organ 2018 Apr 1;96(4):295-296. [doi: 10.2471/BLT.17.202424] [Medline: 29695887]
63. Chang ES, Kannoth S, Levy S, Wang SY, Lee JE, Levy BR. Global reach of ageism on older persons' health: a systematic review. PLOS ONE 2020;15(1):e0220857. [doi: 10.1371/journal.pone.0220857] [Medline: 31940338]
64. Bodner E, Yuval P, Wyman MF. Ageism in mental health assessment and treatment of older adults. In: Ayalon L, Tesch-Römer C, editors. Contemporary Perspectives on Ageism: Springer; 2018:241-262. [doi: 10.1007/978-3-319-73820-8]
65. Lagacé M, Tanguay A, Lavallée ML, Laplante J, Robichaud S. The silent impact of ageist communication in long term care facilities: elders' perspectives on quality of life and coping strategies. J Aging Stud 2012 Aug;26(3):335-342. [doi: 10.1016/j.jaging.2012.03.002]
66. Band-Winterstein T. Health care provision for older persons: the interplay between ageism and elder neglect. J Appl Gerontol 2015 Apr;34(3):NP113-NP127. [doi: 10.1177/0733464812475308] [Medline: 24652870]
67. Bal AC, Reiss AEB, Rudolph CW, Baltes BB. Examining positive and negative perceptions of older workers: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2011 Nov;66(6):687-698. [doi: 10.1093/geronb/gbr056] [Medline: 21719634]
68. Kluge A, Krings F. Attitudes toward older workers and human resource practices. Swiss J Psychol 2008 Mar;67(1):61-64. [doi: 10.1024/1421-0185.67.1.61]
69. Bai X. Images of ageing in society: a literature review. Population Ageing 2014 Sep;7(3):231-253. [doi: 10.1007/s12062-014-9103-x]
70. Loos E, Ivan L. Visual ageism in the media. In: Ayalon L, Tesch-Römer C, editors. Contemporary perspectives on ageism: Springer; 2018:163-176. [doi: 10.1007/978-3-319-73820-8_11]
71. Sánchez-Román M, Autric-Tamayo G, Fernandez-Mayoralas G, et al. Social image of old age, gendered ageism and inclusive places: older people in the media. Int J Environ Res Public Health 2022 Dec 18;19(24):17031. [doi: 10.3390/ijerph192417031] [Medline: 36554910]
72. Bacsu JD, Fraser S, Chasteen AL, et al. Using Twitter to examine stigma against people with dementia during COVID-19: infodemiology study. JMIR Aging 2022 Mar 31;5(1):e35677. [doi: 10.2196/35677] [Medline: 35290197]
73. Daneshvar H, Anderson S, Williams R, Mozaffar H. How can social media lead to co-production (co-delivery) of new services for the elderly population? A qualitative study. JMIR Hum Factors 2018 Feb 12;5(1):e5. [doi: 10.2196/humanfactors.7856] [Medline: 29434014]

XSL•FO
RenderX

74. Jimenez-Sotomayor MR, Gomez-Moreno C, Soto-Perez-de-Celis E. Coronavirus, ageism, and Twitter: An evaluation of tweets about older adults and COVID-19. J Am Geriatr Soc 2020 Aug;68(8):1661-1665. [doi: 10.1111/jgs.16508] [Medline: 32338787]
75. Xi W, Xu W, Zhang X, Ayalon L. A thematic analysis of Weibo topics (Chinese Twitter hashtags) regarding older adults during the COVID-19 outbreak. J Gerontol B Psychol Sci Soc Sci 2021 Aug 13;76(7):e306-e312. [doi: 10.1093/geronb/gbaa148] [Medline: 32882029]
76. Litchman ML, Snider C, Edelman LS, Wawrzynski SE, Gee PM. Diabetes online community user perceptions of successful aging with diabetes: analysis of a #DSMA tweet chat. JMIR Aging 2018 Jun 22;1(1):e10176. [doi: 10.2196/10176] [Medline: 31518231]
77. Fraser KC, Kiritchenko S, Nejadgholi I. Computational modeling of stereotype content in text. Front Artif Intell 2022;5:826207. [doi: 10.3389/frai.2022.826207] [Medline: 35514953]
78. Fraser KC, Kiritchenko S, Nejadgholi I. Extracting age-related stereotypes from social media texts. Presented at: Proceedings of the 13th Language Resources and Evaluation Conference; Jun 21-23, 2022 p. 3183-3194 URL: https://aclanthology.org/2022.lrec-1.341/ [accessed 2025-06-06]
79. Kong D, Chen A, Zhang J, et al. Public discourse and sentiment toward dementia on Chinese social media: machine learning analysis of Weibo posts. J Med Internet Res 2022 Sep 2;24(9):e39805. [doi: 10.2196/39805] [Medline: 36053565]
80. Ng R, Indran N. Role-based framing of older adults linked to decreased ageism over 210 years: evidence from a 600-million-word historical corpus. Gerontologist 2022 Apr 20;62(4):589-597. [doi: 10.1093/geront/gnab108] [Medline: 34323967]
81. Ng R, Indran N. #ProtectOurElders: analysis of tweets about older Asian Americans and anti-Asian sentiments during the COVID-19 pandemic. J Med Internet Res 2024 Mar 29;26:e45864. [doi: 10.2196/45864] [Medline: 38551624]
82. Stracqualursi L, Agati P. COVID-19 vaccines in Italian public opinion: Identifying key issues using Twitter and natural language processing. PLOS ONE 2022;17(11):e0277394. [doi: 10.1371/journal.pone.0277394] [Medline: 36395254]
83. Glogger I, Shehata A, Hopmann DN, Kruikemeier S. The world around us and the picture(s) in our heads: the effects of news media use on belief organization. Commun Monogr 2023;90(2):159-180. [doi: 10.1080/03637751.2022.2149830] [Medline: 37261054]
84. Greenhawt M, Kimball S, DunnGalvin A, et al. Media influence on anxiety, health utility, and health beliefs early in the SARS-CoV-2 pandemic-a survey study. J Gen Intern Med 2021 May;36(5):1327-1337. [doi: 10.1007/s11606-020-06554-y] [Medline: 33629267]
85. Ministry of Internal Affairs and Communications. Digital information flow in Japan. 2023. URL: https://www.soumu.go.jp/main_content/000919817.pdf [accessed 2024-11-25]
86. Mellon J, Prosser C. Twitter and Facebook are not representative of the general population: political attitudes and demographics of British social media users. Research & Politics 2017 Jul;4(3). [doi: 10.1177/2053168017720008]
87. Uppada SK, Manasa K, Vidhathri B, Harini R, Sivaselvan B. Novel approaches to fake news and fake account detection in OSNs: user social engagement and visual content centric model. Soc Netw Anal Min 2022;12(1):52. [doi: 10.1007/s13278-022-00878-9] [Medline: 35573810]
88. Weeg C, Schwartz HA, Hill S, Merchant RM, Arango C, Ungar L. Using Twitter to measure public discussion of diseases: a case study. JMIR Public Health Surveill 2015 Jun 26;1(1):e6. [doi: 10.2196/publichealth.3953] [Medline: 26925459]
89. Matsuo A, Sasahara K, Taguchi Y, Karasawa M. Development and validation of the Japanese moral foundations dictionary. PLOS ONE 2019;14(3):e0213343. [doi: 10.1371/journal.pone.0213343] [Medline: 30908489]
90. Wu Q, Sano Y, Takayasu H, Havlin S, Takayasu M. Twitter communities are associated with changing user's opinion towards COVID-19 vaccine in Japan. Sci Rep 2025 Apr 5;15(1):11716. [doi: 10.1038/s41598-025-95595-6] [Medline: 40188270]
91. Sano Y, Hori A. 12-year observation of tweets about rubella in Japan: a retrospective infodemiology study. PLOS ONE 2023;18(5):e0285101. [doi: 10.1371/journal.pone.0285101] [Medline: 37155655]

Abbreviations

DTM: dynamic topic model
J-LIWC: Japanese version of the Linguistic Inquiry and Word Count Dictionary
NMF: nonnegative matrix factorization

XSL•FO
RenderX

Edited by T Mackey; submitted 26.11.24; peer-reviewed by A Adenwala, R Kobayashi, S Mitra, Y Jiang; revised version received 12.04.25; accepted 22.04.25; published 16.06.25.

Please cite as:
Nakanishi A, Ichikawa M, Sano Y
Public Discourse Toward Older Drivers in Japan Using Social Media Data From 2010 to 2022: Longitudinal Analysis
JMIR Infodemiology 2025;5:e69321
URL: https://infodemiology.jmir.org/2025/1/e69321
doi:10.2196/69321

© Akito Nakanishi, Masao Ichikawa, Yukie Sano. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 16.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Sentiment Analysis Using a Large Language Model–Based Approach to Detect Opioids Mixed With Other Substances Via Social Media: Method Development and Validation

Muhammad Ahmad, MPhil; Ildar Batyrshin, PhD; Grigori Sidorov, PhD

Centro de Investigación en Computación, Instituto Politécnico Nacional, Mexico City, Mexico

Corresponding Author:
Grigori Sidorov, PhD
Centro de Investigación en Computación, Instituto Politécnico Nacional, Mexico City, Mexico

Abstract

Background: The opioid crisis poses a significant health challenge in the United States, with increasing overdoses and death rates due to opioids mixed with other illicit substances. Various strategies have been developed by federal and local governments and health organizations to address this crisis. One of the most significant objectives is to understand the epidemic through better health surveillance, and machine learning techniques can support this by identifying opioid users at risk of overdose through the analysis of social media data, as many individuals may avoid direct testing but still share their experiences online.

Objective: In this study, we take advantage of recent developments in machine learning that allow for insights into patterns of opioid use and potential risk factors in a less invasive manner using self-reported information available on social platforms.

Methods: This study used YouTube comments posted between December 2020 and March 2024, in which individuals shared their self-reported experiences of opioid drugs mixed with other substances. We manually annotated our dataset into multiclass categories, capturing both the positive effects of opioid use, such as pain relief, euphoria, and relaxation, and negative experiences, including nausea, sadness, and respiratory depression, to provide a comprehensive understanding of the multifaceted impact of opioids. By analyzing this sentiment, we used 4 state-of-the-art machine learning models, 2 deep learning models, 3 transformer models, and 1 large language model (GPT-3.5 Turbo) to predict overdose risks to improve health care response and intervention strategies.

Results: Our proposed methodology (GPT-3.5 Turbo) was highly precise and accurate, helping to automatically identify sentiment based on the adverse effects of opioid drug combinations and high-risk drug use in YouTube comments. Our proposed methodology demonstrated the highest achievable F_1-score of 0.95 and a 3.26% performance improvement over traditional machine learning models such as extreme gradient boosting, which demonstrated an F_1-score of 0.92.

Conclusions: This study demonstrates the potential of leveraging machine learning and large language models, such as GPT-3.5 Turbo, to analyze public sentiment surrounding opioid use and its associated risks. By using YouTube comments as a rich source of self-reported data, the study provides valuable insights into both the positive and negative effects of opioids, particularly when mixed with other substances. The proposed methodology significantly outperformed traditional models, contributing to more accurate predictions of overdose risks and enhancing health care responses to the opioid crisis.

(JMIR Infodemiology 2025;5:e70525) doi:10.2196/70525

KEYWORDS

opioid overdose; deep learning; large language models; high dose; NLP; chronic pain; BERT; social media; suicide; ChatGPT; natural language processing; bidirectional encoder representations from transformers; data mining; Reddit

Introduction

Background

Opioid overdose occurs when someone takes an excessive amount of prescribed or illicit drugs, such as heroin or fentanyl, which can cause potentially life-threatening symptoms by interacting with receptors in the brain and nervous system to reduce pain. Chronic pain, one of the leading causes of disability and overall disease burden worldwide [1,2], is a significant factor in the increasing use of opioid drugs mixed with illicit substances, with an estimated 20%-30% of the global population experiencing chronic pain [3,4]. The annual economic impact of chronic pain ranges from US $560 to $635 billion in the United States [5-7]. The challenging nature of chronic pain makes it one of the most persistent medical issues, presenting various diagnostic and treatment difficulties [8]. In the pharmacological management of chronic pain, opioids have long been considered essential medications for patients. Although their effectiveness in treating serious pain is generally

XSL•FO
RenderX

accepted, the use of opioids for chronic pain remains controversial due to long-term side effects such as tolerance and dependence [9,10]. These issues, along with prescription and misuse, have contributed to a significant global health crisis known as the opioid crisis [11], which has resulted in approximately 500,000 overdose deaths in the United States, with nearly 70,000 fatalities reported in 2020 alone [12].

In recent years, sentiment analysis has attracted exponential interest from researchers. The growing number of scientific publications, forums, and related conferences highlights its potential for future development. Social media platforms such as X (formerly known as Twitter), Facebook, Instagram, Reddit, and YouTube play a key role in this expansion, with over 58% of the world's population actively sharing their opinions, experiences, and concerns on these platforms [13]. These platforms provide researchers with valuable insights into health determinants by allowing the analysis of lifestyle choices, habits, and personal experiences. Social media's role in medical research is profound as it enables real-time global observations of important clinical topics, including influenza spread, suicide risk factors, and substance use trends [14-19].

Recent advances in natural language processing (NLP) have facilitated large-scale social media data analysis, making significant contributions to fields such as suicide risk detection, adverse drug reaction identification, and misinformation classification [20-22]. However, there remains a notable gap in applying key phrase extraction techniques to self-reported health-related content on social media, particularly within online health communities. The rise of web-based health care platforms has propelled automatic sentiment analysis of medical reviews into a new era of data-driven insights. This method allows researchers to analyze vast amounts of web-based user-generated data, uncovering hidden patterns about the side effects of opioid drugs. These insights are crucial for refining pharmacovigilance programs, ensuring drug safety and effectiveness. Over time, sentiment analysis in NLP has evolved significantly, enabling more accurate and meaningful interpretations of user experiences with medicines [23,24].

Prior Work

Recent years have witnessed the trend of studying opioid use disorders using social media data such as YouTube comments, X, and Instagram. Social media platforms have become essential for analyzing user-reported experiences with opioid drugs, particularly when mixed with illicit substances, as they offer valuable insights into drug use behaviors and potential overdose risks.

Carabot et al [25] used state-of-the-art machine learning (ML) models on Twitter posts related to opioid drugs. They collected a dataset from January 1, 2019, to December 31, 2020, focusing on user experiences and perceptions of these drugs. They gathered a total of 256,218 Twitter posts. They used preprocessing techniques, and only 27% of the tweets were filtered out, which shows relevancy; after preprocessing, they conducted a manual analysis of 7000 tweets using a detailed codebook. They classified users as patients, health care professionals, or institutions and distinguished between medical and nonmedical content. The findings showed that fentanyl was the most discussed opioid, with patients dominating the conversation, while health care professionals' tweets garnered the most engagement.

Swaileh et al [26] explored sentiment analysis in NLP to improve the understanding of public health and medication experiences. They used a hybrid model that combined traditional methods with advanced ML. Their proposed methodology achieved a high accuracy of 99% in sentiment classification. Their goal was to improve pharmacovigilance and inform public health initiatives by analyzing user feedback on health care and medications.

Chenworth et al [27] conducted a study to analyze public perceptions of methadone and buprenorphine-naloxone (Suboxone) through Twitter posts. They performed manual and automatic analyses, identifying common themes such as access, stigma, and treatment, with limited positive sentiment about the medications. Despite their proven effectiveness, the study suggests that public perceptions may contribute to the underutilization of these treatments for opioid use disorder.

Al-Hadhrami et al [28] explored the performance of deep learning (DL) techniques including bidirectional long short-term memory (BiLSTM) and a hybrid BiLSTM convolutional neural network (CNN) for sentiment analysis of drug-related reviews. They used Global Vectors for Word Representation (GloVe) word embedding methods and achieved an accuracy rate of 96%. The results underscore the enhanced performance of these models in analyzing patient sentiments, demonstrating the value of DL techniques in this context.

Chakrapani et al [29] discussed the challenge of analyzing the mindset of patients affected by acute diseases by introducing a framework that uses a sociomedical dataset of reviews and feedback. They used preprocessing techniques, n-gram tokenization, and polarity scoring to extract sentiments, followed by a probabilistic latent Dirichlet allocation model for review aggregation. They applied various ML models and evaluated the performance of the models in understanding patient perspectives.

Nair et al [30] focused on creating a drug review classification system to label user reviews into multiple classes, such as positive, negative, and neutral, by using publicly available datasets from drugs.com. They applied 3 variants of the pretrained bidirectional encoder representations from transformers (BERT) model, namely mBERT, SciBERT, and BioBERT, to generate embeddings used as features for various ML classifiers, including decision trees (DTs) and DL models. Model performance was assessed using precision, recall, and F_1-score metrics.

Gandy et al [31] assessed the efficacy of 3 automated sentiment analysis tools—VADER, TEXT2DATA, and LIWC-22—against manually labeled datasets of YouTube comments related to opioid epidemics. The LIWC-22 model achieved the highest accuracy with an 88% F_1-score, whereas VADER achieved 83%, and TEXT2DATA achieved 82%. The results suggest that these models can be effectively applied to social media analyses.

XSL•FO
RenderX

Although prior studies have used state-of-the-art ML and NLP models for opioid-related research, they often focused on basic sentiment classification such as positive and negative opinions and did not consider detailed discussions about mixed drug use. Many previous models did not include the many ways drugs can be mixed or their effects, which is essential for fully understanding opioid misuse. Unlike past studies, our research introduces a unique multiclass methodology with 6 different categories, including a mix of opioids and other substances. This classification captures the complexity of real-world drug use, which other studies may overlook. By using a large language model (LLM), we can better study and sort these mixed-drug experiences, detecting subtle feelings and trends that older models cannot. Our approach does more than just basic sentiment analysis. It overcomes the weaknesses of past models and gives a clearer, more complete picture of opioid misuse.

Objective

This study aims to validate a methodology that uses YouTube video comments for sentiment analysis, focusing on instances where people discuss opioid drug use mixed with other substances, increasing the risk of overdose and adverse effects. By using advanced NLP techniques and LLMs such as GPT-3.5 Turbo, this research seeks to uncover hidden patterns and derive meaningful insights from discussions about drug use. Although the information shared on social media platforms can provide valuable insights into individual experiences, it is important to note that these platforms do not directly reflect the cause and usage situations in real-world settings. Despite the high penetration of social media, the data derived from these sources cannot always be used to determine the full context of opioid misuse, overdose, or adverse effects in the real world. Unlike traditional studies that focus solely on sentiment classification, our approach directly contributes to health care by identifying high-risk behaviors and potential opioid misuse patterns, such as the combination of opioids with other substances that significantly increase overdose risk. By analyzing both the emotional tone and detailed drug use experiences, our work aims to empower public health organizations with actionable intelligence to address emerging drug trends proactively and uncover risk factors linked to the misuse of opioids, including adverse physical effects and emotional responses, which could inform public health interventions. The use of ML, DL, and LLMs such as GPT-3.5 Turbo is critical for detecting subtle patterns within large amounts of social media data, which can be difficult to identify manually. Although social media platforms do not directly reflect the full context of opioid misuse or overdose situations in the real world, these advanced techniques enhance our ability to derive accurate and actionable insights from online discussions about opioid misuse, ultimately improving patient outcomes and informing intervention strategies. Although social media data cannot fully capture the complexities of real-world usage, these techniques enable the identification of emerging risks and behavioral trends that might otherwise go unnoticed. This approach facilitates faster responses to public health concerns, enhances community safety, and minimizes reliance on manual intervention by providing comprehensive, data-driven analyses.

To achieve these objectives, we developed a meticulously curated, multilabeled corpus, where each comment was manually annotated to reflect observed adverse effects related to opioid use. The dataset encompasses 6 distinct sentiment categories, including both positive experiences (eg, pain relief, euphoria, relaxation) and negative outcomes (eg, nausea, sadness, and respiratory depression). The selection of these 6 categories was driven by a need to capture the full spectrum of user experiences, both favorable and adverse, when discussing opioid use. By including both subjective emotional states and physical effects, we can gain a more comprehensive understanding of how different opioids impact individuals. This classification approach also supports the creation of precise, targeted interventions aimed at improving health outcomes, as it allows for the identification of both beneficial and harmful patterns in opioid usage.

Contributions

This paper makes the following contributions to the literature.

We applied the schema to build a comprehensive dataset for sentiment analysis that contains opioid mixed with illicit drugs for health care professionals, accurately annotated with high-quality labels able to identify high-risk behaviors and develop targeted interventions.

We trained and tested an LLM (GPT-3.5 Turbo) on YouTube comments where people discuss using opioid drugs mixed with other substances that can cause death. This approach provides health care professionals and policymakers with real-time, data-driven insights into opioid use trends, enabling better response strategies and prevention measures.

We conducted a comprehensive series of experiments that demonstrated that the proposed methodology achieved the best performance compared to the baseline.

The proposed framework (GPT-3.5 Turbo) demonstrated an F_1-score of 0.95 in multiclass to our dataset. This represents performance improvements of 3.26% in F_1-score compared to the baseline model (extreme gradient boosting [XGBoost] demonstrated an F_1-score of 0.92).

By bridging the gap between social media sentiment analysis and health care research, this study highlights how NLP-driven methodologies can contribute to public health strategies, improve patient safety, and enhance health care delivery. However, while NLP models can significantly assist in trend identification and risk assessment, human oversight remains crucial in interpreting results and implementing appropriate public health interventions.

Methods

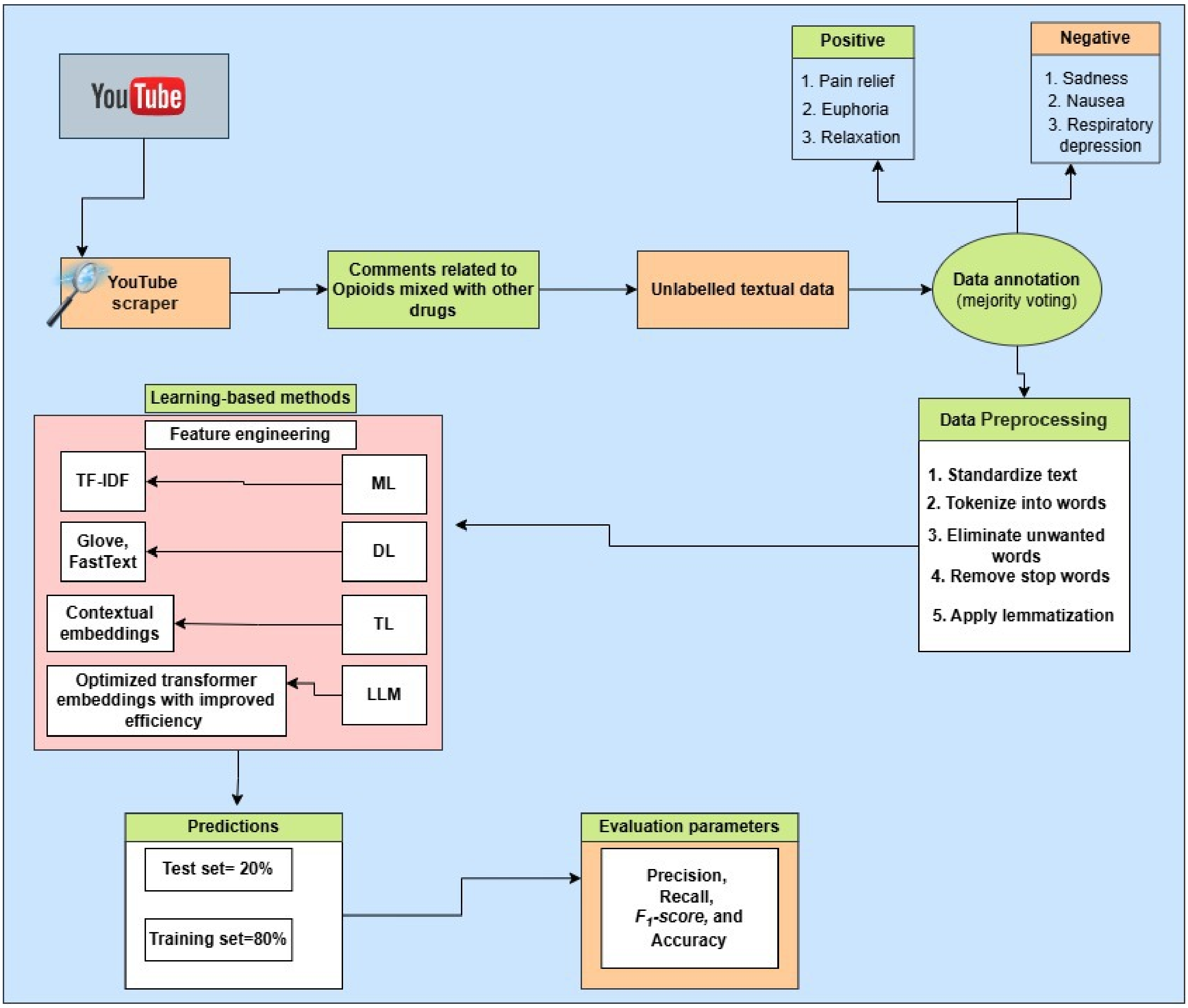
Overview

This section outlines the methodologies used to create a robust sentiment analysis system. Initially, the research design is presented in a descriptive manner, with detailed explanations provided for each component in the flow diagram (Figure 1). The methodology includes multiple phases: (1) construction of dataset, (2) annotation guidelines, (3) annotation selection, (4)

XSL•FO
RenderX

annotation agreement, (5) preprocessing and analysis of the data, (6) features extraction, and (7) application of models and training and testing.

Figure 1. Architecture of proposed solution. DL: deep learning; LLM: large language model; ML: machine learning; TL: transfer learning.



Construction of Dataset

This section outlines the construction of our dataset for sentiment analysis related to opioid overdose discussions on YouTube. First, we selected videos with more than 10 million views that were related to opioid overdose to ensure that the video had a sufficient number of comments discussing the mixing of opioid drugs with other substances. For inclusion, we selected videos based on their relevance to opioid misuse, focusing on videos with clear and significant discussions of opioid drugs mixed with other substances. We excluded videos with irrelevant content, off-topic discussions, or those lacking substantial user comments on opioid misuse and its adverse effects. For data selection, we chose videos from 2020 to 2024. One of the reasons for selecting YouTube comments from this time period was to capture recent discussions in which individuals shared their fresh experiences, especially during the COVID-19 pandemic when opioid misuse surged and individuals turned to social media more frequently to share their personal experiences. We used 20 different opioid-related keywords, such as “kratom,” “fentanyl,” “heroin,” “codeine,” and “buprenorphine,” to filter the relevant samples and drug occurrences and their adverse effects as reported by opioid users. Second, we prepared a code using the YouTube application programming interface in Python, which allowed us to collect approximately 300,000 comments from different videos that reflect self-reported and personal experiences shared by users. For this study, we selected only English-language videos and comments. Third, we manually categorized the dataset into 6 sentiment categories based on the adverse effects shared by the user, ensuring a more accurate and context-sensitive classification than traditional autoannotated methods. Unlike automatic annotation techniques, which often struggle to capture the complexity of user experiences, our manual categorization process allows for a deeper understanding of the nuanced nature of opioid use and its associated effects. By classifying the dataset into sentiment categories, we aim to develop a robust model capable of understanding both the sentiment of user concern and the adverse effects they report. This manual approach ensures high accuracy and precision, which is crucial for identifying patterns related to opioid misuse and overdose risks. An example structure of the dataset, showing sample entries

XSL•FO
RenderX

and classifications, is presented in Table 1 (see annotation guideline section). Figure 1 illustrates the proposed methodology and design used in this study, highlighting the contributions of this more detailed, context-aware classification method.

Table . Samples from the dataset.

Comment text	Sentiment
I felt this amazing rush of happiness, like everything was perfect for a few hours. I know it's risky, but nothing else makes me feel that alive.	Euphoria
The pain was unbearable, so I mixed a little extra with my regular dose. It worked for the pain, but I feel uneasy about it – I know it's dangerous.	Pain relief
After mixing opioids with alcohol, I could barely breathe; it was like my chest was weighed down. Scariest experience of my life.	Respiratory depression
I thought it would help me forget, but all it did was make me feel numb and more alone. It's not worth the spiral I'm in now	Sadness
Just a small dose with some weed, and I felt completely at ease, like I didn't have a care in the world. It's tempting to keep doing it, but I worry about the risks	Relaxation
I thought it would help me unwind, but instead, I felt so sick. I could barely keep anything down, and it just wasn't worth it	Nausea

Annotation Guidelines

After the collection of data, we accurately classified the samples related to opioid overdose drugs to gain insights into public sentiment. Each sample was labeled using predefined criteria, allowing us to classify based on the effects of drugs, including positive (pain relief, euphoria, relaxation) and negative experiences (nausea, sadness, and respiratory depression). Furthermore, the categorizations of posts are presented in Table 1 and the annotation rules are listed here:

1. Full comment reading: Mark only after reading the full comment carefully. Skim-reading will be not allowed.
2. Annotation consistency: Use accurate labels as defined in these guidelines. Any deviation, such as "Maybe" or "Unclear," is not permitted.
3. Data quality check: Annotators must verify their annotated labels before finalizing as it is a necessary step to ensure accuracy and consistency.
4. Out-of-scope content: If a YouTube sample is off-topic, such as spam or irrelevant content, mark it as "Not applicable" and remove it from the corpus.
5. Pain relief: If a sample mentions opioids or mixing other substances with opioids providing relief from physical pain, including chronic pain or injury-related pain, label it as pain relief.
6. Euphoria: If a sample demonstrates a sense of joy, bliss, or intense well-being after using opioids or opioid mixtures, label it as euphoria.
7. Relaxation: If a sample mentions the relaxing, soothing, or sedative effects of opioids or opioids leading to relaxation from stress and anxiety, mark it as relaxation.
8. Nausea: Samples that indicate feeling sick or queasy or vomiting after using opioids or other drugs mixed with opioids should be marked as nausea.
9. Sadness: Samples that indicate feelings of hopelessness or emotional downers linked to opioid use or other mixtures with opioids are marked as sadness.
10. Respiratory depression: Samples that indicate difficulty breathing or a sense of being unable to breathe properly, often as a result of opioid use, should be marked as respiratory depression.

Annotation Selection

Identifying sentiment analysis in multiclass was not an easy task; it presented significant challenges. Each of these classes added another layer of complexity, requiring annotators to carefully interpret and distinguish nuanced information within the text. This made it crucial to select annotators with strong analytical skills and attention to detail. To ensure high-quality labeling for our research, we carefully selected 5 students with strong backgrounds in annotation and ML. The selected candidates were postgraduate students in computer science. We assigned 300 comments to each candidate to label the dataset; separate Google sheets were created for individuals to record their work, which allowed us to track and evaluate their performance individually. After reviewing the results, 3 of the candidates consistently agreed on the same labels across most comments, demonstrating a high level of reliability and accuracy. Based on these results, these candidates were finalized for the full annotation of this dataset.

Annotation Agreement

During the annotation process, variations in opinion arose among annotators. It is essential to analyze these inconsistencies effectively. This evaluation was carried out by calculating the interannotator agreement, which measures the quality and consistency of the annotation process. For our annotation procedure, we used the Fleiss κ statistic to determine this agreement. Fleiss κ is particularly useful when dealing with 3 or more annotators and categorical output labels. In our case, the value of κ was found to be 0.79, suggesting substantial agreement between annotators, as it falls within the range of 0.61 to 0.80. Table 2 provides the full interpretation of κ values.

XSL•FO
RenderX

Table . Interpretation of κ values for agreement between annotators.

κ value	Interpretation
<0	Less than chance agreement
0.10 - 0.20	Slight agreement
0.21 - 0.40	Fair agreement
0.41 - 0.60	Moderate agreement
0.61 - 0.80	Substantial agreement
0.81 - 0.99	Almost perfect agreement

Ethical Considerations

This study used secondary data comprising publicly available, user-generated content collected from Reddit to analyze public sentiment on opioids mixed with other substances. The data were obtained from existing, publicly accessible Reddit posts that do not contain any personally identifiable information. All content was anonymized and analyzed in aggregate to ensure the privacy and confidentiality of individuals.

There was no direct interaction with Reddit users, and no attempt was made to trace or reidentify individuals. Given that the study involved only the analysis of publicly available data, with no human subject intervention or collection of private or identifiable information, institutional review board approval was not required.

Data Preprocessing

YouTube is a video-based social networking platform where video descriptions and comments often contain URLs, hashtags, emoticons, misspelled words, internet slang, and informal grammar expressions. In this context, data preprocessing is crucial to improving text quality, making it suitable for ML models and enhancing overall model performance, especially for sentiment analysis. For traditional ML models such as DTs and XGBoost, we applied standard preprocessing steps, including text normalization by converting all text to lowercase, removing extra spaces and newline characters, tokenizing the text into individual words, and filtering out nonalphanumeric characters. Additionally, stop words were removed using a predefined list, words shorter than 3 characters were discarded, and lemmatization was applied to ensure consistency by reducing words to their base forms. However, for DL models and transformer-based architectures such as GPT-3.5 Turbo, we avoided unnecessary preprocessing steps like tokenization, stop word removal, and term frequency-inverse document frequency (TF-IDF) transformations, as these models are designed to process raw text input directly using their own internal mechanisms for text representation. Instead, we only performed minimal cleaning (removing URLs, special characters, and excessive punctuation) to maintain linguistic integrity while reducing noise. This ensures that transformer-based models fully leverage their contextual embeddings, improving sentiment classification accuracy while preventing the loss of valuable textual information.

Data Augmentation

To enhance the performance and robustness of our proposed models, we used the back translation technique for data augmentation. For the translation process, we used the Google Translate application programming interface, which offers broad language support and high-quality translations. To handle large volumes of text efficiently, we developed custom scripts that automated the translation process. After back translation, we conducted a manual quality check on a sample of the augmented data to ensure that the original meaning was retained and that no significant information was lost during the translation.

Dataset Statistics

Multimedia Appendix 1 depicts a word cloud comprising keywords extracted from posts in the dataset related to the topic of opioid overdose. The word cloud visually highlights the most frequent terms, emphasizing the critical themes discussed in the dataset. Multimedia Appendix 2 illustrates the distribution of labels for each class used in the corpus for sentiment analysis. The chart visually represents the frequency of each sentiment class in the dataset. Multimedia Appendix 3 provides an overview of the text data's structure. It shows that the dataset contained a total of 10,129,795 characters and had a vocabulary size of 31,893 unique words. On average, each sentence had 21.39 words, and each post contained 5.02 sentences. The average post length was 541.32 characters. Additionally, each word had an average of 4.86 characters. These values give a clear picture of the dataset's complexity, showing how detailed and varied the posts are in terms of sentence and word length.

Feature Extraction

After cleaning the text, the next step was feature extraction, where we converted text into numerical form for the ML models. In traditional ML, we used TF-IDF, as shown in Equations 1 and 2, which assigns importance to words based on their frequency in a document and rarity across the dataset. This helps highlight key terms for sentiment analysis. For DL, we used GloVe and FastText embeddings, as shown in Equations 3 and 4. GloVe creates fixed vector representations based on word co-occurrence in large text collections, capturing meaningful relationships between words. FastText improves upon this by considering subword information, which helps in understanding rare and misspelled words, making the model more robust. For transformer-based models and LLMs, we used pretrained embeddings from models like BERT and ChatGPT. These models capture deep contextual meanings by analyzing entire sentences rather than individual words. Unlike traditional

XSL•FO
RenderX

methods, transformers dynamically understand context, improving sentiment analysis accuracy by recognizing complex language patterns.

$$(1)\ TF = \frac{\text{Number of times term t appears in a document}}{\text{Total number of terms in the document}}$$

The inverse document frequency of a term reflects the inverse proportion of documents that contain that term. Terms with technical jargon, for example, hold greater significance compared to words found in only a small percentage of all documents. The inverse document frequency can be computed using Equation 2:

$$(2)\ IDF = \frac{\text{Number of documents in the corpus}}{\text{Number of documents in the corpus containing term}}$$

TF-IDF can be calculated using Equation 3:

$$(3)\ TF-IDF = TF \times IDF$$

FastText extends Word2Vec by representing words as bags of character n-grams. The embedding for a word *w* is calculated using Equation 4:

$$(4)\ Vw = \sum g \in G(w) Vg$$

Where:

- A set of character n-grams in the word *w*.
- *Vg* is the vector representation of each n-gram *g*.

This allows FastText to generate embeddings for out-of-vocabulary words by combining the embeddings of their character n-grams.

GloVe creates word embeddings based on the co-occurrence matrix of words. Equation 5 is derived from the ratio of co-occurrence probabilities.

$$(5)\ Cost = \sum i,jV f(Xi,j)(ViTVj + bi + bj - \log(xi,j))2$$

Where:

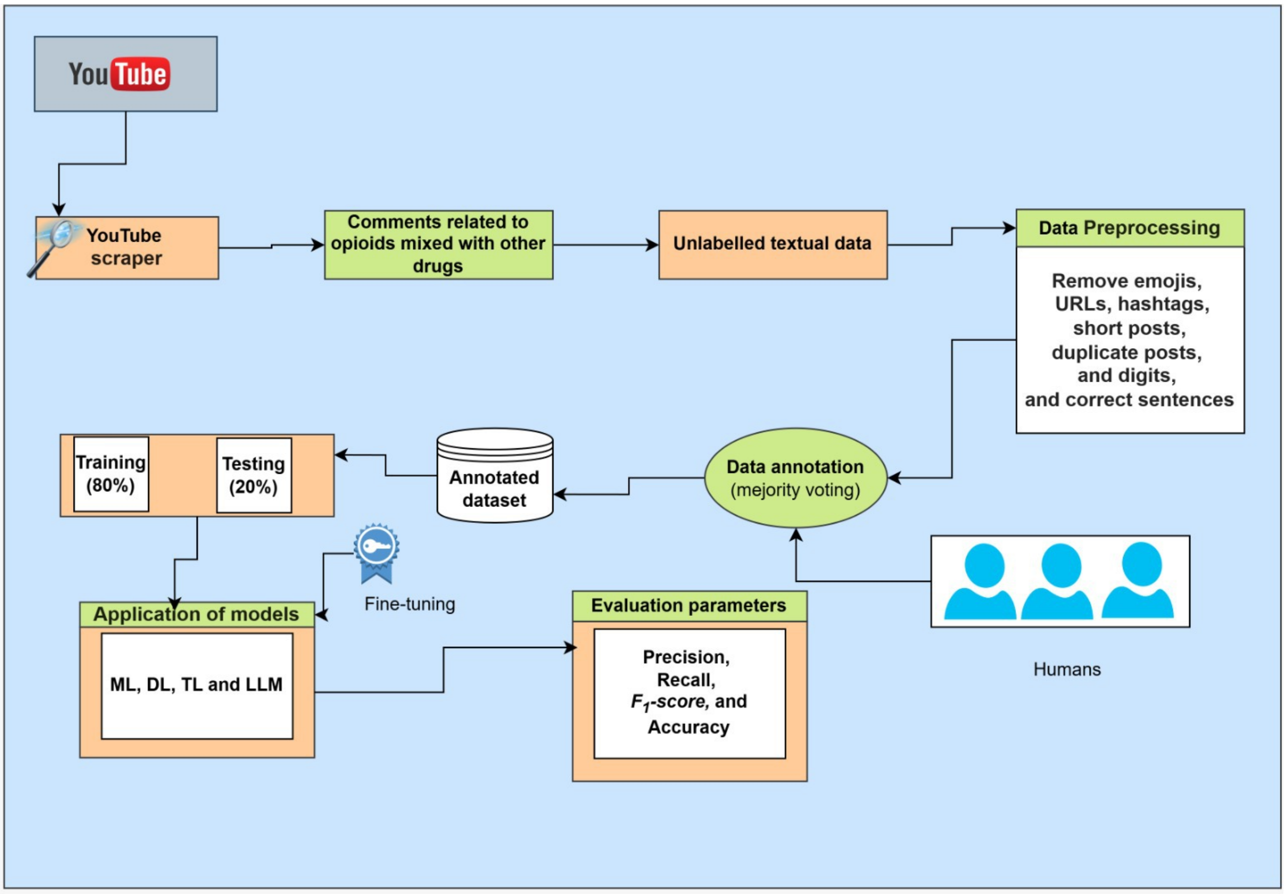
- $X_{i,j}$ is the number of times word *j* occurs in the context of word *i*.
- *V* is the vocabulary size.
- V_i and V_j are the embeddings for words *i* and *j*.
- b_i and b_j are bias terms for the words.
- $f(X_{i,j})$ is a weighting function to downweight the influence of very frequent words.

Application of Models and Training and Testing

In this section, we discuss the application of various models including ML models, DL models, transformer-based models, and LLMs such as GPT-3.5 Turbo. After feature extraction, the data were split into training and testing sets. The training set was processed to train ML algorithms including support vector machine (SVM), logistic regression (LR), k-nearest neighbor (KNN), and XGBoost, as well as 2 DL models (CNNand BiLSTM), 2 pretrained transformer models (BERT and GPT-2), and 1 LLM (GPT-3.5 Turbo). To accomplish this objective, we randomly partitioned the dataset into 80% for training and 20% for testing, as shown in Figure 2, which illustrates the ML-, DL-, and LLM-based model training pipeline for multiclass text classification. These approaches were evaluated using recall, precision, and F_1-score to quantify the performance of the models. We calculated these metrics using the following equations.

XSL•FO
RenderX

Figure 2. ML-, DL-, and LLM-based model training pipeline for multiclass text classification. DL: deep learning; LLM: large language model; ML: machine learning; TL: transfer learning.



Precision: The total number of correct predictions in our model was retrieved during document retrieval.

Recall: This indicates the classifier's ability to identify all relevant instances in the dataset.

F_1-score: The F_1-score is a metric that combines precision and recall.

Equation 8 was used for F_1-score, while Equations 6, 7, and 9 were used for precision, recall, and accuracy, respectively:

(6) $\text{Precision}=\frac{TP}{FP+TP}$

(7) $\text{Recall}=\frac{TP}{FN+TP}$

(8) $F1-\text{score}=2\times\frac{\text{Recall}\times\text{Precision}}{\text{Recall}+\text{Precision}}$

(9) $\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN}$

Where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

Results

Overview

This section discusses the results derived from the methodology, implementation, and experiments outlined earlier. For ML models, we used GridSearchCV for hyperparameter optimization, testing parameters such as regularization (eg, C and gamma for SVM), penalty terms for LR, and boosting-related settings like learning rate, number of estimators, and maximum tree depth for XGBoost. For KNN, we tuned parameters like the number of neighbors and weight functions. In the case of DL models, adjustments were made to epochs, batch sizes, and learning rates to fine-tune the BiLSTM and CNN architectures for optimal performance. For transfer learning models, fine-tuning involved modifying pretrained weights and adapting hyperparameters such as learning rates, sequence lengths, and transformer-specific configurations to improve BERT and GPT-2 on the dataset. Each model's performance was systematically optimized by fine-tuning its parameters to maximize its effectiveness. A comprehensive overview of the hyperparameters and grid search used in the proposed approach is provided in Table 3.

XSL•FO
RenderX

Table . Optimum values identified for the hyperparameters of each learning approach.

Learning approach and models		Hyperparameter	Fine tuning pipeline
Large language model: GPT-3.5 Turbo		Learning rate, epoch, batch size, seed	2, 3, 29, 1414121048
Transformer: bidirectional encoder representations from transformers, RoBERTa, crosslingual language model – RoBERTa		Learning rate, epoch, batch size, optimizer, loss function	2e-5, 3, 32, AdamW, CrossEntropy-Loss
Machine learning			
	K-nearest neighbors	n_neighbors, weights	5, uniform
	Extreme gradient boosting	n_estimators, max_depth, learning_rate	100, 6, 0.3
	Decision tree	random_state, max_depth	42, 10
	Logistic regression	random_state, max_iter, C, solver	42, 1000, 0.1, liblinear
Deep learning: bidirectional long short-term memory and convolutional neural networks		learning rate, epoch, embedding_dim, batch size,	0.1, 3, 300, 32

Software and Hardware

Experiments were conducted on a Lenovo laptop powered by an Intel Core i7, 8th generation processor with 4 cores, bus speed of 8 gigatransfers/second, 24 GB of RAM, and 1 TB of storage. The operating system used was Windows 10 Pro (Microsoft Corp), which provided a stable environment for development and execution. To perform the predictive analysis, Google Colab was selected for programming and easy access to a Python environment. We used Python version 3.12.4. The *Scikit-Learn* [32] package was used for ML models, while *TensorFlow* [33] and *Keras* [34] were used for DL tasks. For transformer-based models, the Hugging Face Transformers library was used. Model training was performed on an NVIDIA Tesla T4 GPU with 2560 CUDA cores and 16 GB GDDR6 memory.

Results for ML

In this section, we will explore the performance of several traditional ML models applied to sentiment analysis, specifically focusing on the complex topics of opioid overdose and drug mixing with other substances. To tackle this, we used 6 models including LR, KNN, random forest, and SVM. Each model was evaluated to understand how well it can detect sentiment in this sensitive area, aiming to identify patterns and nuances within the data related to drug use.

Table 4 shows the performance metrics of 4 different ML models: LR, DT, KNN, and XGBoost. We used 4 different evaluation metrics to assess the performance of these models including precision, recall, F_1-score, and accuracy. Among all models, XGBoost achieved the highest scores on all metrics (0.92 for all 4 metrics), demonstrating that it performs exceptionally well in making correct predictions in our sentiment analysis task. DT follows closely behind, with 0.87 across the board, showing strong performance just slightly lower than that of XGBoost. KNN also performed well, with an F_1-score of 0.85, but LR, while decent, lagged behind with a score of 0.74 in all metrics, suggesting that it may not be a suitable choice for our sentiment analysis task. Overall, XGBoost was the clear winner in terms of accuracy and balanced performance.

Table . Results for machine learning models.

Model	Precision	Recall	F_1-score	Accuracy
Logistic regression	0.74	0.74	0.74	0.74
Decision tree	0.87	0.87	0.87	0.87
K-nearest neighbors	0.85	0.86	0.85	0.86
Extreme gradient boosting	0.92	0.92	0.92	0.92

Results for DL

In text classification tasks, choosing the right model and word embedding technique is essential for achieving accurate results. For this analysis, we compared the performance of 2 popular DL models (CNN and BiLSTM) using 2 different types of word embeddings: FastText and GloVe.

Table 5 compares the performance of different DL models using FastText and GloVe embeddings. When using FastText, CNN performs the weakest, with an F_1-score of 0.72, while BiLSTM performs significantly better at 0.91. However, models trained with GloVe embeddings outperformed those trained with FastText. The CNN model with GloVe achieved the highest performance across all metrics (0.94), followed closely by BiLSTM with 0.93. This suggests that GloVe embeddings provide richer semantic representations for this task, leading to better model performance, especially for CNN. Overall, GloVe-based models outperformed their FastText counterparts, and CNN with GloVe achieved the best results.

XSL•FO
RenderX

Table . Results for deep learning models.

Models	Precision	Recall	F_1-score	Accuracy
FastText: convolutional neural network	0.72	0.72	0.72	0.72
FastText: bidirectional long short-term memory	0.91	0.91	0.91	0.91
Global Vectors for Word Representation: convolutional neural network	0.94	0.94	0.94	0.94
Global Vectors for Word Representation: bidirectional long short-term memory	0.93	0.93	0.93	0.93

Transformer Results

Multimedia Appendix 4 presents the performance comparison of 3 transformer-based models—RoBERTa-base, crosslingual language model (XLM)–RoBERTa-base, and BERT-base-uncased—across 4 key metrics: precision, recall, F_1-score, and accuracy. The RoBERTa-base model (blue bars) consistently outperformed the others, achieving a score of 0.94 in all metrics. The XLM-RoBERTa-base model (red bars) performed equally well in recall and accuracy but lagged slightly in precision and F_1-score. Meanwhile, BERT-base-uncased (green bars) had the lowest performance, with a score of 0.93 across all metrics. Although the differences are small, they highlight how model architecture influences classification performance, with RoBERTa-based models proving to be slightly more effective in this particular task.

Overall, RoBERTa-base outperformed the other models with the highest scores across all metrics, making it the most effective for this task. Although XLM-RoBERTa-base was close, BERT-base-uncased showed slightly lower performance.

LLM Results

LLMs have revolutionized the field of artificial intelligence by enabling machines to understand and generate human-like text with remarkable accuracy. LLM models are trained on a large volume of textual data, allowing them to capture hidden patterns in language, comprehend complex queries, and produce coherent and contextually relevant responses. By using the capabilities of LLMs such as GPT-3.5 Turbo, researchers and developers can unlock innovative solutions, bridging the gap between human communication and machine intelligence. To attain this objective, we have used the power of OpenAI's model for the sentiment analysis task and we evaluated its effectiveness using 4 metrics: precision, recall, accuracy, and F_1-score. Multimedia Appendix 5 presents the performance of GPT-3.5 Turbo across the 4 key metrics, all achieving an impressive 0.95. This indicates that GPT-3.5 Turbo performs exceptionally well in classification tasks, likely benefiting from its large-scale pretraining and contextual understanding. Compared to traditional ML models or even DL architectures, its high and balanced performance across all metrics suggests strong generalization and robustness in text classification.

Overall, GPT-3.5 Turbo excelled, with a perfect balance across all metrics (0.95), making it a highly effective choice for text classification tasks.

Table 6 shows the class-wise performance metrics of our proposed methodology (GPT-3.5 Turbo) on 6 distinct classes, capturing both positive experiences (ie, pain relief, euphoria, relaxation) and negative outcomes (ie, nausea, sadness, respiratory depression), and highlights precision, recall, F_1-score, and support (number of instances per class). Among the classes, euphoria, nausea, and respiratory depression showed the highest performance, achieving nearly perfect scores across all metrics. Euphoria, relaxation, and pain relief also performed well, with slight variations in precision and recall. Sadness, however, had the lowest recall (0.85) and F_1-score (0.89), indicating that the model struggled slightly with detecting this class.

Table . Class-wise score for the GPT-3.5 Turbo model.

Class	Precision	Recall	F_1-score	Support
Euphoria	0.97	0.97	0.97	588
Nausea	0.99	0.99	0.99	601
Pain relief	0.92	0.93	0.92	645
Relaxation	0.92	0.97	0.95	638
Respiratory depression	0.98	1	0.99	628
Sadness	0.94	0.85	0.89	643

Overall, the model performed exceptionally well across most classes, with nausea and respiratory depression achieving

XSL•FO
RenderX

near-perfect classification. However, sadness had the lowest recall, suggesting room for improvement in detecting this category.

Error Analysis

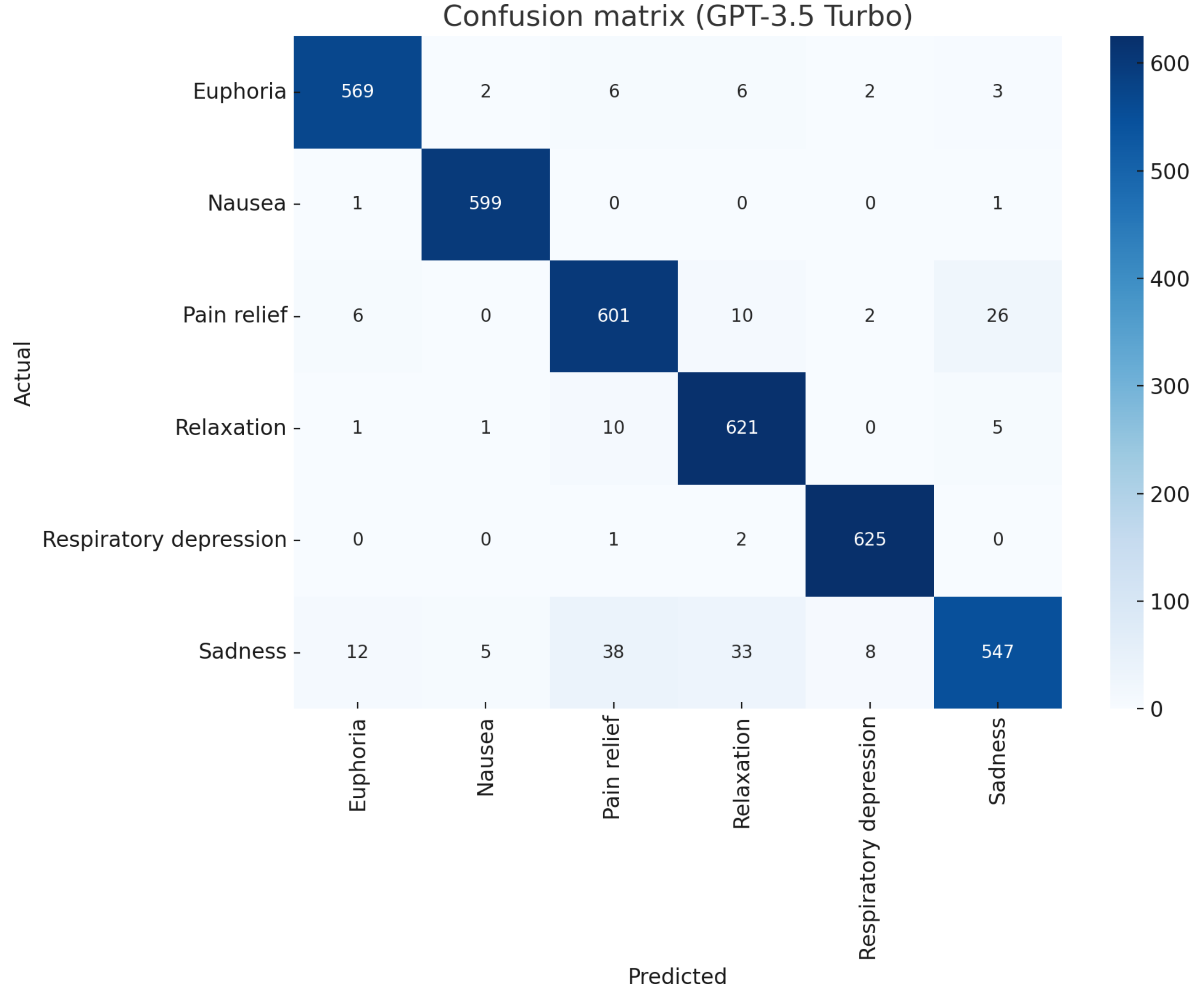
Multimedia Appendix 6 presents the top-performing models across various learning approaches based on their precision, recall, accuracy, and F_1-score metrics. Among ML techniques, the XGBoost model excelled, with solid precision, recall, F_1-score, and accuracy values of 0.92. In DL, the CNN model with GloVe embeddings achieved 0.94 in all metrics. For transfer learning, the roBERTa-base model matched this, achieving a score of 0.94 across the board as well. Finally, GPT-3.5 Turbo (an LLM) took the lead with slightly higher performance, boasting a precision, recall, F_1-score, and accuracy of 0.95, showing its exceptional ability in handling complex tasks. Overall, each approach demonstrated strong performance, but GPT-3.5 Turbo stood out as the highest achiever.

Although RoBERTa-base achieved solid performance with an accuracy, precision, recall, and F_1-score of 0.94, GPT-3.5 Turbo outperformed it with 0.95 across all metrics. This 1.06% performance improvement shows GPT-3.5's superior ability to capture complex, nuanced language patterns and generalize better to diverse user sentiments related to opioid use. Although RoBERTa excels in domain-specific tasks, GPT-3.5 Turbo's versatility allows it to handle a wider range of emotional expressions more effectively. As the dataset size increases, GPT-3.5 Turbo's performance is expected to improve further, reinforcing its edge in predicting overdose risks and understanding nuanced user experiences.

Table 3 summarizes the learning approaches, models, and hyperparameters used across various ML and DL techniques. GPT-3.5 Turbo was fine-tuned with a learning rate of 2, 3 epochs, a batch size of 29, and seed=1,414,121,048, ensuring effective adaptation. Transformer models such as BERT, RoBERTa, and XLM-RoBERTa used a learning rate of 2e-5, 3 epochs, a batch size of 32, AdamW as the optimizer, and CrossEntropyLoss for classification tasks. ML models included KNN (n_neighbors=5, weights='uniform'), XGBoost (n_estimators=100, max_depth=6, learning_rate=0.3), DT (random_state=42, max_depth=10), and LR (random_state=42, max_iter=1000, C=0.1, solver='liblinear'). DL models like BiLSTM and CNN were trained with a learning rate of 0.1, 3 epochs, an embedding dimension of 300, and a batch size of 32. Each model's hyperparameters were fine-tuned to optimize performance for specific tasks, ensuring efficient learning and improved accuracy. Figure 3 shows the confusion matrix of the proposed model (GPT-3.5 Turbo).

XSL•FO
RenderX

Figure 3. Confusion matrix of the proposed GPT-3.5 Turbo model.



Discussion

Principal Findings

This study highlights the effectiveness of sentiment analysis in extracting meaningful insights from self-reported experiences with opioid drugs mixed with illicit substances. By leveraging YouTube comments as a data source, we were able to analyze public discourse on opioid use, uncovering both positive and negative experiences. Our classification system, comprising 6 sentiment-based categories, provided a structured approach to understanding the emotional and physical effects associated with opioid consumption. Notably, this method allowed us to identify key adverse effects such as nausea, respiratory depression, and sadness, alongside reported benefits like pain relief and euphoria.

A significant contribution of this research is the application of OpenAI models such as GPT-3.5 Turbo for sentiment analysis. The model achieved an F_1-score of 0.95 in a multiclass setup, outperforming the baseline XGBoost model by 3.26%. This improvement underscores the utility of advanced NLP techniques in analyzing complex, health-related discussions. By automating the classification process, our approach reduces reliance on manual annotation and offers a scalable solution for monitoring opioid misuse trends. Such insights can enhance pharmacovigilance efforts, enabling real-time analysis of user-generated content to support public health initiatives.

Limitations

Despite its contributions, this study has several limitations. First, the reliance on YouTube as the primary data source may not fully capture the diversity of opioid-related discussions across different social media platforms. Platforms such as X, Facebook, and Reddit have distinct user demographics and language patterns, which could influence sentiment classification outcomes. Expanding data collection to multiple platforms would improve the generalizability of our findings.

Second, the manual annotation process, while aimed at ensuring accuracy, remains inherently subjective. Variability in human interpretation of comments may introduce inconsistencies in the dataset. Future studies could explore semisupervised learning techniques or crowd-sourced annotations to enhance labeling reliability.

Additionally, the 6-class sentiment framework, while comprehensive, may not capture the full spectrum of opioid-related experiences. Refining the classification system to include more granular sentiment categories could provide

XSL•FO
RenderX

deeper insights. Moreover, GPT-3.5 Turbo, despite its strong performance, exhibits occasional errors in interpreting medical terms and context-specific nuances, which may impact classification accuracy.

Conclusions and Future Work

This study demonstrates the effectiveness of ML, DL, and LLMs in analyzing public sentiment surrounding opioid use mixed with other substances. By manually annotating YouTube comments into 6 distinct sentiment-based classes—capturing both positive effects (eg, pain relief, euphoria, relaxation) and negative experiences (eg, nausea, sadness, respiratory depression)—we provided a nuanced understanding of opioid-related discussions.

Our proposed methodology, using GPT-3.5 Turbo, achieved the highest F_1-score of 0.95, outperforming traditional ML models such as XGBoost, which demonstrated an F_1-score of 0.92. This significant improvement underscores the potential of LLMs in accurately identifying high-risk opioid use patterns from user-generated content.

By leveraging social media as a real-time source of self-reported experiences, this approach offers a scalable and less invasive method for opioid surveillance. The findings highlight the potential for artificial intelligence–driven tools to enhance health care interventions and public health strategies by identifying overdose risk more accurately. Future research can expand on this work by incorporating real-time monitoring, larger datasets, and additional language models to further improve predictive performance and intervention strategies.

In future work, we will focus on several key areas. First, we will expand the dataset to include comments from multiple social media platforms, such as Reddit, X, and Facebook, which will enhance the robustness and applicability of the model. Additionally, we plan to expand our dataset to include multilingual content to capture a broader spectrum of experiences across different language groups. Incorporating demographic and geographic metadata could further refine the analysis, providing insights into regional and population-specific trends in opioid use.

Second, refining the classification system by incorporating additional sentiment categories or leveraging hierarchical classification techniques could improve the granularity of sentiment detection. Finally, integrating real-time monitoring capabilities into public health frameworks could facilitate proactive intervention strategies. By developing automated tools for detecting emerging opioid-related trends, policymakers and health care professionals could respond more swiftly to potential risks, ultimately contributing to more effective opioid crisis management.

Overall, this research underscores the potential of sentiment analysis in public health surveillance and emphasizes the need for ongoing advancements in NLP methodologies to improve opioid misuse detection and intervention strategies.

Acknowledgments

The work was done with partial support from the Mexican Government through grant A1-S-47854 of CONAHCYT, Mexico and grants 20241816, 20241819, and 20240951 of the Secretaría de Investigación y Posgrado of the Instituto Politécnico Nacional, Mexico. The authors thank the CONAHCYT for the computing resources brought to them through the Plataforma de Aprendizaje Profundo para Tecnologías del Lenguaje of the Laboratorio de Supercómputo of the INAOE, Mexico, and acknowledge the support of Microsoft through the Microsoft Latin America PhD Award.

Data Availability

The dataset will be made available on request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Word cloud of keywords extracted from the dataset.
[PNG File, 251 KB - infodemiology_v5i1e70525_app1.png]

Multimedia Appendix 2
Class-wise label distribution in the dataset.
[PNG File, 23 KB - infodemiology_v5i1e70525_app2.png]

Multimedia Appendix 3
Statistical overview of the dataset.
[PNG File, 20 KB - infodemiology_v5i1e70525_app3.png]

Multimedia Appendix 4

XSL•FO
RenderX

Transformer results.
[PNG File, 16 KB - infodemiology_v5i1e70525_app4.png]

Multimedia Appendix 5
Performance metrics (precision, recall, and F1-score) for GPT-3.5 Turbo in the sentiment analysis task.
[PNG File, 14 KB - infodemiology_v5i1e70525_app5.png]

Multimedia Appendix 6
Top-performing models across various learning approaches based on their precision, recall, accuracy, and F1-score metrics.
[JPEG File, 42 KB - infodemiology_v5i1e70525_app6.jpeg]

References

1. Jackson TP, Stabile VS, McQueen KAK. The global burden of chronic pain. ASA Monitor 2014;78(6):24-27. [doi: 10.1097/01.ASM.0001071728.63045.b4]
2. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858. [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
3. Zimmer Z, Fraser K, Grol-Prokopczyk H, Zajacova A. A global study of pain prevalence across 52 countries: examining the role of country-level contextual factors. Pain 2022 Sep 1;163(9):1740-1750. [doi: 10.1097/j.pain.0000000000002557] [Medline: 35027516]
4. Mills SEE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population-based studies. Br J Anaesth 2019 Aug;123(2):e273-e283. [doi: 10.1016/j.bja.2019.03.023] [Medline: 31079836]
5. Dahlhamer J, Lucas J, Zelaya C, et al. Prevalence of chronic pain and high-impact chronic pain among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Sep 14;67(36):1001-1006. [doi: 10.15585/mmwr.mm6736a2] [Medline: 30212442]
6. Yong RJ, Mullins PM, Bhattacharyya N. Prevalence of chronic pain among adults in the United States. Pain 2022 Feb 1;163(2):e328-e332. [doi: 10.1097/j.pain.0000000000002291] [Medline: 33990113]
7. Cohen SP, Vase L, Hooten WM. Chronic pain: an update on burden, best practices, and new advances. The Lancet 2021 May;397(10289):2082-2097. [doi: 10.1016/S0140-6736(21)00393-7]
8. Raffaeli W, Arnaudo E. Pain as a disease: an overview. J Pain Res 2017;10:2003-2008. [doi: 10.2147/JPR.S138864] [Medline: 28860855]
9. Hsu ES. Medication overuse in chronic pain. Curr Pain Headache Rep 2017 Jan;21(1):1-7. [doi: 10.1007/s11916-017-0606-z] [Medline: 28097532]
10. Nafziger AN, Barkin RL. Opioid therapy in acute and chronic pain. J Clin Pharmacol 2018 Sep;58(9):1111-1122. [doi: 10.1002/jcph.1276] [Medline: 29985526]
11. Coussens NP, Sittampalam GS, Jonson SG, et al. The opioid crisis and the future of addiction and pain therapeutics. J Pharmacol Exp Ther 2019 Nov;371(2):396-408. [doi: 10.1124/jpet.119.259408] [Medline: 31481516]
12. Lim TY, Stringfellow EJ, Stafford CA, et al. Modeling the evolution of the US opioid crisis for national policy development. Proc Natl Acad Sci U S A 2022 Jun 7;119(23):e2115714119. [doi: 10.1073/pnas.2115714119] [Medline: 35639699]
13. Anoop VS, Asharaf S. Aspect-oriented sentiment analysis: a topic modeling-powered approach. Journal of Intelligent Systems 2019 Dec 18;29(1):1166-1178. [doi: 10.1515/jisys-2018-0299]
14. Sarker A, Gonzalez-Hernandez G, Ruan Y, Perrone J. Machine learning and natural language processing for geolocation-centric monitoring and characterization of opioid-related social media chatter. JAMA Netw Open 2019 Nov 1;2(11):e1914672. [doi: 10.1001/jamanetworkopen.2019.14672] [Medline: 31693125]
15. Curtis B, Giorgi S, Buffone AEK, et al. Can Twitter be used to predict county excessive alcohol consumption rates? PLoS One 2018;13(4):e0194290. [doi: 10.1371/journal.pone.0194290] [Medline: 29617408]
16. Wang S, Ning H, Huang X, et al. Public surveillance of social media for suicide using advanced deep learning models in Japan: time series study from 2012 to 2022. J Med Internet Res 2023 Jun 2;25:e47225. [doi: 10.2196/47225] [Medline: 37267022]
17. Morese R, Gruebner O, Sykora M, Elayan S, Fadda M, Albanese E. Detecting suicide ideation in the era of social media: the population neuroscience perspective. Front Psychiatry 2022;13:652167. [doi: 10.3389/fpsyt.2022.652167] [Medline: 35492693]
18. Jashinsky J, Burton SH, Hanson CL, et al. Tracking suicide risk factors through Twitter in the US. Crisis 2014;35(1):51-59. [doi: 10.1027/0227-5910/a000234] [Medline: 24121153]
19. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature New Biol 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]

XSL•FO
RenderX

20. Mathur P, Sawhney R, Shah RR. Suicide risk assessment via temporal psycholinguistic modeling (student abstract). AAAI 2020;34(10):13873-13874. [doi: 10.1609/aaai.v34i10.7209]
21. Aroyehun ST, Gelbukh A. Detection of adverse drug reaction in tweets using a combination of heterogeneous word embeddings. : Association for Computational Linguistics Presented at: Proceedings of the Fourth Social Media Mining for Health Applications (#SMM4H) Workshop & Shared Task; Aug 2, 2019; Florence, Italy p. 133-135. [doi: 10.18653/v1/W19-3224]
22. Dharawat A, Lourentzou I, Morales A, Zhai C. Drink bleach or do what now? COVID-HeRA: a study of risk-informed health decision making in the presence of COVID-19 misinformation. ICWSM 2022;16:1218-1227. [doi: 10.1609/icwsm.v16i1.19372]
23. Levis M, Leonard Westgate C, Gui J, Watts BV, Shiner B. Natural language processing of clinical mental health notes may add predictive value to existing suicide risk models. Psychol Med 2021 Jun;51(8):1382-1391. [doi: 10.1017/S0033291720000173] [Medline: 32063248]
24. Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on Reddit during COVID-19: observational study. J Med Internet Res 2020 Oct 12;22(10):e22635. [doi: 10.2196/22635] [Medline: 32936777]
25. Carabot F, Fraile-Martínez O, Donat-Vargas C, et al. Understanding public perceptions and discussions on opioids through Twitter: cross-sectional infodemiology study. J Med Internet Res 2023 Oct 31;25:e50013. [doi: 10.2196/50013] [Medline: 37906234]
26. Swaileh A. Alzaidi M, Alshammari A, Almanea M, et al. A text-inception-based natural language processing model for sentiment analysis of drug experiences. ACM Trans Asian Low-Resour Lang Inf Process 2024. [doi: 10.1145/3678470]
27. Chenworth M, Perrone J, Love JS, Graves R, Hogg-Bremer W, Sarker A. Methadone and suboxone mentions on Twitter: thematic and sentiment analysis. Clin Toxicol (Phila) 2021 Nov;59(11):982-991. [doi: 10.1080/15563650.2021.1893742] [Medline: 33821724]
28. Al-Hadhrami S, Vinko T, Al-Hadhrami T, Saeed F, Qasem SN. Deep learning-based method for sentiment analysis for patients' drug reviews. PeerJ Comput Sci 2024;10:e1976. [doi: 10.7717/peerj-cs.1976] [Medline: 38699208]
29. Chakrapani K, Kempanna M, Safa MI, et al. An enhanced exploration of sentimental analysis in health care. Wireless Pers Commun 2023 Jan;128(2):901-922. [doi: 10.1007/s11277-022-09981-8]
30. Nair AB, Jaison DT, Anoop VS. Hey..! This medicine made me sick: sentiment analysis of user-generated drug reviews using machine learning techniques. arXiv. Preprint posted online on 2024 URL: https://arxiv.org/abs/2404.13057 [accessed 2025-05-31]
31. Gandy LM, Ivanitskaya LV, Bacon LL, Bizri-Baryak R. Public health discussions on social media: evaluating automated sentiment analysis methods. JMIR Form Res 2025 Jan 8;9:e57395. [doi: 10.2196/57395] [Medline: 39773420]
32. Pedregosa F. Scikit - learn: machine learning in python Fabian. J Mach Learn Res 2011;12:2825 [FREE Full text]
33. Abadi M, Barham P, Chen J, et al. TensorFlow: a system for {large-scale} machine learning. Presented at: 12th USENIX symposium on operating systems design and implementation (OSDI 16); Nov 2-4, 2016; Savannah, GA p. 265-283.
34. Gulli A, Pal S. Deep Learning with Keras: Packt Publishing Ltd; 2017.

Abbreviations

BERT: bidirectional encoder representations from transformers
BiLSTM: bidirectional long short-term memory
CNN: convolutional neural network
DL: deep learning
DT: decision tree
GloVe: Global Vectors for Word Representation
KNN: k-nearest neighbors
LLM: large language model
LR: logistic regression
ML: machine learning
NLP: natural language processing
SVM: support vector machine
TF-IDF: term frequency-inverse document frequency
XGBoost: extreme gradient boosting
XLM: crosslingual language model

XSL•FO
RenderX

Edited by T Mackey; submitted 23.12.24; peer-reviewed by CH Chan, J Chen, SS Madugula; revised version received 17.03.25; accepted 13.04.25; published 19.06.25.

Please cite as:
Ahmad M, Batyrshin I, Sidorov G
Sentiment Analysis Using a Large Language Model–Based Approach to Detect Opioids Mixed With Other Substances Via Social Media: Method Development and Validation
JMIR Infodemiology 2025;5:e70525
URL: https://infodemiology.jmir.org/2025/1/e70525
doi:10.2196/70525

© Muhammad Ahmad, Ildar Batyrshin, Grigori Sidorov. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 19.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Natural Language Processing and Machine Learning Techniques for Analyzing Conversations About Nutritional Yeasts in the United States and France: Retrospective Social Media Listening Study

Jean-François Jeanne[1], MSc; Joelle Malaab[2], MSc, MPH; Antoine Vanhove[1], MSc; Florian Mourey[1], MSc; Manissa Talmatkadi[2], MSc; Stéphane Schück[2], MD

[1]Gnosis by Lesaffre, Lesaffre Group, Marcq-en-Baroeul, France
[2]Kap Code, 4 rue de Cléry, Paris, France

Corresponding Author:
Joelle Malaab, MSc, MPH
Kap Code, 4 rue de Cléry, Paris, France

Abstract

Background: Nutritional yeast, an inactive form of *Saccharomyces cerevisiae*, has recently become increasingly popular as a food supplement and healthy ingredient, especially among individuals following plant-based diets. It is valued for its health benefits and high content of B vitamins, minerals, and protein. Social media has enabled people to share information and personal experiences at an unprecedented level, further amplifying conversations around health and nutrition. With the rise of social media, data mining techniques like natural language processing and machine learning are increasingly used for analyzing the large amounts of information generated on these platforms.

Objective: This study aimed to analyze social media data from the United States and France to identify the most frequently discussed topics among nutritional yeast consumers. The objective was to fill gaps in our understanding of the perceptions, experiences, and usage trends related to nutritional yeast.

Methods: This study was retrospective, using social media data geolocated in the United States and France, posted by users discussing nutritional yeast between December 2017 and September 2023. Data cleaning and filtering were done using natural language processing methods and specific algorithms. Biterm topic modeling was applied to identify the most frequently discussed topics.

Results: A total of 36,642 posts written by 28,069 users discussing nutritional yeast were identified across 1039 publicly available online sources. This included 34,292 posts from the United States (26,154 users across 994 sources) and 2350 from France (n=1915 users across 45 sources). Twitter was the most commonly used platform in both countries, accounting for 39.6% of posts in the United States (13,587/34,292) and 84.3% in France (1982/2350). In the United States, conversations centered around the role of nutritional yeast role as a vegan nutrient source (n=12,345, 36.0%). Several users highlighted its culinary versatility as a natural seasoning (n=8093, 23.6%) and its health and skin benefits (n=6173, 18.0%). In France, discussions frequently focused on nutritional yeast's use in dietary supplement routines in various forms (n=1177, 50.1%), emphasizing its benefits alongside other supplements such as castor oil, particularly noted for effects on nails and hair (n=928, 39.5%).

Conclusions: This social media listening study identified the perceptions and preferences of nutritional yeast users in France and the United States. Researchers and health care professionals can reflect on these findings to investigate the potential health benefits of nutritional yeast for specific groups and its long-term impact on different diets and lifestyles. Marketers may also use this information to create customized strategies that better align with the preferences and needs of each market.

(JMIR Infodemiology 2025;5:e60528) doi:10.2196/60528

KEYWORDS

nutritional yeast; social media listening; natural language processing; machine learning; infodemiology study

Introduction

For thousands of years, yeast has been used in the preparation of food and beverages due to its fermentation properties in baking and brewing. Nutritional yeast, an inactive form of *Saccharomyces cerevisiae* [1], has been consumed for its health benefits since the early 20th century. More recently, it has become popular as a food supplement and cooking ingredient. Nutritional yeast—also known as brewer's yeast—is rich in B vitamins, minerals, and protein, making it both nutritious and flavorful [2,3]. In today's world, where healthy living is a priority for many, it has become an important part of vegan and vegetarian diets. Individuals following a raw vegan diet and

XSL•FO
RenderX

experiencing B12 deficiency showed significant improvement after supplementing with nutritional yeast and sublingual B12 tablets, whereas probiotic supplements did not yield similar benefits [4]. Additionally, a randomized double-blind clinical trial involving adults with type 2 diabetes mellitus found that brewer's yeast supplementation improved blood pressure [5]. Another study reported beneficial effects on serum triacylglycerol and glucose tolerance [6]. Nutritional yeast has also been shown to support the immune and gastrointestinal systems; it includes probiotics and postbiotics that help in immune modulation, strengthening the body's defense against pathogens [7,8]. Furthermore, it contributes to the balance and growth of gut microbiota, promoting gastrointestinal health and overall well-being [9,10]. As awareness of healthy nutrition continues to grow, nutritional yeast is gaining popularity among a broader population looking for healthier food options or to modify their shopping, cooking, and eating habits as a whole [3,11].

As consumers become more health conscious and search for information about food supplements, they are increasingly turning to social media to gain knowledge about health and nutrition [12]. Users engage online for various purposes, including sharing information and personal experiences, participating in medical education [13], gaining awareness around health campaigns [14,15], and becoming members of online communities [16]. Twitter, Facebook, and health-related forums are increasingly used for information about nutrition, dietary supplements, and related topics [17-19], especially among younger individuals [20].

Given the vast volume of data generated through social media, data mining is being applied to analyze user-generated content [21]. This artificial intelligence technique includes natural language processing, which allows machines to analyze textual data to comprehend human language [22], and machine learning, which creates algorithms that recognize patterns in data and produce predictions [23], both of which are progressively used within this framework [24]. As a result, public perceptions, trends, and behaviors can be identified.

Although nutritional yeast is gaining popularity, knowledge about public opinions and usage patterns is still limited. While previous research has explored its health benefits, little is known about the discussions surrounding it, particularly on social media. Moreover, comparative analyses between different cultural and geographical contexts, such as the United States and France, are scarce. This study, which focuses on online discussions in the United States and France, aims to fill this gap by exploring how nutritional yeast is discussed, used, and perceived. Through data mining, we aim to identify the main topics of discussion on social media among consumers of nutritional yeast.

Methods

Study Design and Population

The present study is retrospective, using data from social media posts by nutritional yeast users geolocated in the United States and France.

Data Extraction

Between December 2017 and September 2023, messages were retrieved from general, publicly accessed sites (eg, Twitter, Reddit) and health-related forums (eg, Doctissimo in France and HealthUnlocked [25] in the United States). Due to restricted data access and closed groups, Facebook and WhatsApp were excluded from this study. An extraction query featuring relevant keywords was first developed to identify pertinent messages. Keywords associated with nutritional yeast were included in English (eg, brewer's yeast, nutritional yeast) and French (eg, levure de bière, *Saccharomyces cerevisiae*). The complete list of keywords was subsequently used in the extraction query (Multimedia Appendix 1).

Using the Brandwatch extractor (Cision Ltd.) [26], we identified and gathered all publicly available posts that contained one of the required keywords, along with their associated metadata (eg, author, publication date). Posts were also geolocated using Brandwatch. When applicable, various spellings of a keyword were considered in the extraction query. For example, the word bière was inserted as biere, bierre, and bier and the words brewer's yeast were included as brewer's yeast and brewer yeast. This approach allowed us to increase exhaustivity by including the various ways an internet user might spell a keyword.

Data Cleaning

First, messages were harmonized to ensure consistency across the dataset. We switched all characters in the messages to lowercase format and removed all accents and apostrophes from words; this approach helped achieve a smoother cleaning, and eliminating duplicates.

The cleaning process then established a list of exclusion criteria, removing messages from sources deemed unsafe or irrelevant (eg, advertising websites, forums related to cars, pets, or animals), duplicate posts, messages containing five words or less, and posts exceeding 10,000 characters. Generally, a message with fewer than five words does not contain enough information to be effectively exploitable and interpreted. Messages exceeding 10,000 characters are rare and nevertheless, are excluded from the analysis dataset due to excessive processing time.

To determine the number of messages for each keyword, we applied a "presence" step. This involved automatically searching the dataset for keywords and identifying the messages that contained them. Not all messages mentioning nutritional yeast were from people who had consumed it—some may have been discussing it without personal use. To address this, we applied a supervised machine learning algorithm [27] to identify messages specifically from nutritional yeast users. This algorithm classified each message by determining whether the user had taken the mentioned "treatment" (nutritional yeast), assigning a value of 0 for no intake and 1 for intake. It analyzed language to make predictions, including first-person pronouns to detect personal use, sentiment words to understand the emotional context, verb tenses for timing, negation words, and possessive pronouns to detect whether a statement denied or confirmed use. The algorithm was trained and tested on a sample

XSL•FO
RenderX

of 1563 messages that spanned a range of pathologies. Its ability to detect whether a treatment was taken generated the following performance results: accuracy, 72%; F_1-score, 74%; sensibility, 78%; specificity, 66%; and precision, 70%. To extend its application to English-language messages, we translated these into French before applying the algorithm, which was initially developed and optimized for French. We then separated the messages into French and US datasets based on the language of the original messages.

Multiple human annotators annotated random samples of messages to further validate the algorithm's output. This step ensured that the algorithm's results were pertinent and accurate.

Topics of Discussion

The main discussion themes among nutritional yeast consumers were identified using Biterm Topic Modeling (BTM). BTM is a natural language processing approach that analyzes large volumes of text and clusters similar text based on common topics [28]. BTM automatically groups messages into different categories—each representing a specific topic—in descending order of frequency. For each category, BTM provides a list of the most recurring words, which helps understand the general focus of each topic. For example, a BTM result about breast cancer may include:

Topic 1 | Proportion of messages = x | List of most frequently mentioned words: body, image, confidence, scars, mastectomy, hair loss, appearance, self-esteem, femininity.

These words suggest a focus on body image issues, which can be further validated when reviewing the messages associated with Topic 1. This allows us to assign a title to the topic, such as *Impact of breast cancer on body image.*

In this study, we applied BTM separately to the US and French datasets without prior knowledge of the topics. For each country, BTM generated distinct categories, providing a list of the most recurrent keywords and the messages associated with each category. Based on the list of keywords, we obtained an initial understanding of each topic. We then reviewed the messages in each category to validate and refine our understanding. This allowed us to assign appropriate titles to the topics.

Ethical Considerations

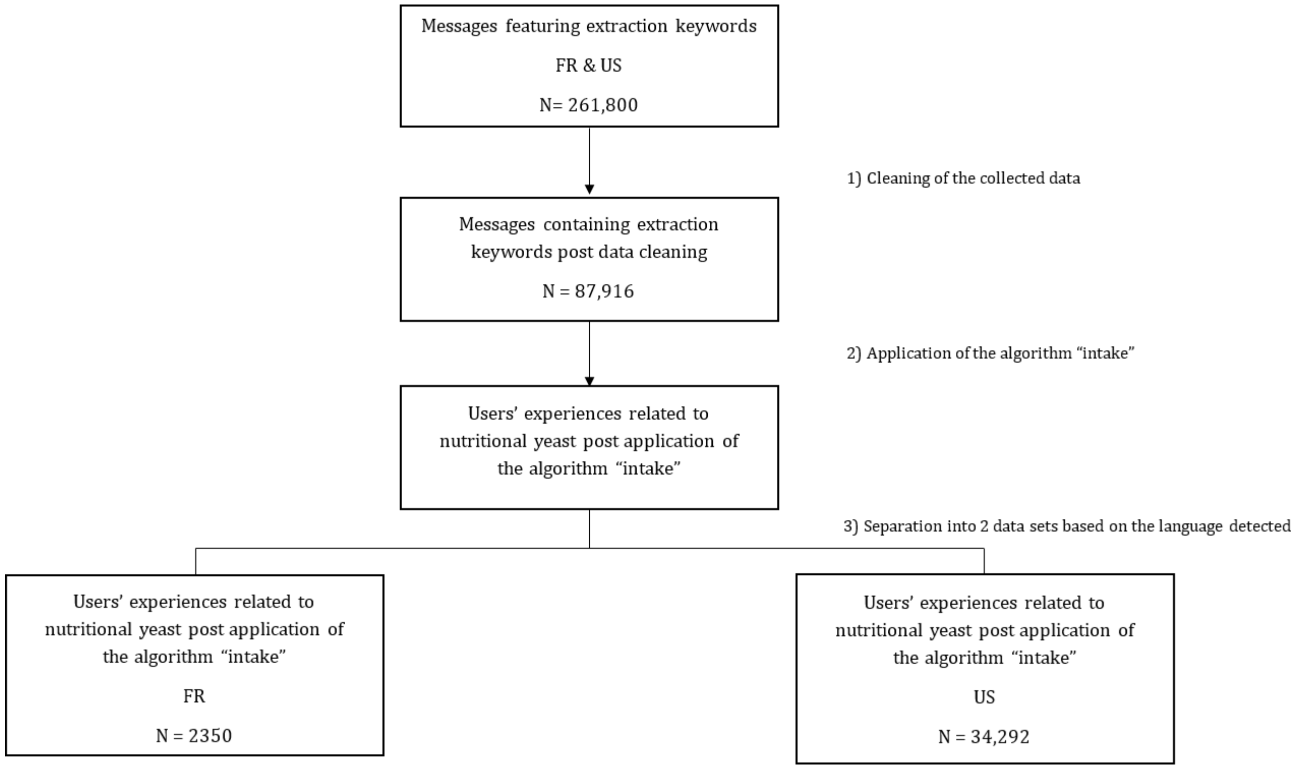
This study used only data from publicly available sources, excluding private groups, forums, and web pages. Given that users posting on public platforms automatically agree to the reuse of their information, we did not seek formal consent for this study. The findings are reported in aggregate, without personally identifiable details such as names, usernames, specific locations, or sensitive information, were deliberately removed.

Results

Population and Posts

A total of 261,800 posts were initially retrieved, written by internet users discussing nutritional yeast in the United States and France. Data cleaning and processing allowed us to identify internet users who had consumed nutritional yeast. As a result, the analysis dataset contained a total of 34,292 posts in the United States written by 26,154 users and 2350 posts in France written by 1915 users (Figure 1).

Figure 1. Flowchart of the data cleaning and sample selection processes showing the number of messages (N) and users discussing nutritional yeast in the United States and France between 2017-2023.



XSL•FO
RenderX

Data Sources and Temporal Evolution

Posts originated from 995 social media platforms in the United States and 45 in France. In the United States, Twitter was the main source of data (13,587/34,292; 39.6%) of posts, followed by Reddit.com (n=5333; 15.6%), and Instagram.com (n=4398, 12.8%). In France, Twitter was also the main source of data (1,982/2,350; 84.3%), followed by jeux vidéo.com (125; 5.3%) and babycenter.fr (n=93; 4.0%) (Table 1). The complete list of sources is found in Multimedia Appendix 2.

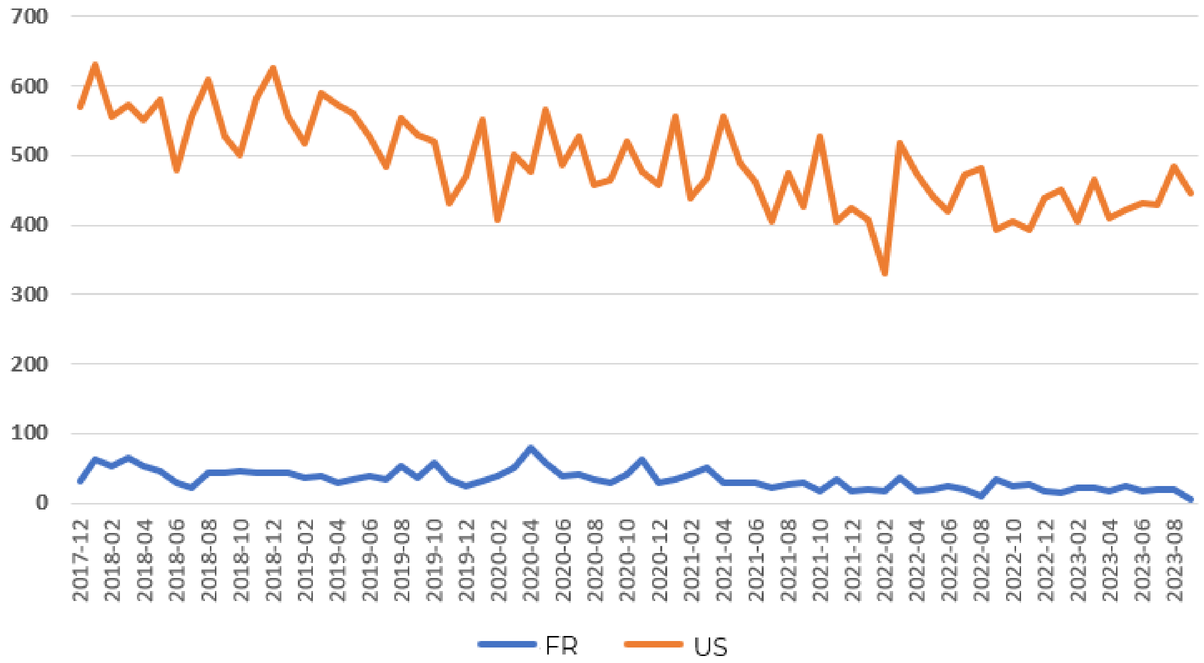
During the analysis period, the number of posts was higher in the United States (n=34,292) than in France (n=2350). Figure 2 shows the temporal evolution of the posts extracted in United States and France.

Table . Top 10 geolocated data sources in the United States and France for messages about nutritional yeast posted between 2017-2023 .

Forum/Social media	Posts, n (%)
United States	
Twitter	13,587 (39.6)
Reddit.com	5333 (15.6)
Instagram.com	4398 (12.8)
Whattoexpect.com	1134 (3.3)
Myproana.com	819 (2.4)
Babycenter.com	773 (2.3)
4channel.org	649 (1.9)
Myfitnesspal.com	483 (1.4)
Edsupportforum.com	307 (0.9)
Shroomery.org	276 (0.8)
France	
Twitter	1982 (84.3)
Jeux vidéos.com	125 (5.3)
Babycenter.fr	93 (4.0)
Instagram.com	32 (1.4)
Sports-sante.com	16 (0.7)
Hardware.fr	15 (0.6)
Au Feminin	14 (0.6)
Beauté test	7 (0.3)
Madmoizelle.com	5 (0.2)
Magic maman	5 (0.2)

XSL•FO
RenderX

Figure 2. Temporal trend in the number of posts on nutritional yeast extracted between December 2017 and September 2023 from social media geolocated in the United States and France.



Topics of Discussion

After applying the BTM, various discussion topics were identified through human interpretation of the most associated terms. It is worth noting that a single message can contain multiple topics. The main revealed topics are shown in Table 2.

Table . Proportions of messages featuring the most frequently discussed topics.

Topics	Posts, n (%)
United States	
Vegan vitamin source	12,345 (36.0)
Seasoning for various recipes	8093 (23.6)
Health and Skin benefits	6173 (18.0)
Taste of cheese	4218 (12.3)
Fermented products	2160 (6.3)
Protein and calorie balance	1269 (3.7%)
France	
Dietary supplement regimens in various forms	1177 (50.1)
Castor oil for nails and hair	928 (39.5)
Personal experience, diet, and taste	223 (9.5)
Reduction of hair loss, effectiveness, and duration	115 (4.9)
Yeast, a dietary supplement rich in vitamins	96 (4.1)
Organic products (purchase and diet)	31 (1.3%)

Main Topics of Discussion in the United States

The most discussed theme in the United States was nutritional yeast as a vegan vitamin source (12,345/34,292; 36.0%). Users from the vegan community mentioned it as a source of vitamins, particularly B vitamins including B12—a nutrient that is challenging to obtain from a diet lacking animal products. Users also shared their experiences of how to incorporate nutritional yeast into their daily lives, such as sprinkling it on meals or blending it into smoothies. An example message reads:

Yesterday's lunch was a salty (and spicy ... Def went too far on the red bell pepper) and hit the spot! Nutritional yeast may sound a bit odd, but it's a great source of B vitamins for eaters (translation to French in Multimedia Appendix 3)

XSL•FO
RenderX

Additionally, online discussions revealed that cooking enthusiasts have used nutritional yeast as a natural and nutritious seasoning (n=8093; 23.6%). Users posted their recipes using it as a flavor enhancer in pasta and casseroles, sharing that it added taste and creaminess without the need for dairy products. For example:

I bought a big carton of egg whites, what I did with the tofu scrabble was to add nutritional yeast, turmeric for color and black pepper, which makes it taste whole and with whole egg and egg white too. (translation to French in Multimedia Appendix 3*)*

Users frequently mentioned the health and skin benefits of nutritional yeast (n=6173; 18.0%) messages, with users highlighting its positive effects on skin health, particularly in managing fungal infections. They also discussed its ability to strengthen nails and improve hair health.

Nutritional yeast was also used as a cheese alternative (n=4218 messages; 12.3%). It was especially popular among vegans and individuals with lactose-intolerant, who praised it as a nutritious, plant-based option that can add flavor to dairy-free diets. Several users described it as both palatable and versatile, therefore important to include in one's balanced diet.

Another topic of discussion was the use of nutritional yeast in home fermentation processes (n=2160; 6.3% messages) such as brewing beer, making bread, and other fermented products. However, since nutritional yeast is an inactive form of *Saccharomyces cerevisiae*, it lacks the enzymes required for fermentation. This means that the conversations discussing fermentation have mistakenly mentioned the use of nutritional yeast.

The high protein content and low-calorie count also made nutritional yeast appealing within fitness and wellness communities (n=1269; 3.7%) messages. Users noted its importance as a complete protein, ie, containing all nine essential amino acids, which is beneficial for muscle development and maintenance.

Main Topics of Discussion in France

In France, internet users primarily discussed the use of nutritional yeast in food supplement routines (1177/2350; 50.1%). Users mentioned how they included brewer's yeast into their routines, sharing the different ways they consumed it in the form of capsules or pills. They reported using it for overall health or for specific concerns like hair loss and mood improvement. One user described their experience with brewer's yeast in the following message:

Personally, brewer's yeast worked really well, no side effects, I also tried Oenobiol, which wasn't bad, and now I'm starting Forcapil, apparently it's a wonder. I can't wait to have Rapunzel's hair (*translated from French, in Multimedia Appendix 3*).

Another frequently discussed topic was the use of castor oil for improving nail and hair health (n=928; 39.5%). Many users paired castor oil with brewer's yeast, highlighting their combined benefits in strengthening and stimulating the growth of hair and nails. Users also shared specific routines, such as applying castor oil directly to their hair or nails and taking brewer's yeast tablets as supplements. An example of a message is shown below:

I used to lose handfuls of them too, that's why they're so damaged, you need to make frequent oil masks that you leave on overnight, rinse them with cold water and take brewer's yeast to strengthen them ☺ (*translated from French, in Multimedia Appendix 3*).

Discussions also focused on personal experience, diet, and sensory attributes of brewer's yeast (n=223; 9.5%). Some users described its texture and taste, while others shared their recipes and methods for incorporating brewer's yeast into their diets. Feedback regarding its taste was generally positive, as shown in the following message:

No problem with taking care of oneself with all kinds of oils (jojoba, coconut, castor) natural shampoo like Liperol, and brewer's yeast in capsules and powder form (I like the taste) (*translated from French, in Multimedia Appendix 3*).

Other messages revolved around the effect of nutritional yeast in reducing hair loss (n=115; 4.9%), as well as its role as a food supplement rich in vitamins (n=96; 4.1%), particularly the B complex. Although less commonly discussed, some users expressed a preference for organic products (n=31; 1.3%), specifically organic brewer's yeast.

Discussion

Principal Findings

The objective of the study was to identify the predominant themes in conversations among internet users discussing nutritional yeast.

A total of 36,642 messages posted by 28,069 users were included in this study. Twitter emerged as the main source of discussion in both countries, accounting for 39.6% of the US dataset with 13,487 posts and 84.3% of the French dataset with 1982 posts. Additional key platforms included Reddit and Instagram in the United States, and jeux vidéo.com and babycenter.fr in France. In the US, discussions mainly focused on its role as a nutrient source for individuals following a vegan diet (36.0% of posts). Many users reported using it as a natural seasoning (23.6%) and for its benefits for health and skin (18.0%). Users also discussed cheese alternatives (12.3%), specifically using nutritional yeast as a substitute for cheese. In France, nutritional yeast was part of supplement regimens in various forms (50.1% of posts) and was used with other supplements such as castor oil, mainly to improve nail and hair health (39.5%). Additional themes included personal experiences regarding dietary habits and flavor preferences (9.5%), the role of nutritional yeast in preventing hair loss (4.9%), and its high vitamin content (4.1%).

Our findings are consistent with previous studies highlighting the role of plant-based foods in today's diets. In the United States, nutritional yeast was already recognized as an important component in vegan nutrition, with a rapid transition to plant-based diets. In fact, veganism in the United States increased by 600% between 2014 and 2018 [29], motivated by

XSL•FO
RenderX

health concerns, environmental sustainability, ethical considerations regarding animal welfare, and media influence [29-31]. Additionally, the plant-based foods market increased by 29% from 2017 to 2019 [32], highlighting the rising demand for plant-based foods among American consumers [33].

Conversely, nutritional yeast was mostly considered a dietary supplement in France, which is consistent with the French tendency to favor holistic health solutions. According to data from the Second Individual and National Study on Food Consumption (INCA), 22% of French adults consume dietary supplements on a regular basis [34], especially women, individuals aged 18 to 44, and those with higher education levels [34]. Although most adults still buy supplements from pharmacies, online purchases have significantly increased from 1% in 2015 to 11% in 2019—a clear shift in consumer habits [34]. In 2022, more than two-thirds of the French population had used dietary supplements at least once, with 32% reporting use in the past three months [35]. The market for nutritional supplements has grown to €2.6 billion and had a+3% yearly growth rate between 2021 and 2022 [36].

Our results align with previous research emphasizing the value of social media in understanding health and nutrition. With the help of data mining, social media has become a powerful source of real-world data, providing insights that traditional research methods such as clinical trials may fail to obtain. It also bypasses some of the limitations of conventional research such as lengthy timelines and complicated participant recruitment processes, making it a useful complementary tool for health research [37-39]. Infodemiological studies allow for the identification of various aspects, including specific populations, their discussion topics, the impact of health on their quality of life, as well as user perceptions and challenges [40-43]. They offer significant data, with 5.35 billion internet users and 5.04 billion actively using social media platforms in 2024 [44]. Consistent with our findings, previous studies have also identified Twitter as an essential source of data and a valuable platform for health-related discussions [38,45-48].

This study may help shape marketing strategies for each country's unique preferences. In France, where consumers were focused on improving their overall health and well-being, marketing strategies could emphasize the holistic benefits of nutritional yeast—perhaps including it as part of multisupplement regimens. In the United States, the growing population of health-conscious, plant-based consumers presents an opportunity to focus on nutritional yeast's role in vegan diets and its versatility in cooking, particularly as a natural flavor enhancer. Strategies could include creating educational content, collaborating with culinary experts, and launching targeted social media campaigns to effectively promote nutritional yeast's benefits in both regions.

Limitations

We recognize several limitations to our study. Our analysis only included openly accessible online sources; private sources such as WhatsApp, private forums, or invitation-only groups were excluded. Furthermore, the level of detail that we obtained and our understanding of the messages' context depended on the information shared by users. Our study included a potential recall bias, as it was based on users' self-reported data, their memory, and subjective interpretation. Additionally, individuals posting on social media may represent certain socioeconomic backgrounds and literacy capacities, which could affect the representativeness of our findings. It is also possible that relevant discussions were incorrectly removed during data cleaning. Another limitation is the variation in the number of users from each country, which may affect data representation. Additionally, since not all social media posts include geolocation data, accurately determining their country of origin can be challenging.

Despite these limitations, this study provides valuable insights into the discussions and perceptions about nutritional yeast.

Conclusions

Nutritional yeast is a natural ingredient valued for its health benefits and culinary versatility by users in France and the United States. Social media allowed us to gain insights into consumer perspectives, experiences, and usage trends related to nutritional yeast. It also allowed the identification of the unique preferences of each country, providing more information about the health-focused French consumers and the growing American vegan population. Future research could include clinical studies to better understand the health benefits of nutritional yeast for specific groups, such as vegans or people interested in natural beauty solutions. Studies could also explore its long-term effects on various diets and wellness habits. Additionally, marketing strategies could be improved by creating tailored communication and messaging that connect more effectively with the preferences and needs of various consumer groups.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the proprietary nature of the algorithms and data used, which are the intellectual property of Kap Code, and in order to comply with the General Data Protection Regulation (Regulation [EU] 2016/679).

Authors' Contributions

Conceptualization: JFJ, JM, AV, FM, MT
Methodology: JFJ, JM, AV, FM, MT
Data curation: MT
Formal analysis: JFJ, JM, AV, FM, MT

XSL•FO
RenderX

Writing – original draft: JM
Writing – review & editing: JFJ, AV, FM, MT

Conflicts of Interest

JM, MT, and SS are employed by Kap Code. JFJ, AV, and FM are employees of Lesaffre.

Multimedia Appendix 1
Extraction query used for Brandwatch.
[DOCX File, 14 KB - infodemiology_v5i1e60528_app1.docx]

Multimedia Appendix 2
List of sources.
[DOCX File, 35 KB - infodemiology_v5i1e60528_app2.docx]

Multimedia Appendix 3
Examples of messages in English and French.
[DOCX File, 16 KB - infodemiology_v5i1e60528_app3.docx]

References

1. Ciudad-Mulero M, Fernández-Ruiz V, Cuadrado C, et al. Novel gluten-free formulations from lentil flours and nutritional yeast: evaluation of extrusion effect on phytochemicals and non-nutritional factors. Food Chem 2020 Jun 15;315:126175. [doi: 10.1016/j.foodchem.2020.126175] [Medline: 31991254]
2. Jach ME, Serefko A. Chapter 9 - Nutritional yeast biomass: characterization and application. In: Holban AM, Grumezescu AM, editors. Diet, Microbiome and Health 2018:237-270. [doi: 10.1016/B978-0-12-811440-7.00009-0]
3. Farid F, Sideeq O, Khan F, Niaz K. Chapter 5.1 - Saccharomyces cerevisiae. In: Nabavi SM, Silva AS, editors. Nonvitamin and Nonmineral Nutritional Supplements 2019:501-508. [doi: 10.1016/B978-0-12-812491-8.00066-7]
4. Donaldson MS. Metabolic vitamin B12 status on a mostly raw vegan diet with follow-up using tablets, nutritional yeast, or probiotic supplements. Ann Nutr Metab 2000;44(5-6):229-234. [doi: 10.1159/000046689] [Medline: 11146329]
5. Hosseinzadeh P, Djazayery A, Mostafavi SA, et al. Brewer's yeast improves blood pressure in type 2 diabetes mellitus. Iran J Public Health 2013;42(6):602-609 [FREE Full text] [Medline: 23967428]
6. Li YC. Effects of brewer's yeast on glucose tolerance and serum lipids in Chinese adults. Biol Trace Elem Res 1994 Jun;41(3):341-347. [doi: 10.1007/BF02917434] [Medline: 7946924]
7. Stier H, Ebbeskotte V, Gruenwald J. Immune-modulatory effects of dietary Yeast Beta-1,3/1,6-D-glucan. Nutr J 2014 Apr 28;13(1):38. [doi: 10.1186/1475-2891-13-38] [Medline: 24774968]
8. Graubaum HJ, Busch R, Stier H, Gruenwald J. A double-blind, randomized, placebo-controlled nutritional study using an insoluble yeast beta-glucan to improve the immune defense system. FNS 2012;03(6):738-746. [doi: 10.4236/fns.2012.36100]
9. Abid R, Waseem H, Ali J, et al. Probiotic yeast Saccharomyces: back to nature to improve human health. J Fungi (Basel) 2022 Apr 24;8(5):444. [doi: 10.3390/jof8050444] [Medline: 35628700]
10. Moslehi-Jenabian S, Pedersen LL, Jespersen L. Beneficial effects of probiotic and food borne yeasts on human health. Nutrients 2010 Apr;2(4):449-473. [doi: 10.3390/nu2040449] [Medline: 22254033]
11. Fuentes M, Fuentes C. Reconfiguring food materialities: plant-based food consumption practices in antagonistic landscapes. Food Cult Soc 2022 May 27;25(3):520-539. [doi: 10.1080/15528014.2021.1903716]
12. O'Leary K, Coulson N, Perez-Vallejos E, McAuley D. Towards understanding how individuals with inflammatory bowel disease use contemporary social media platforms for health-related discourse. Comput Human Behav 2020 Nov;112:106463. [doi: 10.1016/j.chb.2020.106463]
13. Cheston CC, Flickinger TE, Chisolm MS. Social media use in medical education: A systematic review. Acad Med 2013 Jun;88(6):893-901. [doi: 10.1097/ACM.0b013e31828ffc23] [Medline: 23619071]
14. Ahmed W, Bath PA, Sbaffi L, Demartini G. In: Chowdhury G, McLeod J, Gillet V, Willett P, editors. Measuring the effect of public health campaigns on twitter: The case of World Autism Awareness Day: Springer International Publishing; 2018:10-16. [doi: 10.1007/978-3-319-78105-1_2]
15. Chung JE. A smoking cessation campaign on twitter: understanding the use of twitter and identifying major players in a health campaign. J Health Commun 2016 May;21(5):517-526. [doi: 10.1080/10810730.2015.1103332] [Medline: 27115179]
16. Gruzd A, Haythornthwaite C. Enabling community through social media. J Med Internet Res 2013 Oct 31;15(10):e248. [doi: 10.2196/jmir.2796] [Medline: 24176835]
17. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
18. Gupta P, Khan A, Kumar A. Social media use by patients in health care: a scoping review. Int J Healthc Manag 2022 Apr 3;15(2):121-131. [doi: 10.1080/20479700.2020.1860563]

XSL•FO
RenderX

19. Benetoli A, Chen TF, Aslani P. Consumer health-related activities on social media: exploratory study. J Med Internet Res 2017 Oct 13;19(10):e352. [doi: 10.2196/jmir.7656] [Medline: 29030326]
20. Goodyear VA, Armour KM, Wood H. Young people and their engagement with health-related social media: new perspectives. Sport Educ Soc 2019;24(7):673-688. [doi: 10.1080/13573322.2017.1423464] [Medline: 31814804]
21. Adedoyin-Olowe M, Gaber MM, Stahl F. A survey of data mining techniques for social network analysis. Journal of Data Mining & Digital Humanities 2014 Jun 24;2014(2014). [doi: 10.46298/jdmdh.5]
22. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. Multimed Tools Appl 2023;82(3):3713-3744. [doi: 10.1007/s11042-022-13428-4] [Medline: 35855771]
23. Deo RC. Machine learning in medicine. Circulation 2015 Nov 17;132(20):1920-1930. [doi: 10.1161/CIRCULATIONAHA.115.001593] [Medline: 26572668]
24. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc 2011;18(5):544-551. [doi: 10.1136/amiajnl-2011-000464] [Medline: 21846786]
25. A holistic approach to healthcare. HealthUnlocked. 2025. URL: https://healthunlocked.com/ [accessed 2025-04-25]
26. Brandwatch Consumer Research. Brandwatch. URL: https://www.brandwatch.com/products/consumer-research/features/ [accessed 2024-07-25]
27. Bendimerad I, Talmatkadi M, Gedik A, et al. Presented at: 14è Colloque Données de Santé en vie réelle (DSVR); Jun 20, 2023.
28. Yan X, Guo J, Lan Y, Cheng X. A biterm topic model for short texts. 2013 May 13 Presented at: WWW '13; May 13, 2013; Rio de Janeiro Brazil p. 1445-1456 URL: https://dl.acm.org/doi/proceedings/10.1145/2488388 [doi: 10.1145/2488388.2488514]
29. Forgrieve J. The growing acceptance of veganism. Forbes. 2018. URL: https://www.forbes.com/sites/janetforgrieve/2018/11/02/picturing-a-kindler-gentler-world-vegan-month/ [accessed 2024-04-04]
30. The Good Food Institute. U.S. consumer health perceptions of plant-based meat. 2024. URL: https://gfi.org/resource/us-consumer-health-perceptions-of-plant-based-meat-summary-of-recent-research/ [accessed 2024-11-18]
31. Parry J, Szejda K. How to drive plant-based food purchasing: key findings from a Mindlab study into implicit perceptions of the plant-based category. : The Good Food Institute; 2024 URL: https://gfi.org/images/uploads/2019/10/GFI-Mindlab-Report-Implicit-Study_Strategic_Recommendations.pdf [accessed 2025-04-28]
32. Retail sales data: plant-based meat, eggs, dairy. Good Food Institute. 2021 Jan 19. URL: https://gfi.org/marketresearch/ [accessed 2024-04-04]
33. Clem J, Barthel B. A look at plant-based diets. Mo Med 2021;118(3):233-238 [FREE Full text] [Medline: 34149083]
34. Food supplements, the need for informed consumption. Anses. 2019. URL: https://www.anses.fr/en/content/food-supplements-need-informed-consumption [accessed 2024-04-24]
35. Dietary supplements use in France 2022. Statista. 2024. URL: https://www.statista.com/statistics/988681/share-consumers-food-supplements-france/ [accessed 2024-04-24]
36. Food supplement sales by distribution channel in France. Statista. 2024. URL: https://www.statista.com/statistics/1113018/turnover-food-supplement-sector-by-distribution-channel-france/ [accessed 2024-04-24]
37. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
38. Odlum M, Yoon S, Broadwell P, Brewer R, Kuang D. How Twitter can support the HIV/AIDS response to achieve the 2030 eradication goal: In-depth thematic analysis of World AIDS Day tweets. JMIR Public Health Surveill 2018 Nov 22;4(4):e10262. [doi: 10.2196/10262] [Medline: 30467102]
39. Jordan SE, Hovet SE, Fung ICH, Liang H, Fu KW, Tse ZTH. Using Twitter for public health surveillance from monitoring and prediction to public response. Data (Basel) 2019;4(1):6. [doi: 10.3390/data4010006]
40. Renner S, Loussikian P, Foulquié P, et al. Perceived unmet needs in patients living with advanced bladder cancer and their caregivers: infodemiology study using data from social media in the United States. JMIR Cancer 2022 Sep 20;8(3):e37518. [doi: 10.2196/37518] [Medline: 36125861]
41. Schück S, Roustamal A, Gedik A, et al. Assessing patient perceptions and experiences of paracetamol in France: infodemiology study using social media data mining. J Med Internet Res 2021 Jul 12;23(7):e25049. [doi: 10.2196/25049] [Medline: 34255645]
42. Cotté FE, Voillot P, Bennett B, et al. Exploring the health-related quality of life of patients treated with immune checkpoint inhibitors: social media study. J Med Internet Res 2020 Sep 11;22(9):e19694. [doi: 10.2196/19694] [Medline: 32915159]
43. Voillot P, Riche B, Portafax M, et al. Social media platforms listening study on atopic dermatitis: quantitative and qualitative findings. J Med Internet Res 2022 Jan 28;24(1):e31140. [doi: 10.2196/31140] [Medline: 35089160]
44. Internet and social media users in the world 2024. Statista. 2024. URL: https://www.statista.com/statistics/617136/digital-population-worldwide/ [accessed 2024-04-04]
45. Zeraatkar K, Ahmadi M. Trends of infodemiology studies: a scoping review. Health Info Libr J 2018 Jun;35(2):91-120. [doi: 10.1111/hir.12216] [Medline: 29729073]

XSL•FO
RenderX

46. Bian J, Zhao Y, Salloum RG, et al. Using social media data to understand the impact of promotional information on laypeople's discussions: a case study of Lynch syndrome. J Med Internet Res 2017 Dec 13;19(12):e414. [doi: 10.2196/jmir.9266] [Medline: 29237586]
47. Chen T, Dredze M. Vaccine images on Twitter: analysis of what images are shared. J Med Internet Res 2018 Apr 3;20(4):e130. [doi: 10.2196/jmir.8221] [Medline: 29615386]
48. Hswen Y, Naslund JA, Brownstein JS, Hawkins JB. Monitoring online discussions about suicide among Twitter users with schizophrenia: exploratory study. JMIR Ment Health 2018 Dec 13;5(4):e11483. [doi: 10.2196/11483] [Medline: 30545811]

Abbreviations

BTM: Biterm Topic Modeling

Edited by T Mackey; submitted 15.05.24; peer-reviewed by B Senst, S Amil; revised version received 12.03.25; accepted 19.03.25; published 01.05.25.

Please cite as:
Jeanne JF, Malaab J, Vanhove A, Mourey F, Talmatkadi M, Schück S
Natural Language Processing and Machine Learning Techniques for Analyzing Conversations About Nutritional Yeasts in the United States and France: Retrospective Social Media Listening Study
JMIR Infodemiology 2025;5:e60528
URL: https://infodemiology.jmir.org/2025/1/e60528
doi:10.2196/60528

© Jean-François Jeanne, Joelle Malaab, Antoine Vanhove, Florian Mourey, Manissa Talmatkadi, Stéphane Schück. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 1.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Patient Experiences of Vulvodynia Through Reddit: Qualitative Analysis

Aurora J Grutman[1], BS; Sara Perelmuter[2], MPhil; Abigail Perez[3], BS; Janine Meurer[4], BS; Monica Contractor[5], BS; Eva Mathews[6], DO; Katie Shearer[7], BS; Lindsey A Burnett[8,9], MD, PhD; Maria Uloko[9], MD

[1]Johns Hopkins School of Medicine, Baltimore, MD, United States
[2]Weill Cornell Medical College, New York, NY, United States
[3]NYIT College of Osteopathic Medicine, Glen Head, NY, United States
[4]Jacobs School of Medicine and Biomedical Sciences at the University at Buffalo, Buffalo, NY, United States
[5]Loyola University Chicago Stritch School of Medicine, Chicago, IL, United States
[6]Department of Family Medicine, Community Memorial Healthcare, Ventura, CA, United States
[7]University of Tennessee Health Science Center, Memphis, TN, United States
[8]Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Diego, CA, United States
[9]Department of Urology, University of California, San Diego, CA, United States

Corresponding Author:
Aurora J Grutman, BS
Johns Hopkins School of Medicine, Baltimore, MD, United States

Abstract

Background: Vulvodynia is a chronic vulvar pain condition affecting up to 25% of the US population. However, diagnosis and effective treatment remain elusive. Many individuals with vulvodynia face stigma and medical uncertainty, leading them to seek information and web-based support. Reddit is a popular social media platform where patients share health concerns and experiences. The anonymity and accessibility of this platform make it a valuable source of real-world patient perspectives that are often overlooked in clinical settings.

Objective: This study evaluated Reddit content related to vulvodynia to explore how individuals with vulvodynia describe their symptoms, treatments, and personal experiences.

Methods: The subreddits “r/vulvodynia” and “r/vestibulodynia” were selected for analysis in May 2023. Threads were sorted from the most popular to least popular, with “popularity” measured by upvotes. Opening threads from the top 70 posts in each subreddit were extracted and analyzed using inductive qualitative analysis to identify themes and sentiment analysis to evaluate attitudes.

Results: In May 2023, the “r/vulvodynia” and “r/vestibulodynia” subreddits had a total of 7930 members (7245 and 685 members, respectively). Out of 140 analyzed threads, 77 (55%) contained negative attitudes. A total of 50 (35.7%) threads were seeking information or advice and 90 (64.3%) included some form of peer support. Inductive thematic analysis identified 6 core themes: symptoms (n=86, 61.4%), treatments (n=83, 59.3%), sexuality (n=47, 33.6%), erasure or disbelief (n=38, 27.1%), representation or media (n=17, 12.1%), and humor (n=15, 10.7%). Threads that discussed treatments (48/83, 57.8%), sexual experiences (25/47, 53.2%), and representation (14/17, 82.4%) had the highest proportions of positive attitudes, while threads that touched on erasure (21/38, 55.3%), symptoms (51/86, 59.3%), and humor (12/15, 80%), had the highest proportion of negative attitudes. A multivariable logistic regression of valence on the themes revealed that posts referring to treatments (odds ratio 12.5, 95% CI 3.7-42.2; $P<.001$) or representation (odds ratio 21.2, 95% CI 4.2-106.0; $P<.001$) were associated with significantly increased odds of positive valence. Furthermore, it was noted that 3 of the 5 most frequently discussed treatments aligned with clinical guidelines from the American College of Obstetricians and Gynecologists, American Urological Association, and International Society for the Study of Vulvovaginal Disease. Despite this alignment, threads frequently mentioned alternative remedies and frustration with medical professionals related to diagnostic delays and perceived lack of understanding.

Conclusions: This is the first study of Reddit discussions about vulvodynia. Findings suggest a gap between patient experiences and provider understanding, underscoring the need for improved patient education and greater clinician awareness of psychosocial factors in vulvodynia care. While limited by its sample size and lack of demographic data, this study highlights how web-based communities can help identify ways health care providers can better meet patient needs and how patients mutually support each other.

(JMIR Infodemiology 2025;5:e63072) doi:10.2196/63072

XSL•FO
RenderX

KEYWORDS

sexual health; health literacy; vulvodynia; vestibulodynia; pelvic pain; Reddit

Introduction

Vulvodynia is defined as vulvar pain lasting for at least 3 months without an identifiable cause. Vulvodynia is characterized by the location of pain (eg, localized or generalized), triggers (provoked, spontaneous, or mixed), onset (eg, primary or secondary), and temporal nature (eg, intermittent, constant, or delayed) [1,2]. The most common subtype of vulvodynia is vestibulodynia, which is pain isolated to the vulvar vestibule [2].

Vulvodynia is prevalent, affecting up to 1 in 4 women in the United States [3,4]. However, the majority of people living with vulvar pain remain undiagnosed and inadequately treated. It is reported that nearly 40% of people with chronic vulvar pain do not seek treatment, and of those that do, 60% consult at least 3 physicians before receiving a diagnosis, if they receive one at all [3]. Factors contributing to this gap in care may include a lack of knowledge on the part of medical providers, inadequate medical education related to vulvar anatomy and physiology, longstanding dismissal of female pain, and stigma surrounding female reproductive organs and sexuality [5-7].

Since the advent of the internet, individuals have sought web-based medical information, often before consulting health care professionals [8,9]. For individuals with understudied health conditions, digital health forums can be essential sources of information and peer support [10]. Moreover, web-based platforms may allow individuals with chronic diseases to connect with one another and build a social identity that extends beyond the disease itself [11].

Reddit, a popular website with 1.5 billion registered users and over 52 million daily users is a notable platform for exchange and anonymized information-sharing [12]. By design, Reddit facilitates open discussion across various topics, allowing for global information exchange that is not as readily facilitated on other platforms such as Instagram or Facebook. Reddit is organized into various subreddits, which are discussion-based communities devoted to an identified topic or theme. Within a typical subreddit, a user makes a post expressing an opinion or sharing information. Other users can evaluate that post by “upvoting” or “downvoting” it. Users can reply to the initial post or others who have replied to the post. Together, the post and subsequent comments are known as a “thread.”

The anonymous nature of Reddit may be beneficial because it lends itself to open and often vulnerable exchanges. Prior studies have explored Reddit content related to substance use, chronic pain, sexual dysfunction, and mental health, providing key insights into the lived experiences of those who may feel shame [13-17]. Despite relatively extensive investigation of these topics with varied analytic techniques, there is little research on Reddit content related to female sexual dysfunction. Existing studies primarily focus on reproductive conditions, changes in libido, or abortion [18-22]. To date, there has been no study examining how Reddit may be used by patients to obtain or share information about vulvodynia.

While population-level data on patient experiences of vulvodynia exists, patient-centered studies that capture the experiences of individuals living with vulvodynia are rare [23-25]. This study therefore seeks to qualitatively assess patient experiences of vulvodynia as discussed on Reddit, amplifying person-centered perspectives and gauging gaps in medical care for those affected by vulvodynia. By capturing these narratives, this study highlights the importance of understanding patient experiences beyond the clinical setting, which can inform more empathetic and effective health care strategies.

Methods

Subreddit Selection

To evaluate Reddit forum content related to vulvodynia, this study used a cross-sectional design, analyzing publicly available data from “r/vulvodynia” and “r/vestibulodynia,” in May 2023 [26,27]. The “r/vestibulodynia” subreddit was included to provide a more comprehensive understanding of patient experiences, as vestibulodynia is a specific form of vulvodynia.

For each subreddit, threads were sorted from the most popular to the least popular. A thread becomes popular based on the number of upvotes, comments, and overall engagement it receives from others in the Reddit community. We collected and analyzed the most popular 70 threads from each subreddit, as they were deemed most representative of key topics in the community. Comments from other users on each thread were excluded from the analysis.

Thematic and Sentiment Analysis

Quantitative and qualitative data, including the number of upvotes and comments, the post title, and a brief description of the post, were collected by accessing the Reddit website directly and navigating subreddits directly through the site. Data were collected using Google Chrome (version 113) and Safari (version 16.5) web browsers under default settings (cache and cookies enabled, no use of privacy or incognito mode) to simulate a typical user experience. The data were directly downloaded and preserved in an Excel (version 16.82; Microsoft Corp) spreadsheet on May 7, 2023, at approximately 6 PM EST. Four team members (SP, EM, KS, and AP), located in New York, Ventura, Memphis, and Glen Head, served as coders. Extracted data were coded between May 7, 2023, and May 15, 2023. These coders independently assessed all 140 threads, using a thematic analysis approach. All posts were manually annotated, and the themes were derived through iterative review and comparison. Codes were revised as necessary based on commonly identified themes, following established qualitative analysis procedures and an inductive approach in which the analysis is guided by the data itself, allowing for themes to emerge organically [28-30]. Each thread was assigned 1 or multiple themes. Threads were also evaluated for tonal expression. We defined “positive” attitudes to be any expression of optimism, relief, or joy, as well as references to “cures,” and observational or light humor. “Negative” attitudes were defined

XSL•FO
RenderX

as any expression of frustration, despair, fear, or isolation, as well as mentions of pain, pessimism, exhaustion, or unresolved symptoms, and dark humor with elements of bitterness or the macabre or morbid.

Any discrepancies in themes and attitudes based on individual coding were identified by AJG who was not involved in coding. All instances of coding discrepancy were resolved by group consensus, a practice used in qualitative research to constructively arrive at a consistent understanding of the data [31,32]. For example, in cases where coding defined the post as negative because of a user's description of vulvodynia as "terrible," but also noted that the post was "positive" because it used "humor" together, the coding team engaged in a discussion of whether the post should be coded as "negative" due to the word "terrible," or as "positive" because the presence of humor suggested a more complex emotional response, such as a coping mechanism. Examples of representative threads by theme and valence are provided in Multimedia Appendix 1.

A multivariable logistic regression of valence on the 6 identified themes was conducted using STATA/BE (version 18.0; StataCorp LLC). Each post was assigned a binary code for positive or negative valence and the presence or absence of a particular theme. A 2-sided significance level was defined at α=.05.

Any mention of treatments in the subreddit was recorded and compared with guidelines from the American College of Obstetricians and Gynecologists (ACOG), American Urological Association (AUA), and International Society for the Study of Vulvovaginal Disease (ISSVD), as they are viewed as the primary sources of information about managing vulvar pain.

Ethical Considerations

The study was deemed exempt by the Institutional Review Board of the University of California San Diego and Johns Hopkins Institutional Review Board due to its observational nature and analysis of public web-based content. The original data were collected in compliance with Reddit's public content policy, which informs users that researchers can access Reddit's public content for research purposes. All Reddit usernames and any potentially identifiable information were deidentified to protect user privacy. Furthermore, no direct user interactions or private messages were included in the analysis. Only publicly accessible forum posts were analyzed, and efforts were made to ensure that the data could not be traced back to individual users through reverse searchability. In consideration of the potential ethical concerns related to social media–based research, the authors acknowledge the need to engage in ongoing academic debates regarding internet research ethics. While Reddit users agree to the public visibility of their posts, the authors recognize that these ethical discussions, such as those put forth by the Association of Internet Researchers, underscore the need to balance public data use with user privacy in research contexts.

Results

At the time of analysis, the "r/vulvodynia" and "r/vestibulodynia" subreddits had a combined total of 7930 members (7245 and 685 members, respectively). A total of 140 posts were analyzed; these posts received an aggregate of 4166 upvotes. Out of all 140 analyzed threads, 50 (35.7%) were deemed to be seeking information or advice and 90 (64.3%) were deemed to involve peer support discussions of personal experiences related to vulvodynia.

Six core themes emerged from the qualitative analysis: (1) Reddit users' subjective sense of being disbelieved about symptoms or erasure more generally; (2) difficulty managing symptoms; (3) the condition's impact on sexuality and sexual experiences; (4) representation or media; (5) humor as a coping technique or a response to the condition; and (6) treatments sought or tried.

Out of the 140 threads, the most frequently observed themes were symptoms (n=86, 61.4%) and treatments (n=83, 59.3%), followed by sexuality (n=47, 33.6%), erasure or disbelief (n=38, 27.1%), representation or media (n=17, 12.1%), and humor (n=15, 10.7%). Of all 140 analyzed threads, 45% (n=63) of threads were coded as reflecting positive attitudes, and 55% (n=77) of threads were coded as reflecting negative attitudes. The core themes of treatments (48/83, 57.8%), sexual experiences (25/47, 53.2%), and representation (14/17, 82.4%) had the highest proportions of positive attitudes in analyzed threads. The themes of humor (12/15, 80%), erasure or disbelief (21/38, 55.3%), and symptoms (51/86, 59.3%) had the highest proportions of negative attitude threads. Fleiss κ for coded valence was 0.54. The distribution of positive and negative attitudes across themes is illustrated in Table 1.

Results from the multivariable logistic regression revealed that only treatments (odds ratio [OR] 12.5, 95% CI 3.7-42.2; *P*<.001) and representation (OR 21.2, 95% CI 4.2-106.0; *P*<.001) were associated with significantly increased odds of positive valence. Nonsignificant associations were found for themes erasure (OR 1.25, 95% CI 0.5-3.0; *P*=.60), symptoms (OR 0.47, 95% CI 0.2-1.2; *P*=.11), and sexuality (OR 2.2, 95% CI 0.9-5.1; *P*=.07).

There were 119 instances of treatment discussions across the 140 analyzed threads. The most commonly mentioned treatments included topical medications (n=22, 18.5%), physical therapy (n=22, 18.5%), surgery (n=16, 13.4%), dilators (n=14, 11.8%), and stopping oral contraceptive pills (n=11, 9.2%). Three of the 5 most discussed treatments—physical therapy, topical medications, and surgery—aligned with clinical guidelines from ACOG, AUA, and ISSVD.

XSL•FO
RenderX

Table . Positive and negative attitudes by theme.

Theme	Positive	Negative	Total
Erasure or disbelief	17	21	38
Symptoms	35	51	86
Treatments	48	35	83
Sexuality	25	22	47
Representation or media	14	3	17
Humor	3	12	15

Discussion

Principal Findings

This is the first study to analyze Reddit posts about vulvodynia. On Reddit, individuals with vulvodynia shared personal experiences, provided advice, and found communal support. From the qualitative and sentiment analyses, 6 core themes with unique valence distributions were identified, providing insight into the experiences, priorities, and needs of individuals living with vulvodynia.

Reflections on Erasure and Being Disbelieved

A published study exploring the experience of women with vulvodynia in the United Kingdom found that health care professionals often dismiss patients' expressions of concern or physicians lack knowledge about the condition [33]. The substantial percentage of posts mentioning not being taken seriously by a health care provider, which was coded as "erasure and disbelief" indicates health care's inadequate support for patients with vulvodynia, which may explain the prevalence of negative attitude posts. Discussions of erasure and being disbelieved were present in many of the opening threads, and many users described needing to increase self-advocacy in medical settings. Such reports highlight the persistent marginalization and sense of being disbelieved during health care interactions, thereby necessitating substantial self-advocacy. Reddit users shared their disappointment with providers' behavior, attitudes, and expertise: one user shared that her doctor bluntly asked if the patient had tried lubricant, revealing a gap in understanding and empathy about vulvodynia's etiology and treatment. In aggregate, the prevalence of posts mentioning erasure and being disbelieved underscores the critical need for improved medical education and patient-centered care, 2 weaknesses of health care professionals at all levels of training [5,6,34-36].

Physical Symptoms and the Impact on Daily Functioning

Symptoms were the most prevalent theme, and many posts emphasized the wide-ranging impact of symptoms on overall health. Symptom-related posts predominantly had a negative attitude, reflecting the disruptive nature of physical discomfort in all facets of daily life. It is essential to acknowledge, however, that participants posting in these threads may not all have a formal diagnosis of vulvodynia. It is impossible to verify the truth of the contents of any of the posts. Despite this limitation, however, there were notable parallels in the dataset between user-reported symptoms and clinical diagnostic criteria for vulvodynia. Currently, there is no exclusive classification for vulvodynia; rather, a diagnosis is characterized by the description of pain [1,2,16].

Pain was the most discussed symptom, underscoring the debilitating and all-consuming nature of vulvar pain [37]. Reddit users described experiences of burning pain, pain with tampon insertion, pain during sexual intercourse, and irritation from clothing. Some symptoms mentioned, such as swollen tissue associated with tampon insertion, pain with urination, and pelvic floor tightness, do not align with established diagnostic criteria, suggesting current diagnostic tools may not capture the full range of experiences of individuals living with vulvodynia [2].

Vulvodynia can interfere with day-to-day functioning; one user noted vulvar pain made it difficult to ride a bike. Others found it challenging to stay active due to pain. Difficulties related to vulvodynia extended beyond physical discomfort; one user described how finding comfortable and wearable underwear became an unexpected source of financial stress. The heterogeneity of pain associated with vulvodynia suggests that further research is needed to better understand its etiology and develop more effective treatment strategies.

In addition to pain, the subreddit posts included expressions of anger, frustration, anxiety, depression, and even trauma, highlighting the connection between mental health and chronic pain. Although few studies have investigated mental health outcomes in individuals with vulvodynia, current evidence suggests that vulvodynia symptoms contribute to worse quality of life and many individuals living with vulvodynia have comorbid anxiety or depression [38-40]. While further research is needed, comprehensive care for vulvodynia should consider both physical and mental health to improve patient well-being.

Treatment Approaches: Navigating Options and Uncertainty

Treatment-related discussions highlight the range of difficulties individuals face in managing chronic health conditions. Participants exchanged information about various treatment modalities, sharing insights into effectiveness, side effects, and accessibility. Of posts discussing treatments, the higher proportion of positive opening threads suggests that individuals in this Reddit community often shared experiences of treatment that were effective. Three of the 5 most mentioned treatments in the Reddit threads—physical therapy, topical medications, and surgery—aligned with guidelines from ACOG, AUA, and ISSVD. To be sure, not all discussions of these treatments were

XSL•FO
RenderX

positive. However, these discussions indicate that users in this web-based forum are aware of and discuss clinically recommended treatments.

Physical therapy and vaginal creams were the top 2 treatment modalities discussed. While physical therapy is widely recognized as an effective approach for vulvodynia, vaginal creams such as baclofen and amitriptyline, though effective, are still considered novel remedies [41-43]. Surgery and discontinuing oral contraceptives were also commonly discussed. Surgery is considered for cases where conservative methods fail [1,44]. Procedures such as vestibulectomy or neuromodulation aim to alleviate pain by removing affected tissue or modifying nerve signals. Although controversial, the AUA and ACOG recommend discontinuing hormonal contraceptive treatments, as these may worsen symptoms. The literature on this topic is divided, however. Some researchers argue that long-term oral contraceptive pills may contribute to vestibulitis, while others provide evidence that refutes this connection [45,46].

Another notable challenge discussed by Reddit users is the wide variation in rates of treatment success, an observation that is well-documented in the literature [2,3,37]. Success rates for medical interventions are reported to range from 13% to 67% [47]. Note that the AUA, ACOG, and ISSVD provide slightly differing guidelines for treating vulvar pain. This may complicate care for providers already navigating serious time constraints in health care. In light of Reddit users' self-reported challenges in obtaining successful treatment for vulvodynia, harmonizing treatment guidelines would likely benefit clinicians and patients alike.

Sexuality and Relationships: Coping With Intimacy Challenges

It is not surprising that sexual experiences also emerged as a prevalent theme in these subreddit threads, given that vulvodynia directly affects individuals' intimate lives and sexual health [48]. One user shared that vulvar pain complicated their interest in sexual intimacy, demonstrating how the connection between experiences of pain, desire, pleasure, and sexual experiences may be altered by vulvodynia. Some users detailed the frustrations and challenges of finding an understanding partner, while others shared success stories of supportive and accommodating partners. Further research is needed to understand how vulvodynia impacts relationships and sexuality. In subsequent studies, qualitative interviewing would be one way to center the voices of individuals with vulvodynia.

Media Representation and Visibility

The low percentage of posts discussing representation and media highlights the invisibility of vulvodynia to the public. The prevalence of positivity in such posts underscores the urgent need for increased awareness, which can be transformative for an individual's sense of self and confidence. One user shared that representation in media made them feel less isolated in their experience. In this way, media may represent an unexpectedly positive domain in which individuals with vulvodynia can find support and recognition of their experience. Health care providers should be aware of the power of representation to positively impact individuals with vulvodynia who may feel overlooked by the medical system. For others, media can be a reminder of the difficulties associated with pain, sexuality, and daily functioning. Overall, representation was associated with significantly increased odds of positive valence, illustrating the value and importance of representation for individuals with vulvodynia.

Humor as a Coping Strategy

Humor is well-recognized as an adaptive tool for coping with stressful situations. For individuals with chronic pain, in particular, humor has been shown to reduce pain intensity and improve quality of life [49]. Explicit humor therapy, in which individuals engage with materials they find entertaining, is associated with decreased pain and feelings of loneliness [50]. In this way, humor represents a nonpharmacological approach for addressing and even ameliorating pain. Humorous interpersonal interactions have also been noted as a way for individuals with chronic pain to engage with one another and even improve clinical outcomes [51]. Members of the vulvodynia community on Reddit creatively reframed their experiences through memes and conversational threads. Users generated memes and made jokes about symptoms and interactions with doctors; in this way, the separation of body and mind may be a method of relief. Humor therefore represents a unique approach for managing experiences of vulvodynia, and it is one means by which members of the Reddit community express themselves and connect with others.

Limitations

A notable limitation of this study is the lack of access to user demographics due to the anonymous nature of Reddit. As a result, we were unable to interpret the possible effects of factors including race, age, health literacy, socioeconomic status, location, transportation, and access to health care which may have impacted the experiences mentioned by each user. Although we cannot determine whether any user had an official diagnosis or indeed met diagnostic criteria for vulvodynia, the Reddit contributors were driven to the platform for specific reasons. Furthermore, as a cross-sectional study, these results are only representative of the time in which data were collected. Results are not generalizable and should be understood as a snapshot of what anonymous Reddit users reported about vulvodynia.

Conclusions

This study aimed to better understand patient experiences of vulvodynia by analyzing web-based discussions on Reddit. Findings highlight that Reddit serves as a vital platform for sharing personal experiences, accessing peer-to-peer support, and seeking health care–related information. These web-based discussions provide valuable anecdotal evidence underscoring a need for health care providers to be trained on the management of vulvodynia, guided by consensus from professional associations. Such training would help ensure patients receive accurate diagnoses and effective care. By prioritizing and centering the patient perspective, health care providers can gain a deeper understanding of the multifaceted challenges faced by individuals living with vulvodynia. This study contributes to

XSL•FO
RenderX

existing literature by offering insights directly from those affected by vulvodynia or who are experiencing vulvodynia-like symptoms.

Data Availability

The datasets analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

LAB is a consultant for Locus Biosciences and reports funding from the National Institutes of Health. MU is the founder of VULVAi. At the time of this study, VULVAi has not received funding or engaged in commercial activities. This affiliation did not influence the design, execution, or interpretation of the research presented in this manuscript.

Multimedia Appendix 1
Example paraphrased threads by theme and valence.
[PNG File, 38 KB - infodemiology_v5i1e63072_app1.png]

References

1. Committee opinion no 673: persistent vulvar pain. Obstet Gynecol 2016 Sep;128(3):e78-e84. [doi: 10.1097/AOG.0000000000001645] [Medline: 27548558]
2. Stockdale CK, Lawson HW. 2013 Vulvodynia Guideline update. J Low Genit Tract Dis 2014 Apr;18(2):93-100. [doi: 10.1097/LGT.0000000000000021] [Medline: 24633161]
3. Harlow BL, Stewart EG. A population-based assessment of chronic unexplained vulvar pain: have we underestimated the prevalence of vulvodynia? J Am Med Womens Assoc (1972) 2003;58(2):82-88. [Medline: 12744420]
4. Pukall CF, Goldstein AT, Bergeron S, et al. Vulvodynia: definition, prevalence, impact, and pathophysiological factors. J Sex Med 2016 Mar;13(3):291-304. [doi: 10.1016/j.jsxm.2015.12.021] [Medline: 26944461]
5. Beebe S, Payne N, Posid T, et al. The lack of sexual health education in medical training leaves students and residents feeling unprepared. J Sex Med 2021 Dec;18(12):1998-2004. [doi: 10.1016/j.jsxm.2021.09.011] [Medline: 34711518]
6. Pancholy AB, Goldenhar L, Fellner AN, Crisp C, Kleeman S, Pauls R. Resident education and training in female sexuality: results of a national survey. J Sex Med 2011 Feb;8(2):361-366. [doi: 10.1111/j.1743-6109.2010.02117.x] [Medline: 21143411]
7. Templeman L, Eberhardt J, Ling J. Exploring the health care experiences of women diagnosed with vulvodynia. J Sex Med 2023 Jan 14;20(1):97-106. [doi: 10.1093/jsxmed/qdac023] [Medline: 36897241]
8. Hesse BW, Nelson DE, Kreps GL, et al. Trust and sources of health information: the impact of the internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
9. Yeung AWK, Tosevska A, Klager E, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. J Med Internet Res 2022 Jan 25;24(1):e28152. [doi: 10.2196/28152] [Medline: 34951864]
10. Doull M, O'Connor AM, Welch V, Tugwell P, Wells GA. Peer support strategies for improving the health and well-being of individuals with chronic diseases. Cochrane Database Syst Rev 2017 Jun 27. [doi: 10.1002/14651858.CD005352.pub2]
11. Su J, Dugas M, Guo X, Gao G. Building social identity-based groups to enhance online peer support for patients with chronic disease: a pilot study using mixed-methods evaluation. Transl Behav Med 2022 May 26;12(5):702-712. [doi: 10.1093/tbm/ibac008] [Medline: 35312778]
12. Wise J. Reddit users: how many people use Reddit in 2023? EarthWeb. 2023. URL: https://earthweb.com/how-many-people-use-reddit/ [accessed 2024-05-23]
13. Chi Y, Chen HY. Investigating substance use via Reddit: systematic scoping review. J Med Internet Res 2023 Oct 25;25:e48905. [doi: 10.2196/48905] [Medline: 37878361]
14. Thulin EJ, Walton MA, Bonar EE, Fernandez A. Examining the popularity, content, and intersections with the substance abuse and mental health services administration's definition of recovery in a nonclinical online cannabis cessation community: infodemiology study of Reddit posts. J Med Internet Res 2024 Sep 27;26:e47357. [doi: 10.2196/47357] [Medline: 39331460]
15. Park A, Conway M. Harnessing Reddit to understand the written-communication challenges experienced by individuals with mental health disorders: analysis of texts from mental health communities. J Med Internet Res 2018 Apr 10;20(4):e121. [doi: 10.2196/jmir.8219] [Medline: 29636316]
16. Sellke N, Jesse E, Callegari M, et al. Is Reddit a reliable source for information on erectile dysfunction treatment? Int J Impot Res 2023 Aug;35(5):484-489. [doi: 10.1038/s41443-022-00586-0] [Medline: 35597799]
17. Jiang T, Osadchiy V, Mills JN, Eleswarapu SV. Is it all in my head? Self-reported psychogenic erectile dysfunction and depression are common among young men seeking advice on social media. Urology 2020 Aug;142:133-140. [doi: 10.1016/j.urology.2020.04.100] [Medline: 32437776]

XSL•FO
RenderX

18. Goel R, Modhukur V, Täär K, Salumets A, Sharma R, Peters M. Users' concerns about endometriosis on social media: sentiment analysis and topic modeling study. J Med Internet Res 2023 Aug 15;25:e45381. [doi: 10.2196/45381] [Medline: 37581905]
19. Valdez D, Mena-Meléndez L, Crawford BL, Jozkowski KN. Analyzing Reddit forums specific to abortion that yield diverse dialogues pertaining to medical information seeking and personal worldviews: data mining and natural language processing comparative study. J Med Internet Res 2024 Feb 14;26:e47408. [doi: 10.2196/47408] [Medline: 38354044]
20. Belcher R, Sim D, Meykler M, et al. Female sexual dysfunction: a Reddit analysis of the lost libido. J Sex Med 2022 Aug 1;19(Supplement_3):S19-S20. [doi: 10.1016/j.jsxm.2022.05.043]
21. Belcher RE, Sim D, Meykler M, et al. A qualitative analysis of female Reddit users' experiences with low libido: how do women perceive their changes in sexual desire? J Sex Med 2023 Feb 27;20(3):287-297. [doi: 10.1093/jsxmed/qdac045] [Medline: 36763942]
22. Belcher R, Sim D, Meykler M, et al. "O" no: a Reddit analysis of orgasmic dysfunction. Sex Med 2023 Dec;11(6):qfad061. [doi: 10.1093/sexmed/qfad061] [Medline: 38053613]
23. LePage K, Selk A. What do patients want? A needs assessment of vulvodynia patients attending a vulvar diseases clinic. Sex Med 2016 Dec;4(4):e242-e248. [doi: 10.1016/j.esxm.2016.06.003] [Medline: 27484917]
24. Webber V, Miller ME, Gustafson DL, Bajzak K. Vulvodynia viewed from a disease prevention framework: insights from patient perspectives. Sex Med 2020 Dec;8(4):757-766. [doi: 10.1016/j.esxm.2020.07.001] [Medline: 32773263]
25. Leusink P, Steinmann R, Makker M, et al. Women's appraisal of the management of vulvodynia by their general practitioner: a qualitative study. Fam Pract 2019 Nov 18;36(6):791-796. [doi: 10.1093/fampra/cmz021] [Medline: 31074493]
26. Vulvodynia. Reddit. URL: https://www.reddit.com/r/vulvodynia/ [accessed 2025-03-05]
27. Finding cures to sexual pain. Reddit. URL: https://www.reddit.com/r/vestibulodynia/ [accessed 2025-03-05]
5. Braun V, Clarke V, Hayfield N, Terry G. Thematic analysis. In: Liamputtong P, editor. Handbook of Research Methods in Health Social Sciences: Springer; 2019:843-860. [doi: 10.1007/978-981-10-5251-4_103]
29. Vears DF, Gillam L. Inductive content analysis: a guide for beginning qualitative researchers. FoHPE 2022 Mar 31;23(1):111-127. [doi: 10.11157/fohpe.v23i1.544]
30. Bingham AJ. From data management to actionable findings: a five-phase process of qualitative data analysis. Int J Qual Methods 2023 Oct;22:16094069231183620. [doi: 10.1177/16094069231183620]
31. Hill CE, Knox S. Essentials of Consensual Qualitative Research: American Psychological Association; 2021. [doi: 10.1037/0000215-000]
32. Hill CE, Thompson BJ, Williams EN. A guide to conducting consensual qualitative research. Couns Psychol 1997 Oct;25(4):517-572. [doi: 10.1177/0011000097254001]
33. Shallcross R, Dickson JM, Nunns D, Taylor K, Kiemle G. Women's experiences of vulvodynia: an interpretative phenomenological analysis of the journey toward diagnosis. Arch Sex Behav 2019 Apr;48(3):961-974. [doi: 10.1007/s10508-018-1246-z] [Medline: 30047005]
34. Kirubarajan A, Li X, Got T, Yau M, Sobel M. Improving medical student comfort and competence in performing gynecological exams: a systematic review. Acad Med 2021 Sep 1;96(9):1353-1365. [doi: 10.1097/ACM.0000000000004128] [Medline: 33883396]
35. Lurie N, Margolis K, McGovern PG, Mink P. Physician self-report of comfort and skill in providing preventive care to patients of the opposite sex. Arch Fam Med 1998;7(2):134-137. [doi: 10.1001/archfami.7.2.134] [Medline: 9519917]
36. Wimberly YH, Hogben M, Moore-Ruffin J, Moore SE, Fry-Johnson Y. Sexual history-taking among primary care physicians. J Natl Med Assoc 2006 Dec;98(12):1924-1929. [Medline: 17225835]
37. Sadownik LA. Etiology, diagnosis, and clinical management of vulvodynia. Int J Womens Health 2014;6:437-449. [doi: 10.2147/IJWH.S37660] [Medline: 24833921]
38. Iglesias-Rios L, Harlow SD, Reed BD. Depression and posttraumatic stress disorder among women with vulvodynia: evidence from the population-based woman to woman health study. J Womens Health (Larchmt) 2015 Jul;24(7):557-562. [doi: 10.1089/jwh.2014.5001]
39. Ponte M, Klemperer E, Sahay A, Chren MM. Effects of vulvodynia on quality of life. J Am Acad Dermatol 2009 Jan;60(1):70-76. [doi: 10.1016/j.jaad.2008.06.032] [Medline: 19022532]
40. Tribó MJ, Andión O, Ros S, et al. Clinical characteristics and psychopathological profile of patients with vulvodynia: an observational and descriptive study. Dermatology 2008;216(1):24-30. [doi: 10.1159/000109354] [Medline: 18032895]
41. Jahshan-Doukhy O, Bornstein J. Long-term efficacy of physical therapy for localized provoked vulvodynia. Int J Womens Health 2021;13:161-168. [doi: 10.2147/IJWH.S297389] [Medline: 33603495]
42. Keppel Hesselink JM, Kopsky DJ, Sajben N. New topical treatment of vulvodynia based on the pathogenetic role of cross talk between nociceptors, immunocompetent cells, and epithelial cells. J Pain Res 2016;9:757-762. [doi: 10.2147/JPR.S115407] [Medline: 27757050]
43. Ruoss CM, Howard EA, Chan K, Stevenson PG, Vancaillie T. Topical treatment of vulvodynia, dyspareunia and pudendal neuralgia: a single clinic audit of amitriptyline and oestriol in organogel. Aust N Z J Obstet Gynaecol 2021 Apr;61(2):270-274. [doi: 10.1111/ajo.13292] [Medline: 33427301]

XSL•FO
RenderX

44. Welliver C, Gross M. In: Rubin R, editor. Female Sexual Dysfunction: Medical and Surgical Management of Pelvic Pain and Dyspareunia: American Urological Association; 2023.
45. Reed BD, Harlow SD, Legocki LJ, et al. Oral contraceptive use and risk of vulvodynia: a population-based longitudinal study. BJOG 2013 Dec;120(13):1678-1684. [doi: 10.1111/1471-0528.12407] [Medline: 23937077]
46. Bouchard C, Brisson J, Fortier M, Morin C, Blanchette C. Use of oral contraceptive pills and vulvar vestibulitis: a case-control study. Am J Epidemiol 2002 Aug 1;156(3):254-261. [doi: 10.1093/aje/kwf037] [Medline: 12142260]
47. Landry T, Bergeron S, Dupuis MJ, Desrochers G. The treatment of provoked vestibulodynia: a critical review. Clin J Pain 2008 Feb;24(2):155-171. [doi: 10.1097/AJP.0b013e31815aac4d] [Medline: 18209522]
48. Corsini-Munt S, Rancourt KM, Dubé JP, Rossi MA, Rosen NO. Vulvodynia: a consideration of clinical and methodological research challenges and recommended solutions. J Pain Res 2017;10:2425-2436. [doi: 10.2147/JPR.S126259] [Medline: 29070953]
49. Kugler L, Kuhbandner C, Gerum S, et al. Evaluation of a humor training for patients with chronic pain: a randomized clinical trial. J Pain Res 2021;14:3121-3133. [doi: 10.2147/JPR.S313868] [Medline: 34675641]
50. Tse MMY, Lo APK, Cheng TLY, Chan EKK, Chan AHY, Chung HSW. Humor therapy: relieving chronic pain and enhancing happiness for older adults. J Aging Res 2010 Jun 28;2010:1-9. [doi: 10.4061/2010/343574] [Medline: 21151506]
51. Finlay KA, Madhani A, Anil K, Peacock SM. Patient-to-patient interactions during the pain management programme: the role of humor and venting in building a socially supportive community. Front Pain Res (Lausanne) 2022;3:875720. [doi: 10.3389/fpain.2022.875720] [Medline: 35571144]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists
AUA: American Urological Association
ISSVD: International Society for the Study of Vulvovaginal Disease
OR: odds ratio

Edited by T Mackey; submitted 09.06.24; peer-reviewed by B Lyall, F Marankan, J Kaswija, U Sinha; revised version received 16.01.25; accepted 25.01.25; published 06.03.25.

Please cite as:
Grutman AJ, Perelmuter S, Perez A, Meurer J, Contractor M, Mathews E, Shearer K, Burnett LA, Uloko M
Understanding Patient Experiences of Vulvodynia Through Reddit: Qualitative Analysis
JMIR Infodemiology 2025;5:e63072
URL: https://infodemiology.jmir.org/2025/1/e63072
doi:10.2196/63072

© Aurora J Grutman, Sara Perelmuter, Abigail Perez, Janine Meurer, Monica Contractor, Eva Mathews, Katie Shearer, Lindsey A Burnett, Maria Uloko. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 6.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

How Patients With Cancer Use the Internet to Search for Health Information: Scenario-Based Think-Aloud Study

Fiorella Huijgens[1,2,3], MSc; Pascale Kwakman[1,2,3], MSc; Marij Hillen[1,2,3], PhD; Julia van Weert[4,5], Prof Dr; Monique Jaspers[5,6], Prof Dr; Ellen Smets[1,2,3], Prof Dr; Annemiek Linn[4,7], PhD

[1]Department of Medical Psychology, Amsterdam University Medical Centers, location Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands

[2]Quality of Care, Amsterdam Public Health, Amsterdam, Netherlands

[3]Cancer Treatment and Quality of Life, Cancer Center Amsterdam, Amsterdam, Netherlands

[4]Amsterdam School of Communication Research/ASCoR, Department of Communication Science, University of Amsterdam, Amsterdam, Netherlands

[5]Digital Health, Amsterdam Public Health, Amsterdam, Netherlands

[6]Department of Medical Informatics, Amsterdam University Medical Centers, location Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands

[7]Health Behaviors and Chronic Diseases, Amsterdam Public Health, Amsterdam, Netherlands

Corresponding Author:
Fiorella Huijgens, MSc
Department of Medical Psychology
Amsterdam University Medical Centers, location Academic Medical Center
University of Amsterdam
Meibergdreef 9
Amsterdam, 1105 AZ
Netherlands
Phone: 31 623715595
Email: f.l.huijgens@amsterdamumc.nl

Abstract

Background: Patients with cancer increasingly use the internet to seek health information. However, thus far, research treats web-based health information seeking (WHIS) behavior in a rather dichotomous manner (ie, approaching or avoiding) and fails to capture the dynamic nature and evolving motivations that patients experience when engaging in WHIS throughout their disease trajectory. Insights can be used to support effective patient-provider communication about WHIS and can lead to better designed web-based health platforms.

Objective: This study explored patterns of motivations and emotions behind the web-based information seeking of patients with cancer at various stages of their disease trajectory, as well as the cognitive and emotional responses evoked by WHIS via a scenario-based, think-aloud approach.

Methods: In total, 15 analog patients were recruited, representing patients with cancer, survivors, and informal caregivers. Imagining themselves in 3 scenarios—prediagnosis phase (5/15, 33%), treatment phase (5/15, 33%), and survivor phase (5/15, 33%)—patients were asked to search for web-based health information while being prompted to verbalize their thoughts. In total, 2 researchers independently coded the sessions, categorizing the codes into broader themes to comprehend analog patients' experiences during WHIS.

Results: Overarching motives for WHIS included reducing uncertainty, seeking reassurance, and gaining empowerment. At the beginning of the disease trajectory, patients mainly showed cognitive needs, whereas this shifted more toward affective needs in the subsequent disease stages. Analog patients' WHIS approaches varied from exploratory to focused or a combination of both. They adapted their search strategy when faced with challenging cognitive or emotional content. WHIS triggered diverse emotions, fluctuating throughout the search. Complex, confrontational, and unexpected information mainly induced negative emotions.

Conclusions: This study provides valuable insights into the motivations of patients with cancer underlying WHIS and the emotions experienced at various stages of the disease trajectory. Understanding patients' search patterns is pivotal in optimizing

XSL•FO
RenderX

web-based health platforms to cater to specific needs. In addition, these findings can guide clinicians in accommodating patients' specific needs and directing patients toward reliable sources of web-based health information.

(JMIR Infodemiology 2025;5:e59625) doi:10.2196/59625

KEYWORDS

web-based health information seeking; think aloud; scenario based; cancer; patient evaluation; information seeking; web-based information; health information; internet; pattern; motivation; cognitive; emotional; response; patient; survivor; caregiver; interview; scenario; women; men

Introduction

Background

Patients with cancer increasingly use web-based platforms to seek information about their diagnosis, treatment, and implications thereof in the short and long term. In the Netherlands, 85% of patients with cancer use the internet [1,2], a rate comparable to that in most Asian countries [3] and other European countries [4,5]. The internet offers a wealth of information that can be readily accessed. It provides practically limitless opportunities for finding health information and support from both lay and expert perspectives, making it a highly popular source of information for many patients.

Within the context of cancer, patients' web-based health information seeking (WHIS) behaviors have been explained through theories of coping behavior. Most often, cancer literature on information-seeking patterns revolves around coping behaviors such as monitoring and blunting. Studies suggest that most patients manage health threats by proactively seeking information, a behavior referred to as monitoring coping style, whereas others choose to avoid information and opt for distraction, known as blunting coping style [6,7]. However, some studies indicate that the WHIS behaviors of patients with cancer could be explained via a broader range of approaches than merely through theories of coping behavior [8-10]. For instance, patients with cancer could also differ in their choices regarding the kind, quantity, and origins of the sought information, as well as the strategies used for information management. These approaches are based on patients' perceptions of self-care, which means that patients vary in their WHIS based on what they need to adequately take care of themselves [10]. In addition, the reasons behind seeking information and emotional support on the web are contingent on how patients use the internet [9].

Another factor that could explain variations in how people use the internet is patients' disease and treatment stage—which may predict different needs concerning the type and amount of information [11,12]. However, studies investigating WHIS and particularly the motives to engage in WHIS often treat the behavior as a one-time event. By treating WHIS as a one-time event, researchers tend to overlook the dynamic nature of health information needs and fail to capture the evolving motivations that patients experience throughout their disease trajectory. Considering that searching for health information is a rather longitudinal behavior, especially for patients moving through different stages of the disease trajectory, a longitudinal lens is required when studying WHIS [11].

In addition to the different phases in the disease trajectory influencing how patients use the internet, WHIS may also vary depending on patients' *motives* for going on the web. For example, patients may do so to address their cognitive (ie, the need for understanding) and affective (ie, the need to be understood) needs [13]. Cognitive needs (eg, engaging with the internet to enhance preparedness and comprehension of the information provided during a consultation or to validate or challenge the information offered by the provider) will lead to diverse forms of WHIS compared to affective needs (eg, using the internet for peer interaction). In other words, patients' specific goals regarding information seeking could also impact their search queries [13]. However, these motives are often not sufficiently taken into account when studying WHIS behavior.

Finally, in the period between diagnosis and cure or remission, patients often experience a range of emotions, including (but not limited to) uncertainty, hope, fear, and anxiety. These feelings and emotions are important motivators for many patients to seek out information to cope with their illness [14]. For example, when just diagnosed with cancer, individuals might be concerned about the unpredictable aspects of the disease, leading them to search for information to better manage and cope with their newly discovered illness. Apart from instigating patients' WHIS behavior, these emotions may also influence decisions to continue, expand, or terminate WHIS [10,14-16]. Earlier qualitative studies have identified various WHIS patterns and the emotions associated with them, ranging from intense to guarded information seeking [10,16,17]. While all participants in these studies expressed a desire for basic information about their diagnosis, they also exhibited diversity in their motivations for seeking cancer information; the emotions experienced; and the nature, quantity, and sources of the sought information, along with the strategies used to manage this information. However, interviews rely on patients' subjective, retrospective reporting and, therefore, do not provide a comprehensive overview of WHIS behavior.

Hence, it is thus far largely unknown how various motives and emotions guide WHIS behavior in various phases of the cancer disease trajectory, whereas such insights can lead to better designed web-based health platforms catering to patients' changing requirements and supporting them effectively throughout their health journey. In addition, having a comprehensive understanding of how patients navigate information acquisition on the internet is crucial for establishing effective patient-provider communication that accommodates patients' specific needs. These insights may also make health care providers aware of the potential impact that WHIS has on patients and, consequently, on the consultation.

XSL•FO
RenderX

Objectives

Studying the impact of motives and emotions on information-seeking behavior during the disease trajectory poses several challenges that have not been taken into account in previous studies. First, as most WHIS occurs in private settings, such as at home, most of these studies use data collection methods that rely on patients' subjective, retrospective reporting, such as surveys, focus groups, and interviews. Using these retrospective methods presents significant drawbacks, including recall bias, which may lead to inaccurate results [18]. In particular, information collected before or during diagnosis is considered challenging as this often entails a short and stressful period for many patients [19]. New research methods such as the think-aloud method enable participants to verbalize what they are thinking and doing while performing a certain task [20]; this allows researchers to observe patients' WHIS more precisely. This includes assessing attention to web-based information, choices made while selecting information, and people's thoughts and feelings evoked during exposure to information [21]. When combining the think-aloud method with vignettes representing different scenarios at various stages of the disease trajectory, research has the potential to provide a more comprehensive and naturalistic view on the WHIS of patients with cancer. Therefore, this study aimed to explore patterns of motivations and emotions behind the web-based information seeking of patients with cancer at different stages of their disease trajectory, as well as the cognitive and emotional responses evoked via a scenario-based, think-aloud approach. This study adopted a unique explorative approach by observing analog patients (ie, patients or healthy participants putting themselves in the position of a patient [22]) as they engaged in WHIS during different phases of their disease trajectory.

Methods

Study Design, Setting, and Population

We used a scenario-based, think-aloud approach followed by a semistructured interview to obtain more in-depth information regarding analog patients' search strategy, their reasoning and emotions behind this strategy (ie, motives), and the emotions experienced throughout. To increase feasibility and for ethical reasons, we decided to rely on analog patients (patients or healthy participants who are asked to imagine themselves in the role of the patients), who are considered valid proxies for clinical patients [23,24]. The COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines were used to report the methods (Multimedia Appendix 1).

Analog patients were recruited from a local panel of patients with cancer, survivors, and their informal caregivers who were willing to participate in scientific research on patient-provider communication and health information provision [25]. In this way, we ensured that the analog patients had some personal experience with cancer. Via email, panel members were informed about the study purpose and invited to complete a screening questionnaire to establish their eligibility, that is, whether they were aged ≥18 years, had previously used the internet to search for health information, and owned a computer or laptop with internet connection. The screening questionnaire also included panel members' age, gender, and educational attainment to allow for purposive sampling based on these characteristics as research shows that individuals differing in these characteristics navigate the web differently and differ in information needs [26]. In addition, we strived for diversity in relation to cancer experience (eg, "I have (had) cancer" or "My partner has (had) cancer"), cancer type, and frequency of using the internet for health information in the previous year (eg, "1-5 times," "6-10 times," "11-30 times," or "more than 30 times").

In total, 75 panel members indicated an interest in participating. Of these 75 members, we invited 34 (45%) individuals based on purposive sampling to take part in the scenario-based, think-aloud study. Eventually, of the 34 individuals, 5 (15%) participated in the pilot study, and 15 (44%) participated in the think-aloud sessions, 5 (33%) for each scenario. Among the 34 individuals, there were 9 (26%) nonresponses, 1 (3%) failed recording, and 4 (12%) who opted out.

Procedure

The scenario-based, think-aloud sessions were conducted between May 2021 and December 2021 by 3 researchers (PK, FH, and an undergraduate student). PK and the student have a health communication background, and FH has a health science and health care management background. PK is trained in qualitative research. Due to the COVID-19 pandemic, the sessions were held on the web using videoconferencing software (ie, Zoom [Zoom Video Communications] or Microsoft Teams [Microsoft Corp]) and were recorded with video. Analog patients could participate in the sessions from the comfort of their home while using their own devices, thereby enhancing ecological validity.

We used a protocol for the scenario-based, think-aloud sessions, including a semistructured interview guide. This protocol was pilot-tested with 15% (5/34) of the analog patients. On the basis of the pilot, we decided to develop a video tutorial explaining the think-aloud procedure and a written manual explaining the use of the videoconferencing software (eg, "How do I share my screen?"). We also adapted the interview guide by adding questions focusing on analog patients' explanations of and reflections on their WHIS behavior (Multimedia Appendix 2). Participating analog patients received an email including an information letter and the video tutorial.

At the start of each session, the researcher explained the nature of the scenario-based, think-aloud method to the analog patients and asked for their personal experience with WHIS. Then, to become familiar with the process of thinking aloud, the analog patients were presented with a practical task (ie, to find a recipe for a pie or a cake containing apples) [27].

After familiarizing the analog patients with the think-aloud procedure, the researcher asked them to imagine themselves in one of the three following scenarios: (1) being an individual who experienced symptoms that could point toward non-Hodgkin lymphoma (NHL), hereinafter referred to as analog prediagnostic patient; (2) being a patient who is about to receive treatment for NHL, hereinafter referred to as analog patient with cancer; or (3) being a survivor of NHL 2 months after having finished treatment, hereinafter referred to as analog survivor of

XSL•FO
RenderX

cancer (Multimedia Appendix 3). We use the general term *analog patients* when referring to 2 or 3 scenarios.

Each scenario was based on real patient experiences that were reported in blogs and discussion groups of the largest cancer-related website in the Netherlands [28] and was reviewed by a survivor of cancer to optimize external validity [29]. Analog patients were assigned to the scenario that was most appropriate given their health status and relationship to cancer.

To enhance identification, analog patients were asked to report in their own words what they had just heard in the scenario. In addition, the researcher asked analog patients to discuss any thoughts or feelings that were evoked by the scenario and score their stress, anxiety, worries about cancer, hope, and uncertainty on an 11-point thermometer-style scale (0=*not at all*; 10=*an extreme amount*). Next, analog patients were asked to go on the web imagining themselves as the described patient in the scenario. While performing the various tasks, analog patients were asked to share their screen. The researcher instructed analog patients to indicate when they wanted to stop their web-based search. If analog patients fell silent during the session, the researcher reminded them to voice their thoughts.

After the think-aloud process, a short semistructured interview was conducted in which the researchers probed for analog patients' motives (eg, what made them choose particular search terms or why they decided to end their search) and their satisfaction with the content (Multimedia Appendix 2). Each interview session ended with a questionnaire assessing the analog patients' coping style (Dutch Threatening Medical Situations Inventory [30,31]), uncertainty intolerance (Dutch version of the short Intolerance of Uncertainty Scale [32]), information needs [33], and eHealth literacy (Dutch eHealth Literacy Scale [34]). These measures were used to be able to describe the sample.

Data Analysis

In total, 2 coders (FH and PK) first familiarized themselves with the data by watching the recordings and reading the interviewer field notes. Second, they independently selected and transcribed parts of each recording that seemed relevant to the research questions (eg, motives and emotions related to WHIS and search strategies). During the analysis, they focused on the analog patients' actions (observations), their verbalized thoughts during the scenario-based, think-aloud process (what they did vs what they said), and their reflections (interview). What was considered relevant was first discussed with a third team member (AL). Third, the coders independently double coded all relevant fragments. Fragments were coded inductively based on the sensitizing concepts as discussed in the introduction (ie, emotions and motives to seek web-based health information, search strategy used, and type of emotions evoked). During the observations, the coders closely examined the search terms used by the analog patients and the content viewed to deduce the analog patients' underlying motives. Fourth, the coders met and discussed their codes after each session to reach an agreement on the coding scheme together with a third team member (AL). Fifth, after completion of the coding process, the codes were aggregated into potential overarching themes and subthemes through comparisons and discussion between the coders. To improve reliability, validity, and generalizability, the results were substantiated using vivid quotes, and a continuous process of reflection and discussion among the coauthors (FH, PK, AL, and ES) was used. To improve the readability of the overall analysis (N=15), we decided to use the term *most* when the analysis applied to >10 analog patients, *several* when it applied to between 5 and 10 analog patients, and *some* when the analysis applied to <5 analog patients. For scenario-specific analysis (5/15, 33%), we decided to use the term *most* when the analysis applied to 3 or 4 analog patients and the term *some* when the analysis applied to 2 analog patients.

Ethical Considerations

The Amsterdam School of Communication Research Ethical Review Board approved this study at the University of Amsterdam (ethics approval code: 2021-PC-13493). Informed consent was verbally obtained from analog patients at the start of the scenario-based, think-aloud session. Analog patients could withdraw their consent at any time. The data could not be anonymized as the think-aloud interviews were video recorded. The data are saved on a secured drive of the Amsterdam University Medical Center. No compensation was provided to the participants.

Results

Sample Characteristics

Among the 15 participating analog patients (n=9, 60% women and n=6, 40% men), the ages ranged from 28 to 72 years (mean 56.9, SD 12.5 years). Most were former patients with cancer and reported having used the internet for seeking health information >6 times in the foregoing year. In total, the sessions lasted between 25 and 70 minutes, and the web-based search lasted between approximately 6 and 26 minutes. The number of web pages visited ranged from 3 to 15 per session, and changes in search terms ranged from 1 to 16 per session. Table 1 shows the sample characteristics, and Tables 2-4 provide descriptions of the individual search sessions.

XSL•FO
RenderX

Table 1. Analog patient characteristics (N=15).

	Prediagnosis stage (n=5)	Treatment stage (n=5)	Survivor stage (n=5)	Total
Age (y), mean (SD; range)	59.6 (8.1; 51-72)	54.6 (14.9; 28-63)	56.4 (15.7; 29-66)	56.9 (12.5; 28-72)
Gender, n (%)				
Woman	3 (60)	3 (60)	3 (60)	9 (60)
Man	2 (40)	2 (40)	2 (40)	6 (40)
Educational level, n (%)[a]				
Low	1 (20)	1 (20)	1 (20)	3 (20)
Middle	0 (0)	1 (20)	2 (40)	3 (20)
High	4 (80)	3 (60)	2 (40)	9 (60)
Relationship to cancer, n (%)				
Having cancer	0 (0)	2 (40)	1 (20)	3 (20)
Having had cancer	2 (40)	3 (60)	4 (80)	9 (60)
Having a relative with cancer	3 (60)	0 (0)	0 (0)	3 (20)
Frequency of web-based health information seeking in the previous year, n (%)				
1-5 times	3 (60)	2 (40)	1 (20)	6 (40)
6-10 times	1 (20)	0 (0)	2 (40)	3 (20)
11-30 times	0 (0)	3 (60)	1 (20)	4 (27)
>30 times	1 (20)	0 (0)	1 (20)	2 (13)
Uncertainty intolerance score, mean (SD; range)	36.2 (7.9; 25-46)	31.8 (9.4; 24-47)	25.6 (7.4; 15-36)	31.2 (8.9; 15-47)
eHEALS[b] score, mean (SD; range)	34.6 (3.8; 31-40)	34.0 (5.3; 27-40)	36.6 (2.1; 34-39)	35.1 (3.8; 27-40)
Monitoring coping style score, mean (SD; range)	11.8 (2.6; 8-15)	13.0 (2.3; 10-15)	8.2 (1.3; 6-9)	11.0 (2.9; 6-15)
Information preference, n (%)				
"I want to know as much as possible, both positive and negative information."	4 (80)	4 (80)	3 (60)	11 (73)
"I want to know as much as possible, both positive and negative information, but in a dosed way (little by little)."	1 (20)	1 (20)	1 (20)	3 (20)
"I want mainly positive information."	0 (0)	0 (0)	1 (20)	1 (7)
"I don't need to know that much."	0 (0)	0 (0)	0 (0)	0 (0)

[a]Low: secondary education; middle: senior general secondary education, secondary vocational education, and preuniversity education; high: college or university.

[b]eHEALS: eHealth Literacy Scale.

XSL•FO
RenderX

Table 2. Characteristics of the participants and search sessions in the prediagnosis phase.

	Participant S01	Participant S05	Participant S06	Participant S08	Participant S10	Values, mean (SD)
Age (y)	61	51	54	60	72	59.6 (8.1)
Gender	Man	Woman	Man	Woman	Woman	—[a]
Educational level[b]	High	High	High	Low	High	—
Search time	15 min 52 s	8 min 57 s	16 min 51 s	6 min 11 s	7 min 55 s	11 min 9 s (4 min 51 s)
Times changing search terms, N	9	4	8	4	1	5.2 (3.3)
Search engine used	Google	Google	Google	Google	Google	—
Total web pages visited, N	5	9	9	3	5	6.2 (2.7)
Uncertainty intolerance score[c]	35	34	41	46	25	36.2 (7.9)
eHealth literacy score[d]	33	32	37	31	40	34.6 (3.8)
Monitoring coping style score[e]	11	8	15	12	13	11.8 (2.6)
Thermometer score[f]						
Feelings of stress and anxiety	7	6.5	8	7	5	6.7 (1.1)
Worries about cancer	7	7.5	6	5.5	5	6.2 (1.0)
Hope	—	—	—	—	—	—
Uncertainty	7	7.5	8	6	10	7.7 (1.5)

[a]Not applicable.

[b]Low: secondary education; middle: senior general secondary education, secondary vocational education, and preuniversity education; high: college or university.

[c]Uncertainty intolerance was measured using the Intolerance of Uncertainty Scale. Total sum scores can range from 12 to 60. A higher score indicates a higher level of intolerance of uncertainty.

[d]Self-perceived eHealth literacy was measured using the eHealth Literacy Scale (eHEALS). Scores on the eHEALS are summed and can range from 8 to 40, with higher scores representing higher self-perceived eHealth literacy.

[e]Monitoring coping style was measured using the Threatening Medical Situations Inventory. Scores are based on the sum of items and can range from 3 to 16. A higher score means a higher monitoring coping style.

[f]Analog patients were asked before the search session to score feelings of stress and anxiety (thermometer 1), worries about cancer (only for the first scenario), hope (only for the second and third scenarios; thermometer 2), and uncertainty (thermometer 3) on an 11-point thermometer-style scale (*0=not at all*; *10=an extreme amount*).

XSL•FO
RenderX

Table 3. Characteristics of the participants and search sessions in the treatment phase.

	Participant S23	Participant S24	Participant S25	Participant S27	Participant S28	Values, mean (SD)
Age (years)	61	62	63	59	28	54.6 (14.9)
Gender	Woman	Man	Woman	Woman	Man	—[a]
Educational level[b]	Low	High	High	Middle	High	—
Search time	9 min 55 s	13 min 40 s	16 min 34 s	24 min 55 s	16 min 35 s	16 min 19 s (5 min 31 s)
Times changing search terms, N	5	5	9	11	9	7.8 (2.7)
Search engines used	Google	Google and Microsoft Bing	Google, Firefox, and Norton Safe Search	Google	Google	—
Total web pages visited, N	5	3	10	11	9	7.6 (3.4)
Uncertainty intolerance score[c]	32	32	24	47	24	31.8 (9.4)
eHealth literacy score[d]	36	40	30	36	27	34 (5.3)
Monitoring coping style score[e]	15	15	10	14	11	13 (2.3)
Thermometer score[f]						
Feelings of stress and anxiety	7	8	7	9	8	7.8 (0.8)
Worries about cancer	—	—	—	—	—	—
Hope	9	3	9.5	4	4.5	6 (3.0)
Uncertainty	8	8.5	2	9	5.5	6.6 (2.9)

[a]Not applicable.

[b]Low: secondary education; middle: senior general secondary education, secondary vocational education, and preuniversity education; high: college or university.

[c]Uncertainty intolerance was measured using the Intolerance of Uncertainty Scale. Total sum scores can range from 12 to 60. A higher score indicates a higher level of intolerance of uncertainty.

[d]Self-perceived eHealth literacy was measured using the eHealth Literacy Scale (eHEALS). Scores on the eHEALS are summed and can range from 8 to 40, with higher scores representing higher self-perceived eHealth literacy.

[e]Monitoring coping style was measured using the Threatening Medical Situations Inventory. Scores are based on the sum of items and can range from 3 to 16. A higher score means a higher monitoring coping style.

[f]Analog patients were asked before the search session to score feelings of stress and anxiety (thermometer 1), worries about cancer (only for the first scenario), hope (only for the second and third scenarios; thermometer 2), and uncertainty (thermometer 3) on an 11-point thermometer-style scale (0=*not at all*; 10=*an extreme amount*).

XSL•FO
RenderX

Table 4. Characteristics of the participants and search sessions in the survivor phase.

	Participant S32	Participant S34	Participant S35	Participant S36	Participant S37	Values, mean (SD)
Age (y)	66	63	66	29	58	56.4 (15.7)
Gender	Woman	Man	Man	Woman	Woman	—[a]
Educational level[b]	Low	High	High	Middle	Middle	—
Search time	15 min 11 s	21 min 40 s	8 min 40 s	25 min 55 s	23 min 35 s	19 min 00 s (7 min 01 s)
Times changing search terms, N	6	16	4	13	8	9.4 (5.0)
Search engines used	Microsoft Bing	Google and Microsoft Bing	Google	Microsoft Bing	Google	—
Total web pages visited, N	8	13	4	15	12	10.4 (4.4)
Uncertainty intolerance score[c]	25	26	15	26	36	25.6 (7.4)
eHealth literacy score[d]	38	39	34	37	35	36.6 (2.1)
Monitoring coping style score[e]	9	6	9	9	8	8.2 (1.3)
Thermometer score[f]						
Feelings of stress and anxiety	5	8	6	3.5	8	6.1 (1.9)
Worries about cancer	—	—	—	—	—	—
Hope	3	6.5	8	10	8	7.1 (2.6)
Uncertainty	7.5	6	0	5	9	5.5 (3.4)

[a]Not applicable.

[b]Low: secondary education; middle: senior general secondary education, secondary vocational education, and preuniversity education; high: college or university.

[c]Uncertainty intolerance was measured using the Intolerance of Uncertainty Scale. Total sum scores can range from 12 to 60. A higher score indicates a higher level of intolerance of uncertainty.

[d]Self-perceived eHealth literacy was measured using the eHealth Literacy Scale (eHEALS). Scores on the eHEALS are summed and can range from 8 to 40, with higher scores representing higher self-perceived eHealth literacy.

[e]Monitoring coping style was measured using the Threatening Medical Situations Inventory. Scores are based on the sum of items and can range from 3 to 16. A higher score means a higher monitoring coping style.

[f]Analog patients were asked before the search session to score feelings of stress and anxiety (thermometer 1), worries about cancer (only for the first scenario), hope (only for the second and third scenarios; thermometer 2), and uncertainty (thermometer 3) on an 11-point thermometer-style scale (0=*not at all*; 10=*an extreme amount*).

Start of the Search Session

Analog patients reported starting their search session with various associations and reactions evoked by the scenario. For example, in the scenario in which they were experiencing symptoms, analog prediagnostic patients were immediately worried about cancer or felt alarmed by specific symptoms. This was reflected in their search terms, showing a predominant focus on searching for information about these symptoms. This was also reflected in their thoughts as patients expressed concern about the symptoms. Whenever the general practitioner in the scenario showed concern, analog patients more often showed signs of feeling distressed:

> *The word tumor immediately pops into my mind. This is serious. These are symptoms I would not trust.* [S01; analog prediagnostic patient]

> *You do not immediately think the best, especially sweating attacks and weight loss are warning signs.* [S05; analog prediagnostic patient]

Most analog patients with cancer assigned to the scenario of undergoing cancer treatment started their search by expressing fear about the upcoming challenges, particularly the apprehension of chemotherapy, and harboring doubts about the effectiveness of the treatment. The aggressive nature of NHL added to their anxiety, with a lack of optimistic information causing visible distress and confusion about the treatment process:

XSL•FO
RenderX

> *I am scared of what's coming and scared of the chemo. And I am not so hopeful because of my doubt whether the treatment will work.* [S27; analog patient with cancer]

Despite these negative emotions, some analog patients with cancer still remained combative or hopeful:

> *Damn, I have cancer again, now I have to have another treatment, but well I am going for it, because I am far from finished living.* [S25; analog patient with cancer]

This fear was also reflected in their search, with all analog patients with cancer being prone to mainly focus on using search words that were used in the scenario (*(aggressive) non-Hodgkin* and *R-CHOP* [rituximab, cyclophosphamide, hydroxydaunorubicin, oncovin, and prednisone regimen]).

Finally, those who were allocated to the survivor case ("analog survivors of cancer") generally voiced uncertainty at the beginning of the search about whether the cancer was definitely gone. They showed concerns about cancer recurrence and recovery and were somewhat skeptical about recovery:

> *Should I really be happy with being cancer-free? What if it comes back? Before this, I had not felt anything. Now, I do not know what I should and should not feel anymore.* [S32; analog survivor of cancer]

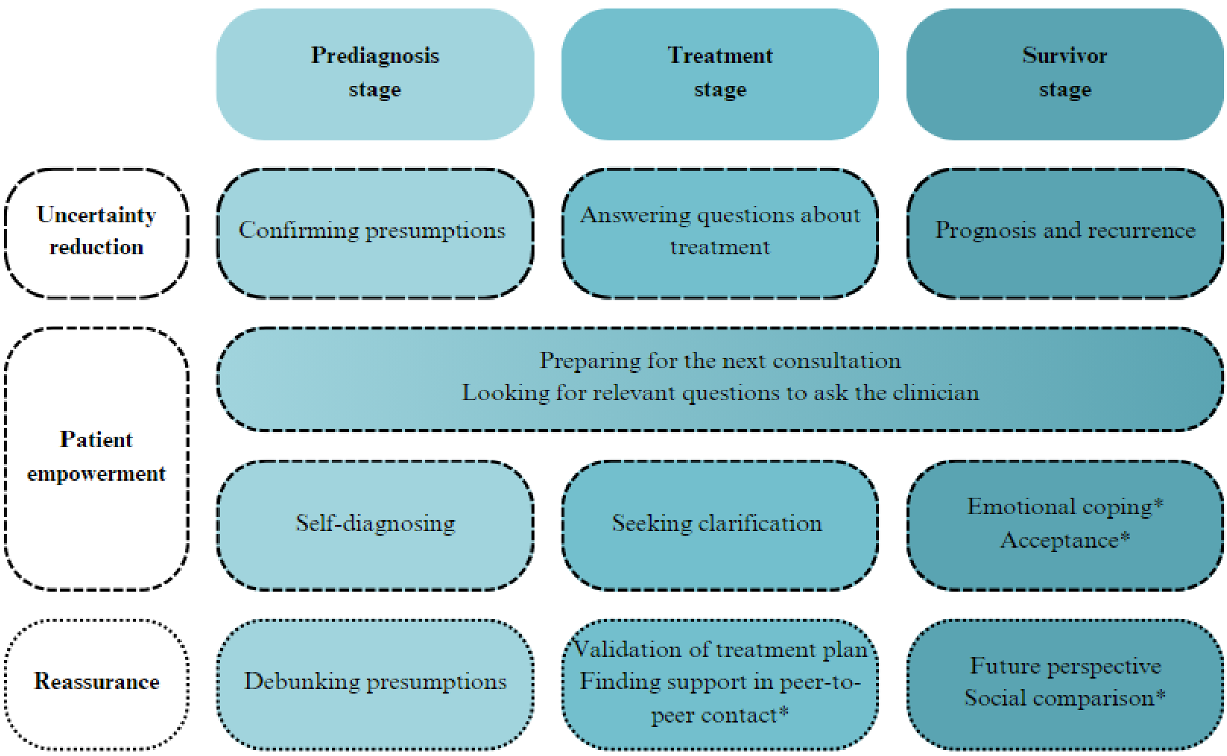
Analog survivors of cancer voiced that, most of all, they wanted to return to their normal lives before the diagnosis and, accordingly, started with search terms related to this desire to get back to the normality of their lives (eg, *out of cancer treatment, what now?*).

Search Motives

Overview

On the basis of the search strategies observed and the thoughts voiced, we were able to distinguish 3 overarching motives guiding patients' WHIS. These overarching motives were prevalent regardless of the allocated stage in the disease trajectory. Each overarching motive was expressed differently throughout the various disease stages (Figure 1). The first motive was *uncertainty reduction* to cope with the anxiety and health threats as most analog patients started their search by expressing uncertainty about what was going to happen to them. The second motive was *empowerment* (ie, "the process of increasing the capacity of individuals (or groups) to make choices and to transform those choices into desired actions and outcomes" [35]) as most analog patients searched content to pursue an active role in their own care process, for example, by actively preparing for the next consultation and looking for relevant questions to ask the clinician. The third motive was *finding reassurance* as analog patients wished to find content that would give them some hope. The 3 overarching motives were not mutually exclusive; they could go hand in hand.

Figure 1. Expressions of the 3 overarching motives for web-based health information seeking—uncertainty reduction, patient empowerment, and reassurance—within the 3 disease stages (prediagnosis stage, treatment stage, and survivor stage). *Affective needs (ie, need to be understood).



Motives in the Prediagnosis Stage

Analog prediagnostic patients wanted to diminish their anxiety and reduce their uncertainty by starting their search with confirming their presumptions. One analog prediagnostic patient immediately used the search term *characteristics cancer*, looked on different websites to compare symptoms, and said the following at search onset:

XSL•FO
RenderX

> *You do not know anything for sure...Apart from the fact that I initially think it is cancer, I still want to confirm that by searching the internet.* [S10; analog prediagnostic patient]

Furthermore, they mostly used the internet to empower themselves by attempting to self-diagnose and prepare for the next consultation. When trying to self-diagnose, they used symptom-related search terms, such as *fatigue*, *swollen glands*, *(unexplained) weight loss*, and *night sweats*. After encountering content about possible diagnoses, some changed their search terms to *symptoms of non-Hodgkin* and *symptoms of cancer* while simultaneously explaining this change:

> *I am actually finding several causes now and cancer is also mentioned. However, I am not quite happy with the information I'm getting yet. But since cancer has come up a few times, I am going to search for symptoms of cancer, so I'm turning it [the search terms] around now [searches for: symptoms of cancer].* [S06; analog prediagnostic patient]

The motive *empowerment* was apparent in one analog prediagnostic patient who used the search terms *preparing consult internist* and read the text *What can you do to prepare for the first visit with an internist?*:

> *What I would do now, because I am going to the hospital, is that I am going to prepare. So, I am now going to search on prepare consult internist. I think I am going to an internist, but obviously I'm not sure yet. [reads text on how to prepare for a visit] I would also like to know, what are useful questions? [clicks on other website] Okay, I have pretty much got everything now I need to consider, only I have to go through the 3 good questions again which I can ask the internist [opens the online brochure about 3 good questions].* [S06; analog prediagnostic patient]

The extent to which analog prediagnostic patients in this phase narrowed down their search to know their exact (possible) diagnosis differed. Some searched various options related to the symptoms, one settled for the likely diagnosis "cancer," whereas others continued their search until they had a specific idea about the type of cancer. Those who searched for various possible diagnoses wanted to be reassured that the symptoms could be anything other than a serious illness such as cancer. They tried to debunk their presumptions, as reflected in the following observation and quote:

> *[reads content about causes of swollen lymph nodes] Infection, which could also be, that makes sense. Then I see here swollen nodes due to a systemic disease. Then I am thinking about Lyme disease, okay. That is different from a tumor. Autoimmune disease is potentially on the table. I already see that swollen nodes can be caused by many factors, which is somewhat reassuring.* [S01; analog prediagnostic patient]

Motives in the Treatment Stage

Analog patients with cancer mostly appeared to use the internet to answer their remaining questions to reduce uncertainty. Reducing uncertainty seemed to be combined with increasing their feeling of empowerment as they appeared to seek for more clarification about diagnosis and treatment. Both uncertainty reduction and empowerment were reflected in search terms such as *What is non-Hodgkin lymphoma?*, *R-CHOP*, *side effects*, and *immunotherapy* (ie, cognitive needs). While searching these terms, they said the following:

> *More than 50% of patients with an aggressive non-Hodgkin lymphoma in an advanced stage are cured after treatment with R-CHOP. Okay, that is quite a lot. But, hmm, yes, the other half does not. It is not clear to me whether the half that does not recover remains chronically ill or simply succumbs to death. I would like to know that in that sense.* [S28; analog patient with cancer]

The motive to obtain reassurance via web-based information was reflected in analog patients with cancer using the internet to validate whether the treatment (as being proposed in the scenario) was indeed the right treatment for them. They specifically searched for websites and information that would convince them of this treatment being the best option:

> *And I would definitely, before starting that treatment, request a second opinion from another institution to ensure that I...um...yes, receive the correct diagnosis or the right treatment [searches for other hospitals].* [S24; analog patient with cancer]

One analog patient with cancer also seemed to use the internet to obtain reassurance via socioemotional content. This was reflected in the search term *experiences with R CHOP*. Of note, none of the analog patients with cancer used search terms indicating a need to know more about the prognosis of NHL.

Motives in the Survivor Stage

Analog survivors of cancer seemed to use the internet to reduce uncertainty only to a limited extent. When they used the internet for that purpose, they wanted to know more about prognosis and recurrence, as reflected in search terms such as *prognosis*, *late effects*, and *what to expect*. While using these search terms, they said the following:

> *Yes, you are quite uncertain about how everything will unfold. There are still quite a few questions, and that diminishes over time, but especially in the beginning after that hospital period, you still have quite a lot of questions.* [S37; analog survivor of cancer]

Analog survivors of cancer mainly used the internet to search for socioemotional content related to pursuing an active role in their own recovery (ie, patient empowerment). This was reflected in search terms regarding feelings, experiences, and emotions (eg, *uncertainty after cancer* and *feelings after non-Hodgkin treatment*). Pursuing an active role in their own recovery mainly encompassed (emotional) coping and finding acceptance (eg, returning to their normal life before diagnosis). Apparently, to satisfy these motives, they often visited blogs of survivors of cancer writing about feelings and experiences and providing advice on coping with survivorship (eg, *how to deal with emotions/fatigue/work/daily life*). Some searched for

XSL•FO
RenderX

psychologists or for recovery programs offered by patient organizations or hospitals, which could also be seen as an expression of empowerment:

> *Not because I do not trust my own hospital, but I just want to look further. What do other hospitals offer their patients? Is there anything I can take advantage of?* [S32; analog survivor of cancer]

To a lesser extent, analog survivors of cancer went on the web to seek reassurance about their future. They seemed to be reassured when encountering people with similar experiences. For example, one survivor stated the following:

> *Okay, I found something here, there are more people like me. Shared sorrow is half sorrow.* [S34; analog survivor of cancer]

Overall WHIS Patterns

The web-based source that analog patients eventually selected seemed to depend on their cancer-specific knowledge, cancer-related experience, and search experience. The use of cancer-specific knowledge and experience was reflected in selecting familiar and well-known websites about cancer. The use of search experience was reflected in analog patients using strategies that they reported to prefer (eg, preferring to use the search bar on specific websites instead of the regular search engine or the other way around). Analog patients mentioned different reasons for selecting content. The most prevalent reasons were familiarity with a website or organization (eg, the Dutch Cancer Society) or previous experience with a website. Some also mentioned that they selected certain websites as part of habitual behavior rather than for specific reasons. Notably, analog patients also visited websites while voicing doubt about their trustworthiness. It seemed that those analog patients thought that it was more important to find information relevant to fulfill their motives than looking for trustworthy information.

WHIS Approaches

In total, 2 overarching WHIS approaches could be identified: explorative and focused. Explorative approaches consisted of spontaneously selecting information seemingly without having an explicit information need. Analog patients who used this explorative approach mainly guided their searches by clicking on referral links and using suggestions made by search features on Google, such as the *autocomplete* (a feature within Google Search that makes it faster to complete searches that users start to type. Google’s automated systems generate predictions that help users save time by allowing them to quickly complete the search they already intended to do) and *people also ask* (a feature within Google Search that provides users with additional questions related to their original search query and quick answers to them) functions. Analog patients were considered to use a focused approach when they seemed to search more purposefully (ie, mainly selecting information aligned with their verbally expressed specific information needs). For instance, an analog prediagnostic patient searched *symptoms of cancer* and exclusively selected content related to these search terms.

Unlike analog patients using an explorative approach, patients using a focused approach only made use of Google features when these explicitly helped them meet their self-reported information needs. For example, an analog patient with cancer searched for and read information about R-CHOP and subsequently encountered the following suggestions from the Google feature *people also ask*: *What does R-CHOP mean?* and *What is a CHOP cure?*

Several analog patients used both explorative and focused approaches. Some started with a clearly focused search strategy based on an information need but appeared to become emotionally distracted by the encountered content and started to use a more explorative approach. Others started with an explorative approach and were triggered by specific content that led them to adopt a new, more focused approach (eg, understanding difficult, complex words or confirming assumptions). In other words, information needs evolved while searching. WHIS approaches seemed independent of the disease stage that analog patients were allocated to.

Dissatisfying Content

All analog patients came across dissatisfying content while searching (in other words, content that did not satisfy the wishes of the patients). Examples of dissatisfying content were difficulty navigating systems on websites, cookies, or information not being in line with search motives. When this dissatisfying content was encountered, analog patients most often changed their search terms or quickly moved on to other web pages (the number of web pages visited ranged from 3 to 15 per session). Search terms were frequently changed during a search session (range 1-16 times per session), mostly because of dissatisfying content:

> *So, I’m not getting anywhere with this either, because I don’t need to know what the cancer looks like...So I guess I’m not getting anywhere with this search term, with the search things. Uhm how am I going to do that?* [S35; analog survivor of cancer]

Impact of WHIS on Emotions and Dealing With Content

Emotions

Regardless of the stage of the disease, emotions were present throughout the entire search process, ranging from anxiety and worry to hope. These emotions fluctuated, and negative emotions were often induced when confrontational, complex, or unwanted information was found. Confrontational content included information on symptoms suggesting cancer or thyroid problems, information on treatment side effects such as hair loss and nausea, or a confronting picture:

> *I am not happy with the image I see here. That photo confirms the nightmare I have about chemotherapy. This is someone surrounded by nurses, being injected, and she has no hair, so that picture embodies for me everything that is wrong with this disease in one image. They have succeeded very skillfully in capturing all of that in one photo, but I do not think that was the intention of the person who took the photo. However, that is how it comes across at me: the embodiment of a mountain of misery.* [S27; analog patient with cancer]

XSL•FO
RenderX

Complex information included content containing medical jargon, such as *malignancies*; *cachexia*; or drug names such as *rituximab*, *cyclophosphamide*, and *hydroxydaunorubicin*. Most analog patients seemed to be affected by complex words:

> *This is getting annoying because I already see a word here that I do not know at all. I'm getting a lot of medical terms here that do not mean much to me...* [S06; analog prediagnostic patient]

Sometimes, positive emotions emerged from information that gave hope (eg, indolent NHL more often has a chance of recurrence than aggressive NHL). Moreover, analog patients who doubted their own navigation skills while searching on the web reported high levels of distress. Some of the analog patients also experienced cognitive dissonance (ie, a mental state of having conflicting beliefs, thoughts, values, or attitudes), as reflected in the following quote:

> *Everything in you says that it is better not to click on it, because you don't want to know it. But if you see the option then you just need to click on it.* [S27; analog patient with cancer]

Dealing With Emotionally Difficult Content

When encountering cognitively or emotionally difficult (or unwanted) information, analog patients with cancer dealt with the content in various ways. They adapted their search strategy, ignored the information by quickly clicking away from it and shifting toward other information, or stopped searching:

> *I immediately find myself with types of cancer, um...all the hits are related to Hodgkin; [scrolling back and forth through search results on the first Google page, but not clicking on anything]. Yes, I find this difficult; I think I will check the next Google pages to see what else comes up, what comes after Hodgkin.* [S01; analog prediagnostic patient]

Several analog patients also mentioned that they would normally seek information multiple times briefly or seek a distraction from the confronting information, such as watching Netflix or having some tea.

End of the Search

As mentioned previously, one of the reasons to stop searching was encountering cognitively or emotionally difficult information (confronting, upsetting, or confusing). This was mostly the case for analog prediagnostic patients and analog patients with cancer. The following quote illustrates this "overload":

> *Nothing [information found] makes me happy. Yeah, you can find information, but I believe I would make a cup of coffee now. I cannot say I'm a lot wiser now.* [S25; analog patient with cancer]

Another reason to stop searching was that analog patients saw their health care provider as a gatekeeper and their primary source of information about their disease and treatment. During the interview, they indicated that they preferred to talk with their clinician to clarify the encountered information instead of looking for more web-based information:

> *I believe that this information is quite overwhelming me right now, so I would put it away for a while. And I would talk it through first at a subsequent consultation with my doctor before I start worrying and assuming things that are not an issue at all...So, I think I will stop looking for now until I have spoken to the doctor again. It is a lot of information, and it is also complicated. So, I want to consult the doctor first.* [S23; analog patient with cancer]

All analog patients with cancer indicated ending their search sessions with many unanswered questions and an increase in uncertainty (compared to the start of the search). Unlike analog patients with cancer, analog prediagnostic patients and analog survivors of cancer ended their search more often with their information needs being fulfilled, as reflected in the following quote during the interview:

> *I do think it is very true. I'm at a point now where I do think: yeah, I'm reading this now, I'm not really getting very comfortable with this. I do not think there is any point in continuing to search now. I think I am now on a trustworthy site, and I find this a very upsetting story now that I see this. I cannot do much but wait and see. I don't know if I'm happy I've figured this out now...* [S01; analog prediagnostic patient]

Compared to analog patients in other disease stages, analog survivors of cancer ended their search most often satisfied and with more positive emotions; they felt less uncertain and found useful (practical) information on ways to cope with the future:

> *I definitely did become a bit wiser, because I can move on: I can go to physio, psychologist and I have a phone line which I can call.* [S36; analog survivor of cancer]

Discussion

Principal Findings

Using a comprehensive scenario-based, think-aloud approach, we were able to show that (1) patients' overarching motives for WHIS were mainly to reduce uncertainty, obtain reassurance, and increase empowerment; (2) these motives differed depending on the disease stage (at the beginning of the disease trajectory, patients mainly showed cognitive needs, whereas this shifted more toward affective needs in the subsequent disease stages); (3) analog patients' WHIS approaches varied from exploratory to focused to a combination of both; and (4) positive (hope and reassurance) and negative (anxiety and worry) emotional responses played an important role in patients' search strategies.

We found 3 overarching motives (ie, reducing uncertainty, obtaining reassurance, and increasing empowerment) for patients to search on the web. With these findings, we not only confirm the problem-solving model in the context of patient motivations to go on the web throughout their illness journey but also extend this model. According to Wilson [36], the process of problem-solving is the result of patients' wishes to reduce uncertainty. Patients' uncertainty at either the prediagnosis or

XSL•FO
RenderX

treatment phase concerned various topics, clearly showing that these motives change over time. However, we also discovered 2 other important motives for patients to engage in a problem-solving process, namely, reassurance and empowerment [36]. In addition, the study's findings revealed a potential conflict between patient empowerment and uncertainty reduction in the context of WHIS. When patients seek web-based information to empower themselves, they gain a better understanding of their situation, which could enable them to ask informed questions to their clinicians. However, this increased knowledge may also give rise to new questions and uncertainties, leading to a potential challenge in fulfilling the motive of uncertainty reduction.

Moreover, our findings provide insights into the search behavior of patients with cancer at various stages of their disease trajectory and how these behaviors vary. In the initial phase of prediagnosis, patients often engaged in self-diagnosis. The results of this study extend those of previous research [9] by showing that patients prepare for a consultation by using the internet not only to help them formulate questions but also to self-diagnose. Despite the popularity of this search approach, research on self-diagnosing remains limited. In the context of web-based self-diagnosis for minor ailments, research shows that using the internet for self-diagnosis can be helpful as 44% of participants achieved accurate final diagnoses after searching the internet compared to 11% before searching the internet [37]. Another study shows that web-based self-diagnosing has the potential to empower patients in appraising and challenging clinicians' advice and requesting further diagnostic procedures [38]. However, web-based self-diagnosis can also be counterproductive if the patient misdiagnoses themselves, leading to unnecessary concerns. In addition, problems may occur if patients visit their clinician with a preconceived diagnosis, potentially causing disagreements about their condition [39]. During the treatment phase, the search strategy of patients with cancer focused on cognitive needs by seeking clarification, gathering more information, and preparing. However, we only observed a shift in search strategies toward affective needs by seeking emotional coping resources for dealing with the disease after patients completed treatments and were declared cancer free. In other words, at the beginning of the disease trajectory, analog patients had mainly cognitive needs, whereas analog survivors also showed affective needs and used the internet for emotional support. The change from more cognitive needs to more affective needs could be explained using the social-cognitive processing model. According to this model, seeking emotional support may facilitate emotional adjustment to traumatic experiences, such as cancer diagnosis and treatment [40]. Potentially, survivors have more mental space to cope with the situation and reflect on what has happened in the past months.

Our results further show that patients tend to use different search strategies: explorative, focused, or a combination of both. Previous research has demonstrated that individuals who are more exploratory seekers tend to tackle unfamiliar problems by using a broader search strategy (symptom exploration), resulting in a broader range of new information [37]. By encountering a broad range of information, patients are possibly confronted with new and unknown content, which could increase their level of uncertainty [41]. Our results also suggest that an exploratory search strategy increased the risk of being confronted with unwanted information. On the other hand, those who are more focused seekers tend to have a clear idea and a specific plan, leading them to research within a limited set of results (hypothesis testing) [37]. Such hypothesis testing can be problematic because it skews the way in which patients process information and distorts their perception of reality—a phenomenon known as confirmation bias [42]. It occurs when patients seek, interpret, or favor information that confirms their existing beliefs while ignoring or downplaying evidence that contradicts those beliefs [43]. Pang et al [41] argue that seekers within one internet visit alternate between exploratory and focused search strategies as new, unknown topics often lead to more exploratory searches. If the topic to be searched becomes clearer, the seeker may use a more focused approach. Our results confirm those of this previous study by showing that patients used both explorative and focused approaches. Some started with a focused search but became emotionally distracted and switched to an explorative approach. Others began exploratively and shifted to a focused search after encountering specific content.

Furthermore, our findings show that positive (hope and reassurance) and negative (anxiety and worry) emotional responses were present before, during, and after the search sessions. On the basis of patients' voiced thoughts and observed behavior, we conclude that these emotions impacted their search behavior. This is in line with the functionalist perspective of emotions, which argues that emotional responses may motivate people to behave in particular ways [44-46]. For instance, hope is seen as a motivating force that helps individuals move toward desired outcomes even in the face of uncertainty [47]. It is a future-oriented emotion as it involves visualization of positive future situations [48], and thus, hope could explain why patients are motivated to seek reassurance. Worry, on the other hand, is seen as an uncertainty-associated emotion and can increase a patient's desire for obtaining additional information [15]. Studies show a positive relationship between worry and the perceived need for additional information [49-51], and thus, worry could explain why patients are motivated to reduce uncertainty by searching for additional information. However, we also observed that patients who were worried ignored or avoided specific information. A possible explanation is that hope and worry are intertwined during WHIS [16]. Confronting or complex information poses a threat to hope, and thus, ignoring certain information may serve as a self-protective behavior to stay hopeful [16].

In our study, patients in the treatment phase were most worried after their search session. This is in contrast to existing literature indicating that perceived knowledge through web-based information seeking decreased patients' worry [15]. WHIS has also been found to help searchers fill information voids and enhance their coping abilities [52]. Although we did find some comparable results for the prediagnosis and survivor phases regarding decrease in worry and enhancing coping abilities, we did not find this for patients in the treatment phase. A possible explanation is that complex or confrontational information (eg,

XSL•FO
RenderX

jargon for medicines and treatments and intense side effects) may have induced worries in analog patients in this phase. This inconsistency with the existing literature could further be explained by our design, which involved one search session only at one specific moment rather than multiple search sessions by one individual patient. Possibly, patients who search for more information at multiple times will eventually be less worried as they become more familiar with the difficult and complex information. Therefore, future research should investigate the longitudinal search behaviors of individual patients during their disease trajectory and the effects of multiple shorter search sessions within a particular disease phase.

Limitations and Strengths

First, a strength of our approach is that we not only observed patients' WHIS behaviors but simultaneously gained insights into their thoughts. During the interview, the interviewer made use of techniques such as paraphrasing and checking to clarify the meaning of the interviewee, thereby enhancing the validity of our findings. This innovative, comprehensive scenario-based, think-aloud approach exhibits strength in its consideration of the intuitive nature of web-based searching while overcoming challenges such as recall bias in retrospective methods. However, certain limitations should be considered. Some remarks suggested that participants may have felt limited in their choice of search engine and might have perceived an obligation to use a specific search platform, such as Google. Furthermore, during the think-aloud sessions, participants did not explore the use of social media channels (eg, Facebook, Instagram, or Twitter [subsequently rebranded X]). Use of social media may have been limited as participants could perceive it as an intrusion into their personal lives. Another reason could be that these communication channels may represent more spontaneous ways through which patients acquire unplanned or unexpected web-based health information while scrolling through their social media timeline [53]. The scenario-based, think-aloud approach as used in this study does not provide any insights in how social media has an effect on patients' WHIS strategies, motives, and emotions. Furthermore, the relatively small sample size used in this study calls for caution when generalizing the findings. It is important to account for variations in patients' (eHealth) literacy, education, and cultural backgrounds [54]. Although previous research demonstrates overlap in WHIS among patients from different countries, it also identifies distinct country-specific differences even when the countries have comparable welfare and health status [5]. As this study was an explorative qualitative study, and despite our relatively small sample size, we believe we achieved thematic saturation during the iterative process as no new codes emerged toward the end of our analysis. Moreover, it is important to bear in mind when interpreting the findings that our sample consisted of analog patients who were presented with a scenario. This may have biased our results as using analog patients is different from using patients with NHL. However, participants in this study possessed preexisting familiarity with cancer; our sample consisted of patients with cancer (other than NHL), survivors of cancer, and informal caregivers of patients with cancer. Thus, this sample's strength lies in their ability to strongly identify with the scenarios presented, which is also reflected in their quotes, the emotions showed during the think-aloud process, and their scores on the thermometers [24]. Furthermore, participants possessed experience in web-based cancer information seeking. Many of them were acquainted with patient advocacy organizations, and a subset even served as administrators for certain web-based platforms dedicated to cancer information and peer support groups. In addition, they had previously encountered medical terminology in the context of their own medical conditions, thus acquiring a degree of familiarity with medical jargon. Consequently, our sample likely possessed a higher level of proficiency in navigating the internet for cancer-related information compared to the average patient with cancer. Despite their advanced familiarity with the subject, the results still indicated that patients encountered difficulties in navigating the internet and understanding medical jargon.

Practical Implications

Knowing how patients with cancer search for web-based health information is a first step toward optimizing web-based health platforms such that patients with cancer can (more) easily find and navigate through information that fits their needs. On the basis of the study results, there are various implications for the development of cancer websites. First, web-based health platforms could use less complex words and show content warnings about confrontational prognostic or side effect–related information on web pages. The latter could warn searchers about unwanted information, which is especially relevant for exploratory searchers. Second, websites should enable users to self-pace and allow for user-initiated tailoring (ie, allowing users to tailor the information themselves based on their information needs). For example, information should be minimalized, with the possibility to read more if wanted (eg, with the use of hyperlinks). Third, it should also be clear to the user whether platforms are expert generated or peer generated as these platforms differ in content focusing on cognitive needs (addressing the needs of analog prediagnostic patients and analog patients with cancer) and affective needs (addressing the needs of analog survivors of cancer) [13]. In the Netherlands, multiple cancer platforms already make use of such features, which patients in our sample experienced as convenient. In addition to these implications for websites, another important finding is that patients see their health care providers as their primary source of information when it comes to their disease and treatment. Patients indicated that they had various remaining questions and considerable uncertainty after their search, which they wanted to resolve during their interaction with their health care provider. Therefore, it is important that, within consultations, there is room for questions arising from WHIS. Furthermore, health care providers can guide patients in the search process by giving tips and tricks on how (not) to use the internet to search for health information and how to cope with any uncertainty that may result from such a search.

Conclusions

This study provides valuable insights into the real-time WHIS strategies of patients with cancer, the motivations behind seeking web-based health information, and the emotions experienced at various stages of the disease trajectory. Understanding patients' search patterns is pivotal in optimizing web-based

XSL•FO
RenderX

health platforms to cater to their specific needs. In addition, these findings can guide clinicians in directing patients toward reliable sources of web-based health information.

Acknowledgments

The authors would like to thank all participants for allowing them to watch how they searched the internet for health information while sharing their thoughts with them. Furthermore, the authors would like to thank Diana Aksan for her work in the selection process and pilot-testing of the protocol. This work was supported by the Dutch Cancer Society. No generative artificial intelligence was used in any portion of manuscript writing.

Data Availability

The datasets generated during and analyzed during this study are not publicly available due to the anonymity of the participants but are available from the corresponding author on reasonable request.

Authors' Contributions

The conceptualization of the study was carried out by ES, MH, JvW, and AL. ES acquired funding for this project. FH and PK gathered and analyzed the data, and AL was responsible for the validation process. The original draft of the manuscript was prepared by FH and AL. All authors reviewed the protocols, contributed to conceptualization and methodology, and reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist.
[DOCX File , 18 KB - infodemiology_v5i1e59625_app1.docx]

Multimedia Appendix 2
Final think-aloud protocol, including semistructured interview guide.
[DOCX File , 15 KB - infodemiology_v5i1e59625_app2.docx]

Multimedia Appendix 3
Think-aloud scenarios.
[DOCX File , 15 KB - infodemiology_v5i1e59625_app3.docx]

References

1. van Eenbergen MC, Vromans RD, Boll D, Kil PJ, Vos CM, Krahmer EJ, et al. Changes in internet use and wishes of cancer survivors: a comparison between 2005 and 2017. Cancer 2020 Jan 15;126(2):408-415 [FREE Full text] [doi: 10.1002/cncr.32524] [Medline: 31580497]
2. van Eenbergen MC, Vromans RD, Tick LW, Vreugdenhil G, Krahmer EJ, Mols F, et al. Comparing survivors of cancer in population-based samples with those in online cancer communities: cross-sectional questionnaire study. JMIR Cancer 2022 Mar 08;8(1):e19379 [FREE Full text] [doi: 10.2196/19379] [Medline: 35258460]
3. He Y, Wang Y, Luan F, Yu Z, Feng H, Chen B, et al. Chinese and global burdens of gastric cancer from 1990 to 2019. Cancer Med 2021 May;10(10):3461-3473 [FREE Full text] [doi: 10.1002/cam4.3892] [Medline: 33931958]
4. Capocaccia R, Colonna M, Corazziari I, de Angelis R, Francisci S, Micheli A, et al. Measuring cancer prevalence in Europe: the EUROPREVAL project. Ann Oncol 2002 Jun;13(6):831-839 [FREE Full text] [doi: 10.1093/annonc/mdf152] [Medline: 12123329]
5. Link E, Baumann E, Linn A, Fahr A, Schulz PJ, Abuzahra ME. Influencing factors of online health information seeking in selected European countries. Eur J Health Commun 2021 May 19;2(1):29-55. [doi: 10.47368/ejhc.2021.002]
6. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. Cham, Switzerland: Springer; 1984.
7. Miller SM. When is a little information a dangerous thing? Coping with stressful events by monitoring versus blunting. In: Levine S, Ursin H, editors. Coping and Health. Cham, Switzerland: Springer; 1980:145-169.
8. Bol N, Linn AJ, Smets EM, Verdam MG, van Weert JC. Tailored communication for older patients with cancer: using cluster analysis to identify patient profiles based on information needs. J Geriatr Oncol 2020 Jul;11(6):944-950 [FREE Full text] [doi: 10.1016/j.jgo.2020.01.004] [Medline: 31974067]

XSL•FO
RenderX

9. Duimel SL, Linn AJ, Smets EM, Smit ES, van Weert JC. Profiling cancer patients based on their motives for seeking informational and emotional support online. Health Commun 2023 Dec;38(14):3223-3237. [doi: 10.1080/10410236.2022.2144287] [Medline: 36415021]
10. Lambert SD, Loiselle CG, Macdonald ME. An in-depth exploration of information-seeking behavior among individuals with cancer: part 1: understanding differential patterns of active information seeking. Cancer Nurs 2009;32(1):11-23. [doi: 10.1097/01.NCC.0000343372.24517.bd] [Medline: 19104197]
11. Bautista JR, Zhang Y, Gwizdka J, Chang YS. Consumers' longitudinal health information needs and seeking: a scoping review. Health Promot Int 2023 Aug 01;38(4):daad066. [doi: 10.1093/heapro/daad066] [Medline: 37432774]
12. Rutten LJ, Arora NK, Bakos AD, Aziz N, Rowland J. Information needs and sources of information among cancer patients: a systematic review of research (1980-2003). Patient Educ Couns 2005 Jun;57(3):250-261. [doi: 10.1016/j.pec.2004.06.006] [Medline: 15893206]
13. Sanders R, Linn AJ, Araujo TB, Vliegenthart R, van Eenbergen MC, van Weert JC. Different platforms for different patients' needs: automatic content analysis of different online health information platforms. Int J Hum Comput Stud 2020 May;137:102386. [doi: 10.1016/j.ijhcs.2019.102386]
14. Savolainen R. Information need as trigger and driver of information seeking: a conceptual analysis. ASLIB J Inf Manag 2017 Jan 16;69(1):2-21 [FREE Full text] [doi: 10.1108/AJIM-08-2016-0139]
15. Lee SY, Hawkins RP. Worry as an uncertainty-associated emotion: exploring the role of worry in health information seeking. Health Commun 2016 Aug;31(8):926-933. [doi: 10.1080/10410236.2015.1018701] [Medline: 26752071]
16. Leydon GM, Boulton M, Moynihan C, Jones A, Mossman J, Boudioni M, et al. Cancer patients' information needs and information seeking behaviour: in depth interview study. BMJ 2000 Apr 01;320(7239):909-913 [FREE Full text] [doi: 10.1136/bmj.320.7239.909] [Medline: 10742000]
17. Ofran Y, Paltiel O, Pelleg D, Rowe JM, Yom-Tov E. Patterns of information-seeking for cancer on the internet: an analysis of real world data. PLoS One 2012;7(9):e45921 [FREE Full text] [doi: 10.1371/journal.pone.0045921] [Medline: 23029317]
18. Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. Expert Rev Pharmacoecon Outcomes Res 2004 Apr;4(2):159-163. [doi: 10.1586/14737167.4.2.159] [Medline: 19807519]
19. Stehl ML, Kazak AE, Alderfer MA, Rodriguez A, Hwang WT, Pai AL, et al. Conducting a randomized clinical trial of an psychological intervention for parents/caregivers of children with cancer shortly after diagnosis. J Pediatr Psychol 2009 Sep;34(8):803-816 [FREE Full text] [doi: 10.1093/jpepsy/jsn130] [Medline: 19091806]
20. Bolle S, Romijn G, Smets EM, Loos EF, Kunneman M, van Weert JC. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Jul 25;18(7):e208 [FREE Full text] [doi: 10.2196/jmir.5618] [Medline: 27457709]
21. Marcu A, Muller C, Ream E, Whitaker KL. Online information-seeking about potential breast cancer symptoms: capturing online behavior with an internet browsing tracking tool. J Med Internet Res 2019 Feb 06;21(2):e12400 [FREE Full text] [doi: 10.2196/12400] [Medline: 30724741]
22. Blanch-Hartigan D. Patient satisfaction with physician errors in detecting and identifying patient emotion cues. Patient Educ Couns 2013 Oct;93(1):56-62. [doi: 10.1016/j.pec.2013.04.010] [Medline: 23669151]
23. van Vliet LM, van der Wall E, Albada A, Spreeuwenberg PM, Verheul W, Bensing JM. The validity of using analogue patients in practitioner-patient communication research: systematic review and meta-analysis. J Gen Intern Med 2012 Nov;27(11):1528-1543 [FREE Full text] [doi: 10.1007/s11606-012-2111-8] [Medline: 22700392]
24. Smets EM, van Vliet LM, Pieterse AH, Visser LN, Labrie NH, Hillen MA. Analogue patients. In: Ho EY, Bylund CL, van Weertj JC, Basnyat I, Bol N, Dean M, editors. The International Encyclopedia of Health Communication. Hoboken, NJ: John Wiley & Sons; 2023:1-10.
25. PanelCom. URL: https://www.panelcom.nl/ [accessed 2021-07-30]
26. Agree EM, King AC, Castro CM, Wiley A, Borzekowski DL. "It's got to be on this page": age and cognitive style in a study of online health information seeking. J Med Internet Res 2015 Mar 24;17(3):e79 [FREE Full text] [doi: 10.2196/jmir.3352] [Medline: 25831483]
27. Young K. Direct from the source: the value of 'think-aloud' data in understanding learning. J Educ Enquiry 2005;6(1):25 [FREE Full text]
28. In de wereld van kanker wijzen wij je de weg. kanker.nl. URL: https://www.kanker.nl/ [accessed 2021-07-30]
29. Kim J. Scenarios in information seeking and information retrieval research: a methodological application and discussion. Libr Inf Sci Res 2012 Oct;34(4):300-307. [doi: 10.1016/j.lisr.2012.04.002]
30. van Zuuren FJ, de Groot KI, Mulder NL, Peter M. Coping with medical threat: an evaluation of the Threatening Medical Situations Inventory (TMSI). Pers Individ Dif 1996 Jul;21(1):21-31. [doi: 10.1016/0191-8869(96)00029-3]
31. Bronner MB, Nguyen MH, Smets EM, van de Ven AW, van Weert JC. Anxiety during cancer diagnosis: examining the influence of monitoring coping style and treatment plan. Psychooncology 2018 Feb;27(2):661-667. [doi: 10.1002/pon.4560] [Medline: 28976610]
32. de Bruin GO, Rassin E, van der Heiden C, Muris P. Psychometric properties of a Dutch version of the intolerance of uncertainty scale. Neth J Psychol 2006 Dec;62(2):87-92. [doi: 10.1007/bf03061055]

XSL•FO
RenderX

33. Thomas R, Kaminski E, Stanton E, Williams M. Measuring information strategies in oncology - developing an information satisfaction questionnaire. Eur J Cancer Care (Engl) 2004 Mar;13(1):65-70. [doi: 10.1111/j.1365-2354.2004.00445.x] [Medline: 14961777]
34. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011 Nov 09;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
35. Aymé S, Kole A, Groft S. Empowerment of patients: lessons from the rare diseases community. Lancet 2008 Jun 14;371(9629):2048-2051. [doi: 10.1016/S0140-6736(08)60875-2] [Medline: 18555918]
36. Wilson T. Exploring models of information behaviour: the 'uncertainty' project. Inf Process Manag 1999 Nov;35(6):839-849. [doi: 10.1016/S0306-4573(99)00029-1]
37. Kwakernaak J, Eekhof JA, de Waal MW, Barenbrug EA, Chavannes NH. Patients' use of the internet to find reliable medical information about minor ailments: vignette-based experimental study. J Med Internet Res 2019 Nov 11;21(11):e12278 [FREE Full text] [doi: 10.2196/12278] [Medline: 31710304]
38. Mueller J, Jay C, Harper S, Todd C. The role of web-based health information in help-seeking behavior prior to a diagnosis of lung cancer: a mixed-methods study. J Med Internet Res 2017 Jun 08;19(6):e189 [FREE Full text] [doi: 10.2196/jmir.6336] [Medline: 28596146]
39. Aboueid S, Liu RH, Desta BN, Chaurasia A, Ebrahim S. The use of artificially intelligent self-diagnosing digital platforms by the general public: scoping review. JMIR Med Inform 2019 May 01;7(2):e13445 [FREE Full text] [doi: 10.2196/13445] [Medline: 31042151]
40. Lepore SJ. A social–cognitive processing model of emotional adjustment to cancer. In: Baum A, Andersen BL, editors. Psychosocial Interventions for Cancer. Washington, DC: American Psychological Association; 2001:99-116.
41. Pang PC, Verspoor K, Chang S, Pearce J. Conceptualising health information seeking behaviours and exploratory search: result of a qualitative study. Health Technol 2015 Feb 10;5(1):45-55. [doi: 10.1007/S12553-015-0096-0]
42. Keselman A, Browne AC, Kaufman DR. Consumer health information seeking as hypothesis testing. J Am Med Inform Assoc 2008;15(4):484-495 [FREE Full text] [doi: 10.1197/jamia.M2449] [Medline: 18436912]
43. Meppelink CS, Smit EG, Fransen ML, Diviani N. "I was right about vaccination": confirmation bias and health literacy in online health information seeking. J Health Commun 2019;24(2):129-140. [doi: 10.1080/10810730.2019.1583701] [Medline: 30895889]
44. Lazarus RS. Emotion and Adaptation. Oxford, UK: Oxford Academic Press; 1991.
45. Frijda NH. The laws of emotion. Am Psychol 1988 May;43(5):349-358. [doi: 10.1037//0003-066x.43.5.349] [Medline: 3389582]
46. Izard CE. Differential emotions theory. In: Izard CE, editor. Human Emotions. Cham, Switzerland: Springer; 1977:43-66. [doi: 10.1007/978-1-4899-2209-0_3]
47. Lazarus RS. Hope: an emotion and a vital coping resource against despair. Soc Res 1999;66(2):653-678 [FREE Full text]
48. de los Santos TM, Nabi RL. Emotionally charged: exploring the role of emotion in online news information seeking and processing. J Broadcast Electron Media 2019 Mar 18;63(1):39-58. [doi: 10.1080/08838151.2019.1566861]
49. Hovick SR, Kahlor L, Liang MC. Personal cancer knowledge and information seeking through PRISM: the planned risk information seeking model. J Health Commun 2014 Apr;19(4):511-527 [FREE Full text] [doi: 10.1080/10810730.2013.821556] [Medline: 24433251]
50. ter Huurne EF, Griffin RJ, Gutteling JM. Risk information seeking among U.S. and Dutch residents. Sci Commun 2009 Feb 27;31(2):215-237. [doi: 10.1177/1075547009332653]
51. Yang ZJ, Kahlor L. What, me worry? The role of affect in information seeking and avoidance. Sci Commun 2013;35(2):189-212. [doi: 10.1177/1075547012441873]
52. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004 Oct;7(5):497-510. [doi: 10.1089/cpb.2004.7.497] [Medline: 15667044]
53. Fergie G, Hilton S, Hunt K. Young adults' experiences of seeking online information about diabetes and mental health in the age of social media. Health Expect 2016 Dec;19(6):1324-1335 [FREE Full text] [doi: 10.1111/hex.12430] [Medline: 26647109]
54. Hofstede G. Dimensionalizing cultures: the Hofstede model in context. Online Read Psychol Cult 2011 Dec 01;2(1):22. [doi: 10.9707/2307-0919.1014]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
NHL: non-Hodgkin lymphoma
R-CHOP: rituximab, cyclophosphamide, hydroxydaunorubicin, oncovin, and prednisone
WHIS: web-based health information seeking

XSL•FO
RenderX

Edited by R Cuomo; submitted 17.04.24; peer-reviewed by PCI Pang, W van Harten; comments to author 05.06.24; revised version received 30.07.24; accepted 21.11.24; published 16.01.25.

Please cite as:
Huijgens F, Kwakman P, Hillen M, van Weert J, Jaspers M, Smets E, Linn A
How Patients With Cancer Use the Internet to Search for Health Information: Scenario-Based Think-Aloud Study
JMIR Infodemiology 2025;5:e59625
URL: https://infodemiology.jmir.org/2025/1/e59625
doi:10.2196/59625
PMID:

©Fiorella Huijgens, Pascale Kwakman, Marij Hillen, Julia van Weert, Monique Jaspers, Ellen Smets, Annemiek Linn. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 16.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Assessment of Reliability and Validity of Celiac Disease–Related YouTube Videos: Content Analysis

Yunus Halil Polat[1*], MD; Rasim Eren Cankurtaran[2*], MD

[1]Department of Gastroenterology, Ankara Training and Research Hospital, Ankara, Turkey

[2]Department of Gastroenterology, Ankara Etlik City Hospital, Ankara, Turkey

[*]all authors contributed equally

Corresponding Author:
Yunus Halil Polat, MD
Department of Gastroenterology
Ankara Training and Research Hospital
bağlıca neighbourhood 1336 street no:7/7 Etimesgut
PO:06870
Ankara, 06870
Turkey
Phone: 90 5556801815
Email: yunushpolat@gmail.com

Abstract

Background: YouTube is an increasingly used platform for medical information. However, the reliability and validity of health-related information on celiac disease (CD) on YouTube have not been determined.

Objective: This study aimed to analyze the reliability and validity of CD-related YouTube videos.

Methods: On November 15, 2023, a search was performed on YouTube using the keyword "celiac disease." This search resulted in a selection of videos, which were then reviewed by 2 separate evaluators for content, origin, and specific features. The evaluators assessed the reliability and quality of these videos using a modified DISCERN (mDISCERN) score, the *Journal of the American Medical Association* (*JAMA*) benchmark criteria score, the usefulness score, video power index (VPI), and the Global Quality Scale (GQS) score.

Results: In the analysis of 120 initially screened CD videos, 85 met the criteria for inclusion in the study after certain videos were excluded based on predefined criteria. While the duration of the videos uploaded by health care professionals was significantly longer than the other group (P=.009), it was concluded that the median scores for mDISCERN (4, IQR 4-5 vs 2, IQR 2-3; P<.001), GQS (4, IQR 4-5 vs 3, IQR 2-3; P<.001), *JAMA* (4, IQR 3-4 vs 2, IQR 2-3; P<.001), and usefulness (8, IQR 7-9 vs 6, IQR 3-6; P<.001) of the videos from this group were significantly higher than those from non–health care professionals. Video interaction parameters, including the median number of views, views per day, likes, dislikes, comments, and VPI, demonstrated no significant difference between the 2 groups.

Conclusions: This study showed that YouTube videos about CD vary significantly in reliability and quality depending on their source. Increasing the production of reliable videos by health care professionals may help to improve patient education and make YouTube a more reliable resource.

(JMIR Infodemiol 2025;5:e58615) doi:10.2196/58615

KEYWORDS

gastroenterology; celiac; YouTube; internet-based information; medical information; health-related; reliability; validity; quality; videos; celiac sprue; sprue; gluten enteropathy; cross-sectional

Introduction

Celiac disease (CD) is an autoimmune disorder that occurs in genetically predisposed individuals as a result of the immune reaction to gluten, primarily affecting the small intestine [1]. Symptoms range from asymptomatic to digestive problems and nutritional deficiencies due to malabsorption of nutrients. Treatment includes a gluten-free diet [1]. Over the past few decades, CD has been estimated to affect around 1% of the world's population [2]. Despite the increasing prevalence of

XSL•FO
RenderX

CD, the majority of the patients with CD remain undiagnosed [1].

In recent years, the internet has become an important source of health information for the public. It has been reported that 80% of internet users use social media (SM) platforms to get information about their disease. Patients with chronic diseases in particular are increasingly relying on SM platforms to manage their conditions [3]. In a recent study investigating the use of SM by patients with CD and parents of patients with CD, it was reported that 96% of participants used SM for disease management [4]. YouTube (Google), is one of the world's most popular video-sharing platforms. Currently, YouTube has more than 1 billion registered users, and billions of videos are watched every day, about 30 million of which are health-related. Health-related videos can be uploaded by anyone, but the content of these videos may contain inaccurate or misleading information without being reviewed by health care professionals.

There are studies in the literature evaluating the reliability and quality of YouTube videos for many diseases [5,6]. There are few studies evaluating CD-related YouTube videos [7,8]. However, one of these studies evaluated non-English language videos [8]. The other study did not measure CD-related YouTube videos with the tests developed for these studies and did not include videometric parameters (such as the number of likes and dislikes) in the evaluation [7]. Unlike previous studies, which either focused on non-English videos or lacked comprehensive quality metrics, this research provides a more robust and comparative analysis of CD-related video content on YouTube.

We could not find any studies in the literature that evaluated the reliability and validity of YouTube videos about CD. This study aims to evaluate the quality and reliability of YouTube videos about CD using validated scoring tools and detailed content analysis.

Methods

Study Design

In this cross-sectional study, videos were collected using the keyword "Celiac Disease" in YouTube's search engine on November 15, 2023. The search was conducted in a Google Chrome browser in incognito mode, logged out of any user account, and using a standard IP address in Turkey. This was chosen because it is the most common keyword that holistically assesses all aspects of the disease, such as clinical, pathogenesis, diet, and nutrition. YouTube's default relevance mode was used to simulate the average consumer's search habits. It is recognized that most viewers rarely venture beyond the first few pages of results. Therefore, the first 120 videos about CD were selected, similar to previous studies. Based on the search results, a total of 120 videos were saved for further analysis, ranging from the most viewed video to the least viewed video. Video sampling criteria were determined with reference to similar studies [5,9].

The following factors were considered as exclusion criteria in the research: (1) videos in languages other than English, (2) videos with muted or poor picture quality, (3) videos containing advertisements, (4) videos with content unrelated to CD, and (5) videos with repetitive content.

Data Review

Data such as video type (real and animation), video length (min), time since upload (d), number of views, number of daily views (number of views/d since upload), number of likes, number of daily likes (number of likes/d since upload), number of dislikes, and number of comments were recorded. In our study, we categorized video sources into two groups: educational content of health care professionals (doctors, academic institutions or professional organizations, and health-related websites) and personal narratives of non–health care professionals (patients, independent users). The videos were independently analyzed by 2 raters (YHP and REC) and coded according to the themes of "Educational content" and "Personal narratives." Discrepancies in coding were resolved through repetitive discussions and consensus, ensuring a reliable and consistent categorization process. This method of assessment has been used in similar studies of other diseases [10].

Video Usefulness

The usefulness score is a usefulness scale defined by Lee et al [11]. Each video is rated with a score between 0 and 10 depending on the content of the video, such as causes, symptoms, diagnosis, diagnosis, and recovery status. According to the total score obtained, it is categorized as follows: 0=not useful, 1-3=less useful, 4-7=useful, and 8-10=very useful.

Video Popularity

The video power index (VPI) developed by Erdem et al [12] shows the popularity of videos and has been used in many studies [9]. The VPI calculation is as follows: VPI=(×100/[number of likes+number of dislikes]) × (number of views/number of d since upload)/100.

Quality and Reliability Evaluation

The Global Quality Scale (GQS) assesses the quality by providing the interpretation and usefulness of the videos for patients based on the flow of information. GQS has a 5-point Likert structure according to the quality, flow, and ease of use of the analyzed videos [13]. As used in similar studies, scores 1-2 were considered as low quality (inadequate in terms of patient information, contains incomplete information), 3 as medium quality (video flow is poor, some information is available but important issues are not addressed), and 4-5 (contains sufficient and useful information for patients) as high quality [14].

The quality assessment included the *Journal of the American Medical Association* (*JAMA*) benchmark criteria for determining authorship, attribution, disclosure, and currency. Each of these criteria was given a score of 1, with a maximum score of 4 [15].

The mDISCERN scale developed by Charnock et al [16] and later adapted to YouTube videos by Singh et al [17] was used to assess the reliability of the videos. The mDISCERN scale consists of 5 questions and is a questionnaire about information sources, purpose, reliability, bias, additional sources, and areas of uncertainty. Each question can be answered yes or no. Each

XSL•FO
RenderX

yes answer is worth 1 point and 5 points represent the highest quality.

The video content was evaluated and graded according to the most recent American College of Gastroenterology guidelines for the management of CD [18]. These guidelines emphasize accurate symptom identification, diagnostic criteria, and effective dietary management strategies. Videos were scored for reliability, usefulness, and consistency with evidence-based practice.

Statistical Analyses

The SPSS (version 25.0 for Windows; IBM Corp) package program was used. Continuous variables were evaluated using the Shapiro-Wilk test to determine whether they were normally distributed. Continuous variables are reported as median and IQR, while categorical variables are presented as counts and percentages. Chi-square tests were used to analyze categorical variables and Mann-Whitney *U* test for numerical variables. The significance level was set at *P*=.05 for all analyses.

Ethical Considerations

The study adhered to the ethical standards outlined in the Helsinki Declaration and complied with national regulations in the respective field. Since the study did not involve the use of human or animal data, ethics committee approval was not necessary. This study analyzed publicly available YouTube videos. No identifiable personal data was used, and all results are presented in aggregate. Therefore, formal ethics approval was not required.

Results

Main Characteristics of Videos and Video Analysis

In total, 120 videos were analyzed and 85 videos met the study criteria and were included. A total of 35 videos were excluded from the study, including 2 non-English language videos, 13 videos with repetitive content, 12 videos with advertising content, and 8 videos with poor picture and sound quality. Most (22/85, 25.9%) were published by universities and other organizations, and most (50/85, 59%) were uploaded by health care professionals. A total of 68.2% (58/85) of the videos consisted of real images. Descriptive statistics of the above characteristics and other variables are shown in Table 1.

Table 1. Main characteristics of the analyzed videos. Categorical variables are expressed as n (%), and numerical variables are expressed as median (Q1-Q3).

Characteristics	Values
Source, n (%)	
Physicians	12 (14
Universities and professional organizations	22 (26)
Health information websites	16 (19)
Independent users	16 (19)
Patient	19 (22)
Source, n (%)	
Health care professionals	50 (59)
Non–health care professionals	35 (41)
Image type	
Real image, n (%)	58 (68)
Animation, n (%)	27 (32)
Number of views, median (IQR)	17,026 (2860-46,358)
Number of likes, median (IQR)	306 (45-820)
Number of dislikes, median (IQR)	6 (1-20)
Duration (min), median (IQR)	6.3 (3.4-12.1)
Days on YouTube, median (IQR)	1381 (572-2290)
Number of comments, median (IQR)	27 (5-130)
Views per day, median (IQR)	13.1 (4-33.2)
Likes per day, median (IQR)	0.2 (0.1-0.7)

Content Analysis and Source Evaluation of Videos

In the health care professional group, most (37/85, 43.1%) of the videos were uploaded by universities and other organizations, whereas in the non–health care professional group, most (19/34, 55.9%) of the videos were uploaded by "patients" (*P*<.001). While the duration of the videos uploaded by health care professionals was significantly longer than the

XSL•FO
RenderX

other group (P=.009), it was concluded that the median scores for mDISCERN (4, IQR 4-5 vs 2, IQR 2-3; P<.01), GQS (4, IQR 4-5 vs 3, IQR 2-3; P<.001), *JAMA* (4, IQR 3-4 vs 2, IQR 2-3; P<.001), and usefulness (8, IQR 7-9 vs 6, IQR 3-6; P<.001) of the videos from this group were significantly higher than those from non–health care professionals. (Tables 2 and 3)

Table 2. The average scales of the analyzed videos.

Video scales	Values, median (IQR)
mDISCERN[a]	3 (3-4)
GQS[b]	4 (3-4)
JAMA [c]	3 (2-4)
VPI[d]	12.8 (4-33)
Usefulness	7 (5-9)

[a]mDISCERN: modified DISCERN score.
[b]GQS: Global Quality Scale score.
[c]*JAMA*: Journal of the American Medical Association.
[d]VPI: video power index.

Table 3. Comparison of videos according to source status. Categorical variables are expressed as n (%), and numerical variables as median (Q1-Q3).

Variables	Source		*P* value
	Health care professionals	Non–health care professionals	
Image			
Real image, n (%)	31 (62)	27 (77.1)	.21
Animation, n (%)	19 (38)	8 (22.9)	
Number of views, median (IQR)	16,657 (4858-57,896)	17,851.5 (1907-43,310)	.87
Number of likes, median (IQR)	297 (52-774)	373 (22-846)	.67
Number of dislikes, median (IQR)	6 (1-24)	8.5 (0-18)	.92
Duration (min), median (IQR)	7.4 (4.2-16.4)	3.9 (2.5-8.2)	.009
Days on YouTube, median (IQR)	1291 (516-2290)	1467.5 (832-2470)	.64
Number of comments, median (IQR)	21 (6-79)	67 (3-170)	.52
View per day, median (IQR)	12.8 (4.6-40.9)	15.6 (2.1-33.2)	.50
Like per day, median (IQR)	0.23 (0.07-1)	0.18 (0.03-0.73)	.39
mDISCERN[a], median (IQR)	4 (4-5)	2 (2-3)	<.001
GQS[b], median (IQR)	4 (4-5)	3 (2-3)	<.001
JAMA[c], median (IQR)	4 (3-4)	2 (2-2)	<.001
VPI[d], median (IQR)	12.3 (4.6-41)	15.3 (2.1-33)	.72
Usefulness, median (IQR)	8 (7-9)	5 (3-6)	<.001

[a]mDISCERN: modified DISCERN score.
[b]GQS: Global Quality Scale score.
[c]*JAMA: Journal of the American Medical Association.*
[d]VPI: video power index.

Themes Identified in Videos

From the 85 included videos, two major themes were identified.

Educational Content

These videos, primarily created by health care professionals, provided detailed information about CD symptoms, diagnosis, treatment, and long-term management. This category accounted for 59% (50/85) of all videos and demonstrated significantly higher scores in quality and reliability metrics (mDISCERN, GQS, *JAMA*, and Usefulness; P<.001).

Personal Narratives

Uploaded by patients or non–health care professionals, these videos focused on personal journeys, sharing challenges, and tips for living with CD. They received moderate interaction

XSL•FO
RenderX

metrics (likes, comments) but were lower in quality and reliability scores (*P*<.001).

Discussion

Principal Findings

In this study, we analyzed YouTube videos about CD, an important disease that can occur at any age. We found that CD videos uploaded by health care professionals were significantly more reliable, adequate, useful, and quality information sources than those uploaded by non–health care professionals. Another striking result of the study was that there was no difference in video interaction parameters between those with and without health care professionals as video sources.

Recently, SM has become a popular way to access medical information and knowledge. Patients with many chronic diseases, including CD, have been reported to use SM as a source of information since adolescence [19]. Especially YouTube, a video sharing website, has become an important source of information in the field of health. In a recent nationally based survey study, it was reported that younger patient groups and patients with chronic diseases such as hypertension, diabetes mellitus, and chronic lung disease were more likely to watch YouTube videos as a source of health-related information [20].

As in other chronic diseases, SM use among patients with CD and their families has become widespread in recent years [4]. When we consider the importance of increasing adherence to a gluten-free diet as well as the diagnosis, risk factors, and clinical presentation of the disease, access to real and adequate information through SM becomes even more important. In a recent survey of patients with CD, two-thirds of the patients used SM every day for an average of 60 minutes per day. The 3 most common reasons for using SM were researching gluten-free diet products, obtaining information about diet, and CD. In the study, it was stated that the most frequently used platform was WhatsApp (Meta), and it was suggested that YouTube usage was 4% [4]. Although this rate may vary according to regional and cultural differences, it is still a relatively low rate and suggests that the use of YouTube may be higher than this data. In another similar survey study conducted in Japan, 27% of more than 2000 participants with chronic diseases stated that they used the YouTube platform related to their disease [20].

One of the studies evaluating YouTube videos on CD was a study in which 100 videos were evaluated in 2019. In this study, it was examined whether there was a difference between sources in 31 different topics such as etiology, symptoms, diagnosis, and treatment of the disease, and it was stated that there was no significant difference in terms of content in all remaining topics except 3 [7]. However, none of the video reliability-efficacy tests used in our study were used in this study. Nevertheless, it differs from our study because it claims that there is mostly no significant difference between videos whose source is health care professionals and other videos in terms of topics. Another study in the literature evaluated Polish-language videos, so it does not seem possible to make a comparison with our study [8].

Among the videos analyzed in our study, the fact that the reliability, usefulness, and quality scores of the videos of health care professionals were significantly higher than those of non–health care professionals was also observed in similar studies evaluating other diseases [21]. One of the most remarkable findings of our study is that there was no significant difference between the groups in terms of views, likes, dislikes, and VPI. There are many factors that can contribute to this, such as the visual presentation of the video, the demographic and cultural make-up of the viewers, the video's viral status, and the influencer's effect [22,23]. In a recent study investigating the influencer effect on SM related to dermatology, it was shown that dermatologists without competence and certification had as high a level of interaction as those with competence and certification [23]. This finding shows us that videos that may be insufficient as a source of information may also have high interaction and accordingly may cause misinformation and negative effects on patients and their families.

Based on these findings, we believe that in order for YouTube to be an accurate source of information about CD, many organizations and institutions, such as professional associations and universities, should provide training for health care professionals to produce high-quality videos that can provide more interaction and raise awareness among health care professionals about this issue. On the other hand, it is also important to raise patient awareness of the possibility that patients may be exposed to misinformation when using YouTube. We think that more use of YouTube and other SM platforms by health care professionals and peer review of health-related video content may reduce misinformation.

Limitations

There were some limitations in our study. The first 120 videos searched with the keyword "Celiac disease'" in the search results were analyzed and the other videos were not analyzed. In addition, since YouTube is a dynamic SM platform, video interaction parameters such as daily views, likes, and dislikes can change every day. Finally, the fact that only English videos were analyzed in our study can be considered among the limitations.

Conclusions

This study showed that YouTube videos about CD vary significantly in reliability and quality depending on their source. Increasing the production of reliable videos by health care professionals may help to improve patient education and make YouTube a more reliable resource.

XSL•FO
RenderX

Authors' Contributions

REC and YHP contributed to conceptualization, resources, data curation, writing (original draft preparation and review & editing), formal analysis, project administration, software, validation, and visualization. REC was responsible for developing the methodology, conducting the statistical analyses, and investigating and supervising the project.

Conflicts of Interest

None declared.

References

1. Lebwohl B, Sanders DS, Green PHR. Coeliac disease. Lancet 2018;391(10115):70-81. [doi: 10.1016/S0140-6736(17)31796-8] [Medline: 28760445]
2. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. Clin Gastroenterol Hepatol 2018;16(6):823-836.e2. [doi: 10.1016/j.cgh.2017.06.037] [Medline: 29551598]
3. Fox S. Online health search 2006. Pew Research Center. 2006. URL: https://www.pewresearch.org/internet/2006/10/29/online-health-search-2006/ [accessed 2025-01-22]
4. Al Sarkhy A. Social media usage pattern and its influencing factors among celiac patients and their families. Saudi J Gastroenterol 2020;26(2):99-104 [FREE Full text] [doi: 10.4103/sjg.SJG_495_19] [Medline: 32031161]
5. Karakoyun A, Yildirim A. YouTube videos as a source of information concerning Behçet's disease: a reliability and quality analysis. Rheumatol Int 2021;41(12):2117-2123. [doi: 10.1007/s00296-021-05009-9] [Medline: 34590188]
6. Tutan D, Kaya M. Evaluation of YouTube videos as a source of information on hepatosteatosis. Cureus 2023;15(10):e46843 [FREE Full text] [doi: 10.7759/cureus.46843] [Medline: 37829652]
7. Basch CH, Hillyer GC, Garcia P, Basch CE. Content of widely viewed YouTube videos about celiac disease. Public Health 2019;167:147-151. [doi: 10.1016/j.puhe.2018.11.004] [Medline: 30682698]
8. Kiedrowski M, Mróz A, Gajewska D, Nurzyński P, Deptała A. Celiac disease on YouTube - a study of the Polish content available on the popular video-sharing website. Pol Merkur Lekarski 2017;43(256):168-171. [Medline: 29084190]
9. Pons-Fuster E, Ruiz Roca J, Tvarijonaviciute A, López-Jornet P. YouTube information about diabetes and oral healthcare. Odontology 2020;108(1):84-90. [doi: 10.1007/s10266-019-00445-3] [Medline: 31396751]
10. Uz C, Umay E, Ballı Uz F, Cankurtaran D, Gundogdu I. Assessment of the quality and reliability of the information on spasticity on YouTube. J Clin Neurosci 2023;113:142-146. [doi: 10.1016/j.jocn.2023.05.018] [Medline: 37269749]
11. Lee JS, Seo HS, Hong TH. YouTube as a source of patient information on gallstone disease. World J Gastroenterol 2014;20(14):4066-4070 [FREE Full text] [doi: 10.3748/wjg.v20.i14.4066] [Medline: 24744597]
12. Erdem MN, Karaca S. Evaluating the accuracy and quality of the information in kyphosis videos shared on YouTube. Spine (Phila Pa 1976) 2018;43(22):E1334-E1339. [doi: 10.1097/BRS.0000000000002691] [Medline: 29664816]
13. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol 2007;102(9):2070-2077. [doi: 10.1111/j.1572-0241.2007.01325.x] [Medline: 17511753]
14. Kocyigit BF, Akaltun MS, Sahin AR. YouTube as a source of information on COVID-19 and rheumatic disease link. Clin Rheumatol 2020;39(7):2049-2054 [FREE Full text] [doi: 10.1007/s10067-020-05176-3] [Medline: 32447603]
15. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the internet. JAMA 1997;277(15):1244-1245. [doi: 10.1001/jama.1997.03540390074039]
16. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
17. Singh AG, Singh S, Singh PP. YouTube for information on rheumatoid arthritis--a wakeup call? J Rheumatol 2012;39(5):899-903. [doi: 10.3899/jrheum.111114] [Medline: 22467934]
18. Rubio-Tapia A, Hill ID, Semrad C, Kelly CP, Greer KB, Limketkai BN, et al. American College of Gastroenterology guidelines update: diagnosis and management of celiac disease. Am J Gastroenterol 2023;118(1):59-76. [doi: 10.14309/ajg.0000000000002075] [Medline: 36602836]
19. De Nardi L, Trombetta A, Ghirardo S, Genovese MRL, Barbi E, Taucar V. Adolescents with chronic disease and social media: a cross-sectional study. Arch Dis Child 2020;105(8):744-748. [doi: 10.1136/archdischild-2019-317996] [Medline: 31941715]
20. Mitsutake S, Takahashi Y, Otsuki A, Umezawa J, Yaguchi-Saito A, Saito J, INFORM Study Group. Chronic diseases and sociodemographic characteristics associated with online health information seeking and using social networking sites: nationally representative cross-sectional survey in Japan. J Med Internet Res 2023;25:e44741 [FREE Full text] [doi: 10.2196/44741] [Medline: 36862482]
21. Onder ME, Zengin O. Quality of healthcare information on YouTube: psoriatic arthritis. Z Rheumatol 2023;82(Suppl 1):30-37 [FREE Full text] [doi: 10.1007/s00393-021-01069-1] [Medline: 34468808]

XSL•FO
RenderX

22. Rao A, Morstatter F, Lerman K. Partisan asymmetries in exposure to misinformation. Sci Rep 2022;12(1):15671 [FREE Full text] [doi: 10.1038/s41598-022-19837-7] [Medline: 36123387]
23. O'Brien KF, Newsom EC, Park JH, Lawrence N. Increasing a dermatologist's footprint on Instagram: an analysis of top influencers performing nonsurgical cosmetic procedures. Dermatol Surg 2021;47(8):1093-1097. [doi: 10.1097/DSS.0000000000003088] [Medline: 33988555]

Abbreviations

CD: celiac disease
GQS: Global Quality Scale
***JAMA*:** *Journal of the American Medical Association*
mDISCERN: modified DISCERN
SM: social media
VPI: video power index

Edited by T Mackey; submitted 20.03.24; peer-reviewed by S Guandalini, H Meteran; comments to author 24.05.24; revised version received 03.06.24; accepted 05.01.25; published 04.02.25.

Please cite as:
Polat YH, Cankurtaran RE
Assessment of Reliability and Validity of Celiac Disease–Related YouTube Videos: Content Analysis
JMIR Infodemiology 2025;5:e58615
URL: https://infodemiology.jmir.org/2025/1/e58615
doi:10.2196/58615
PMID:39756057

©Yunus Halil Polat, Rasim Eren Cankurtaran. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 04.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Unveiling Topics and Emotions in Arabic Tweets Surrounding the COVID-19 Pandemic: Topic Modeling and Sentiment Analysis Approach

Farah Alshanik[1], PhD; Rawand Khasawneh[2], PharmD; Alaa Dalky[3], PhD; Ethar Qawasmeh[4], MS

[1]Department of Computer Science, Faculty of Computer and Information Technology, Jordan University of Science and Technology, Irbid, Jordan

[2]Department of Clinical Pharmacy, Faculty of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan

[3]Department of Health Management and Policy, Faculty of Medicine, Jordan University of Science and Technology, Irbid, Jordan

[4]Department of Computer Science, Faculty of Computer Science and Engineering, The Ohio State University, Columbus, OH, United States

Corresponding Author:
Farah Alshanik, PhD
Department of Computer Science
Faculty of Computer and Information Technology
Jordan University of Science and Technology
Alhusun St
Irbid, 22110
Jordan
Phone: 962 2 7201000 ext 23130
Fax: 962 2 7095123
Email: fmalshanik@just.edu.jo

Abstract

Background: The worldwide effects of the COVID-19 pandemic have been profound, and the Arab world has not been exempt from its wide-ranging consequences. Within this context, social media platforms such as Twitter have become essential for sharing information and expressing public opinions during this global crisis. Careful investigation of Arabic tweets related to COVID-19 can provide invaluable insights into the common topics and underlying sentiments that shape discussions about the COVID-19 pandemic.

Objective: This study aimed to understand the concerns and feelings of Twitter users in Arabic-speaking countries about the COVID-19 pandemic. This was accomplished through analyzing the themes and sentiments that were expressed in Arabic tweets about the COVID-19 pandemic.

Methods: In this study, 1 million Arabic tweets about COVID-19 posted between March 1 and March 31, 2020, were analyzed. Machine learning techniques, such as topic modeling and sentiment analysis, were applied to understand the main topics and emotions that were expressed in these tweets.

Results: The analysis of Arabic tweets revealed several prominent topics related to COVID-19. The analysis identified and grouped 16 different conversation topics that were organized into eight themes: (1) preventive measures and safety, (2) medical and health care aspects, (3) government and social measures, (4) impact and numbers, (5) vaccine development and research, (6) COVID-19 and religious practices, (7) global impact of COVID-19 on sports and countries, and (8) COVID-19 and national efforts. Across all the topics identified, the prevailing sentiments regarding the spread of COVID-19 were primarily centered around anger, followed by disgust, joy, and anticipation. Notably, when conversations revolved around new COVID-19 cases and fatalities, public tweets revealed a notably heightened sense of anger in comparison to other subjects.

Conclusions: The study offers valuable insights into the topics and emotions expressed in Arabic tweets related to COVID-19. It demonstrates the significance of social media platforms, particularly Twitter, in capturing the Arabic-speaking community's concerns and sentiments during the COVID-19 pandemic. The findings contribute to a deeper understanding of the prevailing discourse, enabling stakeholders to tailor effective communication strategies and address specific public concerns. This study underscores the importance of monitoring social media conversations in Arabic to support public health efforts and crisis management during the COVID-19 pandemic.

(JMIR Infodemiology 2025;5:e53434) doi:10.2196/53434

XSL•FO
RenderX

KEYWORDS
topic modeling; sentiment analysis; COVID-19; social media; Twitter; public discussion

Introduction

Background

Throughout history, humanity has faced numerous outbreaks of infectious diseases that have resulted in significant loss of life and economic impact. Toward the end of 2019, the World Health Organization reported a series of pneumonia cases in Wuhan, which were later identified as COVID-19. As a novel infectious disease transmitted through respiratory droplets and contact, COVID-19 quickly spread across the globe, leading to an unprecedented impact on global public health, businesses, and economies. As of February 7, 2023, there have been >676 million confirmed cases and 500,000 reported deaths in >200 countries [1]. Social media platforms, particularly Twitter, have emerged as valuable sources of information for understanding and predicting disease outbreaks. Text mining techniques allow for the extraction of relevant health information from user-generated content on social media platforms. Twitter, in particular, provides researchers with vast amounts of real-time data, enabling early response strategies and enhancing situational awareness. Analyzing Twitter data has become a crucial area of focus in medical informatics research [2,3].

COVID-19 emerged as a prominent and sustained topic on Twitter starting from January 2020, and its discussion has persisted uninterrupted up to the present day [4]. With quarantine measures implemented worldwide, individuals increasingly relied on social media to access news and express their opinions. Twitter data offer valuable insights into public discussions, sentiments, and real-time updates during global pandemics [2,5]. Using Twitter as a data source enables infodemiology studies, providing health authorities with opinions and concerns to inform their responses [6].

Since the outset of the COVID-19 outbreak, an escalating number of studies have been harnessing Twitter data to delve into the public's reactions and discussions surrounding the COVID-19 pandemic. In their respective studies, researchers used distinct methodologies to explore COVID-19–related discussions and sentiments. For instance, Xue et al [4,7] used latent Dirichlet allocation (LDA) for topic identification. Similarly, a study by Alharbi and Alkhateeb [8] investigated the sentiment of the Arabic public on Twitter, using natural language processing (NLP) and machine learning techniques, finding that the long short-term memory model outperformed the naive Bayes model with an accuracy rate of 99% [8]. Another study focused on Arabic sentiment analysis for vaccine-related COVID-19 tweets, introducing the first and largest human-annotated dataset in Arabic for this purpose; it used advanced models such as the stacked gated recurrent unit and AraBERT, achieving a 7% accuracy enhancement [9]. During the COVID-19 pandemic, a separate study analyzed online learning–related tweets in Arabic, using various classification algorithms and achieving a maximum accuracy of approximately 89.6% using the Support Vector Machine classifier to analyze public perceptions of the coronavirus [10]. In addition, research conducted in Saudi Arabia showed a significant increase in negative sentiments during the COVID-19 pandemic, with deep learning algorithms achieving high accuracy rates [11]. Other studies explored sentiment differences between countries and in response to events, using topic modeling and sentiment analysis to reveal previously unreported patterns [12]. Furthermore, a study from Morocco compared different machine learning algorithms for tweet classification, finding logistic regression to yield the best sentiment predictions [13].

Recent advancements in NLP have shown significant potential in transforming various aspects of health care, including clinical decision support, patient management, and automated analysis of health records. Recent studies, such as the one by Tamang et al [14], highlight the use of NLP for optimizing patient outcome predictions and identifying disease patterns through electronic health record data. Similarly, a study by Elbattah et al [15] explores the role of NLP in extracting actionable insights from unstructured medical texts, further underscoring the growing relevance of NLP in enhancing the health care decision-making processes.

COVID-19 remains a scientifically and medically novel disease that requires in-depth and consistent research. Leveraging social media data, particularly from platforms such as Twitter, is essential for syndromic surveillance and understanding public health–related concerns. Twitter, as a prominent communication modality during disease outbreaks, offers valuable insights into public awareness and provides real-time reflections of public sentiment. Despite extensive research on COVID-19, limited studies have used social media data, specifically Twitter, to address conclusive themes and sentiment analysis in Arab regions during the early stages of the COVID-19 pandemic.

While numerous studies have investigated similar themes in different languages and contexts, there remains a notable gap in the analysis of Arabic tweets [16-22]. The Arabic-speaking population plays a significant role in the global discourse on COVID-19, and their perspectives and sentiments warrant dedicated exploration. Building on previous research, and to bridge this gap, our study used a combination of topic modeling techniques, specifically LDA, and sentiment analysis methods to uncover the predominant topics of discussion and the prevailing emotional tones within this corpus.

This Study

This study aims to analyze Twitter posts during the early stages of the COVID-19 pandemic in Arab regions to provide valuable insights into public sentiment, concerns, and awareness regarding COVID-19 in Arab communities. To achieve this, >1 million tweets posted between March 1 and March 31, 2020, were collected and analyzed. Through this analysis, we hope to assist policy makers in making informed decisions, enhancing public health communication, and implementing effective interventions to mitigate the impact of future outbreaks.

XSL•FO
RenderX

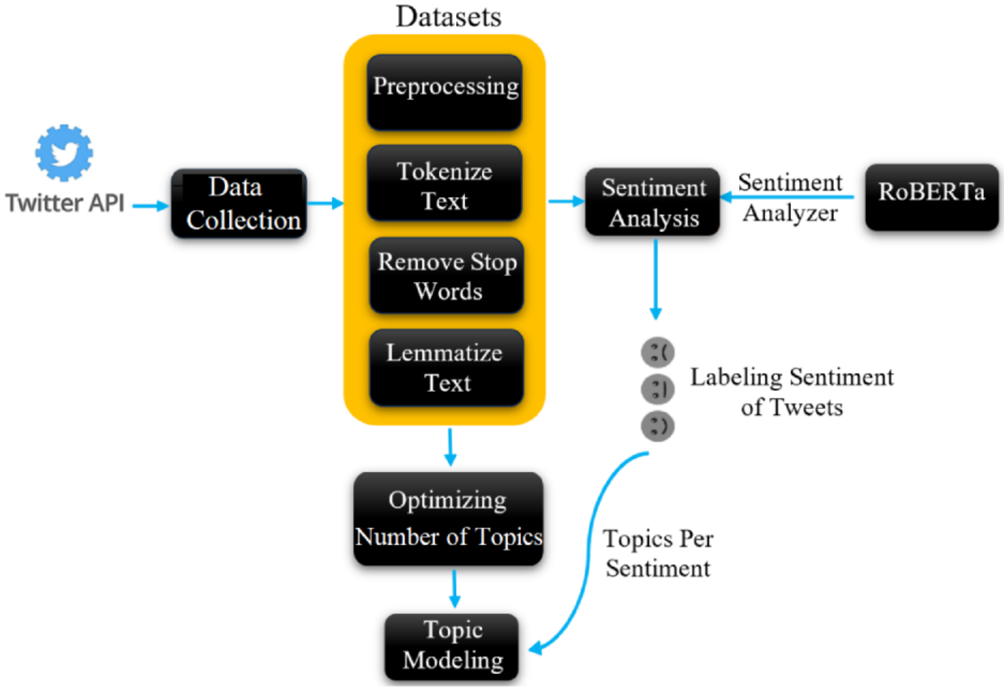
Although this study was conducted during the COVID-19 pandemic, its scope extends beyond the immediate implications of the COVID-19 pandemic. The primary goal of this research is to enhance health care planning and resource allocation in Jordan, which remains a critical issue regardless of pandemic conditions. The findings are designed to inform strategies that could be beneficial in various health care scenarios, whether in routine health care management or in response to other emergent public health challenges. Therefore, the study's relevance persists even in a postpandemic context, making it valuable for long-term health care system improvements.

Methods

Research Design

This study uses LDA for topic modeling and a sentiment analysis emotion detection tool to uncover topics and emotions in Twitter data related to COVID-19 in the Arab region. The methodological flowchart is depicted in Figure 1. Our approach to mining Twitter data adheres to the following 4 primary steps: data collection, data preprocessing, sentiment analysis, and topic modeling. The flowchart in Figure 1 illustrates how these steps are interconnected and carried out in our data analysis pipeline. Through these methods, we aim to gain valuable insights into the topics of discussion and the emotional responses of individuals in the Arab region concerning the COVID-19 pandemic.

Figure 1. Twitter data mining pipeline. API: application programming interface.



Data Collection

In our research, we harnessed the GeoCoV19 dataset, a multilingual COVID-19 Twitter dataset that spans a significant period of 90 days, from February 1, 2020, to May 1, 2020. This extensive dataset comprises hundreds of millions of tweets and is enriched with a diverse set of multilingual hashtags and keywords to ensure its comprehensiveness [23]. The dataset primarily provides tweet IDs, which presented us with the task of retrieving the actual tweet text associated with these IDs. To accomplish this, we made effective use of the Twarc application programming interface (API), a robust and efficient tool explicitly designed for this purpose [24]. The Twarc library was chosen due to its robustness in handling large-scale data collection, effective management of Twitter's API rate limits, seamless integration with existing data pipelines, and support for extended tweet metadata, making it an ideal tool for ensuring the integrity and completeness of the dataset required for this study. The Twarc API streamlined the process of collecting tweet texts corresponding to the tweet IDs provided. As we gathered all the tweets, we applied a language filter to focus exclusively on Arabic tweets. This selective filtering step was crucial for tailoring the dataset to our specific analysis, concentrating on tweets in the Arabic language.

Data Preprocessing

Data preprocessing plays a pivotal role in text mining, and it serves as a fundamental step in this domain. The purpose of this preprocessing is 2-fold: it optimizes the efficiency of prediction algorithms by eliminating potentially detrimental words, and it conserves storage space, contributing to improved computational performance [25]. In our analysis, we worked with Arabic text data, which requires thorough preprocessing to filter out any noise or irrelevant elements. The initial raw Arabic text underwent a series of transformations as part of this preprocessing effort. These transformations involved tokenization and the removal of various elements such as white spaces, punctuation marks, special characters, emojis, and URLs.

XSL•FO
RenderX

To accomplish this, we used a set of established methods for Arabic text preprocessing, including the use of Farasa [26]. Farasa proved invaluable in normalizing Arabic characters, stripping away diacritics, erasing punctuation marks, and eliminating repetitive characters, collectively enhancing the quality and relevance of the text data for our analysis.

Sentiment Analysis

Overview

To classify the primary sentiments expressed in Twitter messages, such as fear and joy, we used sentiment analysis, an NLP technique [27]. Our approach involved deploying the RoBERTa-base model, meticulously trained on a vast corpus of approximately 58 million tweets and further fine-tuned for precise emotion recognition leveraging the TweetEval benchmark [28]. This specific model, known as Twitter-RoBERTa-Base-Emotion [29], has been purposefully tailored for the nuanced task of emotion recognition within Twitter text data. It adeptly classifies text into various emotion categories, including joy, sadness, anger, fear, surprise, disgust, anticipation, and trust. Our sentiment analysis process unfolded in a sequence of four distinct steps, described in the following sections.

Step 1: Translation to English

As a reliable Arabic emotion detection API was not readily available, we initiated the process by translating Arabic tweets to English. To accomplish this, we leveraged the Google Translation API. We established an account and procured the necessary translation service. It is worth noting that the cost associated with using the Google Translation API amounts to US $20 per 1 million characters. Given that we were dealing with a substantial volume of data, encompassing 5.1 million Arabic tweets with a staggering 970,801,329 characters, the estimated cost tallied up to US $19,420. Consequently, we opted to translate 1 entire month of tweets. March was selected as the ideal candidate for translation, primarily due to its status as the month with the highest tweet volume. In addition, March witnessed several pivotal events, including Trump’s declaration of COVID-19 as a national emergency, the implementation of travel bans on non-US citizens traveling from Europe, and the World Health Organization’s formal declaration of the coronavirus as a global pandemic. To verify the quality of the translations, a sample of 5000 tweets was randomly selected and evaluated both before and after translation. Bilingual experts reviewed these tweets, comparing the original Arabic content with the translated English text. This review process focused on ensuring that the translations accurately conveyed the original meaning, context, and sentiment. On the basis of their feedback, we confirmed that the translations were of high quality, making them suitable for further analysis.

Step 2: English Text Preprocessing

Once the translation was complete, we embarked on preprocessing the English text. This entailed removing common stop words such as “and,” “the,” and “to.”

Step 3: Stemming

To further refine the text data, we applied a stemming process, which involves eliminating predefined prefixes and suffixes. This step aids in reducing words to their root form. For instance, it transforms “running” into “run” through stemming.

Step 4: Emotion Determination

The final step involved determining the emotion expressed in the tweets using Twitter-RoBERTa-Base-Emotion.

Table 1 illustrates the distribution of emotions across the analyzed tweets, providing valuable insights into the prevailing sentiments during the specified time frame.

Table 1. Number of tweets per emotion.

Emotion	Tweets, n
Anger	182,105
Disgust	150,022
Joy	141,446
Anticipation	60,449
Sadness	44,591
Surprise	30,666
Fear	28,439

Topic Modeling Using LDA

In our analysis, we harnessed the power of LDA as a formidable tool for uncovering latent topics within our extensive dataset. LDA, a generative probabilistic model, proves exceptionally useful for extracting these hidden themes from a vast collection of documents. Its underlying mechanism involves representing documents as random combinations of latent topics and characterizing each topic as a distribution of words [30]. This framework of the LDA model adheres to a 3-level Bayesian approach to effectively capture the generative process. However, before delving into the application of LDA or any other probabilistic topic modeling techniques, a critical step is to determine and define the number of topics often denoted as “k” [31]. This crucial decision significantly impacts the outcomes of the topic modeling process.

Qualitative Analysis

To strengthen the reliability of our findings obtained through the LDA model, we integrated a qualitative method focused on gaining a more profound insight into the identified themes. In particular, we followed the established 6-step thematic analysis

XSL•FO
RenderX

framework outlined by Braun and Clarke [32] and successfully used by Xue et al [33]. This framework includes the following steps: (1) familiarizing ourselves with the keyword data and reviewing the most representative tweets for each topic, (2) generating initial codes to summarize key themes, (3) searching for thematic patterns by grouping similar topics, (4) reviewing and refining these potential themes to ensure coherence and consistency, (5) defining and naming themes based on their overall significance and contribution to the research question, and (6) reporting and documenting the final themes. This process was iterative and reflexive, involving multiple rounds of discussion and reassessment. Two researchers with extensive experience in social media analysis and public health independently reviewed and documented the initial codes. These codes were then examined by 2 additional researchers to refine the themes, ensuring that they accurately captured the essence of the topics.

Ethical Considerations

This study analyzed publicly available data collected from Twitter. The dataset consisted of tweet IDs, and no personally identifiable information was included in the analysis. All tweet texts were retrieved in compliance with Twitter's terms of service. Ethics approval was not sought, as the study used publicly accessible data, ensuring that no identifiable personal information was involved. To maintain the highest ethical standards, all results are presented in aggregate, guaranteeing the anonymity and privacy of individuals represented in the dataset.

Results

Descriptive Results

A total of 637,718 tweets were included in the final dataset after processing raw data. The analysis focused on identifying the most frequently tweeted bigrams (pairs of words) related to COVID-19. Bigrams are 2 consecutive words, regardless of their grammar structure or semantic meaning. They may not be self-explanatory, as in the case of the bigram "social distancing," which does not convey the meaning of either word on its own. Such an approach was adopted by Xue et al [4], and it was proved that bigrams can be a useful way to identify the most prominent topics and themes in Twitter conversations. The identified bigrams included pairs of words such as "virus corona," "stay home," "home order," "travel curfew," "new coronavirus," "spread virus," "home quarantine," "health quarantine," "coronavirus pandemic," "new infected," and "new case." Among the popular unigrams were words such as "coronavirus," "virus," "home," "new," "health," "world," "visit," "pandemic," "stay," "case," "quarantine," and "curfew." Most common unigrams and bigrams related to COVID-19, and pertinent details are listed in Table 2 (original Arabic tweets are provided in Multimedia Appendix 1).

XSL•FO
RenderX

Table 2. Top 50 unigrams and bigrams and their distributions.

	Values (%)
Top 50 unigrams	
Coronavirus	6.558451
Virus	2.350919
Home	0.921041
New	0.857981
Health	0.614924
Kuwait	0.576566
Condition	0.551307
Saudi Arabia	0.503562
World	0.491143
Country	0.487031
Visit	0.392251
Pandemic	0.391468
Curfew	0.359459
Stay	0.359077
Country	0.352204
Spread	0.34872
Infected	0.340486
Quarantine	0.339662
Case	0.335292
Disease	0.331376
Infected	0.328934
Urgent	0.314949
Egypt	0.313753
Virus	0.288958
People	0.272675
Minister	0.263771
People	0.257506
Health	0.244108
China	0.243201
Good	0.241965
Travel	0.241181
Citizen	0.239945
COVID	0.238966
King	0.238255
New	0.220993
Procedure	0.213274
Lebanon	0.211883
Wanted	0.209183
Confrontation	0.205782
Education	0.205174
In	0.198331

XSL•FO
RenderX

	Values (%)
Infection	0.193302
Thanks	0.187623
Announced	0.186263
Prevention	0.185222
Nation	0.184861
Iran	0.180255
House	0.178111
Italy	0.174504
In house	0.172979
Top 50 bigrams	
Virus, coronavirus	2.029932
Coronavirus, new	0.526419
Stay home	0.325347
Coronavirus, coronavirus	0.302665
Visit, health	0.263658
Virus, coronavirus	0.19593
Coronavirus, Kuwait	0.194992
Coronavirus, new	0.192446
Curfew, travel	0.18009
Spread, virus	0.155542
Coronavirus, virus	0.146133
Quarantine, home	0.138868
Quarantine, health	0.123512
New, virus	0.122492
Coronavirus, Lebanon	0.108992
Pandemic, coronavirus	0.108868
Home, coronavirus	0.107683
Coronavirus, Saudi Arabia	0.105704
Coronavirus, Egypt	0.103818
Infected, virus	0.102376
New, case	0.09342
Coronavirus, COVID	0.091503
Kuwait, coronavirus	0.089236
New, coronavirus	0.088587
Health, global	0.08464
Stay, home	0.083898
Minister, health	0.083743
Crisis, coronavirus	0.083589
Coronavirus, stay	0.076416
Organizer, health	0.073128
Confrontation, coronavirus	0.068563
Condition, in	0.06845
Saudi Arabia, coronavirus	0.064812

XSL•FO
RenderX

	Values (%)
Coronavirus, wanted	0.061967
Coronavirus, urgent	0.060535
Recording, case	0.055537
Confrontation, virus	0.054918
Spread, virus	0.053424
Spread, coronavirus	0.053187
Coronavirus, curfew	0.050755
Curfew, curfew	0.04958
Procedure, precautionary	0.049426
United, State	0.048818
Staying, home	0.048519
Disease, coronavirus	0.047993
Infected, coronavirus	0.047849
Citizen, resident	0.047684
Servant, holy mosque	0.04552
Prevention, travel	0.045458
Coronavirus, visit	0.044582

COVID-19–Related Topics

In our study, we used the LDA technique to identify and categorize frequently co-occurring words associated with COVID-19. The LDA algorithm allowed us to manually determine the number of topics we wanted to generate. In this study, we used 2 widely recognized metrics, CaoJuan2009 and Deveaud2014, available through the R package (R Foundation for Statistical Computing), to determine the optimal number of topics for our dataset. These metrics provided a robust framework for evaluating the coherence and distinctiveness of the topics, ensuring that the final model best captured the underlying structure of the data. The CaoJuan2009 measure is minimized when the number of topics aligns with the data's intrinsic structure, while the Deveaud2014 measure is maximized to indicate topic coherence and separation. These metrics were used to assess and validate the number of topics to ensure they reflect the data's diversity and relevance. By leveraging these 2 complementary metrics, we ensured that the selected number of topics provided meaningful insights and reduced the risk of overfitting. The number of topics was determined when these metrics stabilized, indicating a consistent result.

Upon evaluating the metrics, it was found that the CaoJuan2009 score converged at its minimum value with 16 topics, while the Deveaud2014 score peaked at its maximum value with the same number of topics. On the basis of this, we concluded that the optimal number of topics, denoted as "k," is 16, as shown in Figure 2.

In addition, we calculated the topic distance and visualized the intertopic relationships using a 2D plane [34]. Each circle in the plot represents a distinct topic, ranging from topic 1 to topic k. The positioning of these circles reflects the calculated distances between topics, offering a visual representation of their relationships.

It is also worth noting that cross-validation is less commonly applied in topic modeling for several reasons. These include computational challenges associated with applying cross-validation to unsupervised models, the interpretive nature of topic models, and the emphasis on qualitative coherence over predictive performance. Most studies on LDA and related techniques do not apply cross-validation, as the focus of topic modeling is on the interpretability and coherence of the topics rather than on predictive performance. Instead, topic models are typically evaluated using internal coherence and stability measures, such as the CaoJuan2009 and Deveaud2014 metrics, which prioritize the coherence of the topics and the consistency of the results across multiple runs. This approach is consistent with what is found in most related work on LDA. For example, Blei et al [30] introduced LDA and highlighted that the evaluation of topic models is traditionally done using measures such as coherence scores.

In Table 3 (original Arabic tweets are provided in Multimedia Appendix 2), we present the findings of the 16 LDA topics, revealing the most frequently occurring words within each topic along with the percentage of tweets falling under each respective topic. Among all 16 topics, topic 5 stands out with the highest percentage (9.98%) of tweets associated with it. In topic 5, we observed a significant co-occurrence of specific words, including "coronavirus," "increase," "health," "new," "infected," "death," "recovery," and "case." This combination of words indicates an escalation in the number of COVID-19 infections, leading to unfortunate fatalities and the emergence of new cases. Moreover, the presence of the term "recovery" implies that some individuals who were previously infected are now undergoing healing and improvement. Furthermore, we

XSL•FO
RenderX

calculated the topic distance and illustrated the intertopic distance [35] in a 2D plane, as depicted in Figure 3. Each circle on the plot corresponds to a topic, ranging from topic 1 to topic 16 in this study. The positions of these circles were determined based on the calculated distances between the topics. Notably, in the visualization, the circles were not overlapping, which served as a validation of the 16 topics.

Figure 2. Metrics for estimating the optimal number of topics, ranging from 2 to 25 topics.

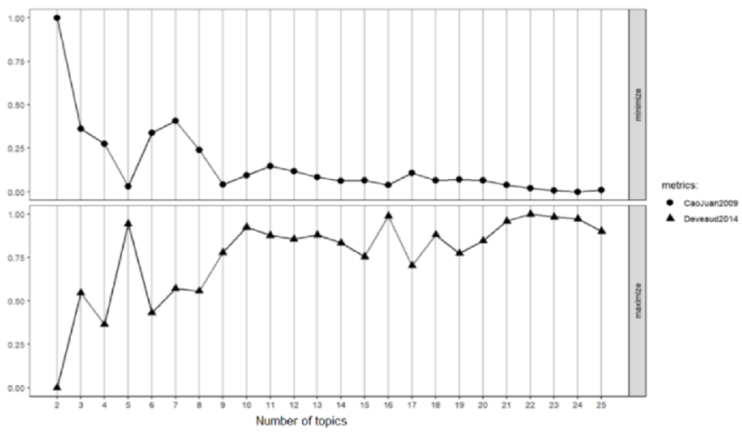
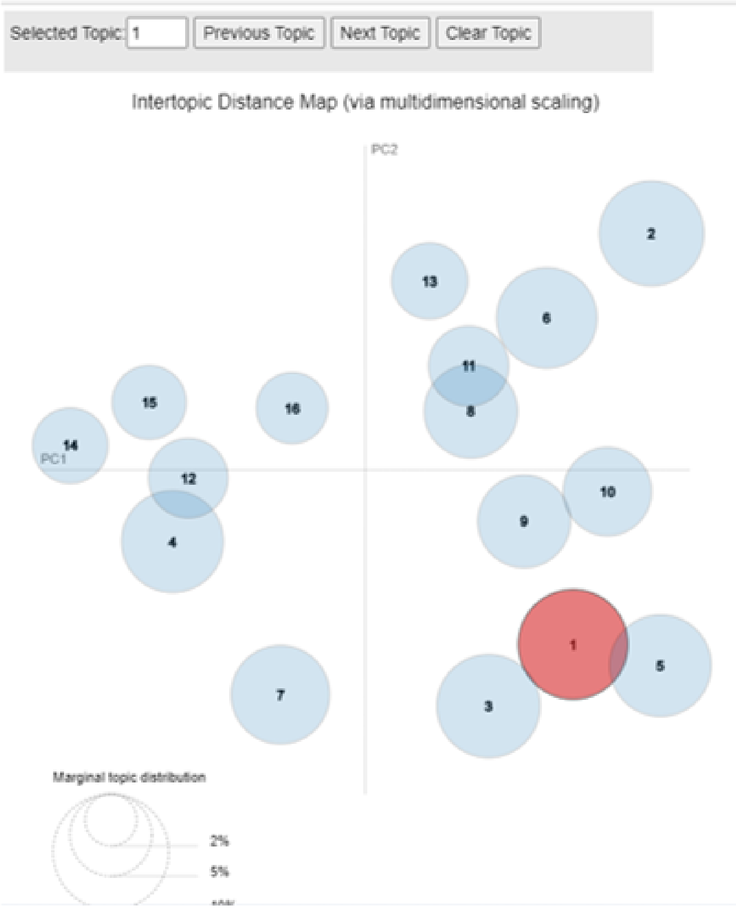


Table 3. Topic, words, and percentage of tweets.

Topic	Words	Values (%)
0	country, corona, Kuwait, praise, protection, gratitude to god, blessing, people, protect, people or nation, state, goodness, world, Saudi Arabia, Muslim, illness, thanks, pandemic, virus, Egypt	6.31
1	corona, affliction, pandemic, goodness, virus, Muslim, mercy, supplication/prayer, new, mind, world, lift or remove, great, illness, heart, raise, evil, people, mercy, Earth	8.5
2	corona, hand, virus, mask, washing, people, new, water, sanitizer, way, discount, knowledge, world, wear, person, soap, usage, glove, mask, beautiful	4.69
3	corona, virus, illness, Iran, medical, infected, hospital, doctor, treatment, Iraq, examination or test, health, person, device, hospital, Bahrain, infected, transmission, Italy, system	7.28
4	corona, virus, Kuwait, Egypt, new, emerging, COVID, health, visited, suspension, Saudi Arabia, corona, statement, Kuwaiti, confrontation, Emirate, study, crew, state, prevention	6.09
5	corona, virus, condition, new, case, infected, health, infected, died, infection, urgent, recording, death, announced, visited, increase, recovery, recorded, total, rose	9.98
6	corona, virus, education, visited, minister, confrontation, support, private, health, student, bank, spread, sector, state, responsible, crisis, communication, community, request, home	8.16
7	corona, China, state, virus, world, pandemic, union, Italy, hate, Europe, league, America, new, spread, presented, condition, European, action, player, east	4.13
8	corona, virus, house, scene, protect, country, Algeria, Egypt, died, rest, detail, video, lead, people, young man, Morocco, new, image, wanted, film	3.83
9	house, corona, stay, curfew, quarantine, wandering, home based, virus, new, Saudi Arabia, home, Kuwait, responsible, effectiveness, roaming, health, wanted, complete, goodness, Zoom	7.18
10	corona, virus, world, Trump, Oman, new, vaccine, president, faced, America, China, treatment, wanted, news, Chinese, partnership, vaccine, COVID, American, Palestine	4.68
11	corona, virus, spread, health, state, pandemic, prevention, illness, enemy, awareness, danger, way, gathering, must, home, country, avoidance, citizen, world, prevention	5.86
12	corona, virus, spread, health, state, pandemic, prevention, illness, enemy, awareness, danger, way, gathering, must, home, country, avoidance, citizen, world, prevention	5.86
13	corona, China, state, virus, world, pandemic, union, Italy, hate, Europe, league, America, new, spread, presented, condition, European, action, player, east	4.13
14	corona, Saudi Arabia, thanks, Kuwait, king, health, protection, country, homeland, virus, citizen, people or nation, visited, effort, state, sanctuary, praise, Salman, pandemic, protect	7.35
15	corona, Lebanon, people, one, age, went out, quarantine, meant, topic, condition, house, what, virus, safety, health, Egypt, people or nation, world	4.98

XSL•FO
RenderX

Figure 3. Latent Dirichlet allocation—intertopic distance.



COVID-19–Related Themes

Through the process of thematic analysis, we were able to categorize the identified topics, bigrams, and representative tweet samples into distinct themes, as shown in Table 4 (original Arabic tweets are provided in Multimedia Appendix 3).

The sample tweets provided in Table 4 are excerpts taken from the original tweets. These 16 topics have been categorized into eight overarching themes, summarized below.

1. Preventive measures and safety ("public health measures"): this theme focuses on various measures to prevent the spread of COVID-19, such as wearing masks, washing hands, using sanitizers, and practicing social distancing.
2. Medical and health care aspects: this theme encompasses topics related to the medical and health care aspects of COVID-19, including hospitals, doctors, treatments, testing, and recovery.
3. Government and social measures: this theme covers government actions, social measures, and policies implemented to address the COVID-19 pandemic, including lockdowns, travel restrictions, home orders, suspending schools, avoiding gatherings, closing shops, staying at home, and support measures.
4. Impact and numbers: this theme involves discussions about the impact of COVID-19, including the number of cases, deaths, recoveries, and updates on the situation.
5. Vaccine development and research: this theme revolves around vaccine development, clinical trials, and scientific research related to finding a solution to COVID-19.
6. COVID-19 and religious practices: this topic discusses how COVID-19 has impacted religious practices and gatherings. It mentions places of worship ( ) and the importance of adhering to prayers ( ) and religious guidelines ( ) during the COVID-19 pandemic, especially during occasions such as Ramadan ( ). The theme also includes expressions of gratitude and good wishes for nations and people (, ,).
7. Global impact of COVID-19 on sports and countries: this topic discusses the spread of COVID-19 in different countries, including China, Italy, and the United States, and its impact on various aspects, such as sports events and leagues in Europe and the Middle East. It also mentions the virus as a global pandemic and its effects on athletes and players ( ) as well as its presence in different regions around the world.
8. COVID-19 and national efforts: this theme focuses on the efforts of different nations, including Saudi Arabia and Kuwait, in combating COVID-19. It mentions leaders (, ,) and their efforts to protect the health and well-being of their citizens (,). The theme includes expressions of gratitude for the nation's efforts in managing the COVID-19 pandemic ( ) and highlights the importance of public health ( ). Textbox 1 provides a comprehensive list of topics, thoughtfully translated into English for better clarity and accessibility.

XSL•FO
RenderX

Table 4. Themes based on topic classification, bigrams, and sample tweets.

Theme and topic	Bigrams	Sample tweets
Preventive measures and safety		
Face mask	Wear mask	A note for your safety from the new coronavirus infection: Avoid social gatherings with more than 1 person. Avoid crowded areas or places where you might interact with individuals who are sick. Avoid handshakes as they are among the primary causes of virus transmission. Wear a mask whenever possible.
Hands	Wash hands, use sanitizers	Avoid gatherings, closed spaces, and crowded areas, along with regularly washing your hands with water and soap or sanitizing them with alcohol-based disinfectants. By God's will, you will be protected from contracting the new coronavirus.
Social distancing	Social distancing	Social distancing means staying away from gatherings and crowded places. If you must leave your home, maintain a distance of at least 2 meters from the people around you. Source: Cleveland Clinic, COVID-19.
Medical and health care aspects		
Health authorities	Precautionary measures, followed the instructions	Home quarantine protects against the risk of a person spreading the coronavirus without showing symptoms, making them a potential source of transmission to various groups. Preventive measures against COVID-19 ease the burden on health care providers, enabling them to fulfill their roles in treating other illnesses and performing preventive tasks, including COVID-19 detection. Voice of the physician.
Recovery	Case recovery	Breaking: The Ministry of Health announces the recovery of the first coronavirus case in the kingdom. This concerns the young man who returned from Italy and was previously announced as the first imported case of the virus in Morocco. COVID-19, Morocco, Recovery, Ministry of Health.
Treatment	Treating the infected	The Minister of Health announces the initiation of treating patients with COVID-19 with the chloroquine vaccine.
Treatment	New drug	The *Washington Post* reports that Chinese experts and physicians have successfully fought COVID-19 using chloroquine, a drug primarily used to treat malaria, and Kaletra, an HIV medication that combines lopinavir and ritonavir. Emirati physician Omar Al Hammadi shares the success of this trial.
Hospital	Field hospital	Starting Sunday, a physician will accompany every ambulance, and a field hospital will be established inside the trade unions complex. Dr Ali Al-Abous, President of the Jordanian Medical Association, comments on the nationwide curfew in Jordan due to the COVID-19 pandemic.
Government and social measures		
Lockdowns and suspending	Closing shops, suspending schools	Precautionary measures in Kuwait against COVID-19: suspension of studies and work, cancellation of weddings, closure of mosques, closure of malls, closure of salons, partial curfew, extension of the suspension of studies, regulation of work in central markets, closure of shops, postponement of installments.
Travel restrictions	Travel ban	Saudi Arabia: Saudi Arabia suspended studies, banned cafes and shisha, prohibited sports gatherings and cinemas, halted entertainment activities, stopped Umrah and travel, and conducted intensive testing to search for patients. All for your benefit—help your government overcome these circumstances with minimal losses.
Home orders	Stay home	Stay home and protect your family from coronavirus. Prevention guidelines. Stay home.
Curfew	Curfew	Breaking: Al Jazeera correspondent reports the sounding of alarm sirens across Jordan as the nationwide curfew begins to combat the spread of COVID-19.
Remote	Remote work	It is everyone's duty to follow the precautionary measures taken by our government, may God protect them, to prevent the spread of COVID-19. At our facility, we have informed the success team to work remotely from their homes until further notice.
Impact and numbers		
New cases	Confirmed cases, increase in cases	The Kuwaiti Ministry of Health has reported new cases of the novel coronavirus, and the total number of patients that have exited quarantine is 20.
Deaths	Coronavirus deaths	A new death has been recorded in Jordan due to COVID-19, bringing the total number of deaths to 5.
Vaccine development and research		
Religious guidelines	Prayer, supplication	Breaking: The Senior Scholars Authority calls on everyone to adhere to the instructions, guidelines, and regulations, to fear God, and to resort to prayer and supplication. COVID-19, Saudi Arabia.

XSL•FO
RenderX

Theme and topic	Bigrams	Sample tweets
Umrah	Suspension of Umrah	It was discovered during the COVID-19 crisis that preserving life is one of the most important objectives of Sharia, and everything is subordinated to it. The suspension of Umrah and prayer in mosques reflects the greatness of Islam and the depth of Sharia's objectives.
Global impact of COVID-19 on sports and countries		
Postponement of matches	Postponement of matches	The Union of European Football Association has decided to postpone all matches scheduled for next week. Sports, COVID-19.
Italy	The situation in Italy	Terrifying numbers in Italy and Iran; a video shows the spread of the coronavirus outside China until March.
COVID-19 and national efforts		
King Salman	Royal support	King Salman bin Abdulaziz and Crown Prince Mohammed bin Salman. The Saudi Arabian Monetary Authority announces support for the private sector with 1 billion Saudi riyals to face the expected financial and economic impacts of the coronavirus.
Thanks	Government gratitude	We thank God for the blessing of Islam and the blessing of Salman. Every Saudi has the right to be proud and boast about Saudi Arabia. May God protect its government and people from all harm. Saudi Arabia. COVID-19. Stay at home.

Textbox 1. Topic and words (English translations) used in the study.

- Topic 0: country, corona, Kuwait, Hamad, preserve, Alhamdulillah, blessing, people, preserve, people, state, good, world, Saudi Arabia, Muslim, disease, thanks, epidemic, virus, and Egypt
- Topic 1: corona, calamity, epidemic, good, virus, Muslim, mercy, prayer, new, by, world, lift, great, disease, heart, raise, evil, people, mercy, and land
- Topic 2: corona, hand, virus, mask, wash, people, new, water, sanitizer, road, discount, know, world, wear, person, soap, use, gloves, mask, and beautiful
- Topic 3: corona, virus, disease, Iran, medical, infected, hospital, doctor, treatment, Iraq, test, health, person, device, hospital, Bahrain, infected, transfer, Italy, and system
- Topic 4: corona, virus, Kuwait, Egypt, new, novel, Covid, health, visit, suspension, Saudi Arabia, core, statement, Kuwaiti, confront, Emirate, study, cure, country, and protection
- Topic 5: corona, virus, condition, new, condition, infected, health, infected, and, infection, urgent, registration, death, announce, visit, rise, recovery, register, total, and rise
- Topic 6: corona, virus, education, visit, minister, confront, support, special, health, student, bank, publish, sector, state, official, crisis, contact, community, request, house
- Topic 7: corona, China, country, virus, world, epidemic, union, Italy, football, Europe, league, America, new, spread, foot, player, and east
- Topic 8: corona, virus, home, scene, protect, country, Algeria, Egypt, die, wind, detail, video, top, people, young, Morocco, new, picture, wanted, and film
- Topic 9: home, corona, stay, ban, quarantine, circulation, homely, virus, new, Saudi Arabia, home, Kuwait, official, activity, circulation, health, wanted, complete, good, and old
- Topic 10: corona, virus, world, Trump, Oman, new, vaccine, president, confront, America, China, treatment, wanted, news, Chinese, company, vaccine, coveted, American, and Palestine
- Topic 11: corona, virus, spread, health, state, epidemic, protection, disease, enemy, awareness, threat, road, gathering, mandatory, country, avoid, citizen, world, and protection
- Topic 12: corona, mosque, people, gathering, prayer, congregation, Lebanon, condition, Ramadan, virus, prayer, I mean, talk, cover, world, Egypt, great, good, people, and peace
- Topic 13: corona, virus, procedure, spread, prevention, decision, sanitization, closure, local, logic, visit, urgent, Saudi Arabia, new, governor, application, shop, Riyadh, precautionary, and system
- Topic 14: corona, Saudi Arabia, thanks, Kuwait, king, health, preserve, country, homeland, virus, citizen, people, visit, effort, state, crisis, blessing, Salman, epidemic, and preserve
- Topic 15: corona, Lebanon, people, and, age, came out, quarantine, from me, subject, condition, house, and, mean, virus, peace, health, Egypt, people, world, and damn

XSL•FO
RenderX

Sentiment Analysis

We conducted sentiment analysis for each of the 16 topics and presented the results in Figure 4 and Table 5. Figure 4 visualized 7 emotions: anger, disgust, joy, anticipation, sadness, surprise, and fear. Across all 16 topics, anger (represented by the red line) was the dominant emotion in 16 topics, followed by disgust (green line), joy (blue line), and anticipation (orange line). To delve deeper into the emotional aspects of the data, we provide a breakdown of the number of tweets associated with each emotion across different topics in Table 5. For example, in topic 5, a substantial number of tweets (n=17,848) expressed anger, reflecting a strong sentiment regarding the need for essential measures and precautions. This high prevalence of anger in topic 5 stands out in comparison to the other topics. It is worth noting that excessive anger, if left unmanaged, can lead to a range of medical problems. Managing emotions such as anger is crucial not only for mental well-being but also for overall physical health.

Figure 4. Sentiment analysis for each of the 16 latent topics.

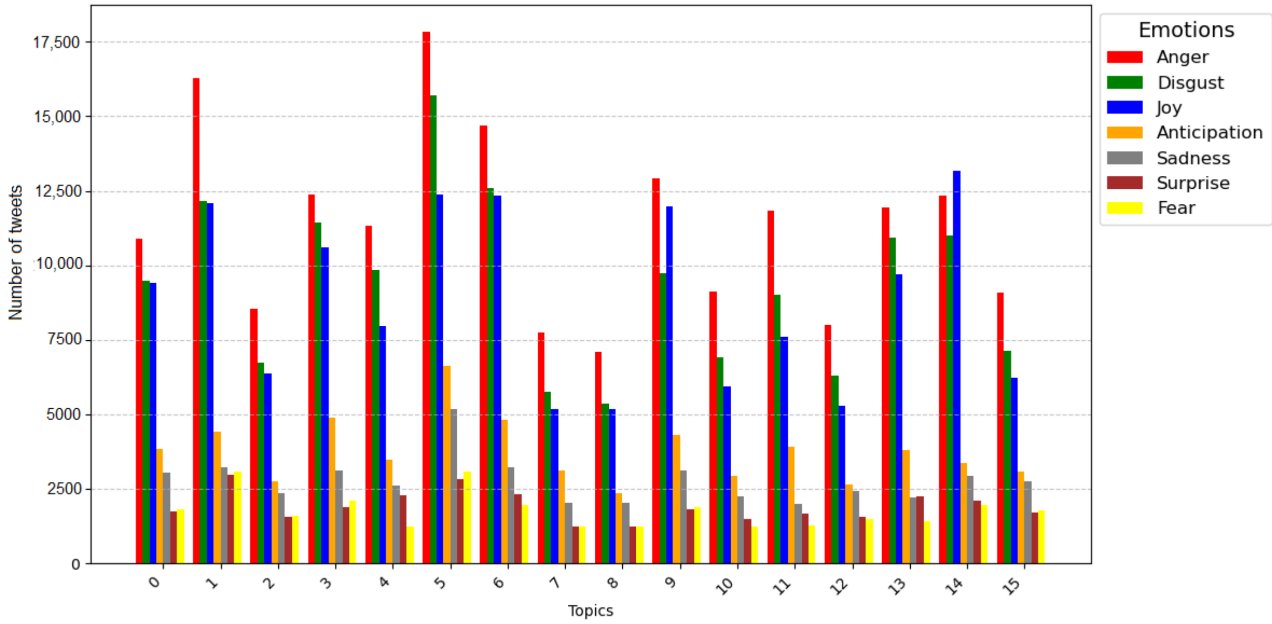


Table 5. The number of tweets for 7 emotions across 16 topics.

Topic	Anger	Anticipation	Disgust	Fear	Joy	Sadness	Surprise
0	10,896	3834	9475	1835	9428	3040	1757
1	16,295	4434	12,164	3079	12,083	3217	2965
2	8559	2772	6729	1585	6366	2346	1554
3	12,391	4894	11,424	2110	10,602	3113	1876
4	11,345	3493	9838	1246	7969	2624	2295
5	17,848	6630	15,688	3078	12,375	5192	2847
6	14,702	4820	12,601	1979	12,329	3239	2342
7	7757	3130	5752	1224	5189	2031	1230
8	7082	2367	5344	1222	5168	2035	1228
9	12,926	4321	9742	1896	11,989	3132	1802
10	9113	2948	6901	1225	5953	2245	1486
11	11,836	3915	9027	1288	7606	2015	1656
12	7988	2646	6286	1494	5282	2416	1547
13	11,942	3801	10,928	1437	9717	2231	2241
14	12,328	3357	11,007	1962	13,172	2953	2124
15	9097	3087	7116	1779	6218	2762	1716

XSL•FO
RenderX

Discussion

Principal Findings

This study delved into public discussion and emotional expressions related to COVID-19 using Arabic Twitter messages. Twitter users engaged in discussions encompassing 8 primary themes regarding COVID-19. Using topic modeling on the tweets proved valuable in uncovering insights into COVID-19–related topics and concerns. The outcomes highlighted several crucial observations.

This analysis concentrates on tweets from March 2020, a pivotal phase in the COVID-19 pandemic's unfolding narrative. During this period, the second stage of the COVID-19 pandemic emerged prominently, marked by a significant milestone as Arabic countries reported their initial cases of COVID-19. Subsequently, a cascade of vital health measures ensued, encompassing the enforcement of quarantine protocols, the temporary cessation of air travel, and the inevitable postponement or cancelation of various events. This time frame aligns logically with the peak frequency of tweets, as previously observed by Taneja et al [22] and Haouari et al [34].

Amidst the array of all 16 topics, a discernible pattern surfaced, characterized by the recurring presence of specific keywords such as "coronavirus," "increase," "health," "new," "infected," "death," "recovery," and "case." This linguistic cluster strongly implies a surge in COVID-19 infections, accompanied by lamentable loss of life and the emergence of new cases during the ongoing COVID-19 pandemic. It is imperative to emphasize that our chosen time frame aligns precisely with the onset of the COVID-19 pandemic's second phase, coinciding with heightened global concern. The substantial spike in COVID-19 cases in Italy during this period ignited a profound sense of alarm on a global scale. This surge in worldwide apprehension may have contributed to the observed increase in tweet frequency, corroborating findings from multiple studies [22,34].

Furthermore, substantial discussions revolving around the COVID-19 pandemic within diverse Arabic nations have drawn significant interest. These conversations are marked by a prevailing sense of indignation. Moreover, public sentiments concerning the spread of COVID-19 unveiled an underlying sense of anticipation toward prospective measures. These sentiments were accompanied by a mix of emotions, including anger and fear; a notable undercurrent of fear was predominant in discussions revolving around the COVID-19 crisis and the resulting fatalities. This trend aligns with global sentiments, as documented by Lwin et al [36], wherein public emotions underwent a noticeable shift from fear to anger throughout the COVID-19 pandemic, with traces of sadness and joy also emerging.

Noteworthy, the appearance of dialogues concerning COVID-19 and religious practices introduced a fresh subject not previously detected in prior research. This indicates a developing connection between COVID-19 and religious matters on the Twitter platform. This is particularly apparent due to the substantial influence of religious identity on attitudes and actions concerning the COVID-19 pandemic and vaccination efforts; the COVID-19 pandemic has significantly reshaped communal worship and gatherings as measures to curb the virus's transmission [37]. Furthermore, religious leaders have assumed a central role in championing COVID-19 vaccination campaigns, effectively addressing and mitigating vaccine hesitancy [38].

In-Depth Analysis of Findings

The application of topic modeling and sentiment analysis in this study provided several valuable insights into public sentiment and thematic discussions during the early stages of the COVID-19 pandemic in Arab regions. The findings largely align with anticipated outcomes, such as the focus on preventive measures and safety and medical and health care aspects, both of which were expected topics given the nature of the COVID-19 pandemic.

However, the emergence of discussions on COVID-19 and religious practices was a unique finding that adds depth to the understanding of public discourse in Arab communities. This theme highlights the intersection of the COVID-19 pandemic with cultural and religious practices, which had not been as thoroughly explored in previous research. It underscores the significant impact that COVID-19 had on religious identity, communal worship, and adherence to religious guidelines during pivotal periods such as Ramadan.

Another notable aspect was the attention given to the global impact of COVID-19 on sports and countries, reflecting the broad international concern and how global events, especially sports, were affected. This indicates that the COVID-19 pandemic's influence went beyond public health and extended into societal and cultural dimensions, impacting activities that are deeply integrated into daily life.

In addition, the sentiment analysis revealed a nuanced distribution of emotions, with a significant proportion of tweets expressing anger and disgust, as expected, given the uncertainty surrounding the COVID-19 pandemic. However, there was also a notable presence of positive emotions, such as hope and solidarity, particularly in tweets discussing community support and coping mechanisms. This suggests that, despite the overwhelming nature of the crisis, many users turned to social media not only to express negative emotions but also to share supportive messages and encourage others.

Overall, the identified themes and their respective discussions provide a comprehensive view of public sentiment, concerns, and priorities during the early COVID-19 pandemic period. These insights not only reflect the immediate response to the health crisis but also highlight the diverse and context-specific aspects that shaped public discourse. Such findings offer a foundation for more effective public health communication and intervention strategies, particularly in culturally sensitive contexts.

Strengths

This study provided valuable insights into the sentiments and concerns of Arabic-speaking Twitter users during the COVID-19 pandemic, underscoring the significance of social media as a means of understanding and addressing public health issues in the digital era. First, the analysis encompassed a substantial

XSL•FO
RenderX

dataset of 1 million Arabic tweets, offering a comprehensive view of the sentiments and topics expressed by Twitter users in Arabic-speaking countries during a specific period of the COVID-19 pandemic. Besides, the study used a combination of machine learning techniques, including topic modeling and sentiment analysis, to uncover and categorize themes and emotions within the dataset, providing a holistic understanding of the data. By identifying and categorizing 16 conversation topics into 8 themes, the study offered a structured view of the discussions surrounding COVID-19 in the Arab region, making it easier to interpret and use the findings. Finally, the inclusion of emotion analysis adds depth to the study, revealing how Twitter users in the Arab world emotionally responded to various aspects of the COVID-19 pandemic.

Limitations

First, at the forefront of our approach, we meticulously aimed to unravel the complexities embedded within the COVID-19 pandemic's second phase. Our focus was sharp and exclusive, centered on harnessing tweets originating exclusively from March 2020. The motivation behind this specific time frame stemmed from our intention to subject translated tweets to a comprehensive sentiment analysis. This intricate process relied upon the Google API translation service, which, although effective, is accompanied by a substantial cost factor. The financial implication associated with translating the entirety of the datasets using this service was a noteworthy consideration that prompted us to make strategic choices in our analysis approach.

Second, it is crucial to recognize that Arabic is a linguistically intricate language characterized by a rich array of dialects and intricate cultural nuances. These unique linguistic qualities can present substantial challenges for automated sentiment analysis tools. While we attempted to apply automated sentiment analysis to Arabic tweets, we encountered difficulties in precisely capturing the subtleties of emotions. Automated tools often grappled with interpreting nuanced sentiments, such as sarcasm, irony, and contextual shifts in sentiment that frequently permeate social media conversations.

Third, a strategic decision was made to exclude non-Arabic tweets from our analyses. As a result, our findings were inherently confined to users who exclusively communicated in Arabic. It is essential to underscore that the fundamental objective of our research revolves around gaining insights into the opinions and reactions of Arabic countries in relation to COVID-19.

Furthermore, while our study leveraged social media data as a proxy for public sentiment, it is essential to recognize the inherent biases associated with using Twitter data. For instance, social media users may not be representative of the general population, as certain demographics might be underrepresented on platforms such as Twitter. A study by Padilla et al [39] has shown that social media content can be biased based on whether individuals are local residents or visitors and the types of activities they engage in throughout the day. Similarly, Gore et al [40] highlighted that the sentiment of tweets is often correlated with the geographical area in which they were composed, suggesting that local context and specific events may have a significant impact on sentiment analysis results. Frank et al [41] also found that emotional expressions, such as happiness, vary significantly by location, further reinforcing the influence of geographic factors on sentiment.

In addition, it is plausible that individual personality traits or political affiliations, as suggested by Auer and Elena [42], could influence whether a user expresses positive or negative sentiments. This raises an open question about the extent to which sentiment reflects variance in psychological traits versus the situational context in which those traits are expressed. These factors could contribute to biases in our dataset and should be considered as potential sources of influence on the study's outcomes.

Future Work

Regarding future studies focusing on COVID-19, first, there arises a noteworthy avenue for exploration comparing the sentiments and opinions of Arabic-speaking populations with those of individuals expressing themselves in other languages. A comprehensive approach might encompass languages such as English, Italian, French, German, and Spanish. Such comparative analyses have the potential to yield valuable insights into the cross-linguistic dynamics of perceptions and responses to the COVID-19 pandemic.

Second, another promising avenue for future research involves conducting a comparative analysis between sentiment analysis using human-labeled data and automated tools specifically tailored for Arabic languages. This comparative study should aim to ascertain the feasibility of leveraging these automated tools as an alternative to translation APIs. By meticulously comparing the results obtained from human-labeled sentiment analysis and those generated by automated tools, researchers can gauge the efficacy, accuracy, and reliability of automated sentiment analysis for Arabic tweets. The outcomes of this research hold the potential for far-reaching implications, potentially presenting a cost-effective and streamlined avenue for sentiment analysis that eliminates the reliance on costly translation APIs.

By providing an accurate and efficient mechanism for measuring sentiments in Arabic tweets, researchers and mental health professionals could identify patterns of emotional distress or psychological well-being. This could be especially pertinent during times of crises, enabling timely interventions and support for individuals experiencing heightened emotional responses. Importantly, the ability to effectively harness sentiment analysis for understanding emotional states has the potential to empower the broader field of mental health research and intervention as well as enhance our understanding of collective emotional dynamics within Arabic-speaking communities.

Third, there is an imminent need for research to unravel the stem of fabricated tweets that emerge during a pandemic. Given that Twitter users experience a heightened sense of fear, which might be exacerbated by the proliferation of misinformation, it becomes a critical endeavor to investigate the prevalence and impact of false tweets. Subsequent studies could significantly benefit from spotlighting the issue of misinformation, with a specific focus on understanding how government officials and

XSL•FO
RenderX

international organizations can effectively manage the dissemination of deceptive messages targeting the public. By comprehensively addressing the challenges posed by misleading content, we can enhance our collective understanding of navigating information dissemination during such critical periods.

Conclusions

This study delves deep into the intricate web of topics and emotions found in Arabic tweets about COVID-19. It highlights how platforms such as Twitter, especially during times of global change, are crucial for capturing the diverse feelings and concerns of Arabic speakers. Through a mix of topic modeling and sentiment analysis, we revealed the basic human emotions in user responses to COVID-19 tweets from March 2020.

We used 2 methods together: topic modeling (specifically LDA) and sentiment analysis tools. These helped us uncover the main themes and feelings within the tweets. Anger was the prominent emotion tied to COVID-19 topics, accompanied by other emotions. Joy was linked to vaccine and education discussions, while authority and politics stirred up anger. Sadness emerged from topics about cases, deaths, and the impacts on families and mental health.

This study connects social media, emotions, and the global scene. It sheds light on the emotional layers of digital conversations, offering insights into COVID-19–related tweets. These findings guide better communication strategies and compassionate responses, strengthening our collective resilience in the face of challenges.

Moreover, the results and workflow of this study present actionable insights for the medical and public health communities. By integrating our findings into official government documentation or public health research, authorities can tailor their communication strategies based on public concerns and emotions. This, in turn, helps in shaping more effective educational campaigns and policy interventions. Our methodology also serves as a robust tool for continuous monitoring of public sentiment in real time, allowing policy makers to stay informed and adapt their strategies accordingly. This approach ensures that responses are not only timely but also grounded in the actual sentiments and needs of the population.

Acknowledgments

The authors would like to acknowledge the deanship of research at the Jordan University of Science and Technology for supporting this research (grant 653/2023).

Data Availability

The data sets generated and analyzed during this study are available in the GitHub repository [43].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Original Arabic versions of tweets shown in Table 2.
[DOCX File , 20 KB - infodemiology_v5i1e53434_app1.docx]

Multimedia Appendix 2
Original Arabic versions of tweets shown in Table 3.
[DOCX File , 18 KB - infodemiology_v5i1e53434_app2.docx]

Multimedia Appendix 3
Original Arabic versions of tweets shown in Table 4.
[DOCX File , 25 KB - infodemiology_v5i1e53434_app3.docx]

References

1. Hasan MN, Islam MA, Sangkham S, Werkneh AA, Hossen F, Haque MA, et al. Insight into vaccination and meteorological factors on daily COVID-19 cases and mortality in Bangladesh. Groundw Sustain Dev 2023 May;21:100932 [FREE Full text] [doi: 10.1016/j.gsd.2023.100932] [Medline: 36945723]
2. Rosenberg H, Syed S, Rezaie S. The Twitter pandemic: the critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. CJEM 2020 Jul 06;22(4):418-421 [FREE Full text] [doi: 10.1017/cem.2020.361] [Medline: 32248871]
3. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]

XSL•FO
RenderX

4. Xue J, Chen J, Gelles R. Using data mining techniques to examine domestic violence topics on Twitter. Violence Gend 2019 Jun;6(2):105-114. [doi: 10.1089/vio.2017.0066]
5. Signorini A, Segre AM, Polgreen PM. The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. PLoS One 2011 May 04;6(5):e19467 [FREE Full text] [doi: 10.1371/journal.pone.0019467] [Medline: 21573238]
6. Rufai S, Bunce C. World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. J Public Health (Oxf) 2020 Aug 18;42(3):510-516 [FREE Full text] [doi: 10.1093/pubmed/fdaa049] [Medline: 32309854]
7. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: using latent Dirichlet allocation for topic modeling on Twitter. PLoS One 2020 Sep 25;15(9):e0239441 [FREE Full text] [doi: 10.1371/journal.pone.0239441] [Medline: 32976519]
8. Alharbi NH, Alkhateeb JH. Sentiment analysis of Arabic tweets related to COVID-19 using deep neural network. In: Proceedings of the 2021 International Congress of Advanced Technology and Engineering. 2021 Presented at: ICOTEN '21; July 4-5, 2021; Taiz, Yemen p. 1-11. [doi: 10.1109/icoten52080.2021.9493467]
9. Alhumoud S, Al Wazrah A, Alhussain L, Alrushud L, Aldosari A, Altammami RM, et al. ASAVACT: Arabic sentiment analysis for vaccine-related COVID-19 tweets using deep learning. PeerJ Comput Sci 2023;9:e1507 [FREE Full text] [doi: 10.7717/peerj-cs.1507] [Medline: 38077546]
10. Ali M. Arabic sentiment analysis about online learning to mitigate COVID-19. Int J Intell Syst 2021;30(1):524-540. [doi: 10.1515/jisys-2020-0115]
11. Alqarni A, Rahman A. Arabic tweets-based sentiment analysis to investigate the impact of COVID-19 in KSA: a deep learning approach. Big Data Cogn Comput 2023 Jan 13;7(1):16 [FREE Full text] [doi: 10.3390/bdcc7010016]
12. AlAgha I. Topic modeling and sentiment analysis of twitter discussions on COVID-19 from spatial and temporal perspectives. J Inf Sci Theory Pract 2021;9(1):35-53 [FREE Full text] [doi: 10.1633/JISTAP.2021.9.1.3]
13. Habbat N, Anoun H, Hassouni L. Sentiment analysis and topic modeling on Arabic Twitter data during COVID-19 pandemic. Indones J Innov Appl Sci 2022;2(1):60-67 [FREE Full text] [doi: 10.47540/ijias.v2i1.432]
14. Tamang S, Humbert-Droz M, Gianfrancesco M, Izadi Z, Schmajuk G, Yazdany J. Practical considerations for developing clinical natural language processing systems for population health management and measurement. JMIR Med Inform 2023 Jan 03;11:e37805 [FREE Full text] [doi: 10.2196/37805] [Medline: 36595345]
15. Mahmoud E, Emilien A, Maxime G, Gilles D. The role of text analytics in healthcare: a review of recent developments and applications. In: Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies. 2021 Presented at: BIOSTEC '21; February 11-13, 2021; Virtual Event p. 825-832. [doi: 10.5220/0010414508250832]
16. Amores JJ, Blanco-Herrero D, Arcila-Calderón C. The conversation around COVID-19 on Twitter—sentiment analysis and topic modelling to analyse tweets published in English during the first wave of the pandemic. Journal Media 2023 Mar 30;4(2):467-484. [doi: 10.3390/journalmedia4020030]
17. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. J Med Internet Res 2020 Apr 21;22(4):e19016 [FREE Full text] [doi: 10.2196/19016] [Medline: 32287039]
18. Aslam F, Awan TM, Syed JH, Kashif A, Parveen M. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. Humanit Soc Sci Commun 2020 Jul 08;7(1):13. [doi: 10.1057/S41599-020-0523-3]
19. Kwok SW, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian twitter users: machine learning analysis. J Med Internet Res 2021 May 19;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: 33886492]
20. Park S, Suh YK. A comprehensive analysis of COVID-19 vaccine discourse by vaccine brand on twitter in Korea: topic and sentiment analysis. J Med Internet Res 2023 Jan 31;25:e42623 [FREE Full text] [doi: 10.2196/42623] [Medline: 36603153]
21. Wang Y, Chen Y. Characterizing discourses about COVID-19 vaccines on Twitter: a topic modeling and sentiment analysis approach. J Commun Healthc 2023 Mar 24;16(1):103-112. [doi: 10.1080/17538068.2022.2054196] [Medline: 36919802]
22. Taneja SL, Passi M, Bhattacharya S, Schueler SA, Gurram S, Koh C. Social media and research publication activity during early stages of the COVID-19 pandemic: longitudinal trend analysis. J Med Internet Res 2021 Jun 17;23(6):e26956 [FREE Full text] [doi: 10.2196/26956] [Medline: 33974550]
23. Qazi U, Imran M, Ofli F. GeoCoV19: a dataset of hundreds of millions of multilingual COVID-19 tweets with location information. SIGSPATIAL Special 2020 Jul 08;12(1):6-15. [doi: 10.1145/3404820.3404823]
24. A command line tool (and Python library) for archiving Twitter JSON. GitHub. URL: https://github.com/DocNow/twarc [accessed 2024-04-29]
25. Alshanik F, Apon A, Herzog A, Safro I, Sybrandt J. Accelerating text mining using domain-specific stop word lists. In: Proceedings of the 2020 IEEE International Conference on Big Data. 2020 Presented at: Big Data '20; December 10-13, 2020; Atlanta, GA p. 2639-2648 URL: https://ieeexplore.ieee.org/document/9378226 [doi: 10.1109/bigdata50022.2020.9378226]

XSL•FO
RenderX

26. Abdelali A, Darwish K, Durrani N, Mubarak H. Farasa: a fast and furious segmenter for arabic. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. 2016 Presented at: NAACL '16; June 12-17, 2016; San Diego, CA p. 11-16 URL: https://aclanthology.org/N16-3003.pdf [doi: 10.18653/v1/n16-3003]
27. Beigi G, Hu X, Maciejewski R, Liu H. An overview of sentiment analysis in social media and its applications in disaster relief. In: Pedrycz W, Chen SM, editors. Sentiment Analysis and Ontology Engineering: An Environment of Computational Intelligence. Cham, Switzerland: Springer; 2016:313-340.
28. Barbieri F, Camacho-Collados J, Neves L, Espinosa-Anke L. TWEETEVAL: unified benchmark and comparative evaluation for tweet classification. In: Proceedings of the 2020 Conference on Association for Computational Linguistics. 2020 Presented at: EMNLP '20; November 16-20, 2020; Virtual Event p. 1644-1650 URL: https://aclanthology.org/2020.findings-emnlp.148.pdf [doi: 10.18653/v1/2020.findings-emnlp.148]
29. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. Huggingface's transformers: state-of-the-art natural language processing. arXiv Preprint posted online October 9, 2019 [FREE Full text]
30. Blei DB, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:4-5 [FREE Full text] [doi: 10.1017/9781009218245.012]
31. Blei D, Carin L, Dunson D. Probabilistic topic models: a focus on graphical model design and applications to document and image analysis. IEEE Signal Process Mag 2010 Nov 01;27(6):55-65 [FREE Full text] [doi: 10.1109/MSP.2010.938079] [Medline: 25104898]
32. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
33. Xue J, Chen J, Hu R, Chen C, Zheng C, Su Y, et al. Twitter discussions and emotions about the COVID-19 pandemic: machine learning approach. J Med Internet Res 2020 Nov 25;22(11):e20550 [FREE Full text] [doi: 10.2196/20550] [Medline: 33119535]
34. Haouari F, Hasanain M, Suwaileh R, Elsayed T. ArCOV-19: the first Arabic COVID-19 twitter dataset with propagation networks. In: Proceedings of the 6th Conference on Arabic Natural Language Processing Workshop. 2021 Presented at: WANLP '21; April 19, 2021; Kyiv, Ukraine p. 82-91 URL: https://aclanthology.org/2021.wanlp-1.9.pdf
35. Chuang J, Ramage D, Manning CD, Heer J. Interpretation and trust: designing model-driven visualizations for text analysis. In: Proceedings of the 2012 SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI '12; May 5-10, 2012; Austin, TX p. 443-452 URL: https://dl.acm.org/doi/10.1145/2207676.2207738 [doi: 10.1145/2207676.2207738]
36. Lwin MO, Lu J, Sheldenkar A, Schulz PJ, Shin W, Gupta R, et al. Global sentiments surrounding the COVID-19 pandemic on twitter: analysis of Twitter trends. JMIR Public Health Surveill 2020 May 22;6(2):e19447 [FREE Full text] [doi: 10.2196/19447] [Medline: 32412418]
37. Sisti LG, Buonsenso D, Moscato U, Costanzo G, Malorni W. The role of religions in the COVID-19 pandemic: a narrative review. Int J Environ Res Public Health 2023 Jan 17;20(3):1691 [FREE Full text] [doi: 10.3390/ijerph20031691] [Medline: 36767057]
38. Chu J, Pink SL, Willer R. Religious identity cues increase vaccination intentions and trust in medical experts among American Christians. Proc Natl Acad Sci U S A 2021 Dec 07;118(49):e2106481118 [FREE Full text] [doi: 10.1073/pnas.2106481118] [Medline: 34795017]
39. Padilla JJ, Kavak H, Lynch CJ, Gore RJ, Diallo SY. Temporal and spatiotemporal investigation of tourist attraction visit sentiment on Twitter. PLoS One 2018;13(6):e0198857 [FREE Full text] [doi: 10.1371/journal.pone.0198857] [Medline: 29902270]
40. Gore RJ, Diallo S, Padilla J. You are what you tweet: connecting the geographic variation in America's obesity rate to twitter content. PLoS One 2015;10(9):e0133505 [FREE Full text] [doi: 10.1371/journal.pone.0133505] [Medline: 26332588]
41. Frank MR, Mitchell L, Dodds PS, Danforth CM. Happiness and the patterns of life: a study of geolocated tweets. Sci Rep 2013 Sep 12;3(1):2625 [FREE Full text] [doi: 10.1038/srep02625] [Medline: 24026340]
42. Auer, Elena ML. Detecting deceptive impression management behaviors in interviews using natural language processing. Old Dominion University. URL: https://digitalcommons.odu.edu/psychology_etds/70/#:~:text=Given%20this%20limitation%2C%20natural%20language,IM%20in%20mock%20employment%20interviews [accessed 2024-04-29]
43. FarahAlshanik / COVID-Paper. GitHub. URL: https://github.com/FarahAlshanik/COVID-Paper [accessed 2024-04-29]

Abbreviations

API: application programming interface
LDA: latent Dirichlet allocation
NLP: natural language processing

XSL•FO
RenderX

Edited by T Mackey; submitted 07.10.23; peer-reviewed by M Elbattah, R Gore, MO Khursheed; comments to author 06.02.24; revised version received 06.05.24; accepted 08.12.24; published 10.02.25.

Please cite as:
Alshanik F, Khasawneh R, Dalky A, Qawasmeh E
Unveiling Topics and Emotions in Arabic Tweets Surrounding the COVID-19 Pandemic: Topic Modeling and Sentiment Analysis Approach
JMIR Infodemiology 2025;5:e53434
URL: https://infodemiology.jmir.org/2025/1/e53434
doi:10.2196/53434
PMID:39928401

©Farah Alshanik, Rawand Khasawneh, Alaa Dalky, Ethar Qawasmeh. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 10.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Visualizing YouTube Commenters' Conceptions of the US Health Care System: Semantic Network Analysis Method for Evidence-Based Policy Making

Lana V Ivanitskaya[1], PhD; Elina Erzikova[2], PhD

[1]Health Administration, School of Health Sciences, The Herbert H. and Grace A. Dow College of Health Professions, Central Michigan University, Mount Pleasant, MI, United States

[2]Strategic Communication, School of Communication, Journalism and Media, College of the Arts and Media, Central Michigan University, Mount Pleasant, MI, United States

Corresponding Author:
Lana V Ivanitskaya, PhD
Health Administration, School of Health Sciences
The Herbert H. and Grace A. Dow College of Health Professions
Central Michigan University
208D Rowe Hall
Mount Pleasant, MI, 48859
United States
Phone: 1 989 774 1639
Email: ivani1sv@cmich.edu

Abstract

Background: The challenge of extracting meaningful patterns from the overwhelming noise of social media to guide decision-makers remains largely unresolved.

Objective: This study aimed to evaluate the application of a semantic network method for creating an interactive visualization of social media discourse surrounding the US health care system.

Methods: Building upon bibliometric approaches to conducting health studies, we repurposed the VOSviewer software program to analyze 179,193 YouTube comments about the US health care system. Using the overlay-enhanced semantic network method, we mapped the contents and structure of the commentary evoked by 53 YouTube videos uploaded in 2014 to 2023 by right-wing, left-wing, and centrist media outlets. The videos included newscasts, full-length documentaries, political satire, and stand-up comedy. We analyzed term co-occurrence network clusters, contextualized with custom-built information layers called overlays, and performed tests of the semantic network's robustness, representativeness, structural relevance, semantic accuracy, and usefulness for decision support. We examined how the comments mentioning 4 health system design concepts—universal health care, Medicare for All, single payer, and socialized medicine—were distributed across the network terms.

Results: Grounded in the textual data, the macrolevel network representation unveiled complex discussions about illness and wellness; health services; ideology and society; the politics of health care agendas and reforms, market regulation, and health insurance; the health care workforce; dental care; and wait times. We observed thematic alignment between the network terms, extracted from YouTube comments, and the videos that elicited these comments. Discussions about illness and wellness persisted across time, as well as international comparisons of costs of ambulances, specialist care, prescriptions, and appointment wait times. The international comparisons were linked to commentaries with a higher concentration of British-spelled words, underscoring the global nature of the US health care discussion, which attracted domestic and global YouTube commenters. Shortages of nurses, nurse burnout, and their contributing factors (eg, shift work, nurse-to-patient staffing ratios, and corporate greed) were covered in comments with many likes. Comments about universal health care had much higher use of ideological terms than comments about single-payer health systems.

Conclusions: YouTube users addressed issues of societal and policy relevance: social determinants of health, concerns for populations considered vulnerable, health equity, racism, health care quality, and access to essential health services. Versatile and applicable to health policy studies, the method presented and evaluated in our study supports evidence-based decision-making and contextualized understanding of diverse viewpoints. Interactive visualizations can help to uncover large-scale patterns and guide strategic use of analytical resources to perform qualitative research.

XSL•FO
RenderX

(JMIR Infodemiology 2025;5:e58227) doi:10.2196/58227

KEYWORDS

social media; semantic network; health system; health policy; ideology; VOSviewer; health care reform; health services; health care workforce; health insurance

Introduction

Background

The US health care system, characterized by high costs [1] and perceived to fall "far short of its potential" [2], has been a focal point for media attention and public commentary over the past decade. Discussions have revolved around topics such as the repeal of Obamacare, presidential health care agendas, the exorbitant costs of health care, comparisons to systems in other nations, and postpandemic health care personnel shortages. Throughout this period, conservative, moderate, and liberal media outlets have produced a variety of content, including newscasts, full-length documentaries, political satire, and stand-up comedy, all centered on the intricacies of the US health care system [3-6]. When disseminated through YouTube (Google Inc), the most popular platform among US social media users [7], select videos have generated millions of views and tens of thousands of comments. To the best of our knowledge, the perspectives of YouTube commenters on the US health care system and its reform, despite their considerable value for policy analysis, remain unexplored.

Objectives

Social media discussions are abundant, but they are often chaotic, noisy, indignant, and hateful [8-11]. There is a need for a method that effectively visualizes large volumes of commentary, filters out the noise, and highlights key patterns, making the information more digestible for stakeholders. The current state of social media research falls short of efficiently and clearly disseminating scientific outputs to diverse audiences. In quantitative social media studies, the constraints are statistical and graphical outputs with low idea density or high decoding requirements, which often require specialized knowledge. In qualitative studies, researchers communicate analytical outputs as summaries of themes and subthemes with representative quotes; however, they are based on limited data samples.

To address these challenges, we propose a mixed methods approach of mapping social media commentary. This approach combines automation and human judgment to create a visual representation of social media comments' contents and structure, presenting them as a semantic network [12]. This methodology is particularly relevant for researchers, policy makers, and the wider public seeking a better understanding of complex social media narratives. We repurpose VOSviewer (Centre for Science and Technology Studies at Leiden University), a user-friendly bibliometric tool, to analyze tens of thousands of social media comments on YouTube regarding the US health care system. In this study, semantic networks are graphical representations of social media comment meanings. Nodes represent terms frequently mentioned in YouTube comments, linked and grouped into clusters based on their co-occurrence.

Since their introduction in 2010, VOSviewer algorithms have been extensively applied to build term co-occurrence networks from the text of article titles and abstracts [13-20]. Visualization of nonbibliometric textual data as semantic networks in VOSviewer was proposed in 2011 [21], followed by early visualizations of Twitter and YouTube discussions ([22-25]). Subsequent explorations of VOSviewer's applications to social media comments and hashtags primarily led to cluster mapping ([26-35]). Notably, some scholars enhanced their cluster maps with informational layers called custom overlays to reveal patterns not visible in the base network [36-38].

Previous research compared VOSviewer semantic networks to networks generated from manually coded Twitter text [26]. However, there have been few systematic evaluations of VOSviewer-generated semantic networks derived from social media data. Consequently, our overarching goal is to evaluate VOSviewer's application to social media data: Can it produce credible semantic networks to be used as analytical and communication tools? We test VOSviewer's term co-occurrence map with custom-built overlays by posing 3 research questions:

1. How well does the VOSviewer network capture the content, context, and structure of social media comments?
2. What does it reveal about a decade-long online public discussion of the US health care system?
3. What is the policy analysis value of VOSviewer visualizations?

Methods

Semantic Network Construction

VOSviewer generates a custom semantic network by processing a corpus text file featuring social media comments. Our corpus comprised the text of primary comments and first-level replies to 53 videos shared by 17 US-based media outlets on their respective YouTube platforms between 2014 and 2023. The videos were sourced from news outlets such as Consumer News and Business Channel, Cable News Network, Fox News, and Public Broadcasting Service Frontline. Detailed criteria for video selection and video characteristics are outlined in the Tables S1 and S2 in Multimedia Appendix 1 [39]. After eliminating 5575 duplicate comments from the initial dataset of primary comments at first-level responses, our final corpus encompassed a total of 179,193 unique comments.

VOSviewer processes YouTube comments by detecting sentences, applying the Apache Software Foundation's OpenNLP library algorithm for part-of-speech tagging, identifying terms as nouns and the longest noun phrases, and unifying terms through various methods [17,18]. From an initial pool of 1948 terms appearing in at least 60 comments, a subset of 323 (16.58%) terms related to the US health care system, such as Obamacare, prescription, and wait time, was selected for the final semantic network. A detailed term selection process,

XSL•FO
RenderX

including manual screening and thesaurus construction, is described in Multimedia Appendix 1.

By distilling 179,193 comments into a network with several hundred nodes, a macro model of YouTube video commentaries was created, providing insight into social media users' discussions on US health care. In this network, terms are interconnected and organized into distinct, nonoverlapping clusters [15,19,20]. A cluster is a group of terms tightly linked within the group and loosely connected with terms outside it. If >1 term was extracted from the text of the comment, it is possible for the same comment to be represented by multiple nodes in multiple clusters. We did a thematic analysis of clusters to gain insights about the US health system discourse.

We addressed limitations observed in previously published semantic networks by enhancing the network's informational value. First, we added custom overlays to VOSviewer's map, which displays the color of network nodes based on selected attributes. To build overlays, we coded each comment to reflect the theme of its YouTube video and added these codes, along with other comment characteristics (eg, comment date), to a scores file, which was uploaded to the VOSviewer together with our corpus file that contained YouTube comments (for more information on building corpus and scores files, refer to Multimedia Appendix 1). Second, we presented findings with hyperlinks to VOSviewer Online for broader accessibility and interactive engagement with our semantic network.

Network Interpretation and Evaluation

The evaluation of the US health care system's semantic network and its overlays was structured as follows. A comparison of 2 networks, before and after the deletion of repeated comments, served as a test of network robustness. Thematic alignment between the network terms, extracted from YouTube comments, and the videos that elicited these comments was a test of network's content representation.

To examine structural relevance, we asked if network relationships reflected the underlying meanings evident in YouTube comments. We examined clusters: Do terms in the same cluster have more similar meanings than terms in different clusters? We also examined pairs and groups of interconnected terms: Are they used together in the source data? Do their relationships align with existing knowledge? A comprehensive analysis of all pairs or term groups is outside of the scope of this study. For practical reasons, we engaged in close reading of a limited number of comments, focusing mainly on smaller nodes. When the number of comments exceeded 200, we randomly sampled 200 comments to cover discussions of different videos, taking care to sample more than once when we encountered heterogeneous ideas that required careful interpretation. When ≥2 nodes were examined, we used close reading of comments that mentioned all selected terms. Following the approach by Eve [40], network visualizations were used to locate "points of interest, which are then resynthesized into close readings."

In addition, we performed tests of semantic accuracy through raw data verification. Specifically, we cross-checked ambiguous or unexpected terms in our network against the comments that mentioned them. The analysis involved multiple readings of each comment to capture nuances of how individuals articulate their experiences or opinions of the US health care system, focusing on the words that were extracted as terms, their meaning, and context. On several occasions, for example, when performing a close reading for ideology, we offered brief summaries of the main ideas expressed by the commenters. Our validation of semantic network findings against extant comments adhered to the principles for quantitative text analysis outlined by Grimmer and Stewart [41].

Finally, we tested the usefulness of semantic network analysis for generating policy-relevant insights. We picked 4 health system design concepts—universal health care, Medicare for All, single payer, and socialized medicine—and examined how the comments mentioning these concepts were distributed across the terms we mapped. For insights into the policy ramifications of public perceptions of health system design, we focused on ideological terms and those with the highest share of comments referring to each concept.

Ethical Considerations

Ethics approval for this study was sought from Central Michigan University's Institutional Review Board (project 2023-1021-Mt. P). The study did not meet the definition of human participant research under the purview of the institutional review board according to federal regulations. The study used publicly accessible user-generated YouTube comments. The data were deidentified and aggregated before analysis. As the results are presented in an aggregate form, individual commenters cannot be identified. Informed consent has not been obtained. No compensation was provided to comment contributors.

Results

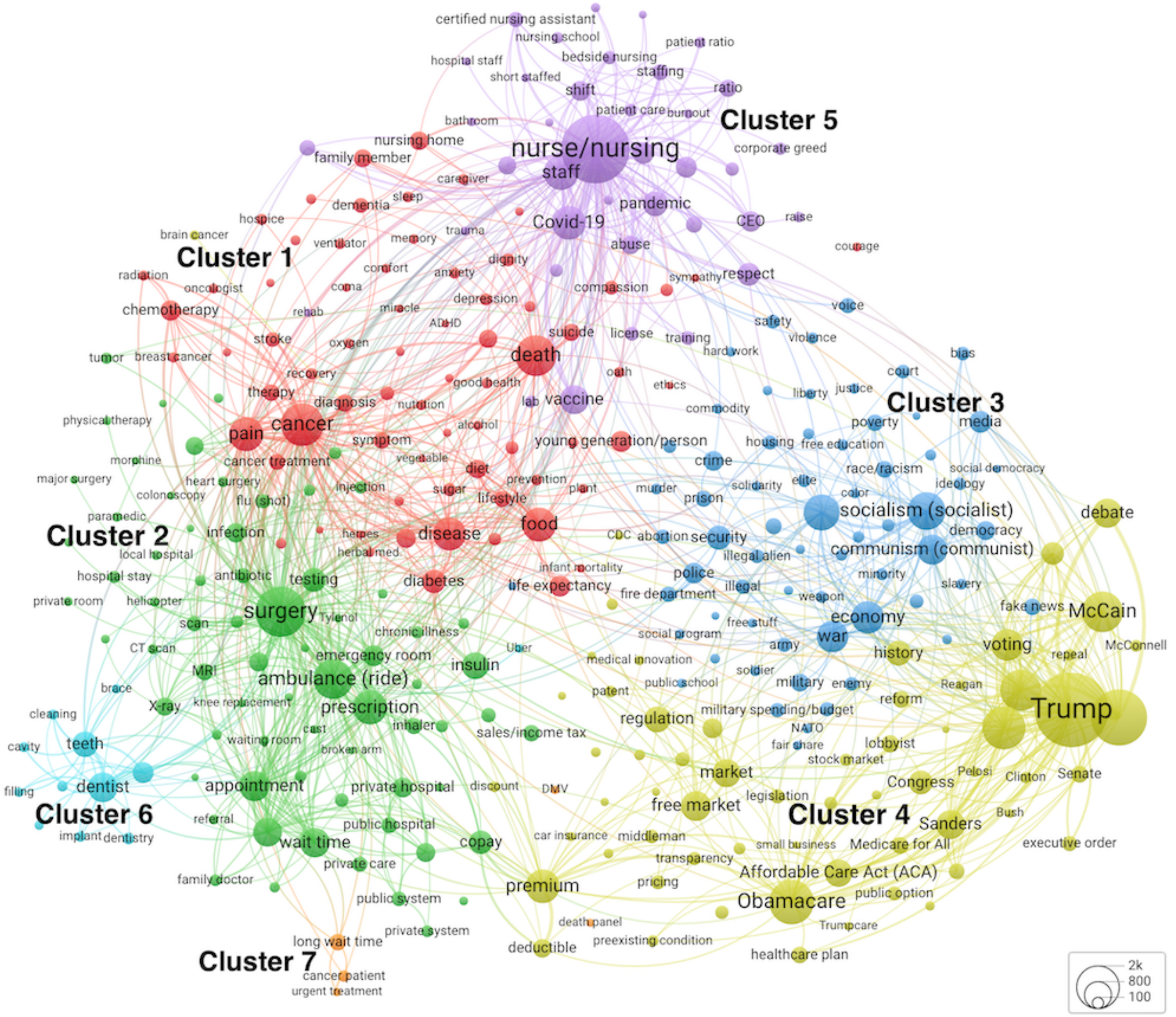
A Semantic Network of Term Co-Occurrence and Clustering

From a manually screened list of 539 terms occurring in our corpus at least 60 times, VOSviewer's algorithm assisted in the selection of 323 (59.9%) most relevant terms [19]. Figure 1 [42] shows a 7-cluster solution for a term co-occurrence network.

On average, each term represented 357.74 (SD 606.88; median 163, IQR 104-321) comments. The longer the comment, the greater the likelihood that multiple terms were extracted from it. VOSviewer assigns cluster numbers based on the quantity of nodes; the same cluster numbers appear in our online interactive maps (URLs are provided in the notes of Figure 1).

XSL•FO
RenderX

Figure 1. A co-occurrence network (cluster map) of terms extracted from the comments on 53 YouTube videos about the US health care system. Binary-counted terms that occurred ≥60 times were mapped. An interactive map is available from Leiden University's VOSviewer app.



Cluster 1 (red) emerged as the largest group of nodes, covering chronic diseases, treatment, pain, and death. Its diverse terms also included topics related to disease prevention (*diet, exercise,* and *smoking*), mental health (*ADHD* [attention-deficit/hyperactivity disorder,] *anxiety*, and *depression*), and end-of-life issues (*hospice, euthanasia*, and *do-not-resuscitate*). Below it, cluster 2 (green) terms covered services, encompassing surgeries, emergency medical services, procedures, diagnostics, wait times, and discussions about public versus private health organizations and prescription medications. On the right, cluster 3 (dark blue) had terms about political ideologies, economic, societal, and cultural issues, surrounded by nodes from cluster 4 (yellow) related to political actors, institutions, the 2010 Patient Protection and *Affordable Care Act* (*ACA* or *Obamacare*), market regulation, and insurance terminology. The top of the map displayed a group of terms (cluster 5, purple) dedicated to health worker shortages, nurse-to-patient ratios, and nurses' burnout. Dental care terms formed a group on the lower left (cluster 6, light blue). Finally, a 5-node group (cluster 7, orange) at the bottom of the map had terms related to long wait times by patients with cancer who required urgent treatments, as well as further away terms *DMV* (Department of Motor Vehicles) and *death panel*.

The network displayed a rather coherent collection of terms, the meaning of which could be intuitively understood within the context of the US health care, with a few exceptions. For instance, as we manually selected terms for map inclusion, we checked the use of an ambiguous term *DMV* in YouTube users' comments. *DMV* was mentioned as a metaphor in a debate of government-managed health care efficiency. It was retained due to its relevance to the health care discourse.

The interpretive value of our network extended beyond a simple list of terms. The network specified links between terms that were often mentioned together, for example, *pricing* and *transparency* in cluster 4. Meaning extraction was further aided by the analysis of spatial proximity, cluster assignment, and cluster boundaries. For example, *preexisting condition*, as a term of interest, was directly and most strongly linked to *Obamacare* and *ACA*, which were mentioned with *preexisting condition* in multiple comments. This finding was consistent with a key ACA provision: insurance companies cannot use applicants' medical history to deny coverage or charge higher

XSL•FO
RenderX

premiums based on their preexisting conditions [43]. Network structure's alignment with existing knowledge speaks to its structural relevance. *Preexisting condition* is located close to *premium*,*deductible*, *pricing*, market-related terms, and *government regulation* from cluster 4 about politics, as well as to *private health insurance* and *copay* on the far right of cluster 2, which is mostly dedicated to health care services. Therefore, when YouTubers discussed the US health care system, they used a noun phrase *preexisting condition* at the semantic intersection of health care politics and legislation, insurance pricing, and health services access.

In summary, the 323 networked terms, identified as most relevant by VOSviewer, unveiled discussions about illness and wellness; health services; ideology and society; the politics of health care agendas and reforms, market regulation, and health insurance; health care workforce; dental care; and concerns such as long wait times.

Before we removed 5575 duplicate comments, our original cluster map (Figure S1 in Multimedia Appendix 1) was quite similar to the cluster map in Figure 1. Our inquiry into the medical debt cluster comments uncovered repeated comments by a single YouTube user. After deletion, this cluster disappeared, but the network's overall structure largely remained intact, demonstrating its robustness.

Next, we examined clusters and nodes using overlays that reflected 2 aspects of the YouTube platform: the videos that elicited comments and the commentary itself. We assessed the usefulness of custom overlays as contextualization tools: Do they improve our understanding of nodes, node groups, and clusters? While we presented data on both video attributes and comment attributes, our analysis prioritized overlays depicting comment characteristics because they are more valuable for understanding digital publics' discussion of the US health care system.

Distribution of Video Groups Across Network Clusters

Thematic alignment between the video content that elicited the commentary and the commentary itself speaks to the content representativeness of the VOSviewer term co-occurrence network. The distribution of comments from 10 thematically diverse YouTube video groups across our term network is shown in overlays in Figure S2 in Multimedia Appendix 1. Our main findings are summarized in Table 1.

We observed substantial thematic congruence between video content and cluster terms. Nodes with above-average concentrations of comments related to the health care workforce were closely grouped in cluster 5, encompassing terms about nurses, staffing shortages, and management. Unlike most nodes in cluster 5, which were associated with health care workforce videos, the term *respect* had an above-average share of comments related to ACA and Obamacare reform videos. Our analysis of comments indicated that commenters mentioned respect for nurses, which explained the placement of *respect* in cluster 5. In addition, many comments on ACA and Obamacare reform videos expressed respect for Senator John McCain, which explained the connection between the term *respect* and *McCain*. *Respect*'s placement within cluster 5 but at its outer boundary, in the direction of node McCain, coupled with video overlay evidence, suggested semantic accuracy and structural relevance of our network.

Videos from 2 groups (health care policies, politics, ACA, and Obamacare reform) generated comments in cluster 4, which consisted of numerous political and reform-related terms. In addition, videos about health costs, one of which was titled "Dollars and Dentists," elicited discussions of dental care (cluster 6). Comments on videos about health care systems in different countries produced terms that appeared in multiple clusters but mostly in cluster 2 about health services and cluster 7 about long wait time concerns. At the same time, a Home Box Office video "Medicare for All" featuring John Oliver and a Netflix video featuring stand-up comedians making jokes about the US health care produced comments in nodes scattered across the map. The Netflix video showcased many comedians and topics, one of whom, Wanda Sykes, spoke about opioids from the perspective of racial and ethnic minority people. A commentary on this topic appeared in nodes *pain* and *prescription* (left side of the map) and *race/racism*, *Black person*, and *White person* (right side of the map), where commenters debated racial disparities in pain medicine access. For race-related nodes, the share of comments on the Netflix video (comedy on the US health care) varied between 1% and 8%, indicating that it was not the only video prompting the discussion. This finding is not unique; it was common for terms to represent commentaries to a wide variety of videos or video groups.

Across all video group overlay legends, the highest scale midpoint was 0.25 for videos about health care costs and financial issues. It means that, on average, 25% (SD 14%) of comments within a term come from that video group. Across 323 map terms and 10 video theme overlays, there were only 11 (0.34%) instances (out of 3230 possible instances) where terms represented >90% of comments from a single video group.

XSL•FO
RenderX

Table 1. Characteristics of videos that elicited comments related to cluster-specific terms.

Cluster number (color)	Topical areas	Cluster's 10 largest terms	Video groups that elicited comments related to most, some, or specific terms within a cluster
1 (red)	Illness and wellness, including mental health and end of life	*Cancer, death, pain, food, disease, diabetes, young generation/person, life expectancy, chemotherapy,* and *cure*	• Children's health care (some terms) • End-of-life health care (some terms) • Health care systems in different countries (*young generation/person* and *life expectancy*) • Comedy on the US health care (*pain*) • Medicare for All video by John Oliver (*pain*)
2 (green)	Health services	*Surgery, ambulance (ride), prescription, appointment, wait time, specialist, insulin, testing, copay,* and *emergency room*	• Health care systems in different countries (most terms) • Medicare for All video by John Oliver (most terms) • Comedy on the US health care (*prescription*)
3 (dark blue)	Ideology and society	*Socialism (socialist), capitalism (capitalist), economy, war, communism (communist), security, media, police, crime,* and *democracy*	• Single-payer health care (most terms) • Health care systems in different countries (some terms) • Medicare for All video by John Oliver (some terms) • Health care costs and financial issues (*capitalism*) • Comedy on the US health care (*race/racism, Black person,* and *White person*) • ACA[a]/Obamacare reform (*race/racism, Black person,* and *White person*)
4 (yellow)	Health care politics, reform, market regulation, and insurance	*Trump, Biden, Obamacare, Republican, Democrat, McCain, premium, voting, free market,* and *debate*	• Health care policies and politics (most terms) • ACA/Obamacare reform (most terms) • Medicare for All video by John Oliver (some terms) • Single-payer health care (some terms) • Health care costs and financial issues (market regulation terms)
5 (purple)	Health care workforce	*Nurse/nursing, staff, Covid-19, vaccine, pandemic, respect, shortage, management, CEO*[b], and *shift*	• Health care workforce (most terms) • Health care systems in different countries (vaccine) • ACA/Obamacare reform (*respect*)
6 (light blue)	Dental care	*Dentist, teeth, dental care, dentistry, implant, dental insurance, cleaning, cavity, filling,* and *brace*	• Health care costs and financial issues (most terms)
7 (orange)	Concerns	*Long wait time, cancer patient, DMV*[c], *urgent treatment,* and *death panel*	• Health care systems in different countries (most terms) • Single-payer health care (*DMV*)

[a]ACA: Affordable Care Act.

[b]CEO: chief executive officer.

[c]DMV: Department of Motor Vehicles.

Comment Date and Ongoing Discussions

When considering the timing of comments, the overall mean for all nodes was December 2020 (mean 2020.99, SD 0.81; range: from early 2018 for *repeal*, referring to the Trump administration and Republican lawmakers' efforts to repeal the ACA, to early 2023 for *do-not-resuscitate*). Clusters 1, 5, and 6 have terms with more recent comments than other clusters (Figure 2, left [42]), which is likely a function of when a video was uploaded on YouTube.

Also shown in Figure 2 are ongoing discussions, conceptualized at the term level as mean posting time since the first comment in the respective video. We calculated time for each comment, based on the video it came from, then averaged across all comments behind each term. The terms that scored above the midpoint of 0.49 years (approximately 6 months) highlighted areas on the map where YouTube users continued to contribute comments long after the videos were posted, serving as a proxy for ongoing interest and engagement. Comment scores were calculated in 2 ways: without standardization, expressed as a fraction of a year (Figure S3 in Multimedia Appendix 1), and with standardization, using the base-10 logarithm to adjust for skewed data. The standardized scores were then normalized so

XSL•FO
RenderX

that the mean is 0 and the scale points represent SDs (Figure 2, right).

Ongoing discussions in cluster 1, "illness and wellness," were about cure (*herbal medicine* and *herpes*), *diabetes*, and life expectancy, and young people persisted, on average, for 11 months. In cluster 2, "health services," ongoing discussions revolved around ambulances, specialist care, prescriptions, appointment wait times, copays, and private (vs public) health insurance or services, roughly covering the same area as high-scoring nodes in an overlay for videos about health care systems in different countries. YouTube commenters demonstrated continued interest in these topics. On average, cluster 2 terms that scored above the mean came from comments posted approximately 9 months after the first comment on a given video.

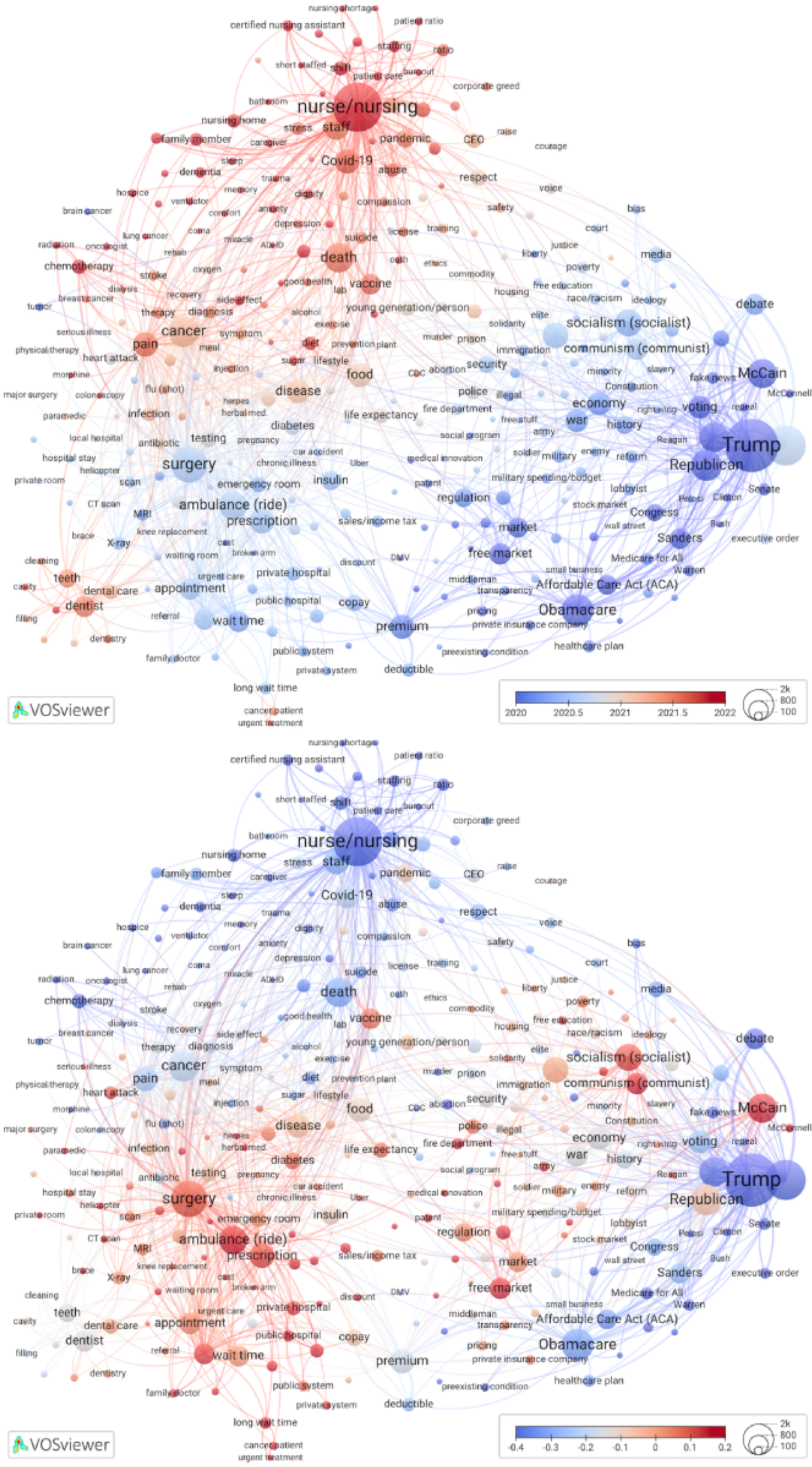
In cluster 3, "ideology and society," YouTube users' comments on political ideologies, police, and military were typically added around the 8-month mark, on average. To better understand an unexpectedly salient group of ideological terms in our map, we analyzed hundreds of comments about communism, socialism, and capitalism. Our analysis confirmed node size and interconnectedness. The discussion of the US health care system was highly politicized, with ideological battles that revolved around dichotomies, such as socialism versus capitalism. Individuals who self-identified as capitalist, conservative, libertarian, or Republican outright rejected any government involvement in health care, calling it socialism, which was often equated with communism (thus confirming node proximity), social democracy, inefficiency, economic decline, and excessive control. Commenters who self-identified as progressive, liberal, social democrat, or left leaning pointed out that health care in the United States was already a mix of capitalism and socialism: publicly funded US police and army were essentially socialized law enforcement, similar to socialized medicine in other countries. They saw no logical reason to reject socialized medicine.

Moreover, several non-US commenters and US residents living abroad shared their positive experiences with health systems in Europe and elsewhere, pointing out that they were affordable to residents with low-income status. Commenters questioned the following: Why do Americans accept *GoFundMe* fundraising to cover medical expenses but not universal health care? Those who defended capitalism praised it for *medical innovation* and high quality of health care but often added that it must be properly regulated. Application of capitalist principles to the US health care system was also discussed in connection to greed, lack of access to health care services, inequities, and poor outcomes. Multiple comments suggested that every economy needed a mix of socialism (relating it to public good or public welfare) and regulated capitalism to counterbalance corporate interests.

Finally, in cluster 4, "health care politics, reform, market regulation, and insurance," we observed ongoing discussions about market-related topics (*monopoly, regulation,* and *market*) and especially the role of John McCain during Obamacare repeal.

XSL•FO
RenderX

Figure 2. Overlays to Figure 1 for mean comment date (top) and ongoing discussions (standardized scores, bottom). High-resolution versions are available in Multimedia Appendix 1 (Figures S4 and S5).



Comment Likes

Comment likes were standardized using the same method as for ongoing discussions. We examined overlays for cluster-specific concentrations of terms that scored above the mean, identified dyads of linked terms that scored high, and summarized the most-liked comments from a specific cluster or term.

In Figure 3 [42], the largest concentrations of above-average liked comments were mostly cluster specific (clusters 1, 2, 5, and 6). Most-liked cluster 5 terms came from comments about shortage of nurses and nurse burnout as well as factors contributing to it (*shift*, *short staffed*, *corporate greed*, *patient ratios*, *abuse*, and *management*). We checked an unexpected connection between *shift* (0.58 SD above the mean for all terms) and *bathroom* (0.48 SD above the mean), which represented

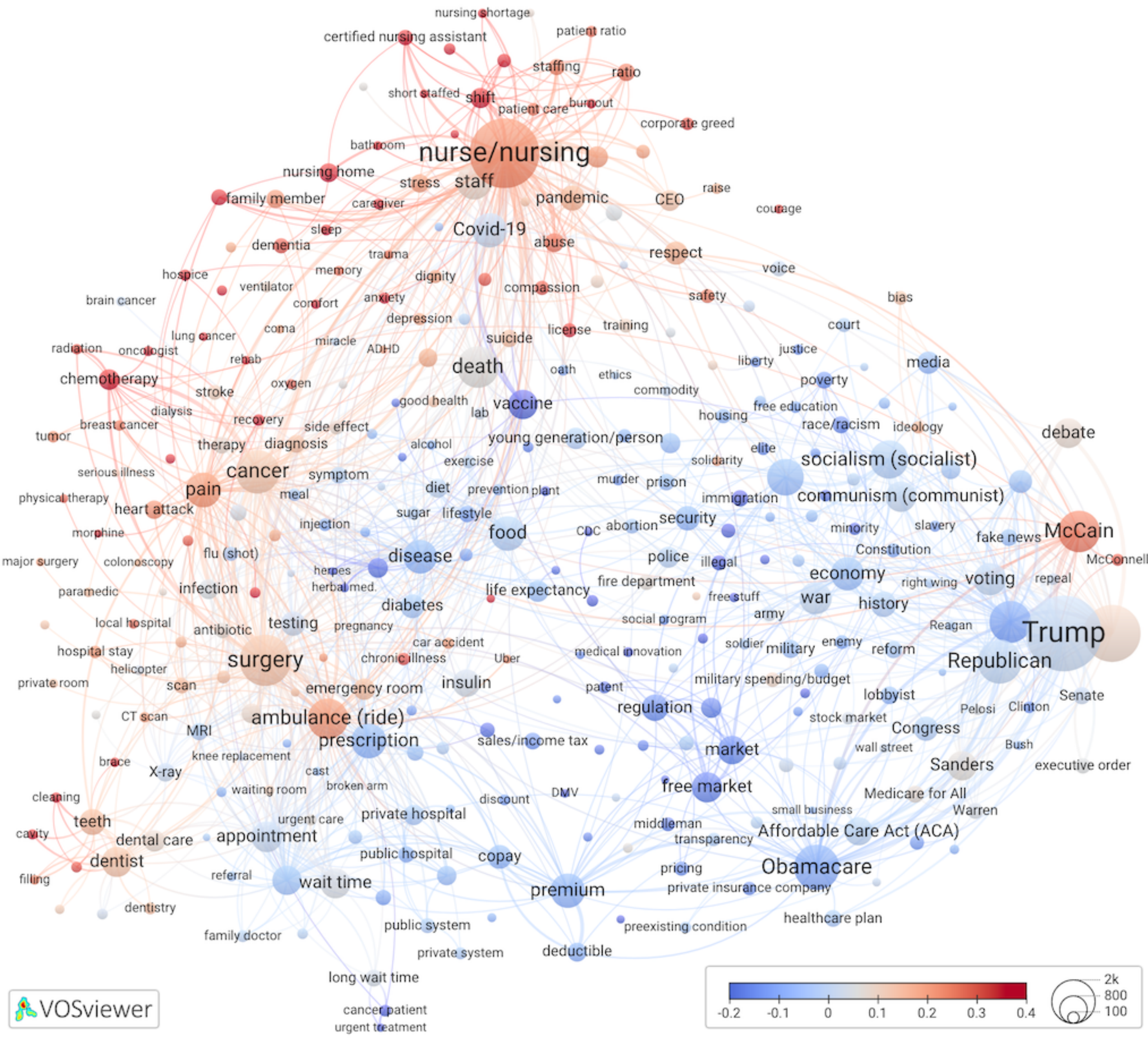
XSL•FO
RenderX

highly liked comments. A total of 20 unique commenters shared stories of extreme job demands, describing how nurses worked long shifts, endured heavy workloads, faced high patient-to-nurse ratios, and had to wait for breaks to address their physiological needs. All but 3 commenters self-revealed their profession. They were experienced nurses, practicing or retired, or nursing students on clinical rotations. Their detail-rich comments described burnout antecedents, such as profits over staffing, mistreatment of nurses, and mandatory overtime, and outcomes, for example, reduced patient care quality and medication errors.

In cluster 1, cancer-related terms, the term *sleep*, and end-of-life terms such as *do-not-resuscitate* were extracted from comments with many likes. Individuals who mentioned “do not resuscitate” (DNR; 0.42) expressed deeply personal desires for autonomy and the avoidance of prolonged distress at the end of life. The commenters identified themselves as older adults, patient advocates, veterans, or health care workers. They discussed the implications of DNR orders, sometimes expressing doubts that an overburdened health care system could handle their implementation in a patient-centered way. Nevertheless, some nurses who witnessed slow deaths of patients without DNR orders chose to create their own advance directives.

Comments about *sleep* were also well liked (0.43) but, unlike the DNR discussion, referred to many different contexts: caregivers, including nurses, experiencing stressors and sleeplessness; sleep as a precondition to wellness; and in the context of passing away peacefully in one’s sleep. The placement of *sleep* within our network, on the boundary of cluster 1 terms (*dementia*, *family member*, *nursing home*, and *caregiver*) and cluster 5 terms (*stress*, *trauma*, and a direct link to *nurse/nursing)*, matched these observations and provided evidence of semantic accuracy and structural relevance.

Figure 3. A mean comment likes (standardized) overlay to Figure 1.



Among dental treatment nodes in cluster 6, *cavity* scored the highest (0.48) on comments with likes. Cavity-related comments came from individuals who revealed the following self-identifications: residence (mostly the United States but also US residents living abroad and foreign nationals from multiple continents), low income (jobless or poor), and medical tourists

XSL•FO
RenderX

(eg, US residents receiving dental treatments in Mexico). Commenters particularly liked quotes of low dental costs in Australia, France, Mexico, and other countries; stories of cost savings after buying airfare and paying for dental treatments abroad; personal accounts of dentists recommending unnecessary procedures; and oral health tips, such as reducing sugar intake. Comments specified systemic problems with US dental care: financial strains, even with dental insurance; potentially superfluous, according to second opinions, or unnecessarily extensive procedures (eg, on baby teeth); worsened conditions due to cost-related treatment delays; and processed sugar industry's influence on consumption of foods, leading to dental decay.

Other clusters also had node groups that were well liked. We explored 2 dyads of linked nodes that scored high on likes: *McCain–McConnell* (0.31-0.34, cluster 4) and *ambulance (ride)–Uber* (0.26 for both, cluster 2), with above-average likes. In first dyad comments, most commenters applauded McCain's vote that helped prevent the repeal of ACA and criticized McConnell and other Republicans. Comments from the second dyad, *ambulance* and *Uber,* were by YouTuber users who expressed concerns about the cost of US ambulances and Americans' reluctance to use specialized emergency transportation. To avoid unpredictable costs, some US commenters planned to use nonmedical transport, such as ride-sharing services like Uber, during health emergencies.

Comments With Select British Spellings

Figure S6 in Multimedia Appendix 1 displays an overlay that approximates contributions from commenters whose backgrounds are associated with regions where British spelling conventions are more common than in the United States. Such spelling was detected in multiple clusters, but the highest-scoring terms were in cluster 2 (*national insurance*, *government hospital*, and *private system*) and cluster 3 (*free education*, *unemployment*, and *justice*).

Commonly Mentioned Health Care Concepts: System Design Ideas

Our last set of overlays demonstrates the distribution of comments that mention policy-relevant ideas on health care system design: universal health care, Medicare for All, a single-payer system, and socialized medicine (Table 2). VOSviewer Online offers a modifiable legend with an option to normalize term scores by subtracting mean and dividing by SD. When term scores are normalized, we can directly compare multiple overlays (Figures 4 and 5 [42]) to identify map areas with terms that are extracted from a high (vs low) share of comments mentioning specific system design ideas. Unlike the standardization of comment scores, normalization is performed at the term level.

Table 2. Mentions of health care system design ideas.

Attributes	Design idea overlay[a]			
	Universal health care	Medicare for All	Single payer	Socialized medicine
Definition[b]	A system where all citizens have access to health care services without financial hardship	A proposed system to expand the US Medicare program to cover all individuals, eliminating private insurance	A system where a single entity (usually the government) pays for all health care costs	A system where the government not only funds but also provides the health care services
Comments, N	3638; "universal health" or "universal healthcare"	2909; M4A or "medicare for all"	1474; "single payer" or "single-payer"	716; "socialized medicine" or "socialised medicine"
Prevalence of comments that mention each design idea within a term-specific comment collection				
Highest-scoring term on a corresponding overlay	*Private room* (12/95, 12.6% of comments also mention universal health care)	*Warren* (116/276, 42% of comments also mention Medicare for All)	*Administrative cost* (16/108, 14.8% of comments also mention single payer)	*Medical innovation* (5/108, 4.6% of comments also mention socialized medicine)
Share of comments within ideological terms[c]				
Socialism/socialist	+1.44 SD	+0.04 SD	−0.16 SD	+0.64 SD
Communism/communist	+3.06 SD	−0.18 SD	−0.65 SD	+0.45 SD
Capitalism/capitalist	−0.33 SD	−0.28 SD	−0.49 SD	−0.53 SD

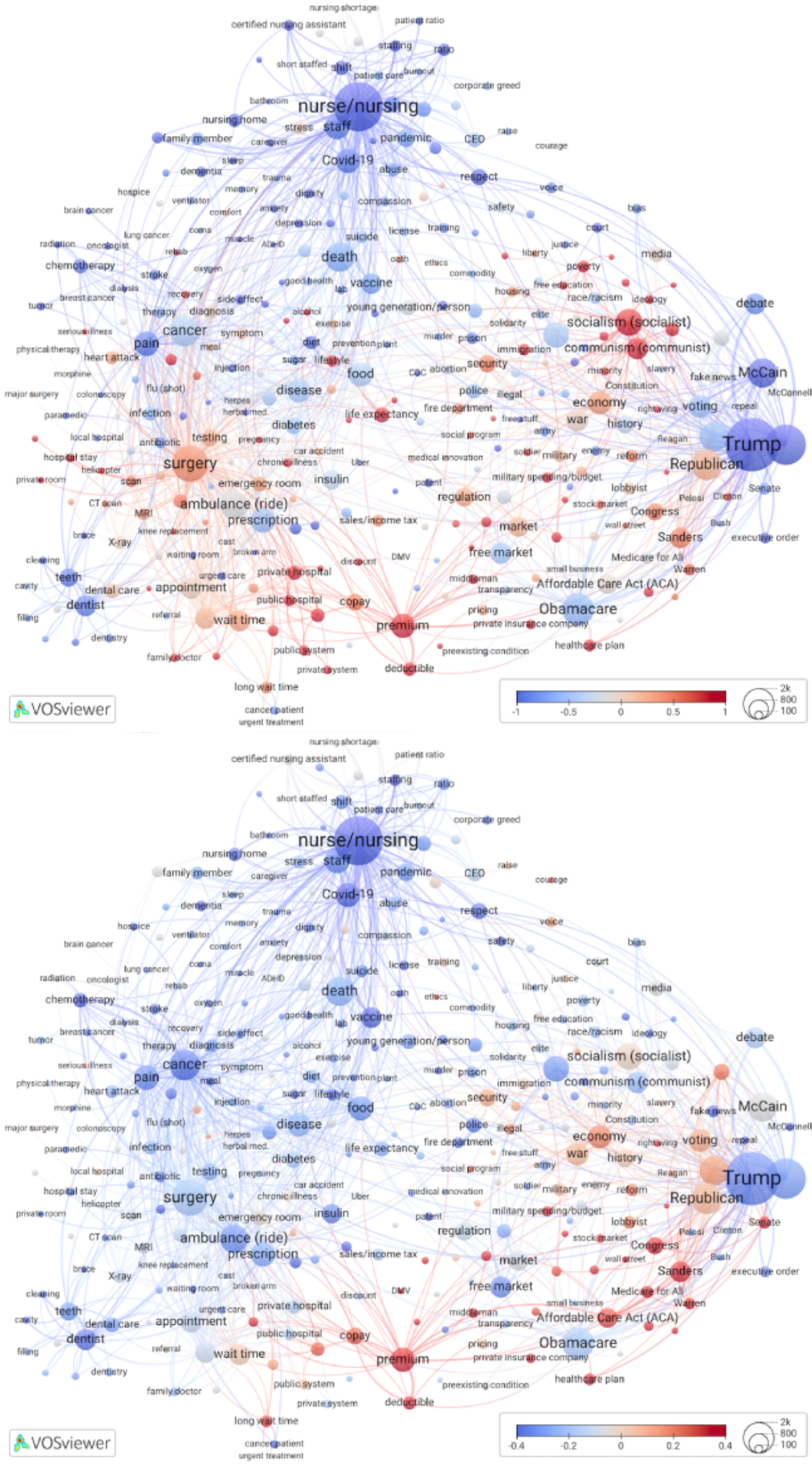
[a]Interactive overlays are available from the left panel (view>items>color >) [42].

[b]Commenters defined health system design ideas in different ways and sometimes used them interchangeably. For example, some commenters talked generally about a state-managed health care system in reference to both single payer and socialized medicine.

[c]Normalized health system design idea overlay scores for 3 ideology nodes are shown relative to all nodes' mean share of comments mentioning that specific health system design idea. Plus or minus signs refer to above or below all terms' mean share, expressed in SD units, within each health system design idea overlay.

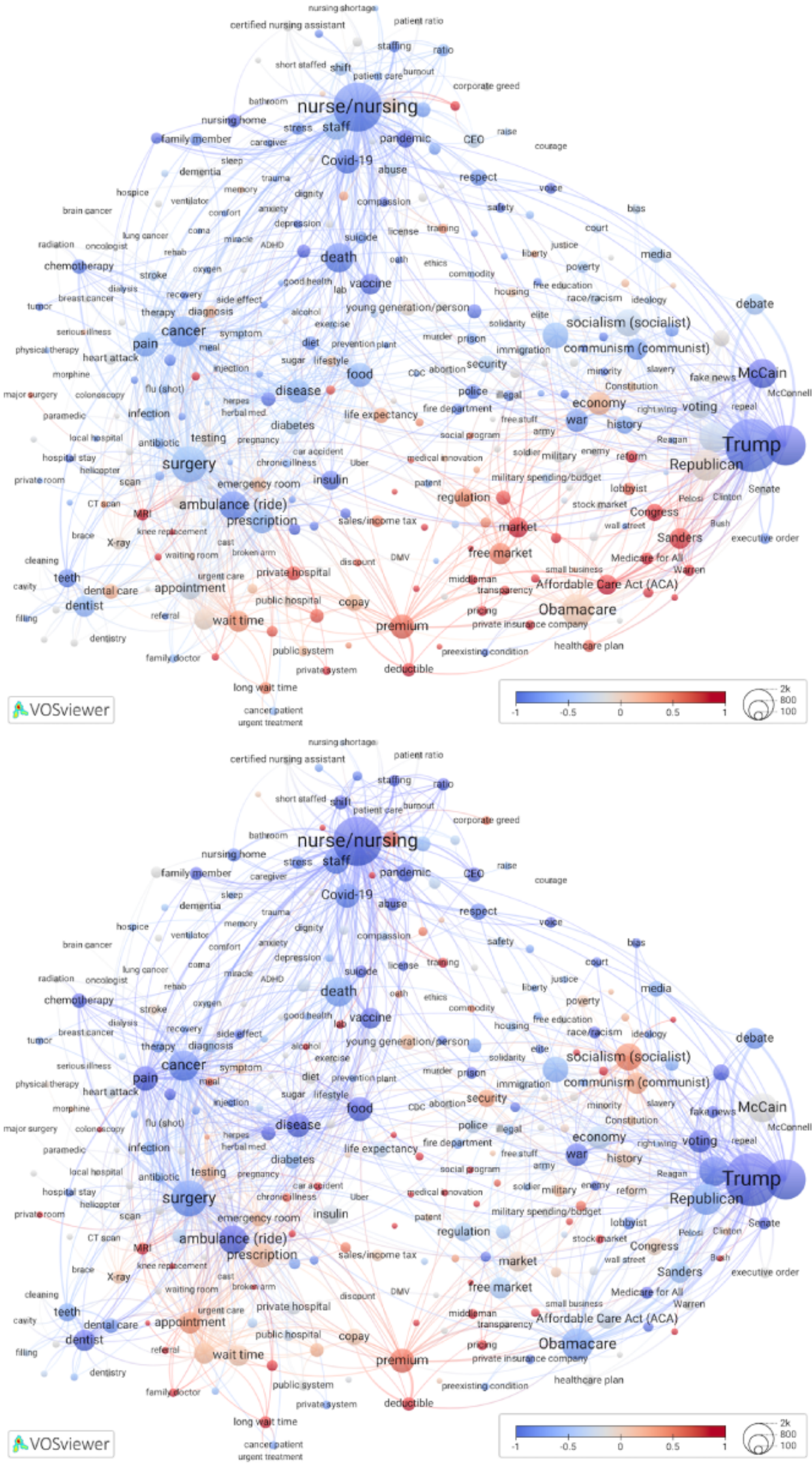
XSL•FO
RenderX

Figure 4. Overlays to Figure 1 depicting the distributions of comments that mention “universal health” (top) and “Medicare for All” (bottom). High-resolution versions are available in Multimedia Appendix 1 (Figures S7 and S8).



XSL•FO
RenderX

Figure 5. Overlays to Figure 1 depicting the distributions of comments that mention "single-payer" (top) and "socialized medicine" (bottom). High-resolution versions are available in Multimedia Appendix 1 (Figures S9 and S10).



As shown in Table 2, the most frequently mentioned health system design idea in our comments—universal health care—was discussed in connection to *private room*, the highest-scoring term on the universal health overlay. The term *private room* also had above-average share (3/95, 3%) of comments, with at least 1 (6%) of 18 British-spelled words. US residents and foreign nationals discussed semiprivate and private hospital rooms as a desirable high standard for hospital stays.

Commenters with experience in universal health systems explained that such systems serve everyone but may not provide extra luxuries unless a patient is also covered by private insurance or pays out of pocket. Several comments expressed preferences for universal health care systems with balanced public and private health care. Private rooms, marble floors, and hotel-like amenities in US hospitals were discussed as

XSL•FO
RenderX

luxuries available to the rich, while care was being denied to the poor.

At the bottom of Table 2, we show 3 ideological terms and compare the extent to which they are linked to each health system design idea. For universal health overlay, the data address the following question: In node *socialism/socialist*, what is the share of comments that mentioned universal health and how far is this share, in SD units, away from the universal health care overlay's mean for all nodes? Compared to 3 other concepts (Medicare for All, a single-payer system, and socialized medicine), universal health care was most strongly linked to discussions of communism and socialism. Specifically, the share of universal health care comments in the node *socialism/socialist* was much greater than that in most other nodes (1.44 SD above all terms' mean). It was even higher for the node *communism/communist* (3.06 SD above the mean).

While discussing Medicare for All in early 2020, YouTube commenters were concerned that it was insufficiently supported by Elizabeth Warren, as compared to Bernie Sanders, which explains why *Warren* was the highest-scoring term in the Medicare for All overlay. In addition to questioning the political viability of Medicare for All, commenters expressed concerns about its funding and tax increases, possible loss of preferred private insurance, unemployment among health insurance workers, increased wait times, diminished quality of care, and fluctuating government or political control over reproductive health.

The highest-scoring term on the single-payer overlay, *administrative cost*, was often mentioned with a term *middleman*, an unnecessary intermediary, for example, private insurance companies and for-profit corporate interests. Discussions of single payer, administrative costs, and middlemen turned into debates. Advocates cited the potential for significant savings and increased efficiency by eliminating the profit-driven insurance model. They pointed to Medicare's low overhead as evidence that a single-payer system could reduce administrative costs. By cutting out middlemen, single-payer systems bring down administrative costs and simultaneously simplify system navigation and transactions for patients, restrain profiteering, reduce health care fraud, and open health care systems to cost control. Critics, however, expressed skepticism about the efficiency of government-run systems, cautioning that replacing one bureaucratic structure with another may not achieve the expected reductions in administrative costs.

Finally, the term *medical innovation* had the highest share of comments that mentioned socialized medicine. The comments often referred to the United States's top position in producing medical innovations. Several US commenters suggested that countries with socialized medicine rely upon US innovations without contributing comparable advancements in new treatments or medical technologies. US medical innovations, according to their comments, come at high cost but also contribute to high quality of care. Others expressed disagreement, saying the United States ranked fourth on medical innovation, behind Switzerland, Germany, and the Netherlands. In addition, hopes were expressed that rising costs of US health care could be controlled through medical innovations, especially in older adult care.

Of the 4 health system design ideas we analyzed, the concept of single-payer health system had the lowest use of ideological terms. The distribution of scores across the single-payer overlay shows that single-payer discussions were less prevalent in ideological terms (*socialism/socialist*, *communism/communist*, and *capitalism/capitalist)* than in other terms we mapped. In the *socialism/socialist* node, an above mean share of comments about Medicare for All (+0.04 SD), socialized medicine (+0.64 SD), and especially universal health care (+1.44 SD) indicated greater use of ideological terms, as compared to single-payer discussions (–0.16 SD). In addition, the universal health care discussion was much more centered around communism or communist (+3.06 SD) compared to the single-payer discussion (–0.65 SD).

Discussion

Overview

We discuss 2 sets of findings. First, we summarize our evaluation of the semantic network. We elaborate on the implications of repurposing VOSviewer to subsequent social media studies and anticipate scientific advances that may result from its broad application. Second, we summarize our US health system insights and discuss their policy implications, pointing out limitations.

VOSviewer Term Co-Occurrence Network as a Social Media Analysis Method

VOSviewer is one of several programs available to researchers for conducting semantic network analysis. For example, previous studies have used the Fruchterman-Reingold algorithm [44], Gephi [45], and R [46] to build semantic networks. At the same time, VOSviewer's user-friendly interface is suitable for users without advanced technical skills. Regardless of the tools used in their construction, semantic networks promise to represent knowledge, while their interconnected nodes likely capture meaning [12], as demonstrated by this analysis.

We used VOSviewer as a data visualization tool to respond to the critical need to decrypt chaotic and extensive social media discussions on a socially important topic. Our analysis suggests that VOSviewer produces visualizations with high information density, interactivity, and interpretive richness. In addition, we obtained evidence regarding the following characteristics of the VOSviewer-generated network: (1) robustness or resilience to variations in data, (2) content representativeness of the diversity of issues related to the US health system, (3) structural relevance defined as meaningful network relationships, and (4) semantic accuracy defined as accurate representation of comment meaning. Our evaluation of the network's decision support usefulness is discussed in the US Health System Insights and Their Policy Implications section.

First, our limited test of robustness confirmed the network's resilience to the removal of approximately 3% of repeated comments from our corpus. If such comments were retained, identical comments by just 1 social media user would have

XSL•FO
RenderX

produced a user-specific map cluster about medical debt and bankruptcy. Striving to build a network reflective of broad conversations, we chose to remove it, but the comments we removed were relevant to the US health system. The person who posted them might have tried to express desperation or draw attention to the seriousness of medical debt.

Second, the network comprehensively covered 10 thematic video groups, representing the entire diversity of video content about the US health care system. In other words, comments from all video groups were represented within the network nodes. Third, we observed a meaningful cluster layout that, overall, could be intuitively interpreted. Structural relevance was confirmed by spatial arrangement of nodes in the network, where the proximity of nodes corresponded to the co-occurring nature of the semantic relationships observed in the text from which the nodes were derived. Moreover, the network's structure aligned with existing knowledge, for example, ACA provisions. Forth, multiple checks confirmed that the mapped terms, including unexpected or ambiguous ones, captured the meanings of posts as well as their context.

Anticipated Scientific Advances of the VOSviewer Application to Social Media Analyses

The VOSviewer's term co-occurrence mapping method and their custom overlays can advance computational social sciences through informative, contextualized semantic networks. Natural language processing enables unbiased extraction of relevant terms, with an option of manual term screening. Revealing large patterns in extensive source data, VOSviewer "visual narratives" [47] can guide researchers to efficiently allocate their analytical resources as they explore salient patterns of societal importance embedded in "context or domain-specific knowledge" [48]. As such patterns involve network terms—nouns and noun phrases that occur in comments—researchers can strategically focus on the most promising subsets of extant data. In addition, VOSviewer-enabled semantic networks bring to light the interdisciplinary nature of social media studies. According to our cluster map, an in-depth analysis of public perceptions of the US health system calls for input from scholars in fields such as communication, economics, health care management, medicine, political science, public health, and others.

Clusters model thematic structure at a macro scale; overlays provide interpretive richness. The method we demonstrated here offers a valuable way for researchers to experience relationships embedded in source data, some of which are hard to document using conventional analyses. Chronological overlays that show video dates, comment dates, and lags in time between the first and the nth comment offer clues on how the discussion progressed over time, enabling a study of unfolding discourses. This is particularly relevant for data from social media platforms, which are "inherently longitudinal" [48]. With additional automation, it would be possible to create dynamic network visualizations that are updated in near–real time as new comments are posted.

Another benefit of semantic map overlays is that they foster cluster exploration and hypothesis testing by combining different data sources. For the YouTube platform, overlays may reflect characteristics of comments, YouTube video channels, videos themselves, or social media users' channels. Therefore, visual overlays represent many opportunities for innovation and experimentation. For example, information excluded during term selection can be brought back in overlays. In this study, we removed geographical references from the cluster model's nodes but created an overlay to highlight discussions with British spelling.

The method we demonstrated in this study can also enhance the value of qualitative research. Resource-intensive qualitative methods can be deployed strategically, guided by the grasp of larger patterns evident in semantic networks. Semantic networks can be contextualized and nuanced through qualitative coding. The qualitative codes can then be incorporated into custom-designed overlays, leading to new hypotheses and qualitative analyses. This iterative approach enables visualization-assisted qualitative inquiry.

Given these methodological strengths, we believe that VOSviewer-enabled semantic network analyses of social media data can advance social science research in the digital era. Thinking even broader, the proposed method can be applied across a variety of contexts and data sources, not limited to social media, and across different disciplines, such as computational humanities.

US Health System Insights and Their Policy Implications

Overview

Health care debates unfold in both in real life and online spheres. We examined digital publics' discourse about the US health care system in response to YouTube videos from right, center, and left media outlets. The YouTube platform allows purposeful selection of videos by varied media outlets on different aspects of an issue. We provided evidence that thematic diversity of videos was passed on to the commentary, opening a door to the policy-relevant analysis of diverse viewpoints. The YouTube platform has emerged as a space for heated debates, thoughtful ideas, misconceptions, and personal narratives in response to the US health care system.

Understanding the viewpoints by social media users provides valuable input for policy makers, health care professionals, and advocates aiming to shape effective reforms. The insights gleaned from the VOSviewer semantic network carry significant implications, which we grouped into 3 categories (concerns about the health care system, domestic and global interconnections in health care discussions, and informing change through key health care discourse insights).

Concerns About the Health Care System

The clusters shed light on a wide range of areas of concern within the US health care system, including those that are likely to be voiced by the public when politicians mention universal health care, Medicare for All, a single-payer system, and socialized medicine. The network analysis was helpful in estimating the use of ideological terms in discussions of various health system design ideas and identifying related concerns, for instance, about continued medical innovation or patients' access to private hospital rooms. The ideology and society cluster

XSL•FO
RenderX

terms, derived from politicized comments, reflect the entrenched ideological conflicts and capitalism-socialism dichotomies within the YouTube discourse about the US health care system.

We observed that comments in the health care workforce cluster, particularly those about staff shortages and burnout, received many likes. This pattern points to a widely shared perception of the urgent need to address challenges faced by nurses and other health professionals. If corroborated across time and other data sources, this sentiment may translate into public support for health care reforms that enhance workforce well-being, improve nurse-to-patient ratios, and support the essential role of health care workers in the system.

Online discussions also highlight ongoing debates about the balance between public and private health care services. Policy makers can use these insights to formulate strategies that optimize the strengths of both sectors, ensuring accessibility, affordability, and quality of care. In sum, a VOSviewer-generated semantic network with overlays shows promise as a decision support tool for policy makers.

Domestic and Global Interconnections in Health Care Discussions

Health care reforms should consider the broader societal and political context of the country to build sustainable and politically viable solutions. The health care discourse we described incorporated widespread debates about political ideologies, societal issues such as racism, and economic considerations. While many of these issues were domestic, there was also a significant international component. Terms such as *national insurance*, *government hospital, private system*, *free education, unemployment*, and *justice* represented 6% to 8% of comments with at least 1 British-spelled word from our list. In much smaller concentrations (2.5%-4%), British-spelled comments appeared in the wellness discussion (*nutrition, vegetable,* and *memory*) and conversations about tax break (or cut), social health care, and private insurance companies. Adding evidence in support of semantic accuracy, several terms extracted from a nonzero share of British-spelled comments (*national insurance* and *social health care*) described societies outside of the United States.

The presence of British-spelled words in our data indicated the global nature of US health care discussions, which is evident in international comparisons of prices and patient experiences. YouTube discussions offered opportunities for US social media users to learn about foreign health systems and explore their benefits, trade-offs, and foundational values. The information was conveyed not by experts or politicians but by laypeople who had encountered foreign systems as taxpayers and patients. Some informants lived in several countries and could compare multiple systems. Informed by global perspectives, the US public may shift its expectations, prompting politicians to incorporate best practices, for example, affordable drugs and predictable costs of emergency patient transportation, into reform initiatives. At the same time, both the public and policy makers stand to benefit from reexamining their own misconceptions and rigid ideological beliefs in light of successful health care models and practices in other countries.

Informing Change Through Key Health Care Discourse Insights

Our semantic network analysis provides insights into the topics that garner the most attention and engagement in ongoing discussions. Health care reforms can be supported by targeted public education and awareness campaigns addressing these key themes, fostering informed public discourse and encouraging active participation in the reform process. Accordingly, policy makers should continuously monitor public sentiments on platforms such as YouTube to inform dynamic, responsive health care policies that adapt to changing societal needs and concerns. Finally, leveraging user engagement patterns, particularly standardized likes and ongoing discussions, can establish effective feedback loops between policy makers and the public. Understanding which aspects of the discourse resonate most strongly with the public allows for the refinement of reform strategies. We provided empirical evidence of links between specific public opinions on health system designs and ideological discourse; comments about universal health care had a much higher use of ideological terms than discussions of single-payer health systems. Overall, the key takeaways drawn from the VOSviewer-generated semantic network analysis provide actionable insights for shaping reforms in health care, which are responsive, inclusive, and aligned with the diverse perspectives expressed by the public on digital platforms.

Finally, we share 2 observations on how VOSviewer maps may support evidence-based policy making and communicating with stakeholders. One consideration is the empirical rootedness of the information we mapped. Decision makers are more likely to accept and act upon information perceived as "evidence based" [48], for example, maps that display intuitively interpretable terms grounded in actual text. In the study by van der Voort et al [47] on big data, decision makers "wanted 'stories to tell' to feed public debate and highlight problems and opportunities," favoring reports at higher resolutions. In our study, clusters communicated broad narratives about the public discourse of the US health system, while overlays enriched and contextualized interpretation of narratives, adding complexity and specificity.

How well decision makers with different levels of education can decode VOSviewer data visualizations remains to be tested. We anticipate that for most decision makers, the learning curve of interpreting maps will be less steep than that for statistical outputs with comparable informational value. While overlays provide a multidimensional understanding of the discourse, they may be harder to decode than clusters. At the same time, the interactive nature of VOSviewer Online is likely to add interest and user engagement, helping to translate research findings into informed decision-making and actionable policy measures.

Limitations

While VOSviewer offers a powerful tool for visualizing and analyzing co-occurrence networks, the algorithm's effectiveness is contingent on the initial selection of terms. The manual screening of a list of terms introduces a potential bias. In addition, the study is limited to English language YouTube comments, which may not fully capture the broader public discourse on health care.

XSL•FO
RenderX

Further research is warranted to validate and expand upon our results. Future studies could use other advanced natural language processing techniques to enhance the accuracy of term selection and clustering. Moreover, a multiplatform analysis that includes other social media platforms and online forums would provide a more comprehensive understanding of public sentiment and discourse surrounding health care.

Acknowledgments

The authors sincerely thank Dr Rodina Bizri-Baryak for her valuable contributions to data collection, which greatly enhanced the depth and breadth of this study.

Data Availability

All data are available in the main text or Multimedia Appendix 1. Map files can be downloaded from map URLs provided in Multimedia Appendix 1. Original YouTube comments (initial comments and first-level replies) can be accessed through YouTube using the video descriptions provided in Multimedia Appendix 1.

Authors' Contributions

LVI conceptualized the study, curated the data, conducted the formal analysis, created the visualizations, provided supervision, and managed the project administration. LVI and EE collaborated on writing the original draft, methodology, investigation, validation of the findings, and contributed to the writing, review, and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary information on video and comment analysis.
[DOC File , 50822 KB - infodemiology_v5i1e58227_app1.doc]

References

1. Blumenthal D, Gumas E, Shah A. The failing U.S. health system. N Engl J Med 2024 Oct 31;391(17):1566-1568. [doi: 10.1056/NEJMp2410855] [Medline: 39383455]
2. Obama B. United States health care reform: progress to date and next steps. JAMA 2016 Aug 02;316(5):525-532 [FREE Full text] [doi: 10.1001/jama.2016.9797] [Medline: 27400401]
3. Lazarus D. Column: The cure for healthcare? An end to selfishness. Los Angeles Times. 2020 Feb 21. URL: https://www.latimes.com/business/story/2020-02-21/column-healthcare-selfishness [accessed 2025-01-27]
4. AAPD responds to PBS frontline's "dollars and dentists". PR Newswire. 2012 Jun 28. URL: https://tinyurl.com/37vckk65 [accessed 2025-01-27]
5. Guo A. What to watch on Tuesday: ‘Frontline: The Healthcare Divide’ on PBS. The Washington Post. 2021 May 17. URL: https://tinyurl.com/mr34ex65 [accessed 2025-01-27]
6. Goldberg J. Who’s to blame for Trump’s failures? Must be Paul Ryan. Los Angeles Times. 2017 May 27. URL: https://www.latimes.com/opinion/op-ed/la-oe-goldberg-paul-ryan-blame-20170327-story.html [accessed 2025-01-27]
7. Gottfried J. Americans’ social media use. Pew Research Center. 2024 Jan 31. URL: https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use/ [accessed 2025-01-27]
8. Matamoros-Fernández A, Farkas J. Racism, hate speech, and social media: a systematic review and critique. Telev New Media 2021 Jan 22;22(2):205-224 [FREE Full text] [doi: 10.1177/1527476420982230]
9. Uwalaka T. Evaluating military use of social media for political branding during online firestorms: an analysis of the Afghan troops withdrawal. J Polit Mark 2023 Jun 12:1-17 [FREE Full text] [doi: 10.1080/15377857.2023.2222683]
10. Johnen M, Jungblut M, Ziegele M. The digital outcry: what incites participation behavior in an online firestorm? New Media Soc 2017 Nov 29;20(9):3140-3160 [FREE Full text] [doi: 10.1177/1461444817741883]
11. Erzikova E, McLean C. Drowning out the message together: analysis of social media comments on a political sex scandal. J Soc Media Soc 2020;9(1):207-234 [FREE Full text]
12. Pereira HB, Grilo M, de Sousa Fadigas I, Souza Junior CT, do Vale Cunha M, Barreto RS, et al. Systematic review of the “semantic network” definitions. Expert Syst Appl 2022 Dec;210:118455. [doi: 10.1016/j.eswa.2022.118455]
13. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2010 Aug;84(2):523-538 [FREE Full text] [doi: 10.1007/s11192-009-0146-3] [Medline: 20585380]
14. Flis I, van Eck NJ. Framing psychology as a discipline (1950-1999): a large-scale term co-occurrence analysis of scientific literature in psychology. Hist Psychol 2018 Nov;21(4):334-362 [FREE Full text] [doi: 10.1037/hop0000067] [Medline: 28726441]

XSL•FO
RenderX

15. van Eck NJ, Waltman L. Citation-based clustering of publications using CitNetExplorer and VOSviewer. Scientometrics 2017 Feb 27;111(2):1053-1070 [FREE Full text] [doi: 10.1007/s11192-017-2300-7] [Medline: 28490825]
16. Waltman L, van Eck NJ, Noyons EC. A unified approach to mapping and clustering of bibliometric networks. J Informetr 2010 Oct;4(4):629-635 [FREE Full text] [doi: 10.1016/j.joi.2010.07.002]
17. van Eck NJ, Waltman L, Noyons EC, Buter RK. Automatic term identification for bibliometric mapping. Scientometrics 2010 Mar 11;82(3):581-596 [FREE Full text] [doi: 10.1007/s11192-010-0173-0] [Medline: 20234767]
18. van Eck NJ, Waltman L. Text mining and visualization using VOSviewer. arXiv Preprint posted online on September 9, 2011 [FREE Full text]
19. van Eck NJ, Waltman L. VOSviewer manual. Universiteit Leiden. 2023 Oct 31. URL: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf [accessed 2025-01-27]
20. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep 2019 Mar 26;9(1):5233 [FREE Full text] [doi: 10.1038/s41598-019-41695-z] [Medline: 30914743]
21. Vlieger E, Leydesdorff L. Content analysis and the measurement of meaning: the visualization of frames in collections of messages. Public J Semiotics 2011 Jun;III(1):28-50 [FREE Full text] [doi: 10.37693/pjos.2011.3.8830]
22. Holmberg K, Hellsten I. Integrating and differentiating meanings in tweeting about the fifth Intergovernmental Panel on Climate Change (IPCC) report. First Monday 2016;21(9) [FREE Full text] [doi: 10.5210/fm.v21i9.6603]
23. Eriksson-Backa K, Holmberg K, Ek S. Communicating diabetes and diets on Twitter - a semantic content analysis. Int J Netw Virtual Org 2016;16(1):8. [doi: 10.1504/ijnvo.2016.075133]
24. Nerghes A, Kerkhof P, Hellsten I. Early public responses to the Zika-Virus on YouTube: prevalence of and differences between conspiracy theory and informational videos. In: Proceedings of the 10th ACM Conference on Web Science. 2018 Presented at: WebSci '18; May 27-30, 2018; Amsterdam, The Netherlands. [doi: 10.1145/3201064.3201086]
25. Garcia-Rudolph A, Laxe S, Saurí J, Bernabeu Guitart M. Stroke survivors on Twitter: sentiment and topic analysis from a gender perspective. J Med Internet Res 2019 Aug 26;21(8):e14077 [FREE Full text] [doi: 10.2196/14077] [Medline: 31452514]
26. Puerta P, Laguna L, Vidal L, Ares G, Fiszman S, Tárrega A. Co-occurrence networks of Twitter content after manual or automatic processing. A case-study on “gluten-free”. Food Qual Prefer 2020 Dec;86:103993 [FREE Full text] [doi: 10.1016/j.foodqual.2020.103993]
27. Zhao Y, Cheng S, Yu X, Xu H. Chinese public's attention to the COVID-19 epidemic on social media: observational descriptive study. J Med Internet Res 2020 May 04;22(5):e18825 [FREE Full text] [doi: 10.2196/18825] [Medline: 32314976]
28. Bashir S, Bano S, Shueb S, Gul S, Mir AA, Ashraf R, Shakeela, et al. Twitter chirps for Syrian people: sentiment analysis of tweets related to Syria Chemical Attack. Int J Disaster Risk Reduct 2021 Aug;62:102397. [doi: 10.1016/j.ijdrr.2021.102397]
29. Tran HT, Lu SH, Tran HT, Nguyen BV. Social media insights during the COVID-19 pandemic: infodemiology study using big data. JMIR Med Inform 2021 Jul 16;9(7):e27116 [FREE Full text] [doi: 10.2196/27116] [Medline: 34152994]
30. Wonneberger A, Hellsten IR, Jacobs SH. Hashtag activism and the configuration of counterpublics: Dutch animal welfare debates on Twitter. Inf Commun Soc 2020 Feb 04;24(12):1694-1711 [FREE Full text] [doi: 10.1080/1369118x.2020.1720770]
31. Noor S, Guo Y, Shah SH, Fournier-Viger P, Nawaz MS. Analysis of public reactions to the novel Coronavirus (COVID-19) outbreak on Twitter. Kybernetes 2020 Nov 09;50(5):1633-1653 [FREE Full text] [doi: 10.1108/k-05-2020-0258]
32. Henson T, Baryak R, Merrill J, Ivanitskaya L. Term network visualization as a means of detecting attempts to sell illicit fentanyl to social media users. In: Proceedings of the American Public Health Association Annual Meeting and Expo. 2023 Presented at: APHA 2023; November 12-15, 2023; Atlanta, GA URL: https://apha.confex.com/apha/2023/meetingapp.cgi/Paper/536121
33. Henson T. Fentanyl, social media, and health policy: a mixed-methods health policy analysis of public comments on YouTube videos about fentanyl. Central Michigan University. 2024. URL: https://tinyurl.com/mwkfpnp6 [accessed 2025-01-27]
34. Nugroho LA, Wuryani E. Exploring public perceptions of ChatGPT in education through netnography and social media analysis. Gagasan Pendidikan Indonesia 2023 Dec;4(2):66 [FREE Full text] [doi: 10.30870/gpi.v4i2.23164]
35. Xiong Y, Pan Z, Yang L. Did online education exacerbate contract cheating during COVID19 in China? Evidence from Sina Weibo. JITE Res 2023;22:295-309 [FREE Full text] [doi: 10.28945/5181]
36. Frasco JM. A social media discussion of COVID-19 emergency use authorization: a mixed methods study of a YouTube influencer. Central Michigan University. 2023. URL: https://tinyurl.com/yeynd3t4 [accessed 2025-01-27]
37. Bizri-Baryak R. Analyzing the overturn of Roe v. Wade: a term co-occurrence network analysis of YouTube comments. Central Michigan University. 2024. URL: https://tinyurl.com/mp7jnnr9 [accessed 2025-01-24]
38. Ivanitskaya LV, Bacon LL, Erzikova E, Gandy L. Visualizing social media discussions of policy-relevant issues surrounding the opioid crisis in the United States. Popul Med 2023 Apr 26;5(Supplement). [doi: 10.18332/popmed/163845]
39. Media bias rating method. AllSides. URL: https://www.allsides.com/media-bias/media-bias-rating-methods [accessed 2024-03-09]
40. Eve MP. Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas. Redwood City, CA: Stanford University Press; 2019.

XSL•FO
RenderX

41. Grimmer J, Stewart BM. Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Polit Anal 2017 Jan 04;21(3):267-297. [doi: 10.1093/pan/mps028]
42. VOSviewer online. VOSviewer. URL: http://tinyurl.com/4v6nafs5 [accessed 2025-01-28]
43. Chatterji P, Brandon P, Markowitz S. Job mobility among parents of children with chronic health conditions: early effects of the 2010 Affordable Care Act. J Health Econ 2016 Jul;48:26-43. [doi: 10.1016/j.jhealeco.2016.03.004] [Medline: 27060524]
44. Fruchterman TM, Reingold EM. Graph drawing by force‐directed placement. Softw Pract Exp 2006 Oct 30;21(11):1129-1164. [doi: 10.1002/spe.4380211102]
45. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. Proc Int AAAI Conf Web Soc Media 2009 Mar 19;3(1):361-362. [doi: 10.1609/icwsm.v3i1.13937]
46. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2022. URL: https://www.R-project.org/ [accessed 2025-01-27]
47. van der Voort H, Klievink A, Arnaboldi M, Meijer A. Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? Gov Inf Q 2019 Jan;36(1):27-38 [FREE Full text] [doi: 10.1016/j.giq.2018.10.011]
48. Kitchin R. Big Data, new epistemologies and paradigm shifts. Big Data Soc 2014 Apr 01;1(1) [FREE Full text] [doi: 10.1177/2053951714528481]

Abbreviations

ACA: Affordable Care Act
DMV: Department of Motor Vehicles
DNR: do not resuscitate

Edited by T Mackey; submitted 10.03.24; peer-reviewed by M Haupt, C Tong; comments to author 15.08.24; revised version received 10.10.24; accepted 08.12.24; published 11.02.25.

Please cite as:
Ivanitskaya LV, Erzikova E
Visualizing YouTube Commenters' Conceptions of the US Health Care System: Semantic Network Analysis Method for Evidence-Based Policy Making
JMIR Infodemiology 2025;5:e58227
URL: https://infodemiology.jmir.org/2025/1/e58227
doi:10.2196/58227
PMID:39932770

©Lana V Ivanitskaya, Elina Erzikova. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 11.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Characterizing Experiences With Hikikomori Syndrome on Twitter Among Japanese-Language Users: Qualitative Infodemiology Content Analysis

Misa Ashley Uchiyama[1]; Hirofumi Bekki[2], MSc, MD, PhD; Tiana McMann[1,3,4], MA; Zhuoran Li[1,4], MS; Tim Mackey[1,3,4], MAS, PhD

[1]Global Health Policy and Data Institute, San Diego, CA, United States

[2]Kyushu Medical Center, Fukuoka, Japan

[3]Global Health Program, Department of Anthropology, University of California San Diego, La Jolla, CA, United States

[4]S-3 Research, San Diego, CA, United States

Corresponding Author:
Tim Mackey, MAS, PhD
Global Health Program
Department of Anthropology
University of California San Diego
9500 Gilman Drive, MC 0505
La Jolla, CA, 92093
United States
Phone: 1 (951) 491 4161
Email: tmackey@ucsd.edu

Abstract

Background: *Hikikomori* syndrome is a form of severe social withdrawal prevalent in Japan but is also a worldwide psychiatric issue. Twitter (subsequently rebranded X) offers valuable insights into personal experiences with mental health conditions, particularly among isolated individuals or hard-to-reach populations.

Objective: This study aimed to examine trends in firsthand and secondhand experiences reported on Twitter between 2021 and 2023 in the Japanese language.

Methods: Tweets were collected using the Twitter academic research application programming interface filtered for the following keywords: "#引きこもり," "#ひきこもり," "#hikikomori," "#ニート," "#脱ひきこもり," "#不登校," and "#自宅警備員." The Bidirectional Encoder Representations From Transformers language model was used to analyze all Japanese-language posts collected. Themes and subthemes were then inductively coded for in-depth exploration of topic clusters relevant to first- and secondhand experiences with *hikikomori* syndrome.

Results: We collected 2,018,822 tweets, which were narrowed down to 379,265 (18.79%) tweets in Japanese from January 2021 to January 2023. After examining the topic clusters output by the Bidirectional Encoder Representations From Transformers model, 4 topics were determined to be relevant to the study aims. A total of 400 of the most highly interacted with tweets from these topic clusters were manually annotated for inclusion and exclusion, of which 148 (37%) tweets from 89 unique users were identified as relevant to *hikikomori* experiences. Of these 148 relevant tweets, 71 (48%) were identified as firsthand accounts, and 77 (52%) were identified as secondhand accounts. Within firsthand reports, the themes identified included seeking social support, personal anecdotes, debunking misconceptions, and emotional ranting. Within secondhand reports, themes included seeking social support, personal anecdotes, seeking and giving advice, and advocacy against the negative stigma of *hikikomori*.

Conclusions: This study provides new insights into experiences reported by web-based users regarding *hikikomori* syndrome specific to Japanese-speaking populations. Although not yet found in diagnostic manuals classifying mental disorders, the rise of web-based lifestyles as a consequence of the COVID-19 pandemic has increased the importance of discussions regarding *hikikomori* syndrome in web-based spaces. The results indicate that social media platforms may represent a web-based space for those experiencing *hikikomori* syndrome to engage in social interaction, advocacy against stigmatization, and participation in a community that can be maintained through a web-based barrier and minimized sense of social anxiety.

(JMIR Infodemiology 2025;5:e65610) doi:10.2196/65610

XSL•FO
RenderX

KEYWORDS

hikikomori; social withdrawal; hikikomori syndrome; mental health; social isolation

Introduction

Background

Hikikomori syndrome, a form of severe social withdrawal largely characterized as experienced among adolescents and young adults in Japan, has recently gained increased attention as a global mental health concern [1]. Importantly, variability in reported *hikikomori* prevalence in countries and regions such as China, Hong Kong, South Korea, Singapore, Nigeria, the United States, and Taiwan may reflect different cultural distinctions of *hikikomori* inclusion criteria, study-specific assessments, and study-specific enrollment methods [2,3]. Nevertheless, increasing prevalence continues to challenge the notion that *hikikomori* is specific to the Japanese context and provides emerging evidence that this phenomenon is widespread cross-nationally [2,4]. Importantly, this form of extreme and persistent social isolation and withdrawal can be viewed as a complex sociocultural mental health phenomenon influenced by a variety of factors, such as economic and employment conditions, social norms and expectations, technology access and use, and changing attitudes toward acceptable social interaction (such as changes in interpersonal dynamics caused by the social isolation experienced during the COVID-19 pandemic) [5,6].

Hikikomori (derived from the verb *hik* [引き], which means to withdraw, and *komori* [籠り], which means to be inside) was originally coined by Japanese psychologist Tamaki Saito in 1998. The term was originally operationalized to refer to an individual who has stopped going to school (*futoukou* [不登校]) or work (*neeto* [ニート]) and has remained at home for a duration of >6 months [7]. A consensus on a standardized definition of *hikikomori* has not been reached, contributing to challenges in measuring the phenomenon, but a commonly used set of criteria was created in 2003 by the Japanese Ministry of Health, Labor, and Welfare (JMHLW) [1,4]. The JMHLW criteria have since been updated with the most recent 2010 definition, which describes *hikikomori* as a result of various factors, such as avoiding social participation (such as schooling, including compulsory education; employment, including part-time jobs; and other interactions outside the home), which in principle has continued under the condition of being housebound for a period of >6 months (this may include leaving the home while still avoiding interactions with others [8]). A later definition in 2020 by Kato et al [9] proposed updated diagnostic criteria for *hikikomori* as a pathological social withdrawal or social isolation in which the essential feature is physical isolation in one's home and for which the person needs to meet the criteria of (1) marked social isolation in one's home, (2) duration of continuous social isolation of at least 6 months, and (3) significant functional impairment or distress associated with the social isolation. Furthermore, many studies have found that patients with *hikikomori* syndrome often had experiences with bullying, peer rejection, or dysfunctional family life and were prone to internet addiction [10,11].

However, until recently, *hikikomori* was understood as a culture-bound phenomenon unique to Japan, reported to affect an estimated 1.2% of the population and over a quarter of students based on household survey data [2,12,13]. Although the causes and risk factors for *hikikomori* are not completely known, many studies have highlighted aspects of Japanese society and culture that enable *hikikomori* features and may account for the especially high number of *hikikomori* cases reported in Japan. Sociocultural factors such as *amae* (甘え), the Japanese term for codependency in parent-child relationships; the tendency of overprotection and indulgence of children by parents; the high-pressure environment created by the Japanese educational system; the need to conform to others and norms; and the challenging job search process for young adults often leading to identity distress have all been hypothesized to be causes of or risk factors for *hikikomori* [14-16]. Furthermore, the idea of isolation has been prominent in Japanese society for centuries. It has been seen as a way of life commonly represented in history with tales of mysterious mountain recluses and hermits [17]. However, numerous *hikikomori*-like situations and the lack of standardized diagnostic methods have made identifying *hikikomori* challenging in Japan [1].

Previous studies have attempted to carry out clinical interviews with families or study individuals who have sought help from public health centers for *hikikomori* syndrome, but the underlining challenges of social reclusion have also made *hikikomori* research and recruitment difficult [12]. Those who experience social and geographic isolation often feel unable to discuss mental illness openly due to the fears of stigma and may feel more comfortable sharing their experiences on the web [18,19]. In response, researchers have leveraged social media platforms as a source of self-reported health information that can be analyzed for stigmatizing issues and topics discussed among hard-to-reach populations, including generating insights specific to certain demographics and geographies [20,21]. Despite this possible application to *hikikomori* research, existing studies using web-based sources of data are limited and have primarily focused on exploring *hikikomori* through Western tweets outside of Japan and tweets in Japanese with limited keywords or have studied *hikikomori* alongside other mental health phenomena [22-24].

Objectives

Importantly, *hikikomori* has evolved since its introduction and original classification in 2003. While initially classified as a cultural syndrome in the 2019 version of the *Diagnostic and Statistical Manual of Mental Disorders*, it has since been included in the appendix of the 2022 *Diagnostic and Statistical Manual of Mental Disorders*, indicating that it will become a formal addition to the volume [25]. These changes may be a result of increasing public awareness, increased willingness to discuss mental health topics, and destigmatization yet can cause an expansion or inflation of the clinical meaning of *hikikomori* [26]. In response, this exploratory study sought to expand knowledge on *hikikomori* syndrome with a focus on

XSL•FO
RenderX

Japanese-language social media posts from Twitter (now known as X), a platform that is popular among Japanese web-based users. Furthermore, no study, to the best of our knowledge, has examined *hikikomori*-related data after 2020 (the start of the COVID-19 pandemic) despite the pandemic contributing to a rise in social isolation due to public health measures mediated by increased use of social media for social interactions [27,28]. We also sought to source more diverse data and web-based discussions by including additional keywords in Japanese, such as more casual terms, closely related words, and synonyms related to *hikikomori*. Finally, this study focused on firsthand and secondhand experiences self-reported by Twitter users and how those experiencing or who have had experience with *hikikomori* interact on the platform. The results of this study can provide insights into how the Japanese *hikikomori* population and their caregivers use social media to discuss this condition and promote a better understanding of primary concerns and behaviors that can help destigmatize this growing condition.

Methods

Data Collection

We first conducted manual searches of *hikikomori* posts on Twitter to identify keywords and hashtags associated with *hikikomori* conversations and mentions in the Japanese language. From this initial search, we identified a set of *hikikomori* keywords that Japanese-language Twitter users commonly used in web-based discussions regarding *hikikomori* syndrome (Multimedia Appendix 1). This initial set of keywords included nonspecific *hikikomori* keywords such as stopping going to school (*futoukou* ["不登校"]) or work (*neeto* ["ニート"]) and staying at home ("自宅警備員"), all terms that are considered a similar social phenomenon to *hikikomori* and often exhibit similar characteristics of social isolation as those included in the definition of *hikikomori* by the JMHLW. This approach was also supplemented by conducting an analysis of Google Trends data for related topics and queries associated with the Japanese-language spelling of *hikikomori* ("引きこもり") from 2004 to the present, which identified additional related topic keywords used in this study.

After the study keywords were finalized, the Twitter application programming interface (API) was used to collect all Twitter posts (ie, tweets) in 50 languages, including Japanese and English. We then limited the data to the Japanese language only (ie, filtered data in the JSON language field *LANG* for the *JA* [Japanese] attribute); removed all retweets; and only included data over a 2-year time frame from January 1, 2021, to January 1, 2023. For the academic API data collection, the API had a limit of 300 queries per a 15-minute window, so the *next_token* feature was used to collect data continually between all queries to ensure that all available data in the given period were collected based on the API settings. The Twitter data field categories analyzed for this study primarily consisted of text, including the following fields: text, link, ID associated with the tweet, username (deidentified and not disclosed in this study), user link, author ID, API type, geolocation (latitude and longitude, if available), and tweet creation time and date. Data collection took place in June 2023.

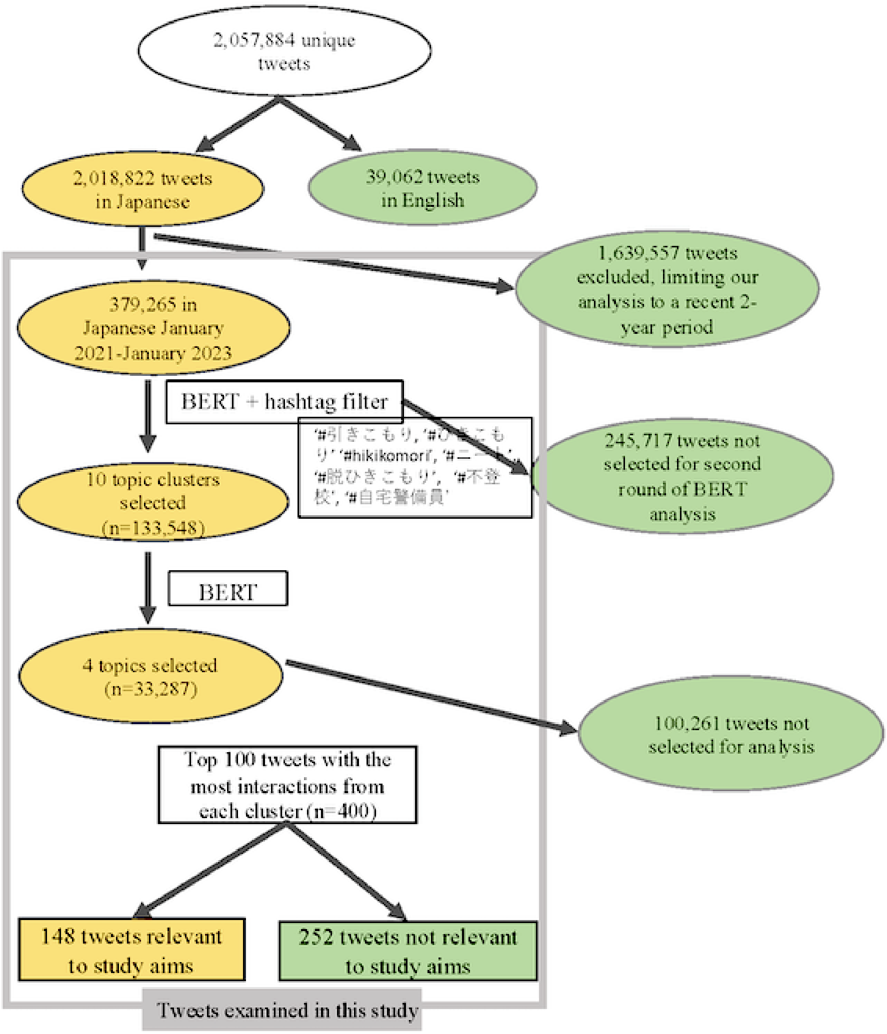
Topic Modeling

Before topic modeling, any retweeted tweets were removed, and only unique tweets with a unique Twitter ID were analyzed. Due to the relatively large volume of data collected, we applied a natural language processing approach to group tweets into relevant thematic clusters. For this corpus of tweets, we used BERTopic, which is a topic modeling technique that leverages Bidirectional Encoder Representations From Transformers (BERT), and class-based term frequency–inverse document frequency, a statistical measure of the importance of words to a particular group of text, to create dense clusters allowing for easily interpretable topics while keeping important words in the topic descriptions. BERTopic then produced an output of data using the k-means algorithm, which includes the sum of the posts into a predetermined number of k clusters (k=10) based on the posts' semantic similarities and groups text containing the same word-related themes into the same clusters. BERTopic was selected due to its use in previous work that has analyzed large-scale Twitter data, its general utility in analyzing unexplored themes that lack existing training data, and utility for the overall exploratory nature of this study's aims [29,30]. Importantly, when compared against other traditional topic models, BERTopic has resulted in a better performance on both topic coherence and topic diversity on Twitter data [31]. Hence, BERTopic methodically has better utility to group tweets that are specifically relevant to *hikikomori* while reducing the potential for noise in selected clusters by providing more accurate and contextually relevant tweet conversational groupings. This study used BERTopic version 0.6.0 with Python version 3.7 (Python Software Foundation).

For the purposes of analyzing tweets specific to the aims of this study, BERTopic was executed in 2 phases: an initial round on the full dataset after data cleaning followed by a second round of focused analysis on relevant, selected topics. Data cleaning performed before the BERTopic analysis included removing punctuation and stop words in posts for optimized BERTopic grouping output. For the initial BERTopic analysis, we ran both 1- and 2-gram analyses on the same dataset to obtain the most visibility of content in our dataset. From both results, we selected 10 clusters in total for review. From those 10 clusters, we limited the data to the following hashtags—"#引きこもり" ("*hikikomori*"), "#hikikomori," "#ニート" ("*neeto*"), "#脱ひきこもり" ("stopping *hikikomori* lifestyle"), "#不登校" ("prolonged absence from school"), and "#自宅警備員" ("home security guard")—to further reduce the data size (see Figure 1 for a summary of the study methodology). With results from these data, a second round of BERTopic analysis was run on the initial 10 topics using a 2-gram BERTopic analysis, and the output topic clusters were reviewed for a final set of 4 clusters that were selected due to their high relevance to *hikikomori* topics. Before manual annotation, we reverted the cleaned posts to their original text with punctuation and stop words to ensure complete comprehension of each post as initially posted. The top 100 tweets with the most interactions from users within the 4 relevant clusters were then selected for manual annotation.

XSL•FO
RenderX

Figure 1. Inclusion criteria and study methodology. BERT: Bidirectional Encoder Representations From Transformers.



Qualitative Content Analysis

The objective of this study was to conduct an in-depth analysis of themes associated with self-reported firsthand and secondhand experiences with *hikikomori* as expressed by Japanese-language Twitter users. For the purpose of study analysis, we relied solely on self-reported *hikikomori* experiences perceived by users and those who observed or interacted with individuals who perceived that they were experiencing *hikikomori* rather than reported verified clinical diagnoses. Hence, there may be variation in the clinical definition of *hikikomori* and self-reported *hikikomori* experiences detected in this study. Our content analysis focused on detecting themes related to firsthand or secondhand knowledge, attitudes, and experiences related to *hikikomori* syndrome or associated characteristics of severe social withdrawal. To classify the content of the collected tweets following topic modeling and topic cluster selection, 2 coders who were native Japanese speakers (the first and second authors, MAU and HB) first independently used a binary coding approach to identify tweets that were relevant to the study aims and excluded tweets that did not fall under the criteria of *hikikomori* syndrome knowledge, attitudes, or experiences self-reported by Twitter users (eg, discussions related to other health or psychological conditions, news articles, statistics or opinions about *hikikomori*, and other topics that used the *hikikomori* term but were not related to the syndrome, hereinafter referred to as *noise*). The primary focus of this study was to identify tweets that met the following conditions: (1) were posted by a Twitter account that appeared to be an individual account (eg, not an organizational, news, or botlike account); (2) self-reported firsthand knowledge of, attitudes toward, or experiences with *hikikomori* syndrome; or (3) self-reported secondhand knowledge of a friend, family member, caregiver, or other social contact that experienced *hikikomori* syndrome.

Discussion of *hikikomori* syndrome or related topics (such as "不登校" or "ニート," which translate to prolonged absence from school or work in English) in the post content, along with pronouns (such as "僕," "私," and "俺," which translate to *I* or first-person pronouns in English, or "彼," "彼女," and "あの子/子供," which translate to *him* or *her*, *that* or *my child*, or second-person pronouns in English) or other form of reference to the user themselves, signaled relevance as a firsthand or secondhand account. After applying this binary coding scheme for inclusion and exclusion, we then used a general inductive coding approach to conduct in-depth qualitative coding of all relevant tweets selected (hereinafter referred to as *signal tweets*). First, a sample of signal tweets were inductively coded by MAU

XSL•FO
RenderX

and HB, and notes were taken on the general themes of posts, from which an initial code list was created focusing on specific *hikikomori* experiences, behaviors, and societal factors associated with *hikikomori* syndrome. Next, formal coding of all signal tweets was conducted using refined codes and developed subcodes. Finally, MAU and HB reviewed the final coded dataset, and the authors reconciled differences in code definitions and application with senior author TKM, also a native Japanese speaker. MAU and HB coded all posts independently and achieved high intercoder reliability for Twitter thematic classification (Cohen κ=0.95).

On the basis of the content of the collected tweets, all detected themes were classified into three major themes: (1) clinical symptoms, with anxiety ("不安障害" and "パニック障害"), social isolation ("社会的孤立" and "ぼっち"), depression ("うつ病"), self-harm ("死にたい," "リスカ," and "自殺"), and developmental and learning disorders ("発達障害" and "学習障害") as subthemes; (2) social determinants, with school ("不登校") and work ("ニート") as subthemes; and (3) awareness, with 1 subtheme detected, education. Descriptive statistics of data characteristics and distribution of the volume of topics coded were also carried out.

Topic Interaction Analysis

To further analyze the levels of user interactivity with different topics related to *hikikomori* experiences self-reported by Japanese-speaking Twitter users, we also examined the volume of users' interaction behavior for all signal tweets. The interactivity with tweets was determined using the number of likes, retweets, comments, and favorites for the tweets analyzed.

Ethical Considerations

This study was exempt from institutional review board approval in accordance with the Common Rule as all data were publicly available and any user-generated data did not include individually identifiable information, and the results are paraphrased and deidentified.

Results

Overview

A total of 2,057,884 tweets were collected from Twitter from February 13, 2009, to June 23, 2023 (n=2,018,822, 98.1% tweets in Japanese and n=39,062, 1.9% tweets in English), based on the method of data collection used. After the exclusion of English-language tweets and limiting our analysis to a recent 2-year period (to examine the more recent discourse concerning *hikikomori* and discussions centered on the general time frame of the COVID-19 pandemic), 18.43% (379,265/2,057,884) of the tweets in Japanese from January 2021 to January 2023 were included for full analysis. Our results are organized into a description of the output topics selected and the qualitative content analysis of specific tweets in each selected cluster.

Topic Selection and Features

The initial 10 topics selected after the first round of BERT analysis all had overlapping themes of mental health and withdrawal from society and high frequency of *hikikomori*-related terms. Frequently mentioned terms included a variation of the term *hikikomori* ("ひきこもり" or "引きこもり"); words associated with mental health conditions such as depression ("うつ病"); and other related terms associated with being socially isolated, such as not being able to go to school ("不登校") or work ("ニート"), which provided a preliminary indication that the cluster included conversational groupings related to *hikikomori* behavior and user perceptions. Other terms identified included *hikikomori*-related services or lifestyle, such as unconventional schooling methods ("free school" ["フリースクール"]), or topics such as gaming, gaming livestreams ("ゲーム 実況"), or web-based platforms (eg, "YouTube") that were suspected as also indicating a *hikikomori* lifestyle. In addition, the presence of more colloquial or casual terms was thought to be more associated with firsthand and secondhand accounts. Following the second round of BERTopic analysis, all *hikikomori*-related themes appeared to be categorized into 1 of the 4 topic clusters selected for the final analysis (topic 1, topic 7, topic 11, and topic 14).

Final Topics for Manual Annotation

The first topic selected as output using BERTopic (topic 1) was selected due the frequency of words such as "お悩み相談" ("consulting for advice"), "話したい" ("want to talk"), and "不登校さんとがりたい" ("seeking connection with someone unable to go to school"), which indicate that tweets in the cluster had a focus on seeking help or connections within the community of those with similar experiences. Phrases such as "必ず返信します" ("I will definitely reply") suggest the practical uses of Twitter as a platform to encourage and facilitate user interaction. Words such as "フリースクール" (schools dedicated to children who fail to fit into conventional school systems in Japan) and "カウンセリング" ("counseling" or "therapy") allude to discussion of the services that are available for those experiencing *hikikomori*, including "精神疾患" ("psychological disorder") and "発達障害" ("developmental disorder"), which further suggests that there is discussion of related disorders and *hikikomori*'s associated impacts. Although there were many possible subtopics in the cluster, there was an overall emphasis on seeking help and exchange of information about the syndrome.

Our second topic selected (topic 7) was focused on mental health, containing topics such as depression ("うつ病") and suicide ("'死にたい"), as well as other words or ideas that are closely related, such as bullying ("いじめ") or feelings of uncertainty or mental instability ("不安"). Words in the cluster were collectively pessimistic or had negative connotations. Phrases such as "どうでもいい報告をする" ("will report something of no use") may indicate that Twitter users in this cluster of tweets feel as if their words have little impact or may be meaningless, which is in concordance with the overall topic of depression and mental health. This cluster contained tweets related to escaping reality and searching for a place to cope, which indicates that platforms such as Twitter may serve as a conversation space for those who are experiencing *hikikomori* syndrome and other related mental health disorders.

The third topic selected (topic 11) centered on secondhand accounts of *hikikomori* experience, most of which came from parents or guardians of youths or minors experiencing "不登校

XSL•FO
RenderX

の親” (“parents of child unable to go to school or hikikomori”). Words in the cluster such as “いじめ” (“bullying”) and “子育て” (“raising children”) indicate caregivers’ concerns about their children regarding their experiences. This cluster is also characterized by a large portion of keywords related to an unwillingness of youth experiencing *hikikomori* to go to school (“学校行きたくない”), which aligns with the social isolation factor that characterizes *hikikomori* syndrome, suggesting the very closely interlinked ideas of *hikikomori* syndrome and “不登校” (“inability to go to school”). More general terms such as “中学生” (“middle schooler”) or “小学生” (“elementary schooler”) that were included in the cluster suggest that education and school are main topics of discussion and indicate at what grade levels children may be first experiencing *hikikomori* syndrome.

The final topic selected (topic 14) was similar to topic 11 in that it also had a focus on caregivers of youth experiencing *hikikomori*. Many similar words, such as “不登校の親” (“parent of a child unable to go to school”) and “子育て” (“raising children”), were also included in this cluster. However, topic 14 had a more specific focus on solutions to struggles, including alternatives to public school—indicated by words such as “フリースクール” (schools dedicated to children who fail to fit into conventional school systems in Japan) and “家庭教師” (“home or private tutoring”). In addition, words such as “メンタルヘルス” (“mental health”) suggested more discussion and awareness associated with *hikikomori* syndrome. Overall, this cluster highlighted the caregivers’ crucial role as the connection between those with *hikikomori* syndrome and the outside world through platforms such as Twitter, discussion and advocacy, and seeking of opportunities for support and services.

Content Analysis

After the initial round of BERTopic analysis, 10 topic clusters (n=133,548 tweets) were selected as relevant to the study aims and underwent an additional round of BERTopic analysis. Following the second round of running BERTopic, 4 topic clusters (n=33,287 tweets from 6403 unique users) were determined to be relevant to the study aims. From these, the top 100 tweets with the most engagement (measured using the sum of the likes, comments, and retweets) from each of the relevant 4 topics (n=400) were extracted and manually annotated for inclusion or exclusion, of which 37% (148/400) of tweets from 89 unique users were identified as relevant to *hikikomori* experiences. Of these 148 relevant tweets, 71 (48%) were identified as firsthand accounts (eg, individuals who currently had or had recently had direct experience with *hikikomori* syndrome), and 77 (52%) were identified as secondhand accounts (eg, parents or guardians of individuals with *hikikomori* syndrome who often lived in the same household). Our qualitative analysis and inductive coding approach derived 8 topics based on our 3 parent categories. All the detected topics were first classified into the 3 parent domains: clinical symptoms (58/148, 39.2%), social determinants (111/148, 75%), and awareness (33/148, 22.3%; see Table 1 for a description and example tweets of the themes and subthemes).

Posts identified within the clinical symptoms domain were characterized by discussions related to explicit or implicit descriptions or mentions of mental health conditions or symptoms related to *hikikomori* syndrome, including both firsthand accounts (52/58, 90%) and secondhand accounts (6/58, 10%). Within this parent domain, frequently discussed symptoms included anxiety (2/58, 3%), social isolation (32/58, 55%), and depression (24/58, 41%). These subthemes represent symptomology most commonly associated with *hikikomori* syndrome. However, additional subthemes within this domain included self-harm (7/58, 12%) and developmental and learning disorders (3/58, 5%) mentioned and discussed alongside *hikikomori* syndrome, suggesting that these other disorders and symptoms may be additional or emerging symptoms associated with *hikikomori*. Frequent examples of posts within the social isolation subtheme included a stated desire for community support, friends, or others with similar experiences on the platform alongside mentions of their physical and social isolation (eg, isolating within their home). For example, users sought friendships with like-minded individuals, often asking to connect with those within a specific age range (eg, “中学生” [“middle schooler”]) or someone with a particular experience (eg, “不登校” [“someone unable to go to school”]). Within the subtheme of depression, individuals often used the platform to rant or openly vent about their depressive symptoms and as a place of expression. Tweets within this subtheme were more of a snapshot of an individual’s emotions rather than a recollection of an event or informational content. These tweets often included the hashtag “#うつ病” (“depression”). Self-harm had overlap with our depression subtheme but diverged in its explicit mentions of self-harm through suicide or wrist cutting. Tweets frequently included the hashtag “#死にたい” (“want to die”). Our final subtheme, developmental and learning disorders, was frequently mentioned specifically by caregivers. Disorders were cited as factors that led to the *hikikomori* lifestyle or the inability to go to school as conditions accompanying *hikikomori* syndrome or as hindrances to daily life. Other disorders that are diagnostically unrelated were also mentioned alongside *hikikomori* syndrome. Tweets within this subtheme frequently included the hashtags “＃発達障害” (“developmental disorder”) and “#学習障害” (“learning disability”).

XSL•FO
RenderX

Table 1. Explanation and paraphrased examples of the identified hikikomori parent domains and topic subcodes detected on Twitter generated from content analysis (N=148).

Parent domain and subtheme	Subtheme description	Example tweet (original+translation)	Tweets, n (%)
Clinical symptoms			58 (39.2)
Anxiety	Content that describes anxious tendencies or uses terms that relate to them (ie, 不安 or “anxious”)	• 不安が増す。(“Anxious thoughts are increasing.”) • 何に追われてるのかわからないが、 • お金、将来、健康 (“I don’t know what is stressing me out, but money, future, health”) • 家にじっとしているともういてもたってもいられなくなる。。(“If I stay at home, I won’t be able to sit still.”) • 読書やテレビを観ても頭に入らない。(“I can’t focus when I read or watch TV.”) • 参ってます。(“I’m exhausted.”) • #鬱 #うつ病 #適応障害 #パニック障害 #不安障害 #不安 #人間関係 #心療内科 #孤独 #ひきこもり #休職 #会社関係 (“#Depressed #Depression #Adjustment disorder #Panic disorder #Anxiety disorder #Anxiety #Human relations #Psychotherapy #Loneliness #Hikikomori #Leave of employment #Co-worker relations”)	2 (1.4)
Social isolation	Content that describes social isolation as part of the user’s lifestyle	• 休職して孤独な年末年始を目前に(“Taking a leave of absence and facing the lonely New Year holidays”) • 家のものを断捨離した (“I got rid of things at home”) • 少しスッキリした社宅の中 (“Inside the slightly empty company housing”) • 少し気分が晴れた (“I feel a little better”)	32 (21.6)
Depression	Content that displays depressive thoughts or episodes that reflect a nonprogressive and pessimistic mindset of the user	• 買いたいもの、欲しいものなんて買えないし、バイトする気力は無いし、そもそも生きたい理由もないから、何もできない。(“I can’t buy anything I want, I don’t have the energy to work part-time, and I have no reason to want to live in the first place, so I can’t do anything.”) • 生きているだけで惨めな思いをする。(“Just being alive makes me feel miserable.”) • 頑張れないし、苦しいし、生きていても迷惑かけるだけだから死にたいって思うのは「甘え」なのだろうか。(“Am I acting like a spoiled child to think that I want to die because I can’t do my best, it’s painful, and even if I live I’ll only cause trouble?”) • #ニート #ひきこもり #死にたい (“#Not in Education, Employment, or Training (NEET) #Hikikomori #Want to die”)	24 (16.2)
Self-harm	Content that includes mentions or descriptions of self-harm and suicide	• 新品のカミソリ気持ちよすぎだろ　(“The brand-new razor blade feels so good　”) • 新品しか勝たん. (“New blades for the win .”) • #自傷行為 #不登校 #od (“#Self-harm #Not going to school #od”) • #不登校と繋がりたい (“#I want to connect with school truants”) • #リスカ #アムカ #レグカ (“#Wrist cutting #Arm cutting #Leg cutting”) • #病み垢 #病み垢女子 (“#Account characterized by mental sickness #Girl with account characterized by mental sickness”) • #病み垢女子さんと繋がりたい (“#Want to connect with girls with accounts characterized by mental sickness”) • #病み垢さんと繋がりたい (“#Want to connect with people with accounts characterized by mental sickness”)	7 (4.7)
Developmental and learning disorders	Content that includes the mention of other developmental and learning disorders alongside “hikikomori” syndrome	• #不登校 なのにも動じないで普通に接してくれるのはとても有難いんだけど #非同期発達 の事は、特に勉強面に関しては言い辛いから一寸困る。 こちらは正直に話しても良いんだけどホントに時々態度急変する人が居るから面倒臭くて試す気にはなれない。 • “I’m very grateful that you don’t get upset and treat me normally even though I’m not going to school, but it’s a bit difficult to talk about #asynchronous development, especially when it comes to studying. I can be honest about this, but there are some people whose behavior changes suddenly after finding out, so I find it annoying and I don’t feel like risking it.”	3 (2)

XSL•FO
RenderX

Parent domain and subtheme	Subtheme description	Example tweet (original+translation)	Tweets, n (%)
Social determinants			111 (75)
School	Content that includes mentions of school and education, often through the act of missing school or unconventional alternatives to public school	• 娘がチャレンジの問題でつまずいて癇癪。学校行ってる子達と比べたら圧倒的に解いてる問題数が違うからすぐにつまずく。適室での勉強も家庭教師も拒否。今自室で『自分だけの力でやってみせる』ってヤケになって勉強しに行った。もう限界でしょうよ. どうすりゃいいの. ("My daughter stumbles over a challenging problem and has a tantrum. Compared to kids who go to school, the number of problems she solves is greatly less, so it's easy to get stuck. She refuses to study in a proper room or have a private tutor. She just said, 'I'll do it on my own' and went to study in her room. It's probably at its limit. What should I do.") • #不登校 #不登校の親 ("#Not going to school #Parents of children not going to school")	103 (69.6)
Work	Content that includes mentions of work, often through the act of missing or quitting work	• 親から「ウーバーイーツでも何でもやれ」「死に物狂いでやるしかないだろう」と言われた。("My parents told me, 'Find a job, work for Uber Eats if you need to' and 'Work as if your life depends on it.'") • 仕事するために"死に物狂い"になる必要がある状況って何だろう。("In what situation would you need to be 'desperate' to do your job") • 仕事するために生きているわけでもないし。("I don't live to work.") • 親から何か言われるたび、死んだほうがマシだとしか思えない。("Every time my parents say something to me, all I can think is that I would be better off dead.") • #ニート #ひきこもり #死にたい ("#NEET #hikikomori #Want to die")	8 (5.4)
Awareness			33 (22.3)
Education	Content that includes active forms of providing education about "hikikomori" syndrome to the public or active advocacy	• #不登校#ひきこもり#ニート 今振り返れば。今の自分なら。そう言えるくらいに全部の経験が今に繋がる。今どこかで悩んでいるおかあさんへ。 当事者さんへ。 あなたは大丈夫。 ひとりじゃない。 あなたはあなた。 他の誰かは他の誰か。 みんな違っていいんだよ。 • "#Not going to school #Hikikomori #NEET Looking back now, I can say that all of my experiences have led me to where I am today. Dear mothers and other people who are worried right now. You are ok. You are not alone. You are you. Someone else is someone else. It's okay for everyone to be different."	3 (2)

The social determinants parent domain was mentioned in 75% (111/148) of the tweets relevant to *hikikomori* experiences. Mentions of school or being "不登校" ("not going to school") were common (103/111, 92.8%), reflecting the younger demographic of those posting about or experiencing *hikikomori* syndrome on Twitter. Although the age of users is difficult to determine, many individuals who sought out a human connection or social interaction on the platform requested relationships within an age range (eg, middle schoolers or high schoolers). Topics such as the lack of friendship due to their isolated lifestyle or their inability to go to school were discussed. Caregivers on Twitter discussing *hikikomori* syndrome were mainly parents of youth who were also "不登校" ("not going to school"). In addition, these caregivers displayed a sense of responsibility to improve their children's lives or ease their difficulties and pain, observed through their active participation in seeking help. As a result, there were many secondhand experiences or caregiver community users detected in these tweets (71/111, 64%). Topics discussed included alternative education opportunities (eg, "フリースクール" [schools dedicated to children who fail to fit into conventional school systems in Japan] or "家庭教師" ["home or private tutoring"]), parenting philosophies, and specific experiences and advice regarding *hikikomori* syndrome. Less common were mentions of work and *hikikomori* (8/111, 7.2%). Tweets identified within this subtheme expressed an even greater disconnect from society and personal accounts of struggling to come to terms with lack of financial independence. Overall, within this parent domain, we found that *hikikomori* syndrome is heavily intertwined with the ideas of "ニート" ("not going to work") and "不登校" ("not going to school") in Japanese society. Even when tweet content alluded to a *hikikomori* lifestyle, many preferred the terms "ニート" and "不登校" ("not going to work" and "not going to school") when self-identifying over explicitly identifying

XSL•FO
RenderX

themselves as having *hikikomori* syndrome. The hashtags "#ニート" and "#不登校" were observed frequently with or even synonymously to *hikikomori* ("#引きこもり" or "＃ひきこもり").

Within the awareness parent domain, individuals sought to reduce stigma regarding *hikikomori* syndrome by spreading awareness about the condition (33/148, 22.3%). Users spread awareness through Twitter primarily in 2 ways. Some tweets portrayed the syndrome positively by clearing misconceptions that previously created apprehensiveness toward *hikikomori* syndrome or by drawing attention to the benefits of the lifestyle (eg, having less disputes between family members and, consequently, having a more peaceful and connected family life in certain circumstances). Other tweets highlighted the negative attitudes toward the syndrome and aimed to reduce stigma by portraying stigmatization of *hikikomori* in a negative context (eg, explaining how a friend's negative comments were morally unacceptable). This was frequently observed through users recalling an experience in which an individual experiencing *hikikomori* syndrome or their caregivers faced shame for their condition. Tweets were not always targeted toward the public and, instead, aimed to reduce internal shame of the syndrome by addressing users with similar experiences. Tweets within the education subtheme (3/33, 9%) were characterized by users providing knowledge to the public through digital flyers, meetings, or other active forms of creating awareness, directly addressing their audience in the process. Tweets within this subtheme were only posted by caregivers and in secondhand accounts as explicit advocacy and education often requires contact with the public. Although it was uncommon, some caregivers used the platform to educate and act in the role of mediator between the isolated population and the uninformed public.

In general, those experiencing *hikikomori* and their caregivers used Twitter to either share experiences and opinions with the public through 1-way communication (personal anecdotes, emotional ranting, and advocacy) or increase social interaction and discussion through 2-way communication (seeking social support and seeking and giving advice). Through 1-way communication, those experiencing *hikikomori* disclosed important and often personal information on their lifestyle or used the platform as a means to discuss and cope with their struggles. Caregivers often shared their own experiences with family members with *hikikomori* syndrome and also worked to directly reduce stigma. Through 2-way communication, those experiencing *hikikomori* found like-minded individuals on the web to connect with. Caregivers also exchanged advice and information to better support individuals experiencing *hikikomori*. We found more 1-way communication (53/71, 75% of firsthand accounts and 54/77, 70% of secondhand accounts) than 2-way communication in the tweets analyzed. However, the audience and motive of the tweets were often unspecified.

Discussion

Principal Findings

This study collected and analyzed 2,018,822 tweets with terms related to *hikikomori* syndrome, a form of severe social withdrawal prevalent in Japan, and after conducting data filtering for more recent posts posted between January 2021 and January 2023 and topic modeling for detection of prevalent themes, we found both first- and secondhand experiences reported among Japanese-language tweets (148/2,018,822, 0.01%) from 89 unique users. Among our sample, we found that 48% (71/148) of tweets discussing their experiences with *hikikomori* syndrome were firsthand accounts of the challenges associated with their daily lives, whereas 52% (77/148) were identified as secondhand accounts mainly from caregivers. Within both first- and secondhand reports, the parent categories identified were clinical symptoms, social determinants, and awareness.

Within the 3 parent domains, we found 8 subthemes, which included users describing firsthand and secondhand experiences with *hikikomori* symptoms, including anxiety, depression, social isolation, self-harm, and developmental disorders, as well as discussion related to missing school or work, a commonly reported manifestation of *hikikomori* [1]. Twitter users in this study also shared advocacy and educational awareness related to the syndrome and sought out connections with other web-based users. Similarly to previous research, this study found a variety of topics. Common use of personal anecdotes and other detected topics such as social support, exchange of advice, and stigma are in line with and further support existing research findings, emphasizing the potential value of social listening–related *hikikomori* discourse on social media platforms where *hikikomori* communities interact [22]. Our findings provide additional novel context by focusing on first- and secondhand experiences of the syndrome to better characterize lived experiences with *hikikomori*. Previous studies have identified topics such as marketing, employment and educational opportunities, and medical and science topics related to the syndrome, which were excluded from this study [22,23].

This study provides additional context to the *hikikomori* literature and provides the first social media–based study to characterize web-based discussions from both the first- and secondhand perspectives in the Japanese language, specifically following the COVID-19 pandemic. Of the social media studies that have characterized lived experiences with *hikikomori*, some have focused on *hikikomori* in Western societies, including European countries, in which individuals who directly experience *hikikomori* were the most active users, in contrast to this study, in which secondhand posts were the most commonly detected overall (eg, caregivers or friends of those with *hikikomori* syndrome) [32]. While COVID-19 restrictions resulted in mandated social isolation to different degrees for people worldwide, there may have been more visibility of *hikikomori* symptoms by caregivers that may not have been otherwise observed before many of the public health restrictions during the pandemic. In turn, caregivers may have turned to social media to connect with others, seek advice about *hikikomori*, or spread awareness of the syndrome.

While our study found more secondhand experiences with *hikikomori* overall, within our clinical symptoms parent topic, we found an overwhelming majority of firsthand reporting of *hikikomori* (52/58, 90%). These findings may indicate that, when an individual is struggling with *hikikomori*, they are more likely to self-report their struggles with the syndrome and its

XSL•FO
RenderX

associated symptoms on the web. Concerningly, as detected in this study, individuals may take to web-based platforms to report more severe symptoms and mental health struggles, such as suicidal ideation [33]. However, social media has increasingly represented a valuable way to detect depression and suicidal ideation and can provide rapid data for policy-level decisions, especially given the rise of mental health conditions during and after the pandemic [33]. As such, this finding may also represent shifting ideas and definitions regarding *hikikomori*, especially after the COVID-19 pandemic, a period characterized by social isolation, remote education, and increasing mental health concerns [26]. Furthermore, firsthand users detected in this study may take to web-based platforms as a way to discuss their own experiences and self-report *hikikomori*-related symptoms but appear to engage in less education, advocacy, or awareness raising compared to those with secondhand experiences based on our observations.

Platforms such as Twitter may be an advantageous and comfortable way for those with *hikikomori* syndrome to interact with others while in a lifestyle that lacks social interaction, especially during the mandated social distancing measures that aligned with the study period. Simultaneously, the results provide updated insights into the lives of those with *hikikomori* syndrome and others who support them, as well as into direct advocacy by those who are affected. The findings indicate that access to information on this syndrome through social media platforms can increase access to other individuals and broader online communities experiencing the syndrome, possibly facilitated by semianonymous and web-based conversations, which may otherwise be inhibited by physical barriers due to the isolating nature of *hikikomori*. By leveraging platforms such as Twitter, greater interactions within the community can potentially reduce internal stigma and shame, whereas greater interactions with the public can reduce external stigma toward the syndrome as a whole [34]. In addition, open discussion about experiences and resources available, both within the community and through interactions with the public, could lead to greater accessibility to those resources and more awareness and acceptance.

Limitations

This study has certain limitations. First, it only evaluated data from publicly available content on Twitter and limited the analysis to Japanese-language tweets and tweets that were in both Japanese and English, which is not representative of general social media *hikikomori*-related discourse, including that occurring on other platforms such as Facebook, Reddit, TikTok, and Instagram. Hence, this study may fail to capture posts from individuals who have additional privacy settings or engage in conversations via private or direct messages due to the stigmatization of mental health issues. Furthermore, this study only analyzed tweets posted by users and not comments or other interactions between Twitter users in response to a tweet, which could have yielded additional discussion related to *hikikomori*. In addition, our period of data collection and analysis coincided with the COVID-19 pandemic, which significantly impacted individuals' way of life and required social isolation. Hence, the volume and nature of *hikikomori* discussions on the web may have also been driven by the COVID-19 restrictions during the study period. This study likely underreported the total amount of *hikikomori*-related content within the dataset as we only coded tweets that were the most highly engaged with within selected topic clusters. This approach streamlines manual coding and allows for more efficient detection of relevant conversations but may exclude tweets that have low engagement. Furthermore, this study may have oversampled what is considered clinical *hikikomori* discussions due to variation in the colloquial meaning of *hikikomori*, the potential expansion of *hikikomori* to refer to less severe symptoms, and the reliance on self-reported accounts from web-based users and their perception of their or someone else's experience with *hikikomori*. Hence, it is crucial to acknowledge that this study's findings are specific to a subset of *hikikomori* accounts and content—those who consider themselves as experiencing *hikikomori* first- or secondhand. As such, it may not capture the diversity of *hikikomori* behaviors and attitudes and lacks generalizability to the overall population of those who experience it. In addition, although Twitter offers users a significant degree of anonymity through features such as customizable usernames and the option to create throwaway accounts for sensitive discussions, self-reported measures remain susceptible to recall bias and social desirability bias, which could lead to over- or underreporting of behaviors. Thus, tweets coded as *hikikomori* may in some instances be less representative of the clinical condition and more associated with the casual use of the term to describe non-*hikikomori* symptoms or may more broadly reflect the collective understanding of *hikikomori* as a concept in Japanese culture by those who do not actually have the condition from a clinical context. Future studies should explore multi-platform analysis for *hikikomori*-related discussions, combine social media data with other data sources (eg, focus groups and surveys), and use other data science approaches (eg, supervised machine learning and large language models) to better characterize *hikikomori* changes over a longer period both before and after the pandemic.

Conclusions

Understanding culturally specific self-reported symptomology through social media studies may offer insights into the convergence and divergence of cross-national *hikikomori* experiences. In addition, commonalities in experiences and rhetoric provide insights into the Japanese public's view of *hikikomori* and its prevalence in Japanese society. The findings of this study also have potential clinical implications. As *hikikomori* is increasingly recognized as a global concern, clinicians may look to web-based platforms and discussion forums to understand modern manifestations of the syndrome and the collective understanding of the concept in different cultural contexts (both in Japan and other cultures experiencing *hikikomori*), especially as a standardized definition and criteria are evolving [4]. This study may also provide additional evidence that online support groups may be well received among those with *hikikomori* and could provide clues on how to help relieve adverse experiences associated with social withdrawal as well as provide social support for those caring for someone with *hikikomori* [32]. These results may also justify the need to increase telehealth consultations in the post–COVID-19 era regarding *hikikomori* screening and possible diagnosis. However, increasing participation in digital care and support

XSL•FO
RenderX

opportunities for patients with *hikikomori* syndrome should be afforded careful consideration to ensure that that same technology does not facilitate further social isolation if not used correctly or in a culturally appropriate manner [35,36].

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

This manuscript has been seen by all authors, who have approved of its content.

Conflicts of Interest

T McMann, ZL, and T Mackey are employees of the start-up company S-3 Research LLC. S-3 Research is a start-up funded with previous and current funding from the National Institutes of Health National Institute on Drug Abuse through a Small Business Innovation Research program for social media research and technology commercialization. T Mackey also holds equity in the start-up company S-3 Research LLC and is the Editor-in-Chief of *JMIR Infodemiology*.

Multimedia Appendix 1
Study keyword selection and rationale for use of the keywords.
[DOCX File , 15 KB - infodemiology_v5i1e65610_app1.docx]

References

1. Teo AR, Gaw A. Hikikomori, a Japanese culture-bound syndrome of social withdrawal?: a proposal for DSM-5. J Nerv Ment Dis 2010 Jun;198(6):444-449 [FREE Full text] [doi: 10.1097/NMD.0b013e3181e086b1] [Medline: 20531124]
2. Eckardt JP. Does the hikikomori syndrome of social withdrawal exist in Denmark? A research request. JMA J 2023 Jan 16;6(1):86-87 [FREE Full text] [doi: 10.31662/jmaj.2021-0217] [Medline: 36793531]
3. Pozza A, Coluccia A, Kato T, Gaetani M, Ferretti F. The 'Hikikomori' syndrome: worldwide prevalence and co-occurring major psychiatric disorders: a systematic review and meta-analysis protocol. BMJ Open 2019 Sep 20;9(9):e025213 [FREE Full text] [doi: 10.1136/bmjopen-2018-025213] [Medline: 31542731]
4. Teo AR, Fetters MD, Stufflebam K, Tateno M, Balhara Y, Choi TY, et al. Identification of the hikikomori syndrome of social withdrawal: psychosocial features and treatment preferences in four countries. Int J Soc Psychiatry 2015 Feb 27;61(1):64-72 [FREE Full text] [doi: 10.1177/0020764014535758] [Medline: 24869848]
5. Yeung J, Karasawa M. Japan was already grappling with isolation and loneliness. The pandemic made it worse. Cable News Network World. 2023 Apr 7. URL: https://edition.cnn.com/2023/04/06/asia/japan-hikikomori-study-covid-intl-hnk/index.html [accessed 2024-07-02]
6. Gent E. The plight of Japan's modern hermits. British Broadcasting Corporation. 2019 Jan 29. URL: https://www.bbc.com/future/article/20190129-the-plight-of-japans-modern-hermits [accessed 2024-07-02]
7. Tamaki S, Angles J. Hikikomori: Adolescence Without End. Minneapolis, MN: University of Minnesota Press; 2013.
8. ひきこもりの評価 支援に関するガイドライン. Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/file/06-Seisakujouhou-12000000-Shakaiengokyoku-Shakai/0000147789.pdf [accessed 2025-02-14]
9. Kato TA, Kanba S, Teo AR. Defining pathological social withdrawal: proposed diagnostic criteria for hikikomori. World Psychiatry 2020 Feb 10;19(1):116-117 [FREE Full text] [doi: 10.1002/wps.20705] [Medline: 31922682]
10. Lee YS, Lee JY, Choi TY, Choi JT. Home visitation program for detecting, evaluating and treating socially withdrawn youth in Korea. Psychiatry Clin Neurosci 2013 May 20;67(4):193-202 [FREE Full text] [doi: 10.1111/pcn.12043] [Medline: 23683149]
11. Stip E, Thibault A, Beauchamp-Chatel A, Kisely S. Internet addiction, hikikomori syndrome, and the prodromal phase of psychosis. Front Psychiatry 2016 Mar 03;7:6 [FREE Full text] [doi: 10.3389/fpsyt.2016.00006] [Medline: 26973544]
12. Koyama A, Miyake Y, Kawakami N, Tsuchiya M, Tachimori H, Takeshima T. Lifetime prevalence, psychiatric comorbidity and demographic correlates of "hikikomori" in a community population in Japan. Psychiatry Res 2010 Mar 30;176(1):69-74. [doi: 10.1016/j.psychres.2008.10.019] [Medline: 20074814]
13. Norasakkunkit V, Uchida Y. To conform or to maintain self-consistency? Hikikomori risk in Japan and the deviation from seeking harmony. J Soc Clin Psychol 2014 Dec;33(10):918-935. [doi: 10.1521/jscp.2014.33.10.918]
14. Neoh MJ, Carollo A, Lim M, Esposito G. Hikikomori: a scientometric review of 20 years of research. Int J Environ Res Public Health 2023 Apr 27;20(9):5657 [FREE Full text] [doi: 10.3390/ijerph20095657] [Medline: 37174175]
15. Furlong A. The Japanese hikikomori phenomenon: acute social withdrawal among young people. Sociol Rev 2008 May 01;56(2):309-325. [doi: 10.1111/j.1467-954x.2008.00790.x]

XSL•FO
RenderX

16. Hihara S, Kambara K, Umemura T, Handa K, Sugimura K. Diverse trajectories of hikikomori symptoms during job search and the role of identity distress: three wave longitudinal research. Front Psychiatry 2022 Jul 7;13:897806 [FREE Full text] [doi: 10.3389/fpsyt.2022.897806] [Medline: 35873266]
17. Kato TA, Kanba S, Teo AR. Hikikomori : multidimensional understanding, assessment, and future international perspectives. Psychiatry Clin Neurosci 2019 Aug 24;73(8):427-440 [FREE Full text] [doi: 10.1111/pcn.12895] [Medline: 31148350]
18. Smith-Merry J, Goggin G, Campbell A, McKenzie K, Ridout B, Baylosis C. Social connection and online engagement: insights from interviews with users of a mental health online forum. JMIR Ment Health 2019 Mar 26;6(3):e11084 [FREE Full text] [doi: 10.2196/11084] [Medline: 30912760]
19. Schnitzler K, Davies N, Ross F, Harris R. Using Twitter™ to drive research impact: a discussion of strategies, opportunities and challenges. Int J Nurs Stud 2016 Jul;59:15-26. [doi: 10.1016/j.ijnurstu.2016.02.004] [Medline: 27222446]
20. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/ajph.2016.303512]
21. Huang X, Li Z, Jiang Y, Li X, Porter D. Twitter reveals human mobility dynamics during the COVID-19 pandemic. PLoS One 2020 Nov 10;15(11):e0241957 [FREE Full text] [doi: 10.1371/journal.pone.0241957] [Medline: 33170889]
22. Pereira-Sanchez V, Alvarez-Mon MA, Horinouchi T, Kawagishi R, Tan MP, Hooker ER, et al. Examining tweet content and engagement of users with tweets about hikikomori in Japanese: mixed methods study of social withdrawal. J Med Internet Res 2022 Jan 11;24(1):e31175 [FREE Full text] [doi: 10.2196/31175] [Medline: 35014971]
23. Pereira-Sanchez V, Alvarez-Mon MA, Asunsolo Del Barco A, Alvarez-Mon M, Teo A. Exploring the extent of the hikikomori phenomenon on Twitter: mixed methods study of western language tweets. J Med Internet Res 2019 May 29;21(5):e14167 [FREE Full text] [doi: 10.2196/14167] [Medline: 31144665]
24. Bozzola E, Spina G, Ruggiero M, Vecchio D, Caruso C, Bozzola M, et al. Media use during adolescence: the recommendations of the Italian Pediatric Society. Ital J Pediatr 2019 Nov 27;45(1):149 [FREE Full text] [doi: 10.1186/s13052-019-0725-8] [Medline: 31775828]
25. Esposito V. The hikikomori phenomenon and its evolution in modern society: a systematic literature review. Italian Sociol Rev 2024;14(3):1021-1044 [FREE Full text] [doi: 10.13136/isr.v14i3.722]
26. Jackson HJ, Haslam N. Ill-defined: concepts of mental health and illness are becoming broader, looser, and more benign. Australas Psychiatry 2022 Aug 14;30(4):490-493. [doi: 10.1177/10398562221077898] [Medline: 35156400]
27. Bonsaksen T, Ruffolo M, Price D, Leung J, Thygesen H, Lamph G, et al. Associations between social media use and loneliness in a cross-national population: do motives for social media use matter? Health Psychol Behav Med 2023 Jan 31;11(1):2158089 [FREE Full text] [doi: 10.1080/21642850.2022.2158089] [Medline: 36618890]
28. Murayama H, Okubo R, Tabuchi T. Increase in social isolation during the COVID-19 pandemic and its association with mental health: findings from the JACSIS 2020 study. Int J Environ Res Public Health 2021 Aug 04;18(16):8238 [FREE Full text] [doi: 10.3390/ijerph18168238] [Medline: 34443988]
29. Naseem U, Razzak I, Musial K, Imran M. Transformer based deep intelligent contextual embedding for Twitter sentiment analysis. Future Gener Comput Syst 2020 Dec;113:58-69. [doi: 10.1016/j.future.2020.06.050]
30. Anwar A, Ilyas H, Yaqub U, Zaman S. Analyzing QAnon on Twitter in context of US elections 2020: analysis of user messages and profiles using VADER and BERT topic modeling. In: Proceedings of the 22nd Annual International Conference on Digital Government Research. 2021 Presented at: DG.O'21; June 9-11, 2021; Omaha, NE. [doi: 10.1145/3463677.3463718]
31. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online on March 11, 2022 [FREE Full text]
32. Correia Lopes F, Pinto da Costa M, Fernandez-Lazaro CI, Lara-Abelenda FJ, Pereira-Sanchez V, Teo AR, et al. Analysis of the hikikomori phenomenon - an international infodemiology study of Twitter data in Portuguese. BMC Public Health 2024 Feb 19;24(1):518 [FREE Full text] [doi: 10.1186/s12889-023-17617-0] [Medline: 38373925]
33. Skaik R, Inkpen D. Using social media for mental health surveillance: a review. ACM Comput Surv 2020 Dec 06;53(6):1-31. [doi: 10.1145/3422824]
34. Naslund JA, Deng D. Addressing mental health stigma in low-income and middle-income countries: a new frontier for digital mental health. Ethics Med Public Health 2021 Dec;19:100719 [FREE Full text] [doi: 10.1016/j.jemep.2021.100719] [Medline: 35083375]
35. Park HE, Yap SF. Technology affordances and social withdrawal: the rise of hikikomori. Psychol Mark 2024 Mar 12;41(7):1469-1488. [doi: 10.1002/mar.21991]
36. Tateno M, Teo AR, Ukai W, Kanazawa J, Katsuki R, Kubo H, et al. Internet addiction, smartphone addiction, and hikikomori trait in Japanese young adult: social isolation and social network. Front Psychiatry 2019 Jul 10;10:455 [FREE Full text] [doi: 10.3389/fpsyt.2019.00455] [Medline: 31354537]

Abbreviations

API: application programming interface
BERT: Bidirectional Encoder Representations From Transformers
JMHLW: Japanese Ministry of Health, Labor, and Welfare

XSL•FO
RenderX

Edited by A Mavragani; submitted 20.08.24; peer-reviewed by K Liew, S Nonaka; comments to author 07.11.24; revised version received 16.12.24; accepted 03.01.25; published 24.02.25.

Please cite as:
Uchiyama MA, Bekki H, McMann T, Li Z, Mackey T
Characterizing Experiences With Hikikomori Syndrome on Twitter Among Japanese-Language Users: Qualitative Infodemiology Content Analysis
JMIR Infodemiology 2025;5:e65610
URL: https://infodemiology.jmir.org/2025/1/e65610
doi:10.2196/65610
PMID:

©Misa Ashley Uchiyama, Hirofumi Bekki, Tiana McMann, Zhuoran Li, Tim Mackey. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 24.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Use of Social Media for Medical Problem Solving by Analyzing the Subreddit r/medical_advice: Quantitative Analysis

Xiyu Zhao[1], BS; Victor Yang[1], BA; Arjun Menta[1], BS; Jacob Blum[1], BS; Padmini Ranasinghe[1], MD

School of Medicine, The Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Xiyu Zhao, BS
School of Medicine
The Johns Hopkins University
733 N Broadway
Baltimore, MD, 21205
United States
Phone: 1 6179352878
Email: xzhao81@jhmi.edu

Abstract

Background: The advent of the internet has transformed the landscape of health information acquisition and sharing. Reddit has become a hub for such activities, such as the subreddit r/medical_advice, affecting patients' knowledge and decision-making. While the popularity of these platforms is recognized, research into the interactions and content within these communities remains sparse. Understanding the dynamics of these platforms is crucial for improving online health information quality.

Objective: This study aims to quantitatively analyze the subreddit r/medical_advice to characterize the medical questions posed and the demographics of individuals providing answers. Insights into the subreddit's user engagement, information-seeking behavior, and the quality of shared information will contribute to the existing body of literature on health information seeking in the digital era.

Methods: A cross-sectional study was conducted, examining all posts and top comments from r/medical_advice since its creation on October 1, 2011. Data were collected on March 2, 2023, from pushhift.io, and the analysis included post and author flairs, scores, and engagement metrics. Statistical analyses were performed using RStudio and GraphPad Prism 9.0.

Results: From October 2011 to March 2023, a total of 201,680 posts and 721,882 comments were analyzed. After excluding autogenerated posts and comments, 194,678 posts and 528,383 comments remained for analysis. A total of 41% (77,529/194,678) of posts had no user flairs, while only 0.1% (108/194,678) of posts were made by verified medical professionals. The average engagement per post was a score of 2 (SD 7.03) and 3.32 (SD 4.89) comments. In period 2, urgent questions and those with level-10 pain reported higher engagement, with significant differences in scores and comments based on flair type ($P<.001$). Period 3 saw the highest engagement in posts related to pregnancy and the lowest in posts about bones, joints, or ligaments. Media inclusion significantly increased engagement, with video posts receiving the highest interaction ($P<.001$).

Conclusions: The study reveals a significant engagement with r/medical_advice, with user interactions influenced by the type of query and the inclusion of visual media. High engagement with posts about pregnancy and urgent medical queries reflects a focused public interest and the subreddit's role as a preliminary health information resource. The predominance of nonverified medical professionals providing information highlights a shift toward community-based knowledge exchange, though it raises questions about the reliability of the information. Future research should explore cross-platform behaviors and the impact of misinformation on public health. Effective moderation and the involvement of verified medical professionals are recommended to enhance the subreddit's role as a reliable health information resource.

(JMIR Infodemiology 2025;5:e56116) doi:10.2196/56116

KEYWORDS

online health information; medical advice; Reddit; r/medical_advice; health information–seeking behavior; user-generated content; subreddits; patient education; virtual environments; information quality; social media; medical problem; quantitative analyses; cross-sectional study; user interactions; online health; decision-making; social news; health information

XSL•FO
RenderX

Introduction

The internet has significantly impacted how individuals access and share health-related information. Online health information–seeking behavior has been a growing area of interest in the medical literature, given its potential impact on patient knowledge, decision-making, and outcomes [1]. As a result, the quality and accuracy of health information shared on the internet have been the subject of numerous studies, which have identified both benefits and risks for users [2,3].

Reddit, a social news forum and discussion website, has emerged as a popular platform for health information sharing [4]. Among its topic-specific communities called "subreddits," r/medical_advice has become a prominent online community where users seek and provide advice related to medical conditions, symptoms, and treatments [5]. r/medical_advice stands out not only for its popularity but also for its extensive user engagement compared with other similar online communities. Despite its popularity, there has been limited research examining the content and user interactions within this online community [6]. As the demand for patient education in internet-based environments continues to grow, it is essential to understand the topics discussed on this subreddit to assess the quality of the information provided, as well as the challenges associated with providing accurate and reliable health information in online spaces.

We define information-seeking behavior as the deliberate pursuit of health-related knowledge by individuals, which differs from information sharing (actively providing knowledge to others) and more general health communication (exchanging health-related messages with various purposes). By focusing on r/medical_advice, we specifically examine users seeking preliminary guidance or reassurance before consulting health care professionals. This study addresses three main research questions: (1) What types of medical questions are asked on r/medical_advice? (2) How do different post flairs, pain levels, and inclusion of media relate to user engagement? and (3) How do verified and nonverified medical professionals contribute to the information ecology of r/medical_advice? The findings of this study will contribute to the growing body of literature on health information–seeking behavior in the digital age and help inform potential strategies for improving the quality and utility of online health information.

Methods

Study Design and Data Collection

This cross-sectional study systematically characterized all posts and their top comments from the r/medical_advice subreddit since its inception on October 1, 2011. Data for this investigation were collected on March 2, 2023, using a public resource created by Jason Baumgartner of pushshift.io [7]. Metadata fields collected for posts included subreddit, post ID, title, self-text, post flair, comment score, post score, author, author flair, URL, image, time stamp, and date (Table 1). Flairs are a feature that allows users to add a label or tag to their posts or usernames. Post flairs categorize post content, while user flairs (also referred to as author flairs) can indicate qualifications or expertise in a specific subject. For comments, the collected metadata fields included subreddit, comment content, score, author, author flair, post ID, URL, image, time stamp, and date. Before analysis, we applied data cleaning steps to remove non–user-generated content and posts that did not represent genuine user inquiries such as automated moderation posts, duplicate entries, or advertisements. We used similar criteria for comments to ensure that both posts and comments represented organic user activity.

Table 1. Definition of metadata fields. This table provides definitions for the common metadata fields encountered in the pushshift.io database.

Metadata field	Definition
Subreddit	The name of the specific Reddit community where the post is made
Post ID	A unique identifier assigned to each post in a subreddit
Title	The heading or title of the Reddit post
Self-text	The main body text of the Reddit post
Post flair	A category or tag assigned to a post to indicate its content or topic
Comment score	A numerical value representing the net upvotes and downvotes a comment receives
Post score	A numerical value representing the net upvotes and downvotes a post receives
Author	The username of the individual who created the post
Author flair	A tag or label next to a user's name that indicates their role, expertise, or affiliation
URL	A direct link to the specific Reddit post
Image	Visual content (photo or graphic) attached to a Reddit post
Time stamp	The exact date and time when the post or comment was made
Date	The date when the post was made, formatted as year-month-day

XSL•FO
RenderX

Subreddit Time Periods and Flair Analysis

The analysis of posts was divided into 3 distinct time periods: October 1, 2011, to March 5, 2019 (period 1); March 6, 2019, to July 31, 2022 (period 2); and August 1, 2022, to March 2, 2023 (period 3). This categorization was necessary due to the varying availability of flairs during these periods. Period 1 had no available flairs, whereas period 2 offered flair options based on pain level or question type. In period 3, flairs were organized using a systems-based approach.

The analysis of author flairs was conducted between May 7, 2019, and March 2, 2023, which corresponds to the implementation of author flairs. Throughout the entire time period, user flair options remained consistent. Flairs related to each post, the account that submitted the post, and comments were analyzed.

Definition of Scores

Scores were defined as the net result of upvotes subtracted by downvotes, with a lower limit set at 0. Upvotes and downvotes on Reddit signify agreement, relevance, or perceived quality of a post or comment. A higher score typically indicates greater community acceptance.

Statistical Analysis and Data Visualization

RStudio (Posit) was used for all statistical analyses, while data visualization was conducted using GraphPad Prism 9.0 (Insight Partners).

Data Analysis

The data analysis process involved the calculation of averages and SDs for posts across the 3 time periods. To comprehensively examine the engagement of the subreddit community with the posts, the study considered several factors, including post flair; the presence of images, galleries (multiple images), or videos; and the combined engagement, which was defined as the sum of scores and comments.

A detailed examination of post flair engagement was conducted, comparing engagement across flairs in periods 2 and 3. The Kruskal-Wallis test was initially applied to assess differences in combined engagement, followed by the Dunn test for pairwise comparisons. During period 2, the analysis was segregated into question type (general, urgent, or other) and pain level (no pain, 1-3, 4-6, 7-9, and 10). Since each post could only be assigned a single flair, posts were exclusively categorized based on either question type or pain level.

In period 3, a similar statistical approach was used to compare combined engagement by the type of medical problem. The analysis in this period focused on the systems-based categorization of post flairs, enabling a more targeted investigation of engagement patterns.

Ethical Considerations

Project data were collected from a publicly accessible online forum. No direct interaction with users occurred, and no personally identifiable information was included in the dataset. In accordance with ethical guidelines for internet research, efforts were made to ensure privacy and confidentiality by excluding usernames and any personally identifiable content from the analysis. The use of Reddit data complies with the platform's terms of service, which allow the analysis of public content for research purposes. Institutional review board approval was not required, as this study exclusively analyzed publicly available, anonymized data and did not involve human participant interventions.

Results

Demographics and Flair Distribution

A total of 201,680 posts (Figure 1A) and 721,882 comments (Figure 1B) were collected from October 2011, the inception of the subreddit, through March 2023. After data cleaning to remove nonmedical inquiries and responses, 194,678 posts and 528,383 comments remained for analysis. The top flairs of periods 2 and 3 are shown in Tables 2 and 3, respectively.

XSL•FO
RenderX

Figure 1. Quarterly trends in posts and comments in r/medical_advice. This bar graph displays the (A) number of posts and (B) comments in the r/medical_advice subreddit over time, with each bar representing a quarter of a year on the x-axis. Qtr: quarter.

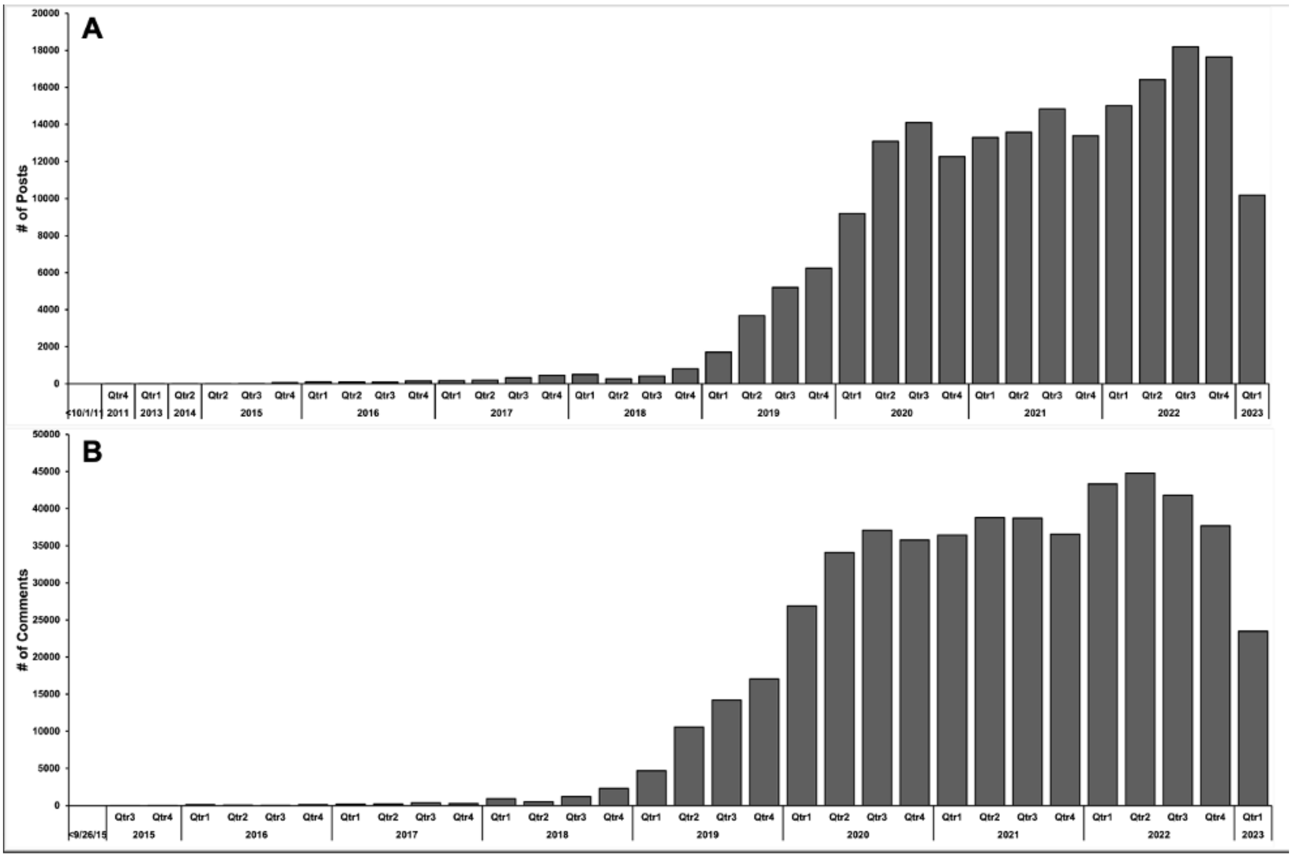


Table 2. Distribution of post flairs in period 2. The table shows the frequency and percentage of post flairs categorized by question type and pain level, illustrating the prevalence of various types of medical questions and reported pain levels in the subreddit during this period.

Post flair type	Post flairs (n=136,486), n (%)
Question type	
General question	50,671 (37.4)
Urgent question	17,739 (13.1)
Other question	6612 (4.9)
Pain level	
No pain	23,844 (17.6)
Levels 1-3	18,337 (13.5)
Levels 4-6	12,252 (9)
Levels 7-9	6055 (4.5)
Level 10	976 (0.7)

XSL•FO
RenderX

Table 3. Distribution of post flairs in period 3. This table presents the frequency and percentage of post flairs across different medical topics during period 3, highlighting the most discussed health issues in the subreddit during this period.

Post flair	Values (n=27,661), n (%)
Skin issues or rashes or freckles or moles	6505 (23.5)
Mouth or gums or throat or cheeks	2648 (9.6)
Genitalia	2486 (9)
Injury	2447 (8.8)
Bones or joints or ligaments	2198 (7.9)
Digestion or stomach or bowels	2179 (7.9)
Illness	2033 (7.3)
Wound care	2015 (7.3)
Medication	1758 (6.4)
Cardiac	1230 (4.4)
Eyes	903 (3.3)
Mental health	741 (2.7)
Parasite concern	264 (1)
Pregnancy	254 (0.9)

User Flair Analysis

Across all time periods, 41% (77,529/194,678) of posts were made by users without user flairs, 42% (81,607/194,678) of posts were made by users who were not verified medical professionals, 18% (35,434/194,678) of posts were made by users who were not verified, and 0.1% (108/194,678) of posts were made by verified medical professionals. The verification process on the subreddit requires the user to upload a picture of their employment badge next to their handwritten username. In other words, 99.9% (194,886/194,691) of the posts were made by Redditors who were not verified medical professionals.

With respect to comments across all three periods, 50% (232,274/528,383) of the comments were made by users tagged “Not a Verified Medical Professional,” 39% (183,470/528,383) of the comments were made by users tagged “Users Not Verified,” and 12% (55,296/528,383) of the comments were made by medical professionals. Table 4 illustrates the breakdown of medical professionals by profession.

Table 4. Breakdown of medical professionals in comments.

Medical profession	Values (n=55,296), n (%)
Nurses[a]	29,838 (54)
Physicians	11,204 (20.3)
Students[b]	6615 (12)
Emergency medical services personnel[c]	1496 (2.7)
Allied health professionals[d]	962 (1.7)
Medical assistants	468 (0.8)
Midlevel providers[e]	215 (0.4)
Nursing support staff[f]	89 (0.2)
Other (moderators, etc)	4409 (8)

[a]Nurses encompass registered nurses, licensed practical nurses, and licensed vocational nurses.

[b]Students involve medical, nursing, and allied health students.

[c]Emergency medical services personnel consist of paramedics and emergency medical technicians.

[d]Allied health professionals include roles such as respiratory therapists, occupational therapists, physical therapists, and radiologic technologists.

[e]Midlevel providers include nurse practitioners and physician assistants.

[f]Nursing support staff includes certified nursing assistants.

XSL•FO
RenderX

Engagement Analysis

Across all posts and time periods on the subreddit, on average, each post received a score of 2 (SD 7.03; range 0-687) and 3.32 (SD 4.89; range: 0-338) comments. To account for the total engagement level of the subreddit over time, the following averages were calculated for each period: (1) score of 1.38 (SD 0.98) and 2.17 (SD 2.87) comments in period 1; (2) score of 2.14 (SD 7.71) and 3.56 (SD 5.05) comments in period 2; and (3) score of 1.48 (SD 3.83) and 2.50 (SD 4.30) comments in period 3.

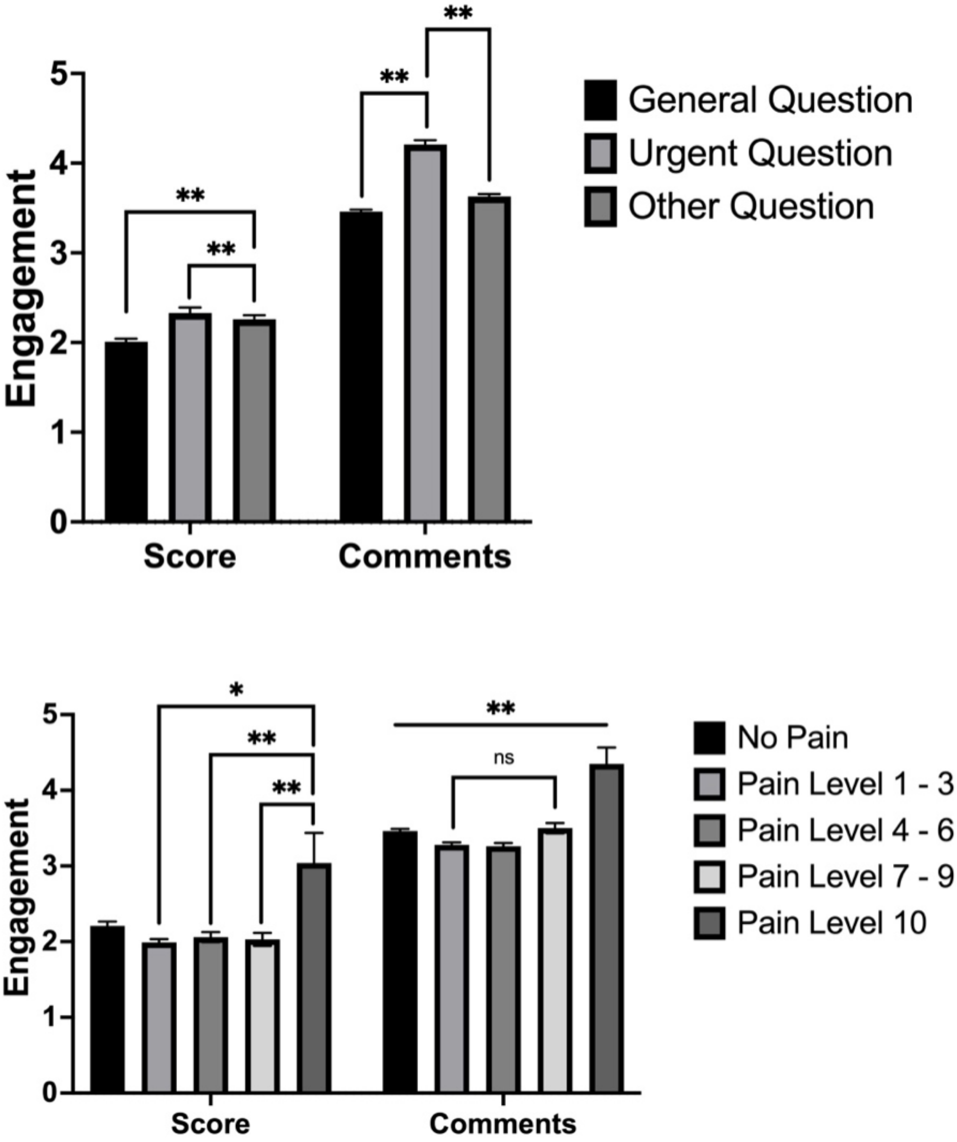
In period 3, posts were divided into system-based flairs. A total of 11,772 posts were removed from the analysis due to the lack of problem-based flairs, leaving 27,661 posts for flair analysis. Engagement, broken down by flair for period 3 is highlighted, is given in Table 5. Posts related to pregnancy had the highest engagement in period 3, while those about bones, joints, or ligaments had the lowest engagement (Figure 2). This pattern was reflected when examining both scores and comments.

Table 5. Engagement by post flair in period 3. This table presents the mean (SD) values of combined engagement (score and comments) for each post flair category during period 3, highlighting the varying levels of engagement across different medical topics within the r/medical_advice subreddit.

Post flair	Combined engagement, mean (SD)	Score, mean (SD)	Comments, mean (SD)
Pregnancy	6.16 (9.18)	1.93 (4.34)	4.23 (5.9)
Wound care	4.60 (9.12)	1.74 (5.53)	2.86 (4.59)
Injury	4.59 (9.97)	1.88 (5.86)	2.71 (5.1)
Parasite concern	4.36 (6.31)	1.51 (2.93)	2.85 (4.23)
Skin issues, rashes, freckles, or moles	4.17 (8.93)	1.60 (4.89)	2.57 (4.58)
Genitalia	4.05 (6.89)	1.42 (3.18)	2.63 (4.32)
Cardiac	3.86 (6.65)	1.34 (2.03)	2.53 (5.01)
Mouth, gums, throat, or cheeks	3.86 (8.13)	1.47 (3.98)	2.40 (4.62)
Illness	3.80 (6.67)	1.39 (2.89)	2.41 (4.46)
Eyes	3.76 (7.41)	1.48 (3.62)	2.28 (4.28)
Mental health	3.65 (5.25)	1.35 (2.66)	2.30 (3.33)
Medication	3.58 (5.09)	1.28 (2.71)	2.30 (3.06)
Digestion, stomach, or bowels	3.45 (4.73)	1.24 (1.97)	2.21 (3.34)
Bones, joints, or ligaments	3.13 (4.34)	1.25 (1.90)	1.88 (2.85)

XSL•FO
RenderX

Figure 2. Engagement analysis by question type and pain level in period 2. This figure presents two separate bar graphs, illustrating the engagement patterns in r/medical_advice during period 2. The top graph displays the engagement by question type, including general question, urgent question, and other question, while the bottom graph shows the engagement by pain level categories (no pain, levels 1-3, levels 4-6, levels 7-9, and level 10). Error bars represent the SEM. Asterisks indicate the level of statistical significance (*P<.05 and **P<.01), with all comparisons in the bottom graph being significant except for the one marked as nonsignificant. These graphs highlight the differences in engagement across various question types and pain levels, shedding light on the patterns of user interaction in the subreddit during period 2. ns: nonsignificant.

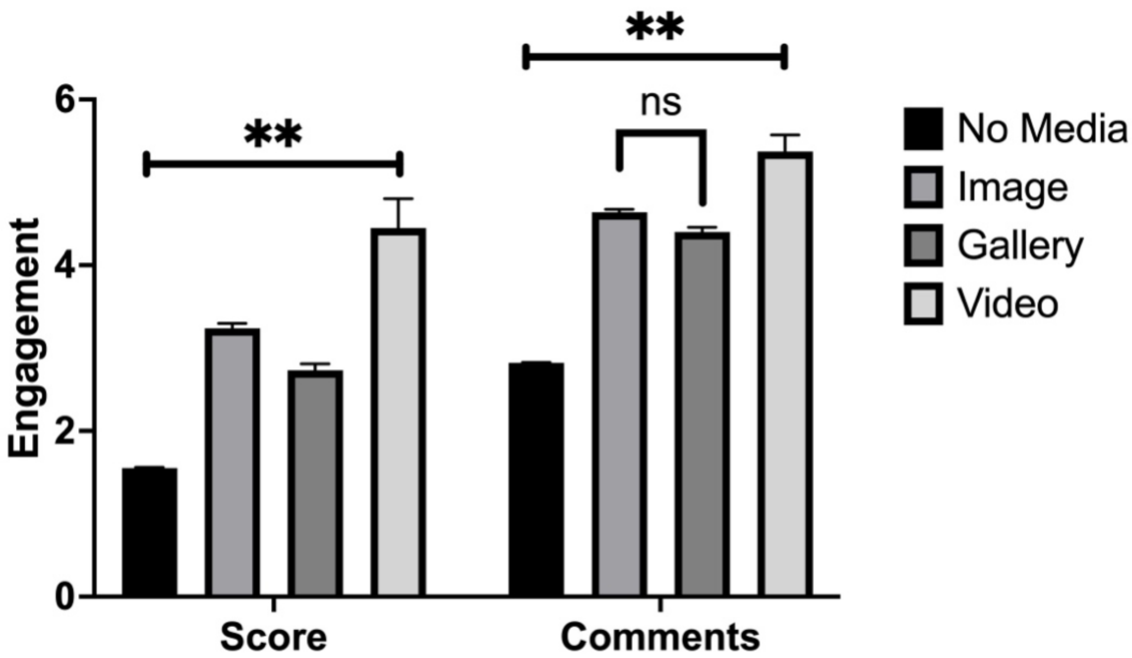


Engagement by Media Inclusion

A total of 28% (56,533/201,904) of posts contained media in the form of images or videos. Of these, 20% (39,776/198,830) included a single image, 8% (15,149/189,363) included multiple images, 0.8% (1608) included a video, and 72% (145,147/201,000) did not include any media. Posts that included a single image received, on average, a score of 3.24 (SD 12.03) and 4.64 (SD 6.91) comments. Posts with multiple images received, on average, a score of 2.73 (SD 9.18) and 4.40 (SD 6.80) comments. Posts with a video received, on average, a score of 4.45 (SD 14.10) and 5.37 (SD 7.98) comments (Figure 3).

XSL•FO
RenderX

Figure 3. Engagement by inclusion of media. This bar graph illustrates the engagement of posts based on the type of media the posts include in r/medical_advice during all periods. Gallery means multiple images are included as part of the post. Error bars represent the SEM. Asterisks indicate the level of statistical significance (*P<.05 and **P<.01). ns: nonsignificant.



Posts with any media received, on average, a score of 3.14 (SD 11.41) and 4.60 (SD 6.92) comments, compared with a score 1.55 (SD 4.15) and 2.82 (SD 3.70) comments for posts without any media. Compared with posts without media, those with media received higher engagement (Dunn test; scores P<.001, comments P<.001). There was a significant difference between engagement of videos, multiple images, and a single image (Dunn test; scores P<.001; comments P<.001). Posts with videos received the highest engagement, followed by posts with images, and posts with no media received the least. Furthermore, posts with multiple images received lower scores (P<.001) but a greater number of comments (P<.001) compared with posts with a single image.

Discussion

Principal Findings

Our study provides an in-depth examination of user dynamics within the subreddit r/medical_advice, illuminating the intricacies of online health information–seeking behaviors. Our findings align with established medical literature on online medical information seeking. Online health forums have been shown to frequently serve as primary sources for addressing nonurgent and less severe medical concerns [8]. The high volume of posts on noncritical health issues suggests a common use of these platforms. It is reasonable to think that users are seeking preliminary advice, or perhaps just reassurance, before consulting a health care professional due to the ease of access to online medical information. Of note, the high level of engagement with pregnancy-related posts is a trend mirroring other online health communities [9], highlighting a consistent public interest in reproductive health.

In addition, our study explored the engagement dynamics of posts containing visual media, an area of study that is lacking in current medical literature. Our results show that posts featuring images or videos, especially concerning dermatological issues such as skin rashes or moles, have attracted higher levels of engagement. This observation not only underscores the effectiveness of visual aids in communicating complex medical information but also hints at a growing user preference for multimedia content [10]. With the rise of telemedicine and digital health communication in the post–COVID-19 era, the importance of visual aids in enhancing both diagnosis and patient understanding cannot be overstated.

Another intriguing aspect of our study is the significant contribution of nonverified medical professionals in providing advice. Our results show that r/medical_advice relies heavily on contributions from laypersons. This may be due to the lack of a robust verification process on the platform as it relies on the user to self-identify. This trend reflects a broader shift in the digital health information landscape, where community-based knowledge exchange is becoming increasingly predominant over traditional expert-driven models. While this democratization of health information has its advantages, it also inevitably raises concerns about the accuracy and reliability of the advice shared—challenges that have been extensively documented [11].

A key limitation of this study is that 41% of posts lacked user flairs, which leaves a significant portion of users' backgrounds unclear. We acknowledge this as a potential source of bias and recommend future investigations using natural language processing or other linguistic analysis methods to characterize these flairless users, which could enhance our understanding of their information-seeking patterns. In addition, by focusing solely on a single subreddit, we acknowledge that our findings

XSL•FO
RenderX

may not fully represent online health-seeking behaviors across various platforms and communities. The unique characteristics of r/medical_advice—including its user demographics, content moderation practices, and engagement patterns—may not perfectly mirror those of other online health forums. Furthermore, the study's reliance on user-generated categorizations for post flairs and the self-identification of medical professionals introduces potential biases and inaccuracies, which could affect our interpretation of the data [12].

In terms of future directions, numerous opportunities for further research present themselves. Comparative studies across various social media platforms could examine unique trends and user behaviors, offering a more comprehensive picture of online health-seeking patterns. Further investigation into the truthfulness and impact of advice provided by online users remains a critical area of exploration [13]. In addition, understanding the motivations behind patients turning to social media for medical advice, and the consequences of acting on potentially incorrect information, is important to assess these platforms' impact on public health and health care costs.

Conclusion

Our investigation into r/medical_advice uncovers a complex and evolving landscape where online platforms serve as significant avenues for medical inquiry and information exchange. This study highlights the role of both professional and nonprofessional users in shaping these interactions and emphasizes the value they bring. While these platforms may offer invaluable opportunities for information sharing and support, the variable quality and reliability of the advice provided require careful consideration from the professional medical community. There is a clear need for increased participation from verified medical professionals and the implementation of effective moderation policies to ensure that online health forums function as reliable and supportive communities for individuals seeking medical guidance. Such measures are vital to mitigate the risks of misinformation and foster a safer, more informed online health ecosystem.

Conflicts of Interest

AM acknowledges support by the Johns Hopkins Institute for Clinical and Translational Research (ICTR), which is funded in part by Grant Number T32 TR004928 from the National Center for Advancing Translational Sciences (NCATS), a component of the National Institutes of Health (NIH) and NIH Roadmap for Medical Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the Johns Hopkins ICTR, NCATS, or NIH.

References

1. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007;64(9):1853-1862. [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]
2. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
3. McMullan M. Patients using the internet to obtain health information: how this affects the patient–health professional relationship. Patient Educ Couns 2006;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006]
4. Park A, Conway M. Longitudinal changes in psychological states in online health community members: understanding the long-term effects of participating in an online depression community. J Med Internet Res 2017;19(3):e71 [FREE Full text] [doi: 10.2196/jmir.6826] [Medline: 28320692]
5. Lu Y, Zhang P, Liu J, Li J, Deng S. Health-related hot topic detection in online communities using text clustering. PLoS One 2013;8(2):e56221 [FREE Full text] [doi: 10.1371/journal.pone.0056221] [Medline: 23457530]
6. Merolli M, Gray K, Martin-Sanchez F. Therapeutic affordances of social media: emergent themes from a global online survey of people with chronic pain. J Med Internet Res 2014;16(12):e284 [FREE Full text] [doi: 10.2196/jmir.3494] [Medline: 25533453]
7. Baumgartner J, Zannettou S, Keegan B, Squire M, Blackburn J. The Pushshift Reddit dataset. 2020 May 26 Presented at: Proceedings of the Fourteenth International AAAI Conference on Web and Social Media (ICWSM 2020); June 8-11, 2019; Virtual event p. 830-839 URL: https://doi.org/10.1609/icwsm.v14i1.7347 [doi: 10.1609/icwsm.v14i1.7347]
8. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. J Med Internet Res 2011;13(4):e122 [FREE Full text] [doi: 10.2196/jmir.1696]
9. Allen C, Vassilev I, Kennedy A, Rogers A. Long-term condition self-management support in online communities: a meta-synthesis of qualitative papers. J Med Internet Res 2016;18(3):e61 [FREE Full text] [doi: 10.2196/jmir.5260] [Medline: 26965990]
10. Tensen E, van der Heijden JP, Jaspers MWM, Witkamp L. Two decades of teledermatology: current status and integration in national healthcare systems. Curr Derm Rep 2016;5:96-104. [doi: 10.1007/s13671-016-0136-7]
11. Grosberg D, Grinvald H, Reuveni H, Magnezi R. Frequent surfing on social health networks is associated with increased knowledge and patient health activation. J Med Internet Res 2016;18(8):e212 [FREE Full text] [doi: 10.2196/jmir.5832]
12. Yan YY. Online health information seeking behavior in Hong Kong: an exploratory study. J Med Syst 2010;34(2):147-153. [doi: 10.1007/s10916-008-9226-9] [Medline: 20433053]

XSL•FO
RenderX

13. Sillence E, Hardy C, Briggs P. Why don't we trust health websites that help us help each other?: an analysis of online peer-to-peer healthcare. 2013 Presented at: Proceedings of the 5th Annual ACM Web Science Conference (WebSci '13); May 2-4, 2013; Paris, France. [doi: 10.1145/2464464.2464488]

Edited by W Ahmed; submitted 06.01.24; peer-reviewed by J Rowley, R Bidkar, Y Ming; comments to author 09.05.24; revised version received 13.12.24; accepted 13.02.25; published 20.03.25.

Please cite as:
Zhao X, Yang V, Menta A, Blum J, Ranasinghe P
Exploring the Use of Social Media for Medical Problem Solving by Analyzing the Subreddit r/medical_advice: Quantitative Analysis
JMIR Infodemiology 2025;5:e56116
URL: https://infodemiology.jmir.org/2025/1/e56116
doi:10.2196/56116
PMID:

©Xiyu Zhao, Victor Yang, Arjun Menta, Jacob Blum, Padmini Ranasinghe. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 20.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Appropriateness of Web-Based Resources for Home Blood Pressure Measurement and Their Alignment With Guideline Recommendations, Readability, and End User Involvement: Environmental Scan of Web-Based Resources

Eleanor Clapham[1], BMedRes; Dean Picone[1,2], PhD; Samuel Carmichael[1], BMedRes; Carissa Bonner[3], PhD; Niamh Chapman[1,2], PhD

[1]School of Health Sciences, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

[2]Menzies Institute for Medical Research, University of Tasmania, Hobart, Australia

[3]Menzies Centre for Health Policy & Economics, School of Public Health, University of Sydney, Sydney, Australia

Corresponding Author:
Niamh Chapman, PhD
School of Health Sciences
Faculty of Medicine and Health
University of Sydney
Susan Wakil Health Building, University of Sydney, Western Avenue, Camperdown
Sydney, 2050
Australia
Phone: 61 473464291
Email: niamh.chapman@sydney.edu.au

Abstract

Background: High blood pressure (≥140/90 mm Hg) is the most prominent mortality risk factor worldwide. Home blood pressure measurement (HBPM) is recommended for blood pressure (BP) management. HBPM is most effective to improve BP management when delivered with patient education. It is unknown whether web-based resources are appropriate for patient education for HBPM. Patient education should provide accurate, evidence-based information, communicate at an eighth grade reading level, and involve end users in development to meet the needs of adults of all health literacy levels. Using these criteria, this study aimed to determine the appropriateness of web-based HBPM resources.

Objective: This study aimed to determine whether web-based resources are appropriate for HBPM education based on three research questions: (1) Do web-based resources provide evidence-based information that aligns with guideline recommendations? (2) Do they communicate at an appropriate reading level? (3) Do they involve end users in their development?

Methods: An environmental scan of web-based resources for HBPM was conducted on Google (October 2022) using search terms developed with consumers (n=6). Resources were included if they were identified on the first page of the search findings, not paywalled, and in English. Resource appropriateness was appraised based on three criteria: (1) alignment of resource content to 23 recommendations for HBPM from 6 international guidelines, (2) being at an appropriate grade reading level as determined by a health literacy assessment software, and (3) having evidence of end user involvement in resource development.

Results: None of the identified resources (n=24) aligned with all 23 of the guideline recommendations. All resources aligned with the recommendation to measure BP when seated, while few aligned with the recommendation to use a validated BP device (n=9, 38%). All resources exceeded the recommended eighth grade reading level (mean 11.8, range 8.8-17.0) and none reported evidence of patient end user involvement in development.

Conclusions: None of the web-based resources met the criteria for appropriate education to support adults to measure BP at home. Resources should be developed with end users using health literacy tools and multimodal communication methods to ensure they are appropriate to meet the needs of patients.

(JMIR Infodemiology 2025;5:e55248) doi:10.2196/55248

XSL•FO
RenderX

KEYWORDS

readability; online resources; blood pressure guidelines; end user; home blood pressure measurement; patient education; educational resource; self-education; hypertension

Introduction

High blood pressure (BP; hypertension: BP≥140/90 mm Hg) is the leading risk factor for death worldwide [1,2]. High BP can be controlled (<140/90 mm Hg) via medication and lifestyle changes to reduce the risk of heart attack and stroke [3]. Home BP measurement (HBPM) is widely recommended to inform hypertension diagnosis and to monitor the control and ongoing management of BP [4-7]. HBPM provides accurate, standardized BP readings, which have greater prognostic value for cardiovascular disease when performed according to guideline recommendations [8]. Adults who measure BP at home are more engaged in BP management and achieve greater BP control [9]. Further, several studies, including a recent meta-analysis, highlight that HBPM is only effective for improving BP management when accompanied by appropriate patient education on how to measure HBPM accurately and act on BP readings [7,10,11]. However, there is a lack of guidance and standardized resources to provide effective education for HBPM in clinical settings [11,12].

In the absence of effective in-clinic education and with the increased use of telehealth, web-based resources are commonly used by adults who seek health-related information for self-education [13]. In addition, recent work in Australia has shown that >35% of adults would prefer to access information about high BP on the web [14]. With this evolution of health care and patient education delivery, government bodies have emphasized the need for web-based resources to provide health information that is evidence-based and understandable [15]. More specifically, several systematic reviews on eHealth, mobile health, and other digital strategies to improve BP management also suggest a growing need to ensure that appropriate education is available on the web to support adults to undertake HBPM [16-18].

Web-based educational resources that are appropriate to support adults to perform HBPM should deliver evidence-based information in a manner that meets the health literacy and learning needs of most adults. To do this, information should be presented at an eighth grade reading level, with the use of visual aids such as graphics to support understandability [19,20]. The use of co-design methods involving target end users during resource development is an effective method to ensure that resources meet the needs of end users for effective patient education [21-24]. However, previous research has shown that web-based educational resources for cardiovascular disease risk management do not provide appropriate information or meet the usability or readability needs of adults [18,25-29], and co-design involving end users (such as community members and medical professionals) is an underused method during resource development [24]. Due to the importance of patient education for HBPM to achieve improved BP control and patient self-efficacy in BP management, patient education resources for HBPM should be appropriate for use by adults who self-monitor BP.

The aim of this study was to determine whether web-based resources are appropriate for HBPM patient education based on three research questions: (1) Do web-based resources provide evidence-based information that aligns with guideline recommendations? (2) Do they communicate at an appropriate reading level? (3) Do they involve end users in their development?

Methods

Study Design

An environmental scan of web-based resources on HBPM was conducted through a Google search designed to emulate the approach taken by adults with lived experience of HBPM when seeking web-based material about HBPM [27]. The resources were characterized according to basic identifying features such as publishing organization and year. Resources were assessed for alignment with 23 recommendations common across 6 international guidelines that encompass HBPM activities including acquiring the BP measurement device, scheduling and preparing for HBPM, selecting and fitting the cuff, BP measurement conditions, and recording and reporting BP readings (Textbox 1 and Multimedia Appendix 1) [6,30-34]. The grade reading level of the content of the resources was determined using the health literacy assessment software Sydney Health Literacy Lab Editor (SHeLL Editor) [19,35,36]. The recommended reading level for maximum comprehension for adults is eighth grade or below [19]. Involvement of community member and/or medical professional end users in resource development was assessed according to whether this was reported within each resource. Data extraction and resource appraisal were undertaken by 2 independent researchers (EC and SC) using a coding framework hosted on the secure web-based platform REDCap (Research Electronic Data Capture; Vanderbilt University) [37].

XSL•FO
RenderX

Textbox 1. Twenty-three key guideline recommendations for home blood pressure measurement.

Acquiring the blood pressure (BP) measurement device:

- Use a validated BP measurement device for home BP measurement (HBPM).
- Finger cuff BP measurement devices should not be used for HBPM.

Scheduling HBPM:

- On a day that HBPM is being conducted, BP should be measured in the morning and the evening.

Preparing for HBPM:

- Do not measure BP if uncomfortable, stressed, or in pain.
- Measure BP before medication.
- Measure BP before eating or 30 minutes or 2 hours after eating.
- Measure BP after emptying the bladder.
- Measure BP before exercise or 30 minutes after exercise.
- Measure BP before consuming caffeine or after 30 minutes or 1 hour of consuming caffeine.
- Measure BP before smoking or 30 minutes or 1 hour after smoking.
- Have 5 minutes, or at least 5 minutes, of seated rest before measuring BP.

Selecting and fitting the cuff:

- Use an appropriately sized arm cuff for HBPM.
- The arm cuff should fit the arm within the accepted range indicated on the cuff.
- Fit the upper arm BP cuff to a bare arm.

Measurement conditions:

- Measure BP in a room at a comfortable temperature.
- Measure BP with the arm fitted with the cuff supported or supported at the heart level.
- Measure BP in a seated position.
- Measure BP with both feet flat on the floor.
- Measure BP with legs uncrossed.
- Measure BP with back supported.
- Take 2 readings 1 minute apart at each HBPM sitting.

Recording and reporting BP:

- Average the BP readings taken over a 7-day period, discarding the first day.
- Take a copy of home BP readings to a doctor.

Search Strategy

The search engine Google Australia was used to identify web-based resources addressing HBPM. Seven search terms were developed with trained research consumer advisors who have lived experience of BP management and using Google Trends (Multimedia Appendix 2). Consumer advisors (n=6) identified the search engine and the top 5 search terms they would use to find information about HBPM on the web. Google Trends was used to identify the search queries related to the term "home blood pressure measurement," which had the highest probability of use worldwide on Google from January 1, 2012, to October 7, 2022. Search terms suggested by consumer advisors, which also had high probability of use on Google according to Google Trends and were relevant to HBPM were used. Search terms included the following: "How to take your blood pressure," "How to check blood pressure at home," "How to take blood pressure at home," "Home blood pressure monitoring," "How to measure blood pressure at home," "How to monitor blood pressure at home," and "Home blood pressure measurement."

Data Extraction From Web-Based Resources

Data extraction was undertaken independently by 2 investigators (EC and SC) on October 17, 2022 (duplicate search). To avoid potential bias attached to the reviewers' Google history, each reviewer conducted the search using default Google search settings within the incognito browser of Google Chrome and cleared the cache before each search. The results obtained with each search term, which were present on the first page of the

XSL•FO
RenderX

search findings on Google, were exported, excluding advertisements. After completing all searches, the resources extracted by each reviewer across all search terms were combined, and duplicate resources were removed (ie, resources identified across >1 search term).

Inclusion Criteria for Web-Based Resources

HBPM resources were included if they met the following inclusion criteria: (1) they were free to access by the public (eg, no paywalls), (2) they were available in English, and (3) they contained content relevant to HBPM (eg, resource mentions "home blood pressure measurement" or "self-measured blood pressure"; Multimedia Appendix 3). The resources extracted from Google were independently analyzed against inclusion criteria by EC and SC, and discrepancies were resolved by third and fourth independent reviewers based on the same criteria (NC and DP). All resources that met the inclusion criteria were included for analysis. Resources were not excluded due to criteria regarding publication date, publication location, or resource format (ie, video, graphic, or blog).

Appraisal of Web-Based Resources

A coding framework hosted on REDCap was used by EC and SC to independently and systematically appraise resources according to three criteria: (1) alignment of resource information with HBPM guideline recommendations, (2) grade reading level of the content of the resources, and (3) end user involvement in resource development. The REDCap appraisal framework captured resource characteristics (type of publishing organization, authorship, year of publication or last review, and location of publication and languages), communication methods used (categorized as written text, visual, video, or audio), the alignment of resource content against HBPM recommendations, and the grade reading level of the content of the resources (Multimedia Appendix 3). Independent reviewers were trained on how to undertake the search, extract data, and appraise resources. During training, the appraisal data of a subset of resources (n=3) were compared to ensure that the correct process was undertaken by both independent reviewers. All data were captured in a framework housed on REDCap. All content (including audio, text, video, and graphical content) included within each HBPM resource was appraised according to 3 main criteria detailed below. Any discrepancies in appraisal were resolved in adjudication sessions where blinded discrepancies were presented to adjudicators (NC and DP) and resolved via discussion until consensus was reached.

Alignment of Resource Information With HBPM Guideline Recommendations

Twenty-three recommendations that encompass activities for HBPM from 6 international guidelines were used to determine the alignment of resource content to the guidelines (Textbox 1) [6,30-34]. Resource content was marked against each recommendation and categorized as "aligned with" if the resource correctly stated the recommendation, "incorrectly stated" if the resource incorrectly or incompletely stated the recommendation, or "not mentioned" if the resource did not include the recommendation (Multimedia Appendix 3). Where resources "incorrectly stated" a guideline recommendation, the incorrect information provided by the resource was recorded on REDCap (Multimedia Appendix 4).

Grade Reading Level of Resource Text

The grade reading level of each resource was calculated using the SHeLL Editor, which is a health literacy assessment tool that calculates the school grade reading score of text according to the Simple Measure of Gobbledygook, and reports other measures such as complex language, uncommon English words, and the use of passive voice [19,35,36]. All text presented within each resource (including written text, image captions, and audio and video transcripts when available) was entered into the SHeLL Editor and the grade reading level, and associated measures were recorded in the REDCap framework (Multimedia Appendix 3).

End User Involvement in Resource Development

End user involvement in resource development was recorded in the REDCap framework as stated within the resource. End users were defined as adults who seek information to measure BP at home (eg, patient, health consumer, service user, carer, or community advisor) or medical professionals due to their role in delivering education for HBPM to adults or directing adults to educational resources for HBPM [38,39].

Data Analysis

Data were analyzed using Stata (version 17; StataCorp). Resources were assigned an identifying number for analysis and the presentation of results (Multimedia Appendix 5). Categorical data are presented as n (%) values.

Results

Resource Characteristics

Twenty-four resources were included in the study (Multimedia Appendices 5 and 6). Not-for-profit organizations (such as the National Heart Foundation) were the most common type of publishing organization (n=6, 25%) followed by websites (such as Healthline), academic journals, and scientific societies (Table 1). Most resources were communicated via a combination of written text, visual (eg, images), and audio and video communication methods (n=17, 71%), and the remaining resources were communicated by written text only (n=7, 29%; Table 1). Most resources were published in Australia (n=10, 42%) or North America (n=9, 38%), and only 3 (13%) were available in languages other than English.

XSL•FO
RenderX

Table 1. Characteristics of the included resources.

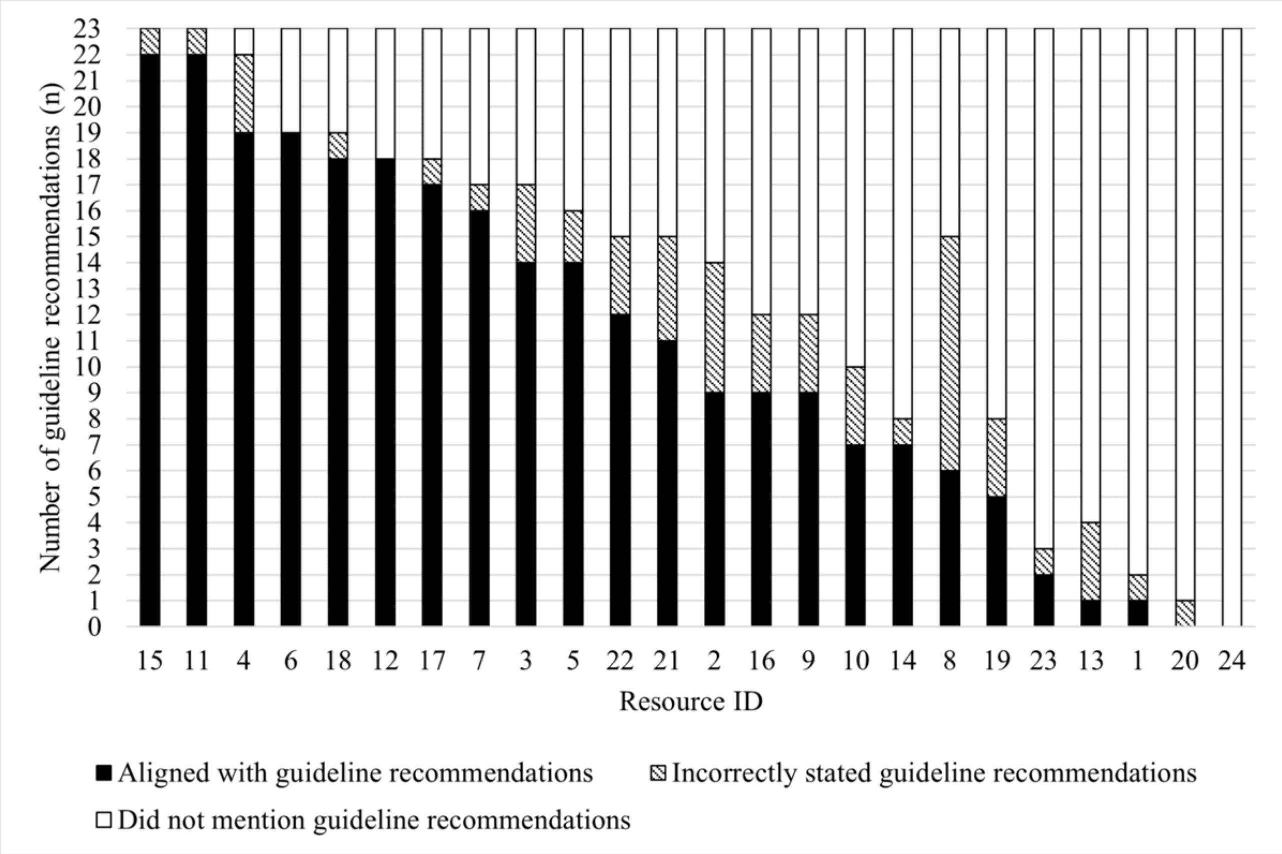
Characteristic	Resources, n (%)
Type of publishing organization	
Commercial entity	3 (12)
Scientific journal	4 (17)
Government body	1 (4)
Not-for-profit organization	5 (21)
Scientific society	5 (21)
Website	4 (17)
Medical research institute	2 (8)
Date of publication or last review	
Last 12 months	3 (13)
1-2 years ago	3 (13)
2-3 years ago	2 (8)
3-4 years ago	1 (4)
4-5 years ago	1 (4)
>5 years ago	4 (16)
Not stated	10 (42)
Location of publication	
Australia	10 (42)
North America	9 (37)
Europe	5 (21)
Communication method	
Written text only	7 (29)
Written text and visual	9 (37)
Written text and video	4 (17)
Written text, visual, and video	3 (12)
Written text, audio, and visual	1 (4)
Languages	
English only	21 (88)
English, Mandarin, and Spanish	1 (4)
English and Spanish	2 (8)

Resource Alignment With HBPM Guideline Recommendations

As shown in Figure 1, none of the resources aligned with all 23 guideline recommendations for HBPM. Almost all (n=22, 92%) of the resources incorrectly stated at least one guideline recommendation. Two (8%) resources did not align with any of the 23 guideline recommendations for HBPM. The alignment of resources with each guideline recommendation is shown in Figure 2, indicating whether the recommendation was "aligned with," "incorrectly stated," and "not mentioned" in each resource. Time- or frequency-bound recommendations were often incorrect within resources. For example, to rest for 5 minutes before measuring BP was incorrectly stated in 25% (n=6) of resources and to take 2 BP readings 1 minute apart at each sitting was incorrect in 46% (n=11) of resources (Figure 2).

XSL•FO
RenderX

Figure 1. Alignment of each resource to 23 guideline recommendations for home blood pressure measurement. Resource content either "aligned with" the HBPM guideline recommendations (black bars), "incorrectly stated" the guideline recommendations (patterned bars), or "did not mention" the guideline recommendations (white bars). The y-axis indicates each of the 23 guideline recommendations, and the x-axis indicates the number of resources (n=24 resources).

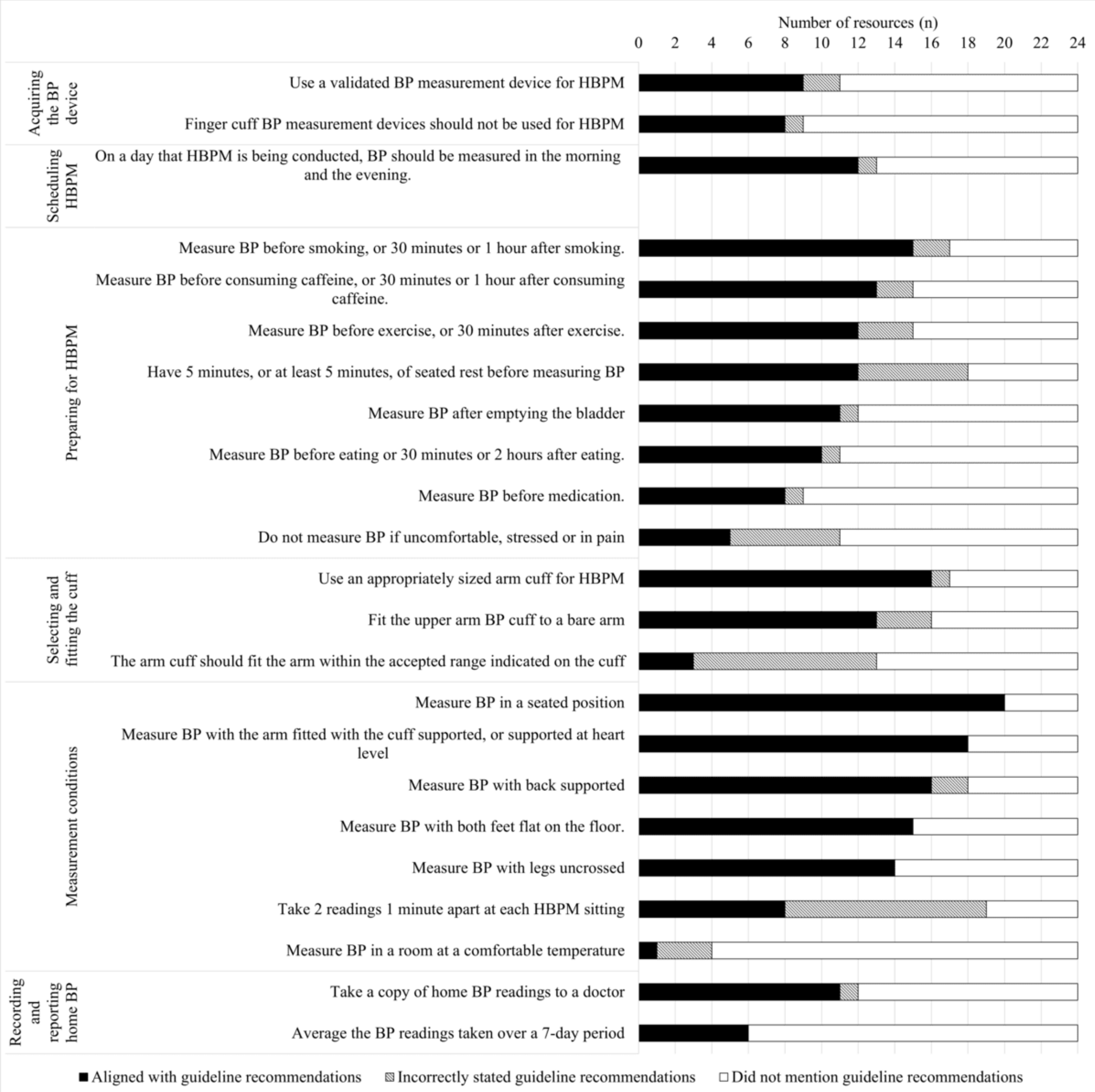


Resources incorrectly stated guideline recommendations because their content was not specific enough to capture the meaning of the guideline recommendation, provided contradictory advice, or stated an alternate rest period, number of measurements, frequency, duration, or other numeric parameters to the guideline recommendations (Multimedia Appendix 4). For example, rather than stating the recommendation to "have five minutes [or at least five minutes] of seated rest before measuring BP," resources that incorrectly stated this recommendation said to "rest for 15 minutes" (resource ID 10) or "rest quietly and wait about one to two minutes before taking another measurement" (resource ID 19). In addition, rather than stating the recommendation to "take two readings one minute apart at each HBPM sitting," a resource that incorrectly stated this recommendation said "if you get a reading that is slightly or moderately higher than normal, take your blood pressure a few more times" (resource ID 6).

Resource alignment to guideline recommendations according to the publishing organization is outlined in Figure 3. Resources published by scientific journals, scientific societies, and not-for-profit organizations aligned with a higher number of HBPM guideline recommendations (14 resources; median 16.5, range 2-22 recommendations) than resources published by websites, commercial entities, and medical research institutes (9 resources; median 6.5, range 0-12 recommendations; Figure 3).

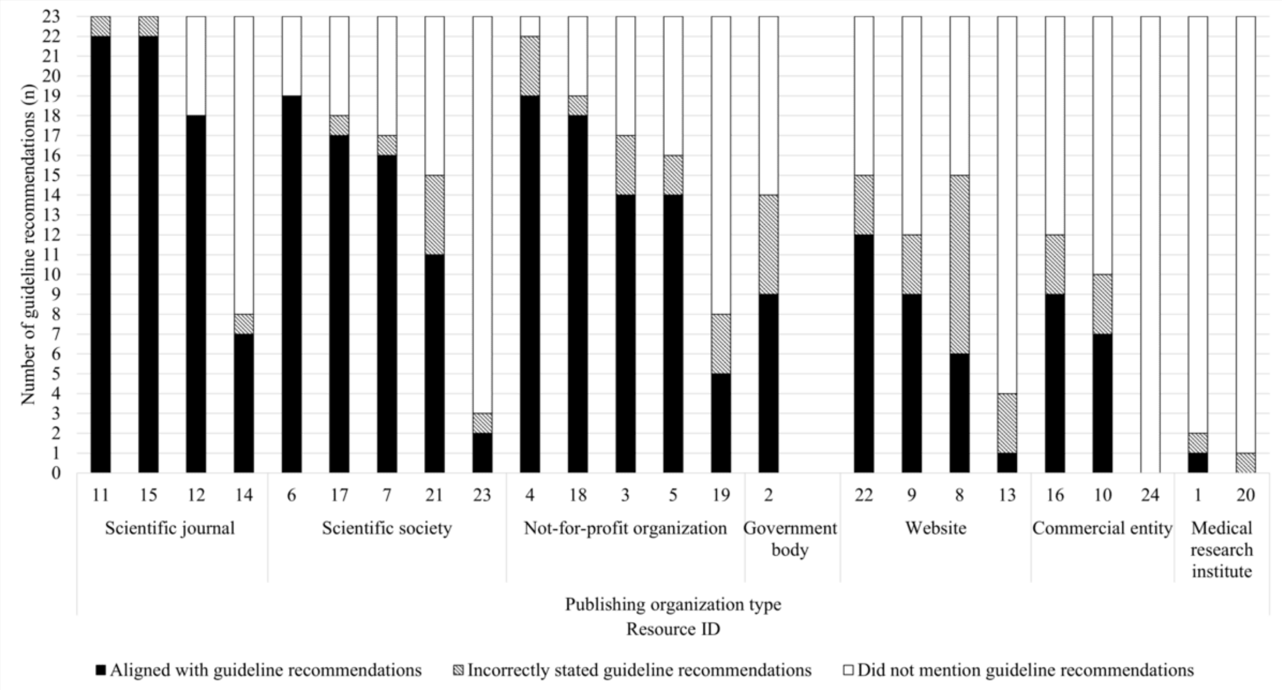
XSL•FO
RenderX

Figure 2. Alignment of all resources to each of the 23 guideline recommendations for key home blood pressure measurement (HBPM) activities. Resource content either "aligned with" the HBPM guideline recommendations (black bars), "incorrectly stated" the guideline recommendations (patterned bars), or "did not mention" the guideline recommendations (white bars). The x-axis indicates the number of HBPM resources. BP: blood pressure.



XSL•FO
RenderX

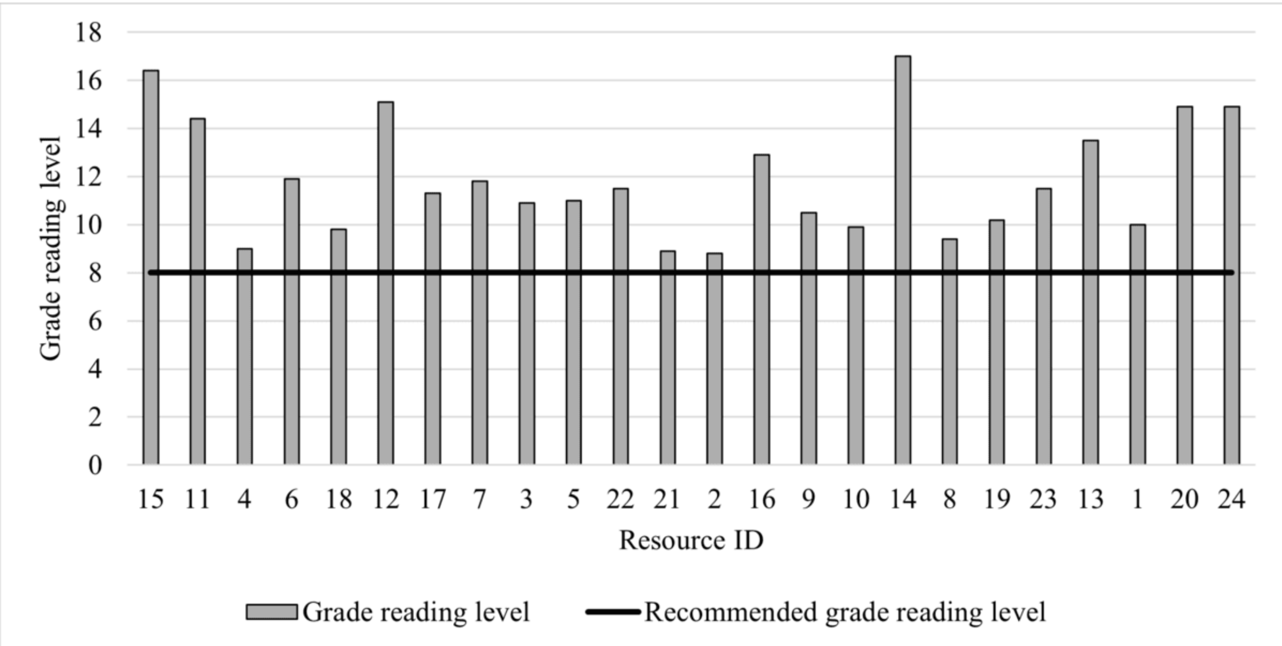
Figure 3. Resource alignment to home blood pressure measurement (HBPM) guideline recommendations according to the type of publishing organization. Resource alignment was determined by appraising resource content against 23 guideline recommendations of core HBPM activities. Resource content either "aligned with" the HBPM guideline recommendations (black bars), "incorrectly stated" the guideline recommendations (patterned bars), or "did not mention" the guideline recommendations (white bars).



Grade Reading Level of Resource Text

All resources exceeded the recommended eighth grade reading level (grade reading level: mean 11.8, range: 8.8-17.0; Figure 4). The grade reading level of resources did not differ according to the level of alignment with HBPM guideline recommendations or communication methods used (Figures 4 and 5). Resources presented through written text only (n=7) had the highest average grade reading level (grade reading level: mean 12.9, range 10.5-16.4; Figure 5). Resources published by scientific journals had the highest average grade reading level (n=4; grade reading level: mean 16.5, range 11.9-17), compared to government bodies (n=1; grade reading level: mean 8.8) and not-for-profit organizations (n=5; grade reading level: mean 10.2, range 9-10.9; Figure 6), which had the lowest average grade reading levels. Multimedia Appendix 7 shows the characteristics of the resource text that contributed to the grade reading level score.

Figure 4. Grade reading level of web-based home blood pressure measurement (HBPM) resources. Resource grade reading level (y-axis) is presented in the order of resources from the highest (left) to the lowest alignment (right) with HBPM guideline recommendations. The grade reading level (y-axis) of resource content was calculated by inputting all resource content into the Sydney Health Literacy Lab Editor.



XSL•FO
RenderX

Figure 5. Resource grade reading level according to the communication method. The average grade reading level (y-axis) of resources according to communication methods used in the resource (x-axis). The grade reading level (y-axis) of resource content was calculated by inputting all resource content into the Sydney Health Literacy Lab Editor.

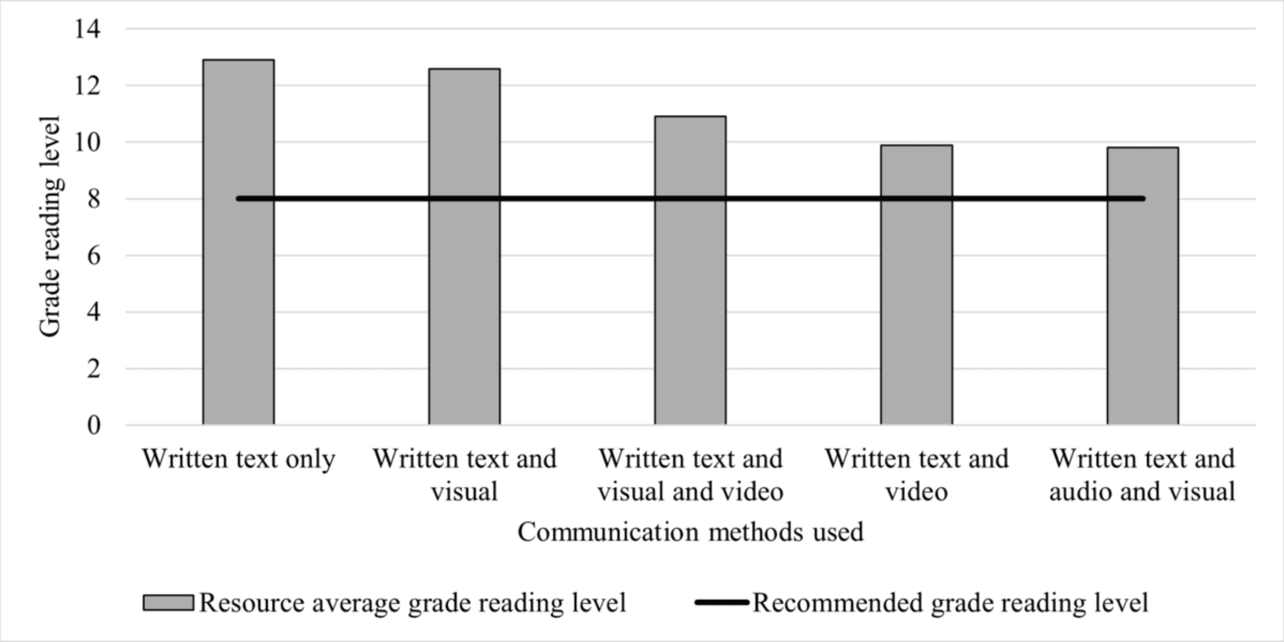
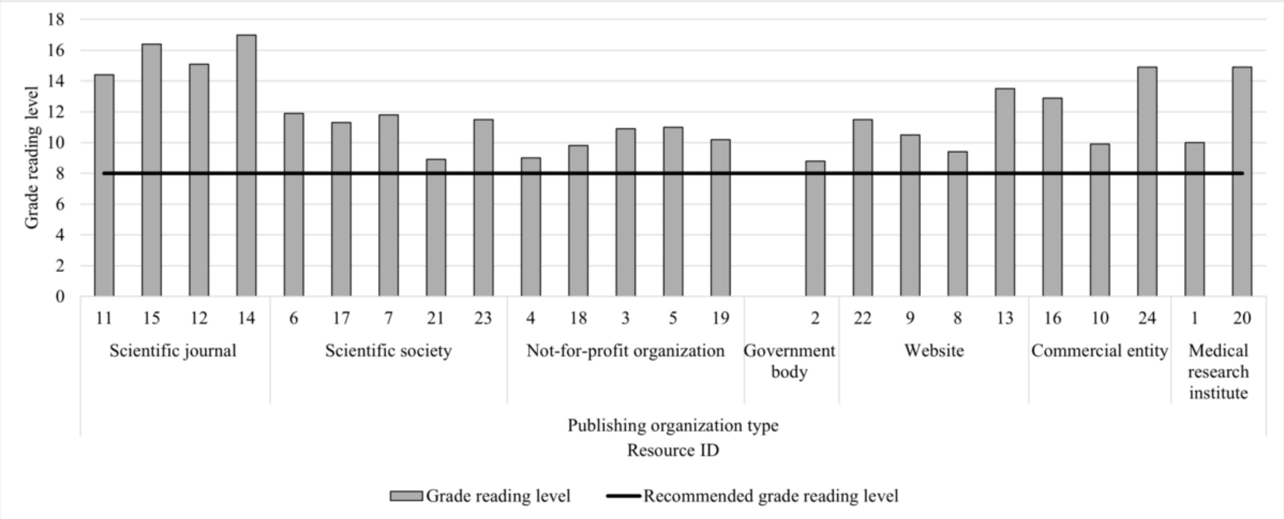


Figure 6. Resource grade reading level according to the type of publishing organization. Grade reading level was calculated by inputting all resource content into the Sydney Health Literacy Lab Editor.



End User Involvement in Resource Development

None of the resources reported involving adults (such as patients, health consumers, or carers) during resource development. Medical professional involvement was reported in 5 (21%) resources. Resources with and those without medical professional involvement during development had similar alignment with HBPM guideline recommendations and grade reading levels (Figures 7 and 8).

XSL•FO
RenderX

Figure 7. Resource alignment to guideline home blood pressure measurement (HBPM) recommendations according to medical professional involvement during resource development. Resource alignment was determined by appraising resource content against 23 guideline recommendations of core HBPM activities. Resource content either "aligned with" the HBPM guideline recommendations (black bars), "incorrectly stated" the guideline recommendations (patterned bars), or "did not mention" the guideline recommendations (white bars).

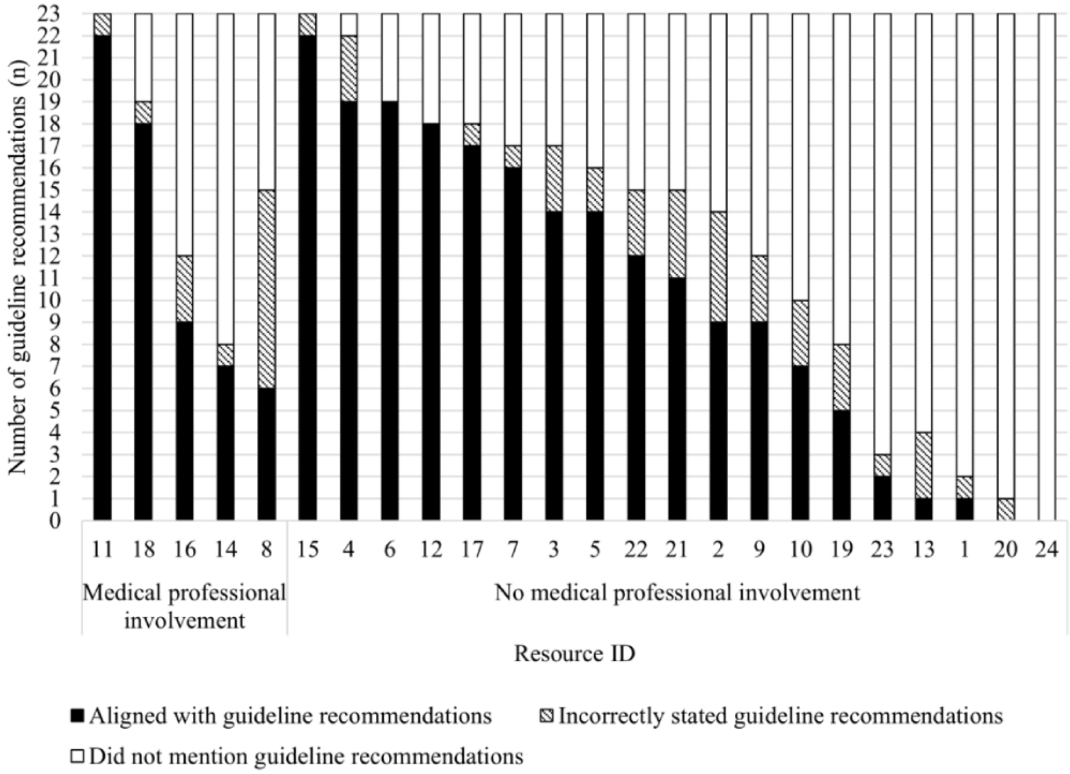
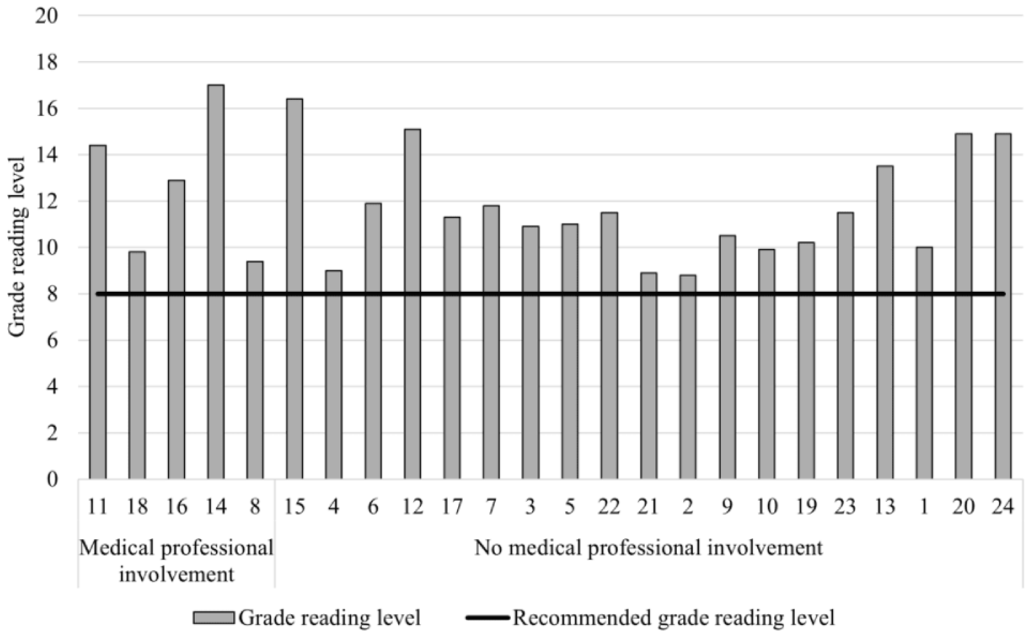


Figure 8. Resource grade reading level according to medical professional involvement during resource development. Grade reading level was calculated by inputting all resource content into the Sydney Health Literacy Lab Editor. Resources are presented in order of the highest (left) to the lowest alignment (right) to home blood pressure measurement guideline recommendations.



Discussion

Principal Findings

This study has demonstrated that web-based resources may not be appropriate to fully support adults to undertake high-quality HBPM because none of them provided sufficient guideline information or communicated at an appropriate reading level. Using methodology that emulated the search strategy of adults with lived experience of HBPM to identify web-based resources, we identified that none of the resources correctly stated all key guideline recommendations for HBPM, and most resources included information that was incorrect according to guideline

XSL•FO
RenderX

recommendations due to incorrectly stating time- and frequency-bound recommendations. The findings of this study highlight the need to design educational materials for key BP management behaviors such as HBPM, which are appropriate for adults who self-monitor BP.

Calls to action on hypertension in the United States and Australia highlight the importance of empowering patients who perform HBPM to improve and monitor BP control [40,41]. At a global level, the World Heart Federation Hypertension Roadmap highlighted that appropriate patient education is an important strategy to improve BP control [20]. Existing research supports this by illustrating that patient education delivered with HBPM improves BP control outcomes [10] and adherence to recommendations for HBPM [11,42]. However, the results of this study suggest that web-based resources may not be appropriate to educate adults about HBPM as the identified resources did not provide guideline aligning information required to support HBPM in a manner that met adult reading needs.

This study found that time- and frequency-based HBPM recommendations, such as the number of BP measurements to take per sitting and the associated rest periods, were incorrectly stated within the most resources, while the recommendation to measure BP when seated was accurately communicated in the most resources. Interestingly, a recent study on BP guideline recommendations followed by adults who measure BP at home found that time-bound recommendations were adhered to by the lowest number of adults, while the recommendation to measure BP when seated was performed by the highest number of adults. Additionally, adults who reported to have previously sought information to support HBPM did so via web-based sources; however, these adults did not perform higher-quality HBPM than those who had not used web-based information to support HBPM [43]. Altogether, these findings highlight that current educational resources are not appropriate to support adults to measure BP at home as recommended by guidelines and illustrate a possible synergy between the inaccurate information delivered within web-based HBPM resources and the practice of adults when measuring BP at home. This emphasizes the need for web-based HBPM resources to accurately and clearly deliver guideline recommendations to enable proper HBPM practice among adults, which is an important behavior for BP management.

This study found that resources published in scientific journals, scientific societies, and not-for-profit organizations stated more guideline recommendations correctly than resources published by websites, commercial entities, and medical research institutes. This suggests that some organizations and resource developers may have low awareness of or access to guideline recommendations for HBPM and may not recognize the importance of standardized BP measurement practices for achieving and maintaining BP control. International BP guidelines should consider the importance of using consistent, unambiguous, and plain language for HBPM recommendations to support the accurate translation of recommendations into educational resources for HBPM. To ensure that guideline information is disseminated to the general public, guideline developers should share guidelines with organizations that publish health information on the web and partner with peak organizations to enable resource developers from outside of the scientific and clinical community to create guideline-informed, evidence-based resources.

Apart from correctly delivering evidence-based guideline information, HBPM resources must deliver information in a format accessible for adults to achieve effective education. Previous evidence has shown that web-based health information is not appropriate to inform patient decisions surrounding cardiovascular disease because the reading level is too high and the information is not adapted to meet the learning needs of adult patients [17,26,27,44]. This is consistent with the findings of our study where all resources were at a reading level that was too high (≥8 grade) for adult comprehension and over a quarter (n=7, 29%) of resources only presented information via written text only.

Strategies to deliver patient education that meet the literacy levels of adult patients should be implemented to ensure that educational resources can support adults to perform key cardiovascular disease risk management behaviors such as HBPM. As highlighted by the World Heart Federation Hypertension Roadmap, the delivery of education via graphical means is a more appropriate communication method to meet the needs of those with lower health literacy levels [20]. This is supported by the findings of this study, where web-based resources with multimodal communication methods achieved a lower average grade reading level than those that communicated via written text alone. Supporting audiovisuals, such as graphs, diagrams, images, videos, and the read-aloud function should be used to aid understandability, comprehensibility, and actionability of web-based health information. Additionally, the use of readability and grammar editing tools when developing resources may help to ensure that resource information is presented at a grade reading level that is appropriate to all adults, and resources such as the Agency for Healthcare Research and Quality's Health Literacy Universal Precautions Toolkit may provide actionable methods to maximize the understandability of patient education strategies [45]. Finally, artificial intelligence (AI) could be used to tailor web-based information to meet patient literacy needs, selectively deliver information most relevant to the unique information needs of patients, and support chat box functions enabling adults to ask clarifying questions [46]. However, although AI-generated content is accurate and retains key meaning, caution should be exercised to ensure that information used by AI generators is sourced from guidelines.

Direct end user involvement in resource development is an increasingly well-recognized strategy to ensure that health products and services, including health information, meet end user needs to deliver quality care and education [21-24]. However, end user involvement is not commonplace in resource development [24], which is consistent with the findings of our study. Although some resources of this study involved medical professionals in their development, this did not improve the resource grade reading level or the number of correctly stated guideline recommendations. While medical professionals play a central role in patient education, they may not be aware or have sufficient resources to meet the health literacy needs of

XSL•FO
RenderX

all patients [47-49]. Additionally, some medical professionals have general distrust in BP guidelines [50] and do not use current guidelines recommendations for HBPM in clinical practice, such as the recommendation to use different cutoffs for a hypertension diagnosis using in-clinic versus at-home BP readings [12,51]. This further highlights the need for adults with lived experience of BP management to be involved in resource development to identify unfamiliar medical jargon, recommend culturally and linguistically sensitive adaptations, and advise on the appropriate use of images. For existing resources, such as those identified in this study, adults could be involved in appraising these resources to identify how they could better meet the needs of adults seeking information on HBPM. Implementing the strategies suggested would ensure that information provided by web-based resources is suitable for use by all end users to support high-quality HBPM among adults.

Strengths and Limitations

A strength of this study was the involvement of consumer advisors in the development of the search strategy to emulate the experience of adults seeking information for HBPM. In addition, a rigorous framework analysis approach was used by 2 independent researchers for resource identification and appraisal. This study was strengthened by the guideline-informed appraisal process; however, guideline recommendations included in this analysis were not exhaustive of all recommendations for HBPM due to inconsistency in recommendations across guidelines. The incognito mode was used to eliminate the impact of cookies and search history unique to the reviewer. However, as a result of using default Google search engine settings and including only the first page of search results, some web-based HBPM resources would have been missed. The location at which this study was conducted has likely impacted the search results, as 42% of included resources were from Australia. This suggests that the resources that an adult seeking HBPM information is exposed to depends on the location from which the search is conducted. This method should be replicated in other locations to assess resources that may not have been identified in this study. The scope of this study was narrow, with highly specific appraisal criteria used to evaluate resources. Other important considerations of web-based resources such as ease of access and usability should be included in future studies for a more complete understanding of resource appropriateness to support HBPM. Further, given the proliferation in use of AI, mobile health, and eHealth for health interventions and patient education, HBPM resources found on these information sources should also be appraised for appropriateness.

Conclusion

This study found that the web-based resources identified herein are not appropriate to fully support adults to measure their BP at home according to HBPM guideline recommendations. None of the resources identified provided sufficient guideline information to support adults to perform high-quality HBPM, were presented at an appropriate reading level, or involved end users in their design. Resources that deliver health information should use strategies such as the use of multimodal communication methods, literacy editor tools, and co-design methods with adult end users to ensure that the information delivered is appropriate to support adults. Due to the recognized importance of effective education in achieving standardized HBPM and improving BP control, creating appropriate educational resources for key BP management behaviors such as HBPM should be considered a priority.

Acknowledgments

We sincerely thank John Stevens, Carol Batt, Michael Whittle, Josephine Castillo, Lesley Hall, and Heather Thurstans for their contributions to this work as consumer advisors with lived experience. NC was supported by an Australian Heart Foundation Postdoctoral Research Fellowship for the duration of this work. CB was supported by a Heart Foundation Future Leader Fellowship. DP was supported by a National Health and Medical Research Council Investigator Grant (GNT2018077) and was an Honorary Future Leader Fellow of the Heart Foundation of Australia (106618).

Data Availability

The data that support the findings of this study are available from the corresponding author (NC) upon reasonable request. NC had full access to all data in the study and takes responsibility for its integrity and the data analysis.

Conflicts of Interest

CB is a director of Health Literacy Solutions, a company set up to fund the future development of the SHeLL Editor.

Multimedia Appendix 1

The guideline recommendations used for resource appraisal.

[DOCX File , 18 KB - infodemiology_v5i1e55248_app1.docx]

Multimedia Appendix 2

Development of the search strategy. Consumer advisors (n=6) and Google Trends data were used to develop the search strategy to identify web-based home blood pressure measurement resources. Search terms suggested by consumer advisors that also had a high probability of use on Google (January 1, 2012, to October 7, 2022) were used.

XSL•FO
RenderX

[DOCX File , 20 KB - infodemiology_v5i1e55248_app2.docx]

Multimedia Appendix 3

Home blood pressure measurement (HBPM) resource eligibility and appraisal form. HBPM resources were appraised for eligibility and appraised for alignment to HBPM guidelines, grade reading level, and end user involvement in development according to the questions in the form house on REDCap (Research Electronic Data Capture).

[DOCX File , 27 KB - infodemiology_v5i1e55248_app3.docx]

Multimedia Appendix 4

Incorrectly stated guideline recommendations. Resource information marked as "incorrectly stated" during resource appraisal "step 1 alignment to guideline recommendations" was recorded in the REDCap (Research Electronic Data Capture) appraisal framework.

[DOCX File , 22 KB - infodemiology_v5i1e55248_app4.docx]

Multimedia Appendix 5

Resources included in the study.

[DOCX File , 18 KB - infodemiology_v5i1e55248_app5.docx]

Multimedia Appendix 6

Search strategy and results.

[DOCX File , 27 KB - infodemiology_v5i1e55248_app6.docx]

Multimedia Appendix 7

Sydney Health Literacy Lab Editor results of home blood pressure measurement resources. All text within resources, including written text and transcripts of audio and video material, was input to the Sydney Health Literacy Lab Editor. Grade reading level was calculated using the Simple Measure of Gobbledygook method.

[DOCX File , 21 KB - infodemiology_v5i1e55248_app7.docx]

References

1. Xu X, Islam SMS, Schlaich M, Jennings G, Schutte AE. The contribution of raised blood pressure to all-cause and cardiovascular deaths and disability-adjusted life-years (DALYs) in Australia: Analysis of global burden of disease study from 1990 to 2019. PLoS One 2024;19(2):e0297229 [FREE Full text] [doi: 10.1371/journal.pone.0297229] [Medline: 38381709]
2. Nguyen TN, Chow CK. Global and national high blood pressure burden and control. Lancet 2021 Sep 11;398(10304):932-933. [doi: 10.1016/S0140-6736(21)01688-3] [Medline: 34450082]
3. Blood Pressure Lowering Treatment Trialists' Collaboration. Pharmacological blood pressure lowering for primary and secondary prevention of cardiovascular disease across different levels of blood pressure: an individual participant-level data meta-analysis. Lancet 2021 May 01;397(10285):1625-1636 [FREE Full text] [doi: 10.1016/S0140-6736(21)00590-0] [Medline: 33933205]
4. Shimbo D, Artinian NT, Basile JN, Krakoff LR, Margolis KL, Rakotz MK, American Heart Associationthe American Medical Association. Self-measured blood pressure monitoring at home: a joint policy statement from the American Heart Association and American Medical Association. Circulation 2020 Jul 28;142(4):e42-e63. [doi: 10.1161/CIR.0000000000000803] [Medline: 32567342]
5. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018 Jun;71(6):e13-e115 [FREE Full text] [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]
6. Stergiou GS, Palatini P, Parati G, O'Brien E, Januszewicz A, Lurbe E, European Society of Hypertension Councilthe European Society of Hypertension Working Group on Blood Pressure MonitoringCardiovascular Variability. 2021 European Society of Hypertension practice guidelines for office and out-of-office blood pressure measurement. J Hypertens 2021 Jul 01;39(7):1293-1302. [doi: 10.1097/HJH.0000000000002843] [Medline: 33710173]
7. Mancusi C, Bisogni V, Maloberti A, Manzi MV, Visco V, Biolcati M, ACCURAPRESS investigators, ACCURAPRESS Investigator:. Accuracy of home blood pressure measurement: the ACCURAPRESS study - a proposal of Young Investigator Group of the Italian Hypertension Society (Società Italiana dell'Ipertensione Arteriosa). Blood Press 2022 Dec 02;31(1):297-304 [FREE Full text] [doi: 10.1080/08037051.2022.2137461] [Medline: 36321781]

XSL•FO
RenderX

8. Ward AM, Takahashi O, Stevens R, Heneghan C. Home measurement of blood pressure and cardiovascular disease: systematic review and meta-analysis of prospective studies. J Hypertens 2012 Mar;30(3):449-456. [doi: 10.1097/HJH.0b013e32834e4aed] [Medline: 22241136]
9. Tran KC, Mak M, Kuyper LM, Bittman J, Mangat B, Lindsay H, et al. Home blood pressure telemonitoring technology for patients with asymptomatic elevated blood pressure discharged from the emergency department: pilot study. JMIR Form Res 2024 Jan 30;8:e49592 [FREE Full text] [doi: 10.2196/49592] [Medline: 38111177]
10. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: a systematic review and individual patient data meta-analysis. PLoS Med 2017 Sep;14(9):e1002389 [FREE Full text] [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
11. Dymek J, Gołda A, Polak W, Lisowski B, Skowron A. Effect of education for hypertensive patients with correctly performed self-blood pressure monitoring (SBPM). Pharmacy (Basel) 2019 Jun 27;7(3) [FREE Full text] [doi: 10.3390/pharmacy7030075] [Medline: 31252513]
12. McManus RJ, Wood S, Bray EP, Glasziou P, Hayen A, Heneghan C, et al. Self-monitoring in hypertension: a web-based survey of primary care physicians. J Hum Hypertens 2014 Feb 04;28(2):123-127. [doi: 10.1038/jhh.2013.54] [Medline: 23823583]
13. Senecal C, Mahowald M, Lerman L, Lopes-Jimenez F, Lerman A. Increasing utility of Google Trends in monitoring cardiovascular disease. Digit Health 2021;7:20552076211033420 [FREE Full text] [doi: 10.1177/20552076211033420] [Medline: 34873449]
14. Chapman N, Marques FZ, Picone DS, Adji A, Broughton BRS, Dinh QN, et al. Content and delivery preferences for information to support the management of high blood pressure. J Hum Hypertens 2024 Jan 10;38(1):70-74 [FREE Full text] [doi: 10.1038/s41371-022-00723-8] [Medline: 35948655]
15. Devine T, Broderick J, Harris LM, Wu H, Hilfiker SW. Making quality health websites a national public health priority: toward quality standards. J Med Internet Res 2016 Aug 02;18(8):e211 [FREE Full text] [doi: 10.2196/jmir.5999] [Medline: 27485512]
16. Mackert M, Champlin SE, Holton A, Muñoz II, Damásio MJ. eHealth and health literacy: a research methodology review. J Comput-Mediat Comm 2014 Apr 12;19(3):516-528. [doi: 10.1111/jcc4.12044]
17. Bonner C, Patel P, Fajardo MA, Zhuang R, Trevena L. Online decision aids for primary cardiovascular disease prevention: systematic search, evaluation of quality and suitability for low health literacy patients. BMJ Open 2019 Mar 13;9(3):e025173 [FREE Full text] [doi: 10.1136/bmjopen-2018-025173] [Medline: 30872547]
18. Bonner C, Fajardo MA, Hui S, Stubbs R, Trevena L. Clinical validity, understandability, and actionability of online cardiovascular disease risk calculators: systematic review. J Med Internet Res 2018 Feb 01;20(2):e29 [FREE Full text] [doi: 10.2196/jmir.8538] [Medline: 29391344]
19. Ayre J, Bonner C, Muscat DM, Dunn AG, Harrison E, Dalmazzo J, et al. Multiple automated health literacy assessments of written health information: development of the SHeLL (Sydney Health Literacy Lab) Health Literacy Editor v1. JMIR Form Res 2023 Feb 14;7:e40645 [FREE Full text] [doi: 10.2196/40645] [Medline: 36787164]
20. Jeemon P, Séverin T, Amodeo C, Balabanova D, Campbell N, Gaita D, et al. World Heart Federation Roadmap for Hypertension - a 2021 update. Glob Heart 2021;16(1):63 [FREE Full text] [doi: 10.5334/gh.1066] [Medline: 34692387]
21. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17 [FREE Full text] [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
22. Nilsen ES, Myrhaug HT, Johansen M, Oliver S, Oxman AD. Methods of consumer involvement in developing healthcare policy and research, clinical practice guidelines and patient information material. Cochrane Database Syst Rev 2006 Jul 19;2006(3):CD004563 [FREE Full text] [doi: 10.1002/14651858.CD004563.pub2] [Medline: 16856050]
23. Moise N, Cené CW, Tabak RG, Young DR, Mills KT, Essien UR, et al. Leveraging implementation science for cardiovascular health equity: a scientific statement from the American Heart Association. Circulation 2022 Nov 08;146(19). [doi: 10.1161/cir.0000000000001096]
24. Talevski J, Kulnik S, Jessup R, Falls R, Cvetanovska N, Beauchamp A. Use of co-design methodology in the development of cardiovascular disease secondary prevention interventions: A scoping review. Health Expect 2023 Feb;26(1):16-29 [FREE Full text] [doi: 10.1111/hex.13633] [Medline: 36366855]
25. Fitzmaurice D, Adams J. A systematic review of patient information leaflets for hypertension. J Hum Hypertens 2000 Apr 30;14(4):259-262. [doi: 10.1038/sj.jhh.1001003] [Medline: 10805051]
26. McKinn S, Batcup C, Cornell S, Freeman N, Doust J, Bell KJL, et al. Decision support tools for coronary artery calcium scoring in the primary prevention of cardiovascular disease do not meet health literacy needs: a systematic environmental scan and evaluation. Int J Environ Res Public Health 2022 Sep 16;19(18):11705 [FREE Full text] [doi: 10.3390/ijerph191811705] [Medline: 36141978]
27. Kapoor K, George P, Evans MC, Miller WJ, Liu SS. Health literacy: readability of ACC/AHA online patient education material. Cardiology 2017;138(1):36-40. [doi: 10.1159/000475881] [Medline: 28571004]
28. Fajardo M, Weir K, Bonner C, Gnjidic D, Jansen J. Availability and readability of patient education materials for deprescribing: an environmental scan. Br J Clin Pharmacol 2019 Jul;85(7):1396-1406 [FREE Full text] [doi: 10.1111/bcp.13912] [Medline: 30848837]

XSL•FO
RenderX

29. Daraz L, Morrow AS, Ponce OJ, Farah W, Katabi A, Majzoub A, et al. Readability of online health information: a meta-narrative systematic review. Am J Med Qual 2018;33(5):487-492. [doi: 10.1177/1062860617751639] [Medline: 29345143]
30. Muntner P, Shimbo D, Carey RM, Charleston JB, Gaillard T, Misra S, et al. Measurement of blood pressure in humans: a scientific statement from the American Heart Association. Hypertension 2019 May;73(5). [doi: 10.1161/hyp.0000000000000087]
31. Rabi DM, McBrien KA, Sapir-Pichhadze R, Nakhla M, Ahmed SB, Dumanski SM, et al. Hypertension Canada's 2020 comprehensive guidelines for the prevention, diagnosis, risk assessment, and treatment of hypertension in adults and children. Can J Cardiol 2020 May;36(5):596-624. [doi: 10.1016/j.cjca.2020.02.086] [Medline: 32389335]
32. Sharman JE, Howes FS, Head GA, McGrath BP, Stowasser M, Schlaich M, et al. Home blood pressure monitoring: Australian Expert Consensus Statement. J Hypertens 2015 Sep;33(9):1721-1728 [FREE Full text] [doi: 10.1097/HJH.0000000000000673] [Medline: 26136205]
33. Umemura S, Arima H, Arima S, Asayama K, Dohi Y, Hirooka Y, et al. The Japanese Society of Hypertension guidelines for the management of hypertension (JSH 2019). Hypertens Res 2019 Sep 2;42(9):1235-1481. [doi: 10.1038/s41440-019-0284-9] [Medline: 31375757]
34. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357. [doi: 10.1161/hypertensionaha.120.15026]
35. Ayre J, Muscat D, Bonner C, Mouwad D, Dalmazzo J, Eliza H. Sydney Health Literacy Lab (SHLL) Health Literacy Editor. The University of Sydney. 2021. URL: https://ses.library.usyd.edu.au/handle/2123/24642 [accessed 2025-02-26]
36. Ayre J, Bonner C, Muscat DM, Cvejic E, Mac O, Mouwad D, et al. Online plain language tool and health information quality: a randomized clinical trial. JAMA Netw Open 2024 Oct 01;7(10):e2437955 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.37955] [Medline: 39378036]
37. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
38. Barnason S, White-Williams C, Rossi LP, Centeno M, Crabbe DL, Lee KS, et al. Evidence for therapeutic patient education interventions to promote cardiovascular patient self-management: a scientific statement for healthcare professionals from the American Heart Association. Circ: Cardiovascular Quality and Outcomes 2017 Jun;10(6). [doi: 10.1161/hcq.0000000000000025]
39. Brunström M, Ng N, Dahlström J, Lindholm LH, Lönnberg G, Norberg M, et al. Association of physician education and feedback on hypertension management with patient blood pressure and hypertension control. JAMA Netw Open 2020 Jan 03;3(1):e1918625 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.18625] [Medline: 31913490]
40. Substance Abuse and Mental Health Services Administration, Office of the Surgeon General. The Surgeon General's Call to Action to Control Hypertension. Washington, DC: US Department of Health and Human Services; 2020.
41. Schutte A, Webster R, Jennings G, Schlaich M. Uncontrolled blood pressure in Australia: a call to action. Med J Aust 2022 Feb 07;216(2):61-63 [FREE Full text] [doi: 10.5694/mja2.51350] [Medline: 34865237]
42. Mondal S, Mondal H, Dutta R, Pal A, Acharya SS, Baidya C. Competency in home blood pressure monitoring and effect of training program on competency. JCDR 2018. [doi: 10.7860/jcdr/2018/32111.11427]
43. Clapham E, Picone DS, Carmichael S, Stergiou GS, Campbell NR, Stevens J, et al. Home blood pressure measurements are not performed according to guidelines and standardized education is urgently needed. Hypertension 2025 Jan 25;82(1):149-159. [doi: 10.1161/hypertensionaha.124.23678]
44. Niksadat N, Rakhshanderou S, Negarandeh R, Ramezankhani A, Farahani AV, Ghaffari M. Concordance of the cardiovascular patient education with the principles of Andragogy model. Arch Public Health 2022 Jan 04;80(1):4 [FREE Full text] [doi: 10.1186/s13690-021-00763-5] [Medline: 34983640]
45. Brach C. AHRQ Health Literacy Universal Precautions Toolkit, 3rd Edition. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/sites/default/files/publications2/files/health-literacy-universal-precautions-toolkit-3rd-edition.pdf [accessed 2025-02-26]
46. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J 2019 Jun;6(2):94-98 [FREE Full text] [doi: 10.7861/futurehosp.6-2-94] [Medline: 31363513]
47. Fletcher BR, Hinton L, Bray EP, Hayen A, Hobbs FR, Mant J, et al. Self-monitoring blood pressure in patients with hypertension: an internet-based survey of UK GPs. Br J Gen Pract 2016 Aug 30;66(652):e831-e837. [doi: 10.3399/bjgp16x687037]
48. Khatib R, Schwalm J, Yusuf S, Haynes RB, McKee M, Khan M, et al. Patient and healthcare provider barriers to hypertension awareness, treatment and follow up: a systematic review and meta-analysis of qualitative and quantitative studies. PLoS One 2014;9(1):e84238 [FREE Full text] [doi: 10.1371/journal.pone.0084238] [Medline: 24454721]
49. Kobayashi M, Obara T, Ohkubo T, Fukunaga H, Satoh M, Metoki H, et al. Practice and awareness of physicians regarding casual-clinic blood pressure measurement in Japan. Hypertens Res 2010 Sep;33(9):960-964. [doi: 10.1038/hr.2010.89] [Medline: 20535112]

XSL•FO
RenderX

50. Howes F, Hansen E, Williams D, Nelson M. Barriers to diagnosing and managing hypertension - a qualitative study in Australian general practice. Aust Fam Physician 2010 Jul;39(7):511-516. [Medline: 20628667]
51. Hyman DJ, Pavlik VN. Self-reported hypertension treatment practices among primary care physicians: blood pressure thresholds, drug choices, and the role of guidelines and evidence-based medicine. Arch Intern Med 2000 Aug 14;160(15):2281-2286. [doi: 10.1001/archinte.160.15.2281] [Medline: 10927724]

Abbreviations

AI: artificial intelligence
BP: blood pressure
HBPM: home blood pressure measurement
REDCap: Research Electronic Data Capture
SHeLL Editor: Health Literacy Lab Editor

Edited by T Mackey; submitted 06.12.23; peer-reviewed by L Daraz, J Yang, SS Palve; comments to author 11.09.24; revised version received 29.11.24; accepted 11.01.25; published 03.04.25.

Please cite as:
Clapham E, Picone D, Carmichael S, Bonner C, Chapman N
Appropriateness of Web-Based Resources for Home Blood Pressure Measurement and Their Alignment With Guideline Recommendations, Readability, and End User Involvement: Environmental Scan of Web-Based Resources
JMIR Infodemiology 2025;5:e55248
URL: https://infodemiology.jmir.org/2025/1/e55248
doi:10.2196/55248
PMID:

©Eleanor Clapham, Dean Picone, Samuel Carmichael, Carissa Bonner, Niamh Chapman. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 03.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Online Information About Side Effects and Safety Concerns of Semaglutide: Mixed Methods Study of YouTube Videos

Andy Wai Kan Yeung[1*], PhD; Fabian Peter Hammerle[2,3], MD; Sybille Behrens[2], MD; Maima Matin[4], MSc; Michel-Edwar Mickael[4], PhD; Olena Litvinova[2,5], DSc; Emil D Parvanov[2,6], PhD; Maria Kletecka-Pulker[2,7], PhD; Atanas G Atanasov[2,4*], PhD

[1]Oral and Maxillofacial Radiology, Applied Oral Sciences and Community Dental Care, Faculty of Dentistry, The University of Hong Kong, Hong Kong, China

[2]Ludwig Boltzmann Institute Digital Health and Patient Safety, Medical University of Vienna, Vienna, Austria

[3]Department of Anaesthesia, Intensive Care Medicine and Pain Medicine, Medical University of Vienna, Vienna, Austria

[4]Institute of Genetics and Animal Biotechnology of the Polish Academy of Sciences, Jastrzebiec, Magdalenka, Poland

[5]Department of Management and Quality Assurance in Pharmacy, National University of Pharmacy of the Ministry of Health of Ukraine, Kharkiv, Ukraine

[6]Department of Translational Stem Cell Biology, Research Institute of the Medical University of Varna, Varna, Bulgaria

[7]Institute for Ethics and Law in Medicine, University of Vienna, Vienna, Austria

[*]these authors contributed equally

Corresponding Author:
Andy Wai Kan Yeung, PhD
Oral and Maxillofacial Radiology, Applied Oral Sciences and Community Dental Care, Faculty of Dentistry, The University of Hong Kong
34 Hospital Road, Sai Ying Pun
Hong Kong, 852
China
Phone: 86 28590403
Email: ndyeung@hku.hk

Abstract

Background: Social media has been extensively used by the public to seek information and share views on health issues. Recently, the proper and off-label use of semaglutide drugs for weight loss has attracted huge media attention and led to temporary supply shortages.

Objective: The aim of this study was to perform a content analysis on English YouTube (Google) videos related to semaglutide.

Methods: YouTube was searched with the words semaglutide, Ozempic, Wegovy, and Rybelsus. The first 30 full-length videos (videos without a time limit) and 30 shorts (videos that are no longer than 1 minute) resulting from each search word were recorded. After discounting duplicates resulting from multiple searches, a total of 96 full-length videos and 93 shorts were analyzed. Video content was evaluated by 3 tools, that is, a custom checklist, a Global Quality Score (GQS), and Modified DISCERN. Readability and sentiment of the transcripts were also assessed.

Results: There was no significant difference in the mean number of views between full-length videos and shorts (mean 288,563.1, SD 513,598.3 vs mean 188,465.2, SD 780,376.2, *P*=.30). The former had better content quality in terms of GQS, Modified DISCERN, and the number of mentioned points from the custom checklist (all *P*<.001). The transcript readability of both types of videos was at a fairly easy level and mainly had a neutral tone. Full-length videos from health sources had a higher content quality in terms of GQS and Modified DISCERN (both *P*<.001) than their counterparts.

Conclusions: The analyzed videos lacked coverage of several important aspects, including the lack of long-term data, the persistence of side effects due to the long half-life of semaglutide, and the risk of counterfeit drugs. It is crucial for the public to be aware that videos cannot replace consultations with physicians.

(JMIR Infodemiol 2025;5:e59767) doi:10.2196/59767

XSL•FO
RenderX

KEYWORDS

YouTube; semaglutide; social media; Ozempic; Wegovy; Rybelsus; safety; knowledge exchange; side effects; online information; online; videos; health issues; drugs; weight loss; assessment; long-term data; consultation

Introduction

Public health information has traditionally been disseminated through printed media. However, with the rise of online social media platforms [1], the internet has become increasingly influential in spreading information and misinformation, notably during the COVID-19 pandemic [2-5]. Safety concerns in health care are frequently discussed on social media platforms such as YouTube (Google) [2,6].

Currently, obesity and being overweight are urgent health issues that reduce quality of life and increase the risk of cardiovascular diseases, type 2 diabetes mellitus, cancers, and reproductive system disorders, among others [7]. While lifestyle changes are essential for managing obesity, many people struggle with adherence [8]. Consequently, medical organizations are developing clinical guidelines for the long-term use of pharmacological therapy for obesity in adults. For instance, the American Gastroenterological Association recommends the use of pharmacological agents (strong recommendation, moderate certainty evidence) for adults with obesity or overweight who have insufficient results from lifestyle changes [8].

Unfortunately, some patients take their own antiobesity medication based on social media information, which can be dangerous without professional guidance. In particular, many people have watched YouTube videos on weight loss. It was found that the 98 most viewed weight loss videos on YouTube were viewed more than 365 million times in total [9]. In recent years, the injectable antidiabetic drug with weight loss property, branded Ozempic, has gained significant attention on platforms such as TikTok (ByteDance) and YouTube. Between 2018 and 2023, online searches for Ozempic surged in the United States [10]. Celebrity endorsements have driven its popularity, with 100 TikTok videos garnering over 70 million views [11]. However, this trend is concerning as many social media posts focused on the off-label use of Ozempic for weight loss, ignoring the potential health hazards [12]. Analyses of Reddit posts and social media comments have revealed discussions about off-label uses, struggles with insurance coverage, interest in compounded formulations, and unwanted side effects such as insomnia, anxiety, and depression [13,14].

Semaglutide, the active ingredient in Ozempic, was developed in 2012 to treat type 2 diabetes [15]. It is a glucagon-like peptide-1 (GLP-1) receptor agonist. GLP-1 receptors are expressed in many organs (pancreas, gastrointestinal tract, heart, brain, kidneys, lungs, and thyroid). This is associated with the pleiotropy and benefits of semaglutide in type 2 diabetes mellitus, weight loss, and cardioprotection. It can lower blood sugar levels through numerous means, including increasing insulin production, inhibiting glucagon secretion, and slowing gastric emptying [16]. Semaglutide is marketed as Ozempic and Rybelsus for treating diabetes and as Wegovy for chronic weight management. Ozempic and Wegovy are injectable, whereas Rybelsus is an oral tablet.

In 2017, the SUSTAIN (Semaglutide Unabated Sustainability in Treatment of Type 2 Diabetes) 1 trial demonstrated that weekly injections of semaglutide significantly improved body weight as well as glycated hemoglobin (HbA_{1c}) levels in type 2 diabetes patients [17]. The STEP (Semaglutide Treatment Effect in People with obesity) 1 trial published in 2021 showed that semaglutide, combined with lifestyle changes, significantly reduced body weight in overweight or obese nondiabetic patients [18]. These trials and their subsequent trials facilitated the United States Food and Drug Administration (FDA) to approve injectable semaglutide for treating diabetes and weight management. Besides, semaglutide was also approved in Europe [19]. Meanwhile, the PIONEER (Peptide Innovation for Early Diabetes Treatment) 1 trial published in 2019 found that daily oral semaglutide, versus placebo, significantly improved HbA_{1c} levels in type 2 diabetes patients managed by diet and exercise [20]. This supported the FDA approval of oral semaglutide to treat diabetes.

During the development of semaglutide (*la semaine* in translation from French, "week"), researchers sought to increase its duration of action. The half-life of oral semaglutide is approximately 1 week [21]. Ozempic (in strengths of 2 mg/1.5 mL, 2 mg/3 mL, 4 mg/3 mL, and 8 mg/3 mL), Rybelsus (3, 7, and 14 mg), and Wegovy (0.25 mg/0.5 mL, 0.5 mg/0.5 mL, 1 mg/0.5 mL, 1.7 mg/0.75 mL, and 2.4 mg/0.75 mL) have different dosages depending on the treatment purposes and patient characteristics. Ozempic, Rybelsus, and Wegovy are prescription medicines [22]. Common side effects include a slowdown in the digestive process from the stomach, nausea, and vomiting, which can be mitigated by gradually increasing the dose. Semaglutide is associated with increased risks of pancreatitis, gallbladder disease, and retinopathy, including vitreous hemorrhage and vision loss [8,23]. Besides, the rapid decrease in glucose levels can also worsen retinopathy in type 1 diabetes patients. Semaglutide is contraindicated in patients with a personal or family history of medullary or multiple thyroid cancer or endocrine neoplasia syndrome type 2. Serious side effects include abdominal pain, constipation, diarrhea, nausea, vomiting, dizziness, cholelithiasis, cholecystitis, acute myocardial infarction, gastroenteritis, and suicidal ideation [8].

The use of semaglutide, particularly the famous Ozempic, by nonsevere overweight individuals might pose safety issues [24]. Off-label drug use is legal and common, though it means a drug is being used for an unapproved indication or population, at an unapproved dosage, or via an unapproved route of administration [24,25]. Off-label users may have a higher safety risk if they obtain semaglutide via online vendors or beauty spas without a proper medical consultation [26]. With this background, this study aimed to investigate whether YouTube videos mentioned or discussed the side effects and safety concerns of semaglutide. It was hypothesized that full-length videos should be more informative than shorts (limited to 1 minute) and that full-length videos uploaded by YouTube-verified health source channels should be more informative than their counterparts. For readers'

XSL•FO
RenderX

information, a video coming from YouTube-verified health source channels would have an information panel underneath the video stating that it comes "from a channel with a licensed doctor (or health professional)" in a particular country, such as the United States.

Methods

Data Source and Search Strategy

On January 5, 2024, a search was performed on YouTube for semaglutide videos in English. Using Google Chrome with Incognito mode, YouTube was searched with the words semaglutide, Ozempic, Wegovy, and Rybelsus, respectively. For each search word, the first 30 full-length videos (videos without a time limit) and 30 shorts (videos that are no longer than 1 minute) resulting from the search, sorted by relevance, were recorded. The number of 30 videos was chosen according to a recent study, which claimed that very few YouTube users searched beyond the 33rd video [27]. After discounting duplicates and excluding unsuitable videos, a total of 96 full-length videos and 93 shorts were analyzed.

Outcome Measures

We recorded the basic video metrics, such as the duration, number of views, number of comments, number of likes, number of channel subscribers, and the age of the video since upload (number of days until March 14, 2024). The readability and sentiment of the video transcripts were assessed. The quality of video content was evaluated. Further details are described in further sections.

Data Extraction

To evaluate the readability and overall sentiment, video transcripts were generated and analyzed. To evaluate the quality of video content, the entire videos were watched and analyzed.

The video transcripts were generated by Whisper (with the "large" version, edition 20231117), an artificial intelligence automatic speech recognition system developed by OpenAI [28]. It has been used in previous research to transcribe educational videos [29] and had the best performance compared with similar automatic transcription tools [30]. The readability of the transcripts was evaluated by the Flesch Reading Ease (FRE) score [31], calculated via an online platform (Readability Formulas website). In brief, the score ranged from 0 to 100, with 90 to 100 being very easy, 0 to 29 being very difficult, and 60 to 69 being standard. Meanwhile, the sentiment of the transcripts was evaluated by ChatGPT 3.5, an artificial intelligence large language model developed by OpenAI that has been demonstrated to be very effective in sentiment analysis across multiple languages [32]. Referring to the method by Fu et al [32], the prompt was set as "Is the sentiment of this text positive, neutral, or negative? Respond with the sentiment label only." The "temperature" of the ChatGPT model was set at 0 to ensure the consistency of the answers with the least creativity. Temperature is a variable that changes the degree of randomness of the output generated by the model [33].

Next, the quality of video content was evaluated by 3 tools, Global Quality Score (GQS) [34], Modified DISCERN [35], and a custom checklist. Manual evaluations were independently performed by two authors (AWKY and AGA). Disagreements were resolved through mutual discussion. During these evaluations, the overall audiovisual content of the videos, not merely limited to verbal narrative, was examined. The GQS is a 5-point Likert scale designed to evaluate online health information. A score of 1 means "poor quality, poor flow, most information missing, not at all useful for patients," whereas a score of 5 means excellence and high usefulness for patients [34]. Meanwhile, the Modified DISCERN contains 5 evaluative items and gives 1 point for every positive answer and 0 points for negative answer. It was designed to evaluate YouTube videos on health care information [35]. The 5 items are as follows: (1) Are the aims clear and achieved? (2) Are reliable sources of information used? (3) Is the information presented balanced and unbiased? (4) Are additional sources of information listed for patient reference? (5) Are areas of uncertainty mentioned? For readers' information, the Modified DISCERN is based on an original version of DISCERN, which was designed to evaluate written health information and used a 5-point Likert scale to answer 15 questions plus an overall rating [36]. Since GQS and Modified DISCERN could only give a more general evaluation of the videos, a custom 12-point checklist was devised to evaluate the video content based on some specific aspects of side effects and safety concerns related specifically to semaglutide.

The custom 12-point checklist was compiled by the authors' team with reference to Smits and Van Raalte [37]. It recorded whether the following contents were mentioned in the videos or not: (1) form of application (injection, exception: oral for Rybelsus); (2) safe dosage; (3) need for long-term usage, or change in lifestyle and eating habits to avoid rebound back to original weight after drug cessation; (4) serious side effects (eg, retinopathy and pancreatitis); (5) gastrointestinal symptoms (eg, nausea, diarrhea, vomiting, gastric reflux, and gastritis); (6) prevalence or frequency of such side effects; (7) increased risk of aspiration during the induction of anesthesia; (8) contraindications; (9) long half-life (7 days) so that potential side effects persist for multiple days after drug cessation; (10) lack of long-term data; (11) potential alternatives (eg, diet and bariatric surgery); and (12) risk of counterfeit drugs.

Statistical Analysis

Following descriptive analysis of the video contents, 2-sample t tests were performed were performed to evaluate if there were significant differences between the full-length videos and shorts in terms of the mean FRE, GQS, and Modified DISCERN scores, as well as the mean number of mentioned points from the custom checklist. To supplement, the same tests were performed among the full-length videos, to evaluate if there were differences between those uploaded by YouTube-verified health source channels and those without this verification.

Ethical Considerations

Ethical approval was not applicable, as this study only analyzed publicly available data from existing datasets, and results were presented in aggregate that did not contain any identifiable information.

XSL•FO
RenderX

Results

The viewing metrics and content quality between full-length videos and shorts are compared in Table 1. All 189 videos (Figure 1) were collectively viewed 45,040,855 times. There was no significant difference in the mean number of views between full-length videos and shorts (288,563.1 vs 188,465.2, *P*=.30). However, full-length videos received thrice the number of comments than shorts on average (669.6 vs 200.4, *P*=.003). Full-length videos were usually older (ie, uploaded earlier) than shorts (468.0 vs 350.0, *P*=.01). The former had better content quality in terms of GQS, Modified DISCERN, and number of mentioned points from the custom checklist than the latter (all *P*<.001). The readability of the transcripts of the 2 types of videos did not significantly differ and both were at the fairly easy level. Meanwhile, 1 full-length video did not have a narration, whereas 16 shorts played music without a narration (Figure 2). Besides, sentiment analysis showed that full-length videos mainly had a neutral tone (n=59), followed by positive (n=24) and negative (n=12) tones. One full-length video did not have a verbal narrative. Meanwhile, shorts mainly had a neutral tone (n=40), rather than positive (n=19) and negative (n=18) tones. There were 16 shorts without a verbal narrative.

Among the full-length videos, those from YouTube-verified health source channels had a higher average view count than their counterparts (Table 2), though the difference did not reach statistical significance (401,867.2 vs 258,746.3, *P*=.41). Readability analysis suggested that the transcripts from health source videos were generally less readable than their counterparts (63.7 vs 73.8, ie, standard vs fairly easy, *P*<.001). However, health source videos had a higher content quality in terms of GQS and Modified DISCERN (both *P*<.001). On average, they also contained a larger number of mentioned points from the custom checklist than their counterparts, though that difference was not significant (3.5 vs 2.7, *P*=.17)

Next, the reporting of the points from the custom checklist was examined. As stated in previous sections, full-length videos were generally more informative than the shorts. Gastrointestinal symptoms and form of application were the 2 mostly reported points among the full-length videos as well as the shorts (Figure 3). Among full-length videos, the most deficient aspect was the lack of mention about an increased risk of aspiration during the induction of anesthesia associated with the use of semaglutide. Only 1 full-length video (1/96) warned the audience about this potential side-effect. Less neglected aspects were the persistence of side effects due to the long half-life of semaglutide (4/96, 4%) and the risk of counterfeit drugs (4/96, 4%). On the other hand, the shorts had generally omitted several important aspects. Apart from the increased risk of aspiration during the induction of anesthesia (1/93), none of the shorts mentioned the prevalence or frequency of side effects, persistence of side effects due to the long half-life, lack of long-term data, and risk of counterfeit drugs.

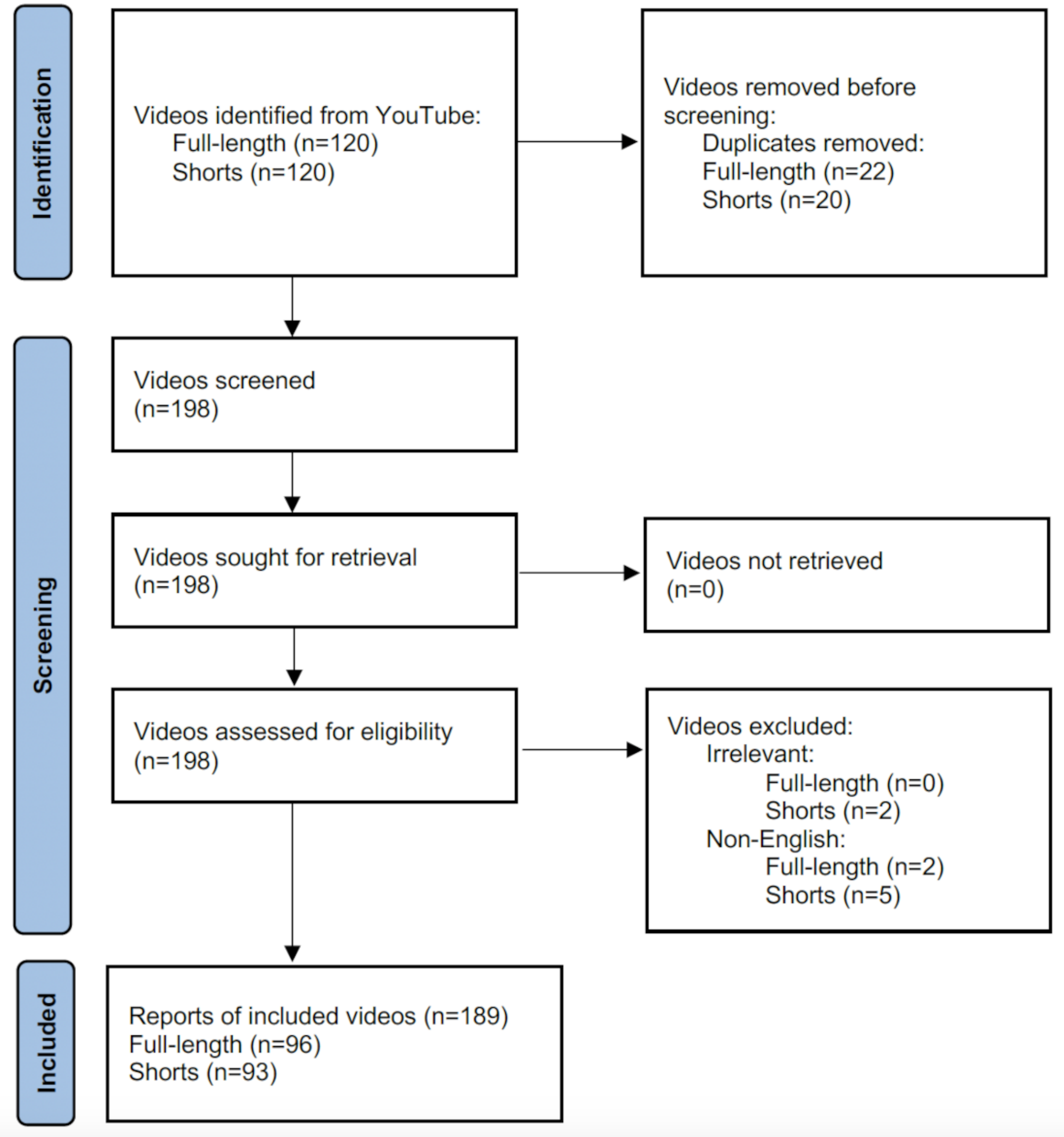
Table 1. Viewing metrics and content quality between full-length videos and shorts.

Metric	Mean (SD)		*P* value
	Full-length videos	Shorts	
Duration (s)	610.5 (542.6)	39.0 (17.2)	<.001
View count	288,563.1 (513,598.3)	188,465.2 (780,376.2)	.30
Comment count	669.6 (1281.0)	200.4 (571.4)	.003
Like count	5363.2 (11,842.7)	7412.4 (39,692.8)	.63
Channel subscriber count	1,721,826.1 (3,500,160.1)	531,761.7 (1,783,695.8)	.004
Video age (days)	468.0 (411.4)	350.0 (152.4)	.01
Flesch Reading Ease score	71.7 (11.6)	72.9 (14.3)	.54
GQS[a] score	2.6 (0.8)	1.1 (0.4)	<.001
Modified DISCERN	2.3 (0.8)	1.1 (0.2)	<.001
Number of content points from custom checklist	2.9 (1.7)	0.7 (0.9)	<.001

[a]GQS: Global Quality Score.

XSL•FO
RenderX

Figure 1. The screening process of YouTube videos on semaglutide.



XSL•FO
RenderX

Figure 2. Percentage of full-length videos and shorts with a narration.

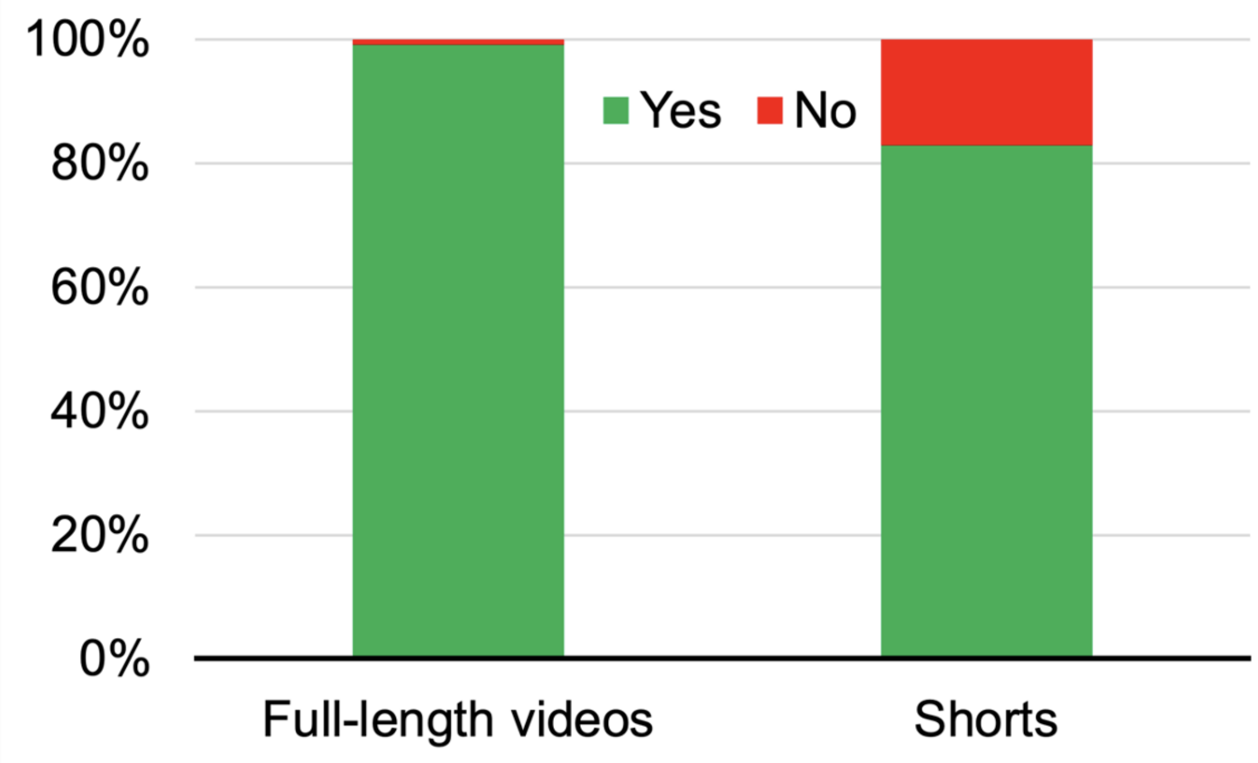


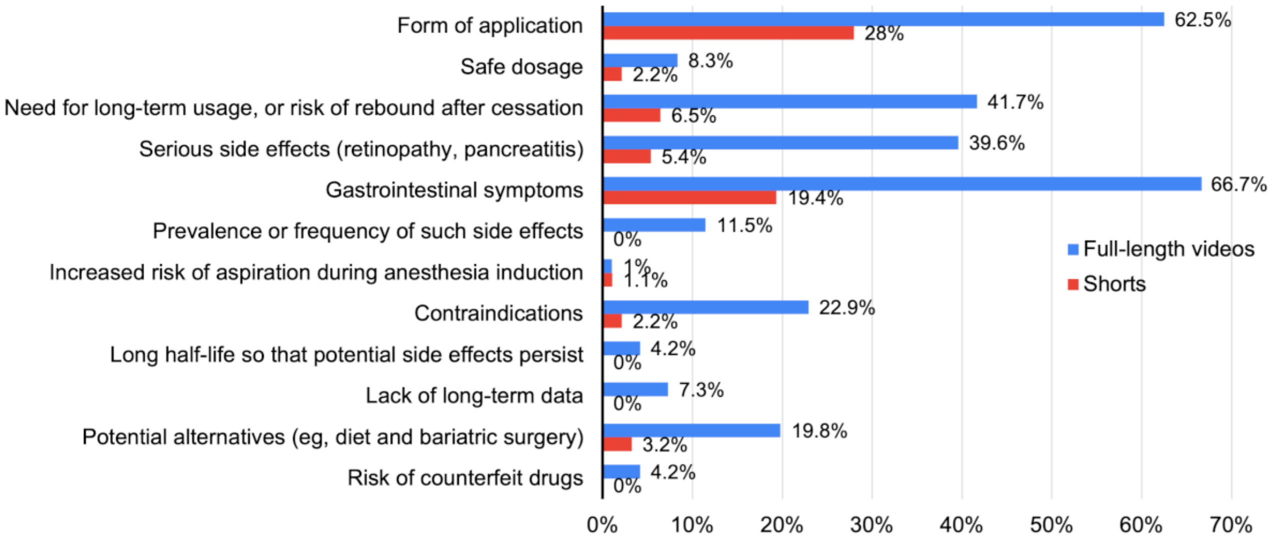
Table 2. Viewing metrics and content quality between full-length videos uploaded by YouTube-verified health source channels and other channels.

Metric	Mean (SD)		*P* value
	Health sources	Other channels	
Duration (s)	587.7 (664.0)	616.6 (510.9)	.83
View count	401,867.2 (733,982.5)	258,746.3 (439,684.8)	.41
Comment count	990.2 (1,631.5)	580.1 (1,163.6)	.32
Like count	10,371.5 (20,053.2)	4009.6 (8079.6)	.18
Channel subscriber count	1,873,152.0 (3,581,554.9)	1,682,003.5 (3,501,532.6)	.83
Video age (days)	630.7 (614.9)	425.1 (331.2)	.16
Flesch Reading Ease score	63.7 (13.3)	73.8 (10.2)	<.001
GQS[a] score	3.3 (0.6)	2.4 (0.8)	<.001
Modified DISCERN	3.1 (0.6)	2.1 (0.7)	<.001
Number of content points from custom checklist	3.5 (2.3)	2.7 (1.5)	.17

[a]GQS: Global Quality Score.

XSL•FO
RenderX

Figure 3. Percentage of full-length videos and shorts that covered the points from the custom checklist regarding the side effects and safety concerns of semaglutide.



Discussion

Principal Findings

This study found that semaglutide videos on YouTube have reached a broad audience. The 189 analyzed videos had a total view count of over 45 million. For comparison, the 98 most viewed YouTube videos on diabetic retinopathy were collectively viewed 1 million times [38]. It implied that these videos might potentially affect the perception and even health care decisions of the general public regarding the use of semaglutide. As expected, full-length videos were generally more informative than shorts. For patient education, shorts would serve better to grab the attention of the patients or make them aware of 1 or 2 particular issues related to semaglutide; whereas selected full-length videos might be more suitable to be incorporated as part of a panel discussion or public forum.

Alarmingly, the videos seldom covered the risk of counterfeit drugs. For instance, the antidiabetic drug, Ozempic, contains semaglutide that is synthesized by yeast fermentation and subsequent synthetic modification [39]. Without proper quality control of the synthetic processes, some falsified Ozempic products were found to contain contaminants such as glass particles and filler substances [39]. Meanwhile, some pharmacies would produce compounded versions of semaglutide to circumvent the patent issue. The high demand, high cost, and limited supply have led to a period of time when some patients and drug providers switched to compounded semaglutide [10]. Subsequently, the FDA stated that the compounded drugs, such as semaglutide sodium and semaglutide acetate, do not have their approval due to lack of testing, may not possess the same drug effect as semaglutide, and may even cause adverse effects (unspecified) [40,41]. Another safety issue of using compounded semaglutide is an increased risk of overdose due to suboptimal drug packaging. A recent case series reported that compounded semaglutide might be dispensed in vials instead of prefilled manufactured injection pens such as those by Ozempic and Wegovy [42]. Vials that contain large volumes of semaglutide and vials dispensed together with subpar syringes might allow for overdose much more easily during self-administration [42]. There seemed to be yet a case series on the overdose of oral semaglutide, but it would be reasoned that compounded semaglutide in the form of tablets for oral intake could have the same increased risk if each tablet did not conform to the amount of semaglutide contained in approved branded semaglutide products such as Rybelsus.

Another not uncommon risk of taking semaglutide that was often neglected by the videos was the risk of aspiration during the induction of anesthesia. One effect of semaglutide is gastroparesis, that is, reduced bowel motility and gastric emptying without any physical obstruction. This would increase the gastric volume and increase the risk of regurgitation and pulmonary aspiration of gastric contents even with the usual recommended fasting time before anesthesia [43,44]. According to a recent study, aspiration of gastric contents accounted for 5% of closed anesthesia malpractice claims in the United States during 2000-2014 [45]. Among these claim cases related to aspiration, 57% (66/115) of patients died and another 14% (16/115) suffered from permanent severe injury [45]. It implied that aspiration during anesthesia was not uncommon, and it could lead to very serious consequences. While medical organizations develop recommendations regarding the use of GLP-1 receptor agonists before operations, the optimal approach to patient data management is still being specified [46,47]. Hence, it is important for researchers and clinicians to conduct subsequent studies to optimize the fasting time and airway management strategy for patients on semaglutide who need to undergo anesthesia.

Although there have been many clinical trials on the efficacy and safety of semaglutide, most of them (if not all) had a study period of up to 2 years (104 weeks) only, such as the STEP, SUSTAIN, and PIONEER trials [37,48]. Since the use of semaglutide could last beyond 2 years and there could be a possibility of life-long usage, a lack of long-term data could mean potential risks yet to be elucidated, such as the discovery of more side effects, and potential development of drug dependence or drug resistance [49,50]. The COVID-19 vaccines

XSL•FO
RenderX

may illustrate the situation of lack of long-term data. After the COVID-19 pandemic began near the end of 2019, pharmaceutical companies put huge efforts into creating vaccines that could lower the infection rate and reduce the symptoms or morbidity. By the end of 2020, at least 10 vaccines have already been introduced into the global market and authorized by governments for emergency use [51]. After several years in use, millions of people vaccinated and billions of doses administered, data have accumulated. By analyzing the retrospective data, researchers have recently found 2 new rare but potentially severe side effects of COVID-19 vaccines, namely acute disseminated encephalomyelitis and transverse myelitis [52]. This new finding may influence the decision-making of some people in the public on whether they should be vaccinated, take a booster, or choose which type of available vaccines to use. In the case of semaglutide, more data besides its weight loss and antidiabetic properties would facilitate better-informed decisions from clinicians and patients. For instance, there were cases reported on the development or recurrence of depression 1 month after taking semaglutide, which was subsequently relieved after discontinuing the drug [53]. On the other hand, a retrospective study that covered over 1.5 million patient records reported that semaglutide users had a lower risk of incident and recurrent suicidal ideation [54]. At the same time, an analysis of over 40,000 user comments posted on social media platforms has found that users of GLP-1 receptor agonists (including semaglutide) felt that the drugs have mixed effects on their mood, anxiety, and insomnia conditions [14]. Patients should be aware of the fact that many weight-loss drugs, mainly appetite suppressants, were withdrawn from the market in the past due to adverse drug reactions [55]. Therefore, in the future when long-term data become available, the safety and side effects of semaglutide could be better assessed and established.

Findings from this study echoed previous studies on Reddit content on Ozempic or semaglutide, in the sense that there is generally a lack of discussion or elaboration on the potential health risks and hazards associated with the use of Ozempic, not to mention its off-label use [12,13]. However, the risks and hazards indeed exist, such as concerns with depression and anxiety raised by social media users on YouTube and TikTok [14]. Therefore, actionable recommendations included better public health awareness campaigns to educate the public on the proper use of Ozempic or semaglutide including the potential side effects and how to manage them. Pharmaceutical companies, governments, and health authorities can organize exhibitions in shopping malls, provide easy-to-understand information on their web pages, and develop user-friendly mobile apps to engage members of the public who are interested. Future research should evaluate the effects of such activities on the knowledge level of the public on semaglutide.

This study has several limitations. First, only the first 30 full-length videos and 30 shorts were initially screened for each search word, rendering a final analysis of 189 videos. This might only represent a small proportion of the entire relevant video collection. Second, only videos in English were analyzed. The themes and foci could be different for videos produced in other languages. Third, errors might exist during computational and manual evaluations. Besides, there was not a detailed analysis of video sources, which might further enhance the findings, for example, whether videos produced by clinicians were more informative than those from general influencers. It was said that there was a tendency for the YouTube algorithm to place more “reliable” videos at the top of the search results [56], hence the presented results in this study might be biased.

Conclusions

The 189 analyzed YouTube videos on semaglutide have attracted more than 45 million views. Full-length videos were much more informative than shorts in terms of side effects and safety issues. The analyzed videos lacked coverage of several important aspects, including the increased risk of aspiration during the induction of anesthesia, the persistence of side effects due to the long half-life of semaglutide, the risk of counterfeit drugs, and the lack of long-term data. Patients should be aware that these videos may not be comprehensive enough even if they were uploaded by YouTube-verified health source channels, and hence cannot replace the consultation from the physician, who may make tailor-made recommendations and treatment plans for each patient.

Conflicts of Interest

None declared.

References

1. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
2. Li HOY, Bailey A, Huynh D, Chan J. YouTube as a source of information on COVID-19: a pandemic of misinformation? BMJ Glob Health 2020;5(5):e002604 [FREE Full text] [doi: 10.1136/bmjgh-2020-002604] [Medline: 32409327]
3. Inoue M, Shimoura K, Nagai-Tanima M, Aoyama T. The relationship between information sources, health literacy, and COVID-19 knowledge in the COVID-19 infodemic: cross-sectional online study in Japan. J Med Internet Res 2022;24(7):e38332 [FREE Full text] [doi: 10.2196/38332] [Medline: 35839380]
4. Moon H, Lee GH. Evaluation of Korean-language COVID-19-related medical information on YouTube: cross-sectional infodemiology study. J Med Internet Res 2020;22(8):e20775 [FREE Full text] [doi: 10.2196/20775] [Medline: 32730221]
5. Lordan R, Devkota HP. Turbulence at Twitter with leadership change: implications for health research and science communication. Explor Digit Health Technol 2023 Nov 1;1:4-10. [doi: 10.37349/edht.2023.00002]

XSL•FO
RenderX

6. Yeung AWK, Parvanov ED, Horbańczuk JO, Kletecka-Pulker M, Kimberger O, Willschke H, et al. Are dental x-rays safe? Content analysis of English and Chinese YouTube videos. Digit Health 2023;9:20552076231179053 [FREE Full text] [doi: 10.1177/20552076231179053] [Medline: 37312949]
7. Obesity and overweight. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight#:~:text=Worldwide%20adult%20obesity%20has%20more,16%25%20were%20living%20with%20obesity [accessed 2024-04-22]
8. Grunvald E, Shah R, Hernaez R, Chandar AK, Pickett-Blakely O, Teigen LM, AGA Clinical Guidelines Committee. AGA clinical practice guideline on pharmacological interventions for adults with obesity. Gastroenterology 2022;163(5):1198-1225. [doi: 10.1053/j.gastro.2022.08.045] [Medline: 36273831]
9. Basch CH, Fung ICH, Menafro A, Mo C, Yin J. An exploratory assessment of weight loss videos on YouTube. Public Health 2017;151:31-38. [doi: 10.1016/j.puhe.2017.06.016] [Medline: 28710925]
10. Han SH, Safeek R, Ockerman K, Trieu N, Mars P, Klenke A, et al. Public interest in the off-label use of glucagon-like peptide 1 agonists (Ozempic) for cosmetic weight loss: a google trends analysis. Aesthet Surg J 2023;44(1):60-67. [doi: 10.1093/asj/sjad211] [Medline: 37402640]
11. Basch CH, Narayanan S, Tang H, Fera J, Basch CE. Descriptive analysis of TikTok videos posted under the hashtag #Ozempic. Journal of Medicine, Surgery, and Public Health 2023;1:100013. [doi: 10.1016/j.glmedi.2023.100013]
12. Fong S, Carollo A, Lazuras L, Corazza O, Esposito G. Ozempic (glucagon-like peptide 1 receptor agonist) in social media posts: unveiling user perspectives through Reddit topic modeling. Emerging Trends in Drugs, Addictions, and Health 2024;4:100157. [doi: 10.1016/j.etdah.2024.100157]
13. Javaid A, Baviriseaty S, Javaid R, Zirikly A, Kukreja H, Kim CH, et al. Trends in glucagon-like peptide-1 receptor agonist social media posts using artificial intelligence. JACC Adv 2024;3(9):101182 [FREE Full text] [doi: 10.1016/j.jacadv.2024.101182] [Medline: 39372460]
14. Arillotta D, Floresta G, Guirguis A, Corkery JM, Catalani V, Martinotti G, et al. GLP-1 receptor agonists and related mental health issues; insights from a range of social media platforms using a mixed-methods approach. Brain Sci 2023;13(11):1503 [FREE Full text] [doi: 10.3390/brainsci13111503] [Medline: 38002464]
15. Lau J, Bloch P, Schäffer L, Pettersson I, Spetzler J, Kofoed J, et al. Discovery of the once-weekly glucagon-like peptide-1 (GLP-1) analogue semaglutide. J Med Chem 2015;58(18):7370-7380. [doi: 10.1021/acs.jmedchem.5b00726] [Medline: 26308095]
16. Goldenberg RM, Steen O. Semaglutide: review and place in therapy for adults with type 2 diabetes. Can J Diabetes 2019;43(2):136-145 [FREE Full text] [doi: 10.1016/j.jcjd.2018.05.008] [Medline: 30195966]
17. Sorli C, Harashima S, Tsoukas GM, Unger J, Karsbøl JD, Hansen T, et al. Efficacy and safety of once-weekly semaglutide monotherapy versus placebo in patients with type 2 diabetes (SUSTAIN 1): a double-blind, randomised, placebo-controlled, parallel-group, multinational, multicentre phase 3a trial. Lancet Diabetes Endocrinol 2017;5(4):251-260. [doi: 10.1016/S2213-8587(17)30013-X] [Medline: 28110911]
18. Wilding JPH, Batterham RL, Calanna S, Davies M, Van Gaal LF, Lingvay I, STEP 1 Study Group. Once-weekly semaglutide in adults with overweight or obesity. N Engl J Med 2021;384(11):989-1002. [doi: 10.1056/NEJMoa2032183] [Medline: 33567185]
19. Public health - union register of medicinal products. European Commission. 2024. URL: https://ec.europa.eu/health/documents/community-register/html/h1251.htm [accessed 2024-04-22]
20. Aroda VR, Rosenstock J, Terauchi Y, Altuntas Y, Lalic NM, Morales Villegas EC, PIONEER 1 Investigators. PIONEER 1: randomized clinical trial of the efficacy and safety of oral semaglutide monotherapy in comparison with placebo in patients with type 2 diabetes. Diabetes Care 2019;42(9):1724-1732. [doi: 10.2337/dc19-0749] [Medline: 31186300]
21. Granhall C, Donsmark M, Blicher TM, Golor G, Søndergaard FL, Thomsen M, et al. Safety and pharmacokinetics of single and multiple ascending doses of the novel oral human GLP-1 analogue, oral semaglutide, in healthy subjects and subjects with type 2 diabetes. Clin Pharmacokinet 2019;58(6):781-791. [doi: 10.1007/s40262-018-0728-4] [Medline: 30565096]
22. Orange book: approved drug products with therapeutic equivalence evaluations. Food and Drug Administration. 2024. URL: https://www.accessdata.fda.gov/scripts/cder/ob/index.cfm [accessed 2024-04-22]
23. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jódar E, Leiter LA, SUSTAIN-6 Investigators. Semaglutide and cardiovascular outcomes in patients with type 2 diabetes. N Engl J Med 2016;375(19):1834-1844. [doi: 10.1056/NEJMoa1607141] [Medline: 27633186]
24. Chiappini S, Vickers-Smith R, Harris D, Papanti Pelletier GD, Corkery JM, Guirguis A, et al. Is there a risk for semaglutide misuse? Focus on the Food and Drug Administration's FDA Adverse Events Reporting System (FAERS) pharmacovigilance dataset. Pharmaceuticals (Basel) 2023;16(7):994 [FREE Full text] [doi: 10.3390/ph16070994] [Medline: 37513906]
25. Stafford RS. Regulating off-label drug use--rethinking the role of the FDA. N Engl J Med 2008;358(14):1427-1429. [doi: 10.1056/NEJMp0802107] [Medline: 18385495]
26. Vermaak I. Semaglutide's slimming properties shifts the scales towards scarcity and shams. SA Pharmaceutical Journal 2024;91(1):74-77 [FREE Full text]
27. Narayan RR, Kadri H, Muhammad HA, Qadan M. Critical appraisal of content quality of YouTube videos on the Whipple pancreaticoduodenectomy. J Surg Res 2024;295:690-698. [doi: 10.1016/j.jss.2023.11.066] [Medline: 38134739]

XSL•FO
RenderX

28. Radford A, Kim J, Xu T, Brockman G, McLeavey C, Sutskever I. Robust speech recognition via large-scale weak supervision. : ACM; 2023 Presented at: ICML'23: International Conference on Machine Learning; 2023 July 23-29; Honolulu Hawaii USA p. 28492-28518.
29. AlShaikh R, Al-Malki N, Almasre M. The implementation of the cognitive theory of multimedia learning in the design and evaluation of an AI educational video assistant utilizing large language models. Heliyon 2024;10(3):e25361 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e25361] [Medline: 38352730]
30. Wollin-Giering S, Hoffmann M, Höfting J, Ventzke C. Automatic transcription of English and German qualitative interviews. Forum: Qualitative Social Research 2024;25(1):8 [FREE Full text] [doi: 10.17169/fqs-25.1.4129]
31. Flesch R. A new readability yardstick. J Appl Psychol 1948;32(3):221-233. [doi: 10.1037/h0057532] [Medline: 18867058]
32. Fu Z, Hsu YC, Chan CS, Lau CM, Liu J, Yip PSF. Efficacy of ChatGPT in Cantonese sentiment analysis: comparative study. J Med Internet Res 2024;26:e51069 [FREE Full text] [doi: 10.2196/51069] [Medline: 38289662]
33. Davis J, Van Bulck L, Durieux BN, Lindvall C. The temperature feature of ChatGPT: modifying creativity for clinical research. JMIR Hum Factors 2024;11:e53559 [FREE Full text] [doi: 10.2196/53559] [Medline: 38457221]
34. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol 2007;102(9):2070-2077. [doi: 10.1111/j.1572-0241.2007.01325.x] [Medline: 17511753]
35. Singh AG, Singh S, Singh PP. YouTube for information on rheumatoid arthritis--a wakeup call? J Rheumatol 2012;39(5):899-903. [doi: 10.3899/jrheum.111114] [Medline: 22467934]
36. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
37. Smits MM, Van Raalte DH. Safety of semaglutide. Front Endocrinol (Lausanne) 2021;12:645563 [FREE Full text] [doi: 10.3389/fendo.2021.645563] [Medline: 34305810]
38. Basch CH, Fung ICH, Berdnik A, Basch CE. Widely viewed English language YouTube videos relating to diabetic retinopathy: a cross-sectional study. JMIR Diabetes 2016;1(2):e6 [FREE Full text] [doi: 10.2196/diabetes.6450] [Medline: 30291055]
39. Janvier S, De Spiegeleer B, Vanhee C, Deconinck E. Falsification of biotechnology drugs: current dangers and/or future disasters? J Pharm Biomed Anal 2018;161:175-191. [doi: 10.1016/j.jpba.2018.08.037] [Medline: 30165334]
40. Medications containing semaglutide marketed for type 2 diabetes or weight loss. United States Food and Drug Administration. 2023. URL: https://www.fda.gov/drugs/postmarket-drug-safety-information-patients-and-providers/medications-containing-semaglutide-marketed-type-2-diabetes-or-weight-loss [accessed 2024-03-25]
41. Aschenbrenner DS. Shortage of weight loss drugs creates problems for patients. Am J Nurs 2024;124(4):22-23. [doi: 10.1097/01.NAJ.0001010560.28176.11] [Medline: 38511704]
42. Lambson JE, Flegal SC, Johnson AR. Administration errors of compounded semaglutide reported to a poison control center-case series. J Am Pharm Assoc (2003) 2023;63(5):1643-1645. [doi: 10.1016/j.japh.2023.06.017] [Medline: 37392810]
43. Fezza R, Rains B, Fezza T, Fezza JP. Emerging anesthesia risks with semaglutide. Plast Reconstr Surg Glob Open 2023;11(11):e5427 [FREE Full text] [doi: 10.1097/GOX.0000000000005427] [Medline: 38025626]
44. Pai SL, Smith MA, Beam WB, Harbell MW. Perioperative considerations for patients on semaglutide. Curr Anesthesiol Rep 2024;14(2):152-160 [FREE Full text] [doi: 10.1007/s40140-024-00611-6]
45. Warner M, Meyerhoff K, Warner M, Posner K, Stephens L, Domino K. Pulmonary aspiration of gastric contents: a closed claims analysis. Anesthesiology 2021;135(2):284-291. [doi: 10.1097/ALN.0000000000003831] [Medline: 34019629]
46. Joshi G, Abdelmalak B, Weigel W, Soriano S, Harbell M, Kuo C. American Society of Anesthesiologists consensus-based guidance on preoperative management of patients (adults and children) on glucagon-like peptide-1 (GLP-1) receptor agonists. 2023. URL: https://www.asahq.org/about-asa/newsroom/news-releases/2023/06/american-society-of-anesthesiologists-consensus-based-guidance-on-preoperative [accessed 2024-04-22]
47. Pfeifer KJ, Selzer A, Mendez CE, Whinney CM, Rogers B, Simha V, et al. Preoperative management of endocrine, hormonal, and urologic medications: Society for Perioperative Assessment and Quality Improvement (SPAQI) consensus statement. Mayo Clin Proc 2021;96(6):1655-1669. [doi: 10.1016/j.mayocp.2020.10.002] [Medline: 33714600]
48. Bergmann NC, Davies MJ, Lingvay I, Knop FK. Semaglutide for the treatment of overweight and obesity: a review. Diabetes Obes Metab 2023;25(1):18-35 [FREE Full text] [doi: 10.1111/dom.14863] [Medline: 36254579]
49. Ilan Y. Next-generation personalized medicine: implementation of variability patterns for overcoming drug resistance in chronic diseases. J Pers Med 2022;12(8):1303 [FREE Full text] [doi: 10.3390/jpm12081303] [Medline: 36013252]
50. Lin SK. Pharmacological means of reducing human drug dependence: a selective and narrative review of the clinical literature. Br J Clin Pharmacol 2014;77(2):242-252 [FREE Full text] [doi: 10.1111/bcp.12163] [Medline: 23701272]
51. King ML. How manufacturing won or lost the COVID-19 vaccine race. Vaccine 2024;42(5):1004-1012 [FREE Full text] [doi: 10.1016/j.vaccine.2023.12.031] [Medline: 38228438]
52. Faksova K, Walsh D, Jiang Y, Griffin J, Phillips A, Gentile A, et al. COVID-19 vaccines and adverse events of special interest: a multinational Global Vaccine Data Network (GVDN) cohort study of 99 million vaccinated individuals. Vaccine 2024;42(9):2200-2211 [FREE Full text] [doi: 10.1016/j.vaccine.2024.01.100] [Medline: 38350768]

XSL•FO
RenderX

53. Li JR, Cao J, Wei J, Geng W. Case Report: semaglutide-associated depression: a report of two cases. Front Psychiatry 2023;14:1238353 [FREE Full text] [doi: 10.3389/fpsyt.2023.1238353] [Medline: 37706035]
54. Wang W, Volkow ND, Berger NA, Davis PB, Kaelber DC, Xu R. Association of semaglutide with risk of suicidal ideation in a real-world cohort. Nat Med 2024;30(1):168-176 [FREE Full text] [doi: 10.1038/s41591-023-02672-2] [Medline: 38182782]
55. Onakpoya IJ, Heneghan CJ, Aronson JK. Post-marketing withdrawal of anti-obesity medicinal products because of adverse drug reactions: a systematic review. BMC Med 2016;14(1):191 [FREE Full text] [doi: 10.1186/s12916-016-0735-y] [Medline: 27894343]
56. Pons-Fuster E, Ruiz Roca J, Tvarijonaviciute A, López-Jornet P. YouTube information about diabetes and oral healthcare. Odontology 2020;108(1):84-90. [doi: 10.1007/s10266-019-00445-3] [Medline: 31396751]

Abbreviations

FDA: United States Food and Drug Administration
FRE: Flesch Reading Ease
GLP-1: glucagon-like peptide-1
GQS: Global Quality Score
HbA$_{1c}$: glycated hemoglobin
PIONEER: Peptide Innovation for Early Diabetes Treatment
STEP: Semaglutide Treatment Effect in People with obesity
SUSTAIN: Semaglutide Unabated Sustainability in Treatment of Type 2 Diabetes

Edited by T Mackey; submitted 22.04.24; peer-reviewed by S Arsić, F Lamy; comments to author 20.11.24; revised version received 25.11.24; accepted 25.01.25; published 08.04.25.

Please cite as:
Yeung AWK, Hammerle FP, Behrens S, Matin M, Mickael ME, Litvinova O, Parvanov ED, Kletecka-Pulker M, Atanasov AG
Online Information About Side Effects and Safety Concerns of Semaglutide: Mixed Methods Study of YouTube Videos
JMIR Infodemiology 2025;5:e59767
URL: https://infodemiology.jmir.org/2025/1/e59767
doi:10.2196/59767
PMID:

©Andy Wai Kan Yeung, Fabian Peter Hammerle, Sybille Behrens, Maima Matin, Michel-Edwar Mickael, Olena Litvinova, Emil D Parvanov, Maria Kletecka-Pulker, Atanas G Atanasov. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 08.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evolutionary Trend of Dental Health Care Information on Chinese Social Media Platforms During 2018-2022: Retrospective Observational Study

Zhiyu Zhu[1,2], BM; Zhiyun Ye[2], BM; Qian Wang[3], BCM; Ruomei Li[1,2], DDS, MD, PhD; Hairui Li[1,2], MDS; Weiming Guo[1,2], DDS, MD, PhD; Zhenxia Li[1,2], DDS, MD, PhD; Lunguo Xia[1,2], DDS, MD, PhD; Bing Fang[1,2], DDS, MD, PhD

[1]Department of Orthodontics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]College of Stomatology, Shanghai Jiao Tong University, National Center for Stomatology, National Clinical Research Center for Oral Diseases, Shanghai Key Laboratory of Stomatology, Shanghai Research Institute of Stomatology, Shanghai, China

[3]School of Computing, National University of Singapore, Singapore, Singapore

Corresponding Author:
Bing Fang, DDS, MD, PhD
Department of Orthodontics
Shanghai Ninth People's Hospital
Shanghai Jiao Tong University School of Medicine
No 500, Quxi Rd
Shanghai, 200011
China
Phone: 86 021 23271699
Email: fangbing@sjtu.edu.cn

Abstract

Background: Social media holds an increasingly significant position in contemporary society, wherein evolving public perspectives are mirrored by changing information. However, there remains a lack of comprehensive analysis regarding the nature and evolution of dental health care information on Chinese social media platforms (SMPs) despite extensive user engagement and voluminous content.

Objective: This study aimed to probe into the nature and evolution of dental health care information on Chinese SMPs from 2018 to 2022, providing valuable insights into the evolving digital public perception of dental health for dental practitioners, investigators, and educators.

Methods: This study was conducted on 3 major Chinese SMPs: Weibo, WeChat, and Zhihu. Data from March 1 to 31 in 2018, 2020, and 2022 were sampled to construct a social media original database (ODB), from which the most popular long-text posts (N=180) were selected to create an analysis database (ADB). Natural language processing (NLP) tools were used to assist tracking topic trends, and word frequencies were analyzed. The DISCERN health information quality assessment questionnaire was used for information quality evaluation.

Results: The number of Weibo posts in the ODB increased approximately fourfold during the observation period, with discussion of orthodontic topics showing the fastest growth, surpassing that of general dentistry after 2020. In the ADB, the engagement of content on Weibo and Zhihu also displayed an upward trend. The overall information quality of long-text posts on the 3 platforms was moderate or low. Of the long-text posts, 143 (79.4%) were written by nonprofessionals, and 105 (58.3%) shared personal medical experiences. On Weibo and WeChat, long-text posts authored by health care professionals had higher DISCERN scores (Weibo P=.04; WeChat P=.02), but there was a negative correlation between engagement and DISCERN scores (Weibo tau-b [τb]=–0.45, P=.01; WeChat τb=–0.30, P=.02).

Conclusions: There was a significant increase in the dissemination and evolution of public interest in dental health care information on Chinese social media during 2018-2022. However, the quality of the most popular long-text posts was rated as moderate or low, which may mislead patients and the public.

(JMIR Infodemiology 2025;5:e55065) doi:10.2196/55065

XSL•FO
RenderX

KEYWORDS

social media; dental health education; natural language processing; information quality assessment; dental care; dental hygiene; dentistry; orthodontic; health care information; retrospective study; observational study; user engagement; Chinese; dental practitioner; WeChat; health information; preventive care

Introduction

Social media usage has been extensively integrated into modern life. As of early April 2024, there were 5.1 billion social media users around the world, equating to 62.6% of the worldwide population [1]. As for Mainland China, active social media users numbered approximately 1.1 billion in January 2024, constituting 74.2% of its total population [2]. Social media provides a platform for everyone to disseminate health knowledge, demonstrate cases, and promote themselves [3]. The general public, especially patients, increasingly turns to social media to obtain health information and join communities that exchange medical experiences [4]. A survey conducted in China revealed that 71.9% of participants obtained health education through the internet, with 30.0% of them frequently seeking health information online [5]. The internet is the most important source of information for patients with cancer, and 80% of them use social media to communicate with others about their condition [6]. The content related to health on social media creates a rich repository of information, dynamically reflecting the public's health perceptions in real time.

The growing reliance on social media for health information is a double-edged sword. On the one hand, social media offers a wealth of readily accessible information and a platform for sharing experiences, enhancing public knowledge and communication. On the other hand, studies have shown that the quality of health information on these platforms is highly variable, which may mislead the public, including patients, and potentially cause adverse outcomes [7,8]. The National Institutes of Health explicitly encourages medical professionals to share accurate health information and curb the spread of misinformation online [9].

Social media as a library for real-world studies has garnered attention from academia. Existing studies have focused on tracking topic trends, analyzing group emotions [10,11], searching for health care development directions, predicting disease spread [12], evaluating network information quality, and highlighting the harm of misinformation dissemination [13,14].

Despite the extensive user engagement and voluminous content in Chinese social media, there remains a conspicuous gap in methodical investigations into dental health care information. This gap is particularly pronounced, given China's unique digital landscape, which is dominated by platforms such as Weibo, WeChat, and Douyin (TikTok) [8,15]. Investigating the information on these platforms could fill this void and provide Eastern insights into contemporary public perceptions and concerns regarding dental health. Although a few studies have surveyed COVID-19–related dental posts on Weibo, primarily focusing on the impact of the pandemic on patients [16,17], there is a lack of comprehensive analyses of information quality and topic trends, possibly due to limitations in research tools.

The application of artificial intelligence (AI) tools has made it possible to analyze and monitor the massive amount of information on social media [18]. Among them, natural language processing (NLP), an important branch of AI, is useful for analyzing social media content for text mining purposes [19]. Another burgeoning branch, sentiment analysis tools, can be used for public opinion analysis, such as epidemic trends [20], willingness for vaccination [10], and even presidential elections [21]. Since the emergence of Chat Generative Pretrained Transformer (ChatGPT), it has also been applied to social media research, such as popular hashtag generation algorithms [22] and the construction of lexica for online pharmacovigilance [23]. AI advancements have increased the popularity of social media research in various industries and provided guidance for their development [3,12].

This study was designed to leverage AI tools to shed light upon the changing patterns and standards of dental health care information on mainstream Chinese social media platforms (SMPs). The collected data were meticulously analyzed to identify evolving trends and assess the quality of information pertaining to dental health care. Additionally, this study probed into the determinants of audience engagement and the influence of social media information, aspiring to trace the shifting contours of public perception and the demand for dental health care. The ultimate goal was to furnish dental health care professionals with actionable insights and strategies to enhance their clinical practice and the quality of doctor-patient interactions.

Methods

Data Sources

This study was conducted on 3 major text-based SMPs in China: Weibo, WeChat, and Zhihu. Data from Match 1 to 31 in 2018, 2020, and 2022 were sampled. Weibo is a microblogging website. The WeChat public platform is a self-media platform based on the short-message service application WeChat. Zhihu is a knowledge question-and-answer (Q&A) community, as well as an original content platform.

Data Extraction

To extract relevant data for this study, a social media scraping program was developed using Python's Selenium module. The inclusion criteria specifically focused on 3 key elements: platform, time, and keywords. The time periods were divided into 3 distinct intervals: March 1-31, 2018; March 1-31, 2020; and March 1-31, 2022. The search for data extraction was conducted using a series of Chinese keywords on April 1, 2023. The translated keywords are presented in Boolean logic format in Table 1 and were divided into 3 predetermined themes (general dentistry, orthodontics, and prosthodontics) based on established classifications of subspecialties within the field of dentistry. The general dentistry section included "dental

XSL•FO
RenderX

fillings," "dental cleaning," "tooth extraction," "root canal treatment," and "teeth whitening." The orthodontics section included "orthodontics," "teeth straightening," "orthodontic treatment," "braces," "dental braces," "get braces," and "retainers." The prosthodontics section included "dental crown," "overlay," "porcelain tooth," "dental implant," "implanted tooth," "tooth veneer," and "porcelain veneer." The bilingual translation table is provided in Table S1 in Multimedia Appendix 1, and the translation was based on Chinese official textbooks and the Chinese edition of authoritative English textbooks [24-26]. These keywords covered a broad spectrum of nearly all commonly used Chinese expressions related to dental practices across various specialties, ensuring that the search was comprehensive and inclusive. Notably, all posts meeting the inclusion criteria were included, irrespective of authorship or topic and without manual intervention. When a post was retrieved, it was automatically assigned to the relevant theme based on the search term. For instance, posts retrieved using the keyword "dental fillings" were classified under the general dentistry theme.

Table 1. Keywords for searching posts (translated from Chinese).

Theme	Keywords
General dentistry	"dental fillings" OR "dental cleaning" OR "tooth extraction" OR "root canal treatment" OR "teeth whitening"
Orthodontics	"orthodontics" OR "teeth straightening" OR "orthodontic treatment" OR "braces" OR "dental braces" OR "get brace" OR "retainer"
Prosthodontics	"dental crown" OR "overlay" OR "porcelain tooth" OR "dental implant" OR "implanted tooth" OR "tooth veneer" OR "porcelain veneer"

Weibo's open application programming interface (API) allowed access to all content that met the inclusion criteria. In contrast, for WeChat and Zhihu, we could rely only on the built-in search functions, which do not provide access to all the data. Zhihu required searches based on time conditions within the website. For the WeChat public platform, Sogou's WeChat search was conducted to retrieve the top 10 pages of posts based on the website's own sorting logic, representing the content that people were most likely exposed to.

The exclusion criteria were primarily based on 4 aspects. "Duplicated" content referred to posts that were entirely identical due to duplicate publication, plagiarism, or other similar reasons. Only the earliest published post was retained, and the others were removed. "Irrelevant" content referred to posts in which dental-related keywords were mentioned only briefly, while the main content focused on unrelated subjects. "Unavailable" content referred to posts where the title or abstract was accessible but the full text was no longer available, possibly due to voluntary removal or deletion by the platform. Finally, "meaningless" content included posts that were composed of incoherent or garbled text consisting of nonsensical strings of words or symbols without any thematic relevance. Upon data collection, the content from the 3 SMPs was rigorously screened according to the exclusion criteria. Posts that could potentially disrupt the integrity of the subsequent analysis were removed, and the remaining posts were included to construct a social media original database (ODB).

Database Construction

Following data collection and screening based on the inclusion and exclusion criteria, the ODB was constructed. Posts with 600 characters or more were defined as "long-text posts." Previous studies suggest that long-text posts tend to provide more comprehensive information and exhibit higher engagement levels [27]. The top 20 long-text posts with the highest level of popularity each month on the three platforms were selected to establish an analysis database (ADB) for lexical analysis with NLP tools and information quality assessment.

Evaluation Strategy

Preliminary analysis of the ODB data involved capturing author information and popularity indicators for each platform, including the number of likes or reads. The metric for popularity varied by platform. For Weibo and Zhihu, the number of likes was used. For WeChat, the number of reads was used, as WeChat users tend to use "like" functions less frequently. The top 20 long-text posts with the highest levels of popularity during each observation period were selected from each platform, totaling 180 posts constituting the ADB for text mining and information quality evaluation. The classification of author account types followed subsequently and was based primarily on usernames. Specifically, accounts with names containing medical-related terms, such as "doctor," "dentist," "hospital," or "clinic," were classified as health care professionals, whereas those without such identifiers were categorized as non–health care professionals.

Text mining analysis and information quality evaluation were conducted on the ADB. NLP tools were used for text mining analysis. Using the Jieba segmentation tool, the first step involved cleaning the text data by removing punctuation, converting all text to lowercase, and eliminating common Chinese stop words, such as "of." The cleaned text was then tokenized into individual words, or tokens, to facilitate the analysis of word frequency across the entire ADB. Python's Counter module was used to compile a frequency distribution of the keywords. The most frequently mentioned words were visualized as word clouds, and the frequencies of the top 30 words for each period or platform were visualized as heatmaps. The font size of the word cloud and the color depth of the heatmap represented the frequency of word occurrence, providing a clear representation of the prevalent topics and their temporal variation. For information quality evaluation, the DISCERN questionnaire (Multimedia Appendix 2), which is widely used in research on the quality of health material information online [28,29], was used to assess the reliability of health material and the quality of information for treatment plan selection. The DISCERN score was rated on a 5-point Likert

XSL•FO
RenderX

scale, with 1 indicating low quality, 5 indicating high quality, and 3 indicating moderate quality [30]. To ensure the objectivity of evaluation, all posts were anonymized during the review and quality assessment process. The determination of account type was conducted separately from the quality scoring, ensuring that the account type did not influence the quality assessment. Given the greater complexity of the DISCERN information quality assessment, this work was conducted by 2 practicing dental professionals following the Cohen kappa consistency test. Among the 180 long-text posts in the ADB, 18 (10%) were randomly selected for independent assessment by 2 evaluators, resulting in a Cohen kappa coefficient of 0.84, indicating reliable consistency between the results of the 2 evaluators.

Statistical Analysis

Statistical analysis was performed using SPSS software version 25 (IBM Corp). Descriptive results are presented as the median (25th-75th percentile) for quantitative data. The Shapiro-Wilks test and the Levene test were performed to determine the normality and homogeneity of variance, respectively. The Mann-Whitney U test and the Kruskal-Wallis test, followed by Bonferroni correction, were used to compare nonparametric data among the groups. The Kendall tau-b (τb) correlation coefficient was calculated to evaluate the potential relationships between the parameters. The significance level was set at $P<.05$.

Ethical Considerations

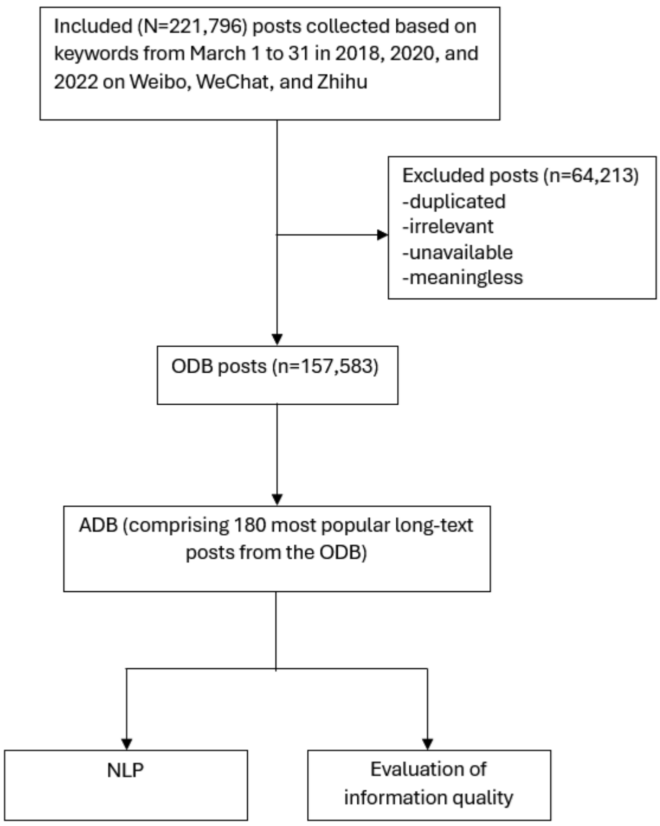
This study did not seek ethical approval, as it exclusively analyzed publicly available data from SMPs, which were voluntarily shared by users in the public domain. All data has been deidentified; account information collected was used solely for research analysis, and relevant results are presented in aggregate to ensure that no personally identifiable information is disclosed.

Results

Post Details

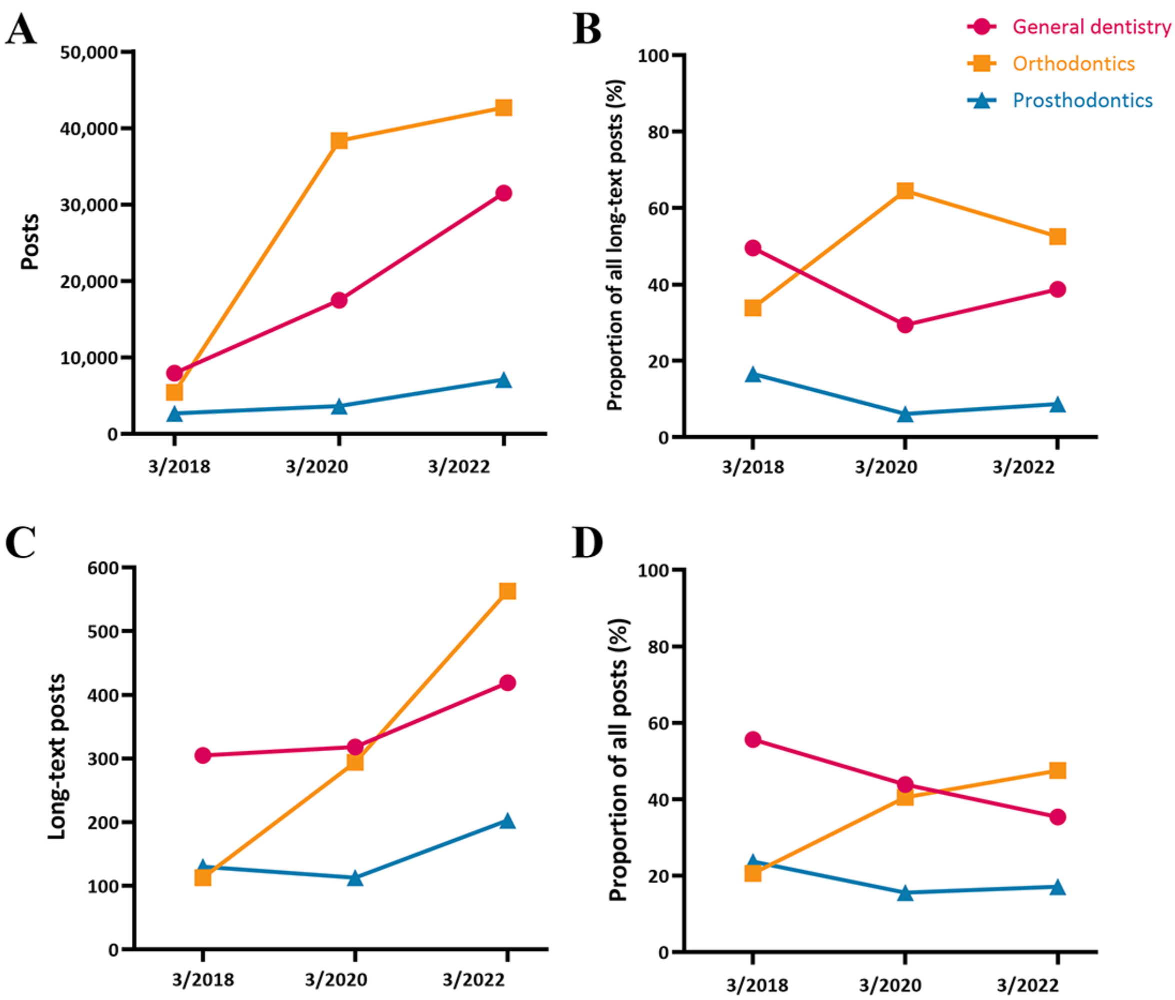
The research methodology is depicted in Figure 1. According to the search keywords, we retrieved a total of 220,869 Weibo posts. After applying the exclusion criteria, 64,039 (29%) posts were deleted, and the remaining 156,830 (71%) posts were included in the ODB. Among Weibo posts, 2458 (1.6%) were long-text posts. The distribution and proportion of posts and long-text posts for each theme in each observation period are shown in Table S2 in Multimedia Appendix 1. Results showed that over time, there has been significant growth in the content related to dental health care on Weibo, with the number of posts increasing by more than 4 times (Figure 2A). In terms of themes, in 2018, there was a greater proportion of posts and long-text posts discussing general dentistry–related topics, accounting for 49.5% (n=7952) and 55.7% (n=305) of the total, respectively. In 2022, 52.5% (n=42,701) of the posts and 47.5% (n=563) of the long-text posts discussed orthodontics topics (Figure 2D), surpassing the proportion of general dentistry topics. In addition, 543 (0.34%) posts from WeChat and 210 (0.13%) posts from Zhihu were included in the ODB.

Figure 1. Overview of data-processing flowchart. ADB: analysis database; NLP: natural language processing; ODB: original database.



XSL•FO
RenderX

Figure 2. Numbers and proportions of Weibo posts (A, B) and long-text Weibo posts (C, D) on each theme in March 2018, 2020, and 2022.



The ADB comprised 180 most popular long-text posts (n=60, 33.3%, posts from each platform) during the observation period. There were significant differences in the engagement of long-text posts published on Weibo, WeChat, and Zhihu between March 2018, 2020, and 2022 (*P*<.001; Figure 3 and Table S3 in Multimedia Appendix 1). The median (IQR) of likes on Weibo increased from 5.5 (IQR 3.25-12.5) in 2018 to 149.5 (IQR 87.75-454.75) in 2022. Similarly, the median number of likes on Zhihu increased from 63 (IQR 29.25-161.25) in 2018 to 214 (IQR 126-435.5) in 2022. In contrast, the median (IQR) of WeChat reads was 30,000 (IQR 20,000-54,250) in 2018, decreased to 13,000 (IQR 11,000-18,000) in 2020, and then increased to 25,500 (IQR 15,250-48,000) in 2022. Interestingly, 143 (79.4%) long-text posts were written by non–health care professionals, including patients and some medical self-media accounts. In terms of topics, there were 105 (58.3%) posts about personal medical experiences, a number significantly greater than the content of health education and popular science provided by dental health care professionals.

In the ADB, the median (IQR) of the DISCERN score for WeChat long-text posts was 3 (IQR 2-4), that for Weibo was 2 (IQR 2-3), and that for Zhihu was 2 (IQR 2-3). There was a significant difference in the DISCERN scores between WeChat and Weibo (*P*=.03; Figure 4 and Table S4 in Multimedia Appendix 1). The scores for each question in the DISCERN questionnaire corresponding to each platform are presented in Table S5 in Multimedia Appendix 1. Among the 180 long-text posts in the ADB, only 37 (20.6%) were authored by dental health care professionals, while the DISCERN scores of these long-text posts on Weibo and WeChat were significantly higher than those of long-text posts written by non–health care professionals. Specifically, for Weibo the median (IQR) of the DISCERN score of health care professionals' long-text posts was 3 (2.5-3.5), while that of non-health care professionals' was 2 (2-3); for WeChat the median (IQR) of the DISCERN score for health care professionals' long-text posts on WeChat was 4 (IQR 4-4), while that for non–health care professionals was 3 (IQR 2-3). The difference was statistically significant (Weibo *P*=.04; WeChat *P*=.02). Furthermore, there was a significant negative correlation between information quality (DISCERN score) and engagement (Weibo τb=–0.45, *P*=.01; WeChat τb=–0.30, *P*=.02). No similar significant negative correlation was observed for the Zhihu long-text posts.

XSL•FO
RenderX

Figure 3. Violin plots of engagement indicators for long-text posts on Weibo (A), WeChat (B), and Zhihu (C) in March 2018, 2020, and 2022.

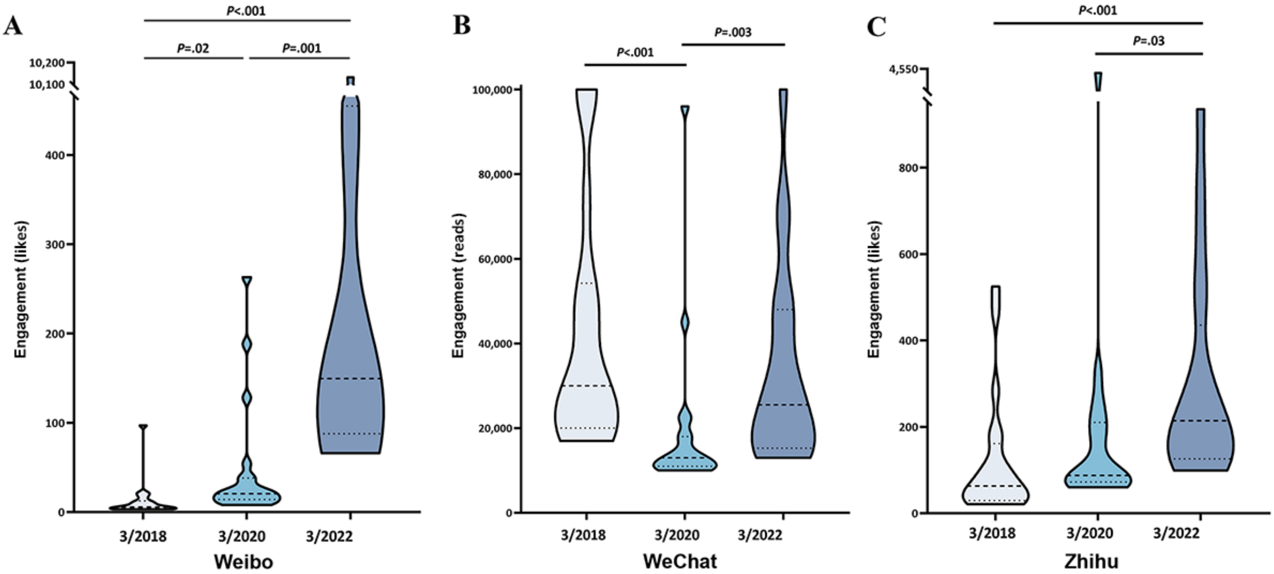
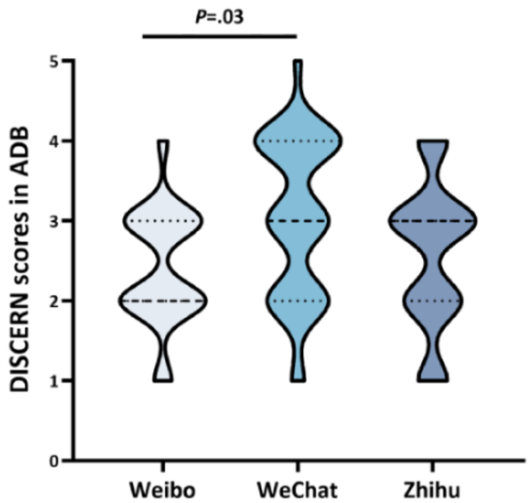


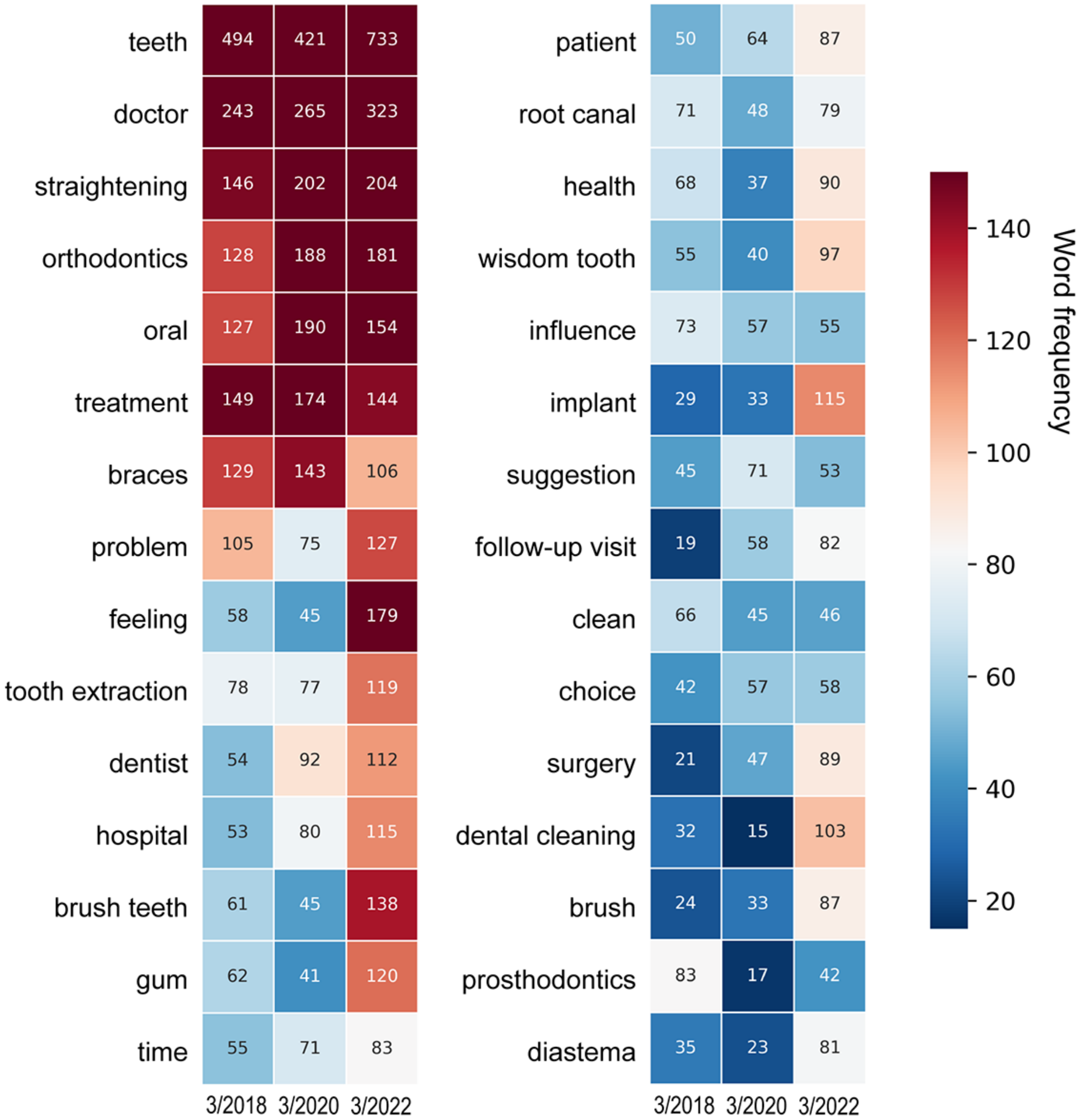
Figure 4. Violin plot of the DISCERN scores of long-text posts on the 3 SMPs in the ADB during the observation period. ADB: analysis database; SMP: social media platform.



Text mining analysis of the ADB using NLP tools was visualized as word clouds and heatmaps of the 3 observation periods (Figure 5) and the 3 platforms (Figure 6). For optimal readability, the bilingual word cloud figures are shown in Multimedia Appendices 3 and 4. The frequencies of the top 30 words in each time period and platform are presented in Tables S6 and S7, respectively, in Multimedia Appendix 1. In the word clouds, "teeth" and "doctor" were the most frequently mentioned core keywords. The most prominent terms in the 2018 word cloud were "straightening," "braces," "orthodontics," "prosthodontics," "tooth extraction," "root canal," "health," and "metal" (crown). In the 2020 word cloud, the most prominent terms were "straightening," "orthodontics," "braces," "tooth extraction," "invisible," "deciduous teeth," "follow-up visit," and "retention." In the 2022 word cloud, the most prominent terms were "straightening," "orthodontics," "feeling," "brush teeth," "gums," "tooth extraction," "implant," and "dental cleaning." This trajectory suggested that over time, orthodontics and tooth extraction have consistently been the most mentioned terms in Chinese social media in regard to dental health care. However, it could be seen from the word clouds and heatmaps that there was a shift from topics such as root canal treatments and metal crown restoration, which were discussed more often in 2018, to topics such as teeth brushing, gingival health, and dental cleaning by 2022, indicating the concentration on periodontal health and early prevention of dental diseases. There was also an increase in discussions related to dental implantation and orthognathic surgery. The most prominent terms in the Weibo word cloud were "straightening," "orthodontics," "wisdom tooth," "tooth extraction," and "gum." The most prominent terms in the WeChat word cloud were "straightening," "brush teeth," "health," "deciduous teeth," "orthodontics," and "gum." In Zhihu, the most prominent terms were "orthodontics," "braces," "feeling," "straightening," "tooth extraction," and "follow-up visit."

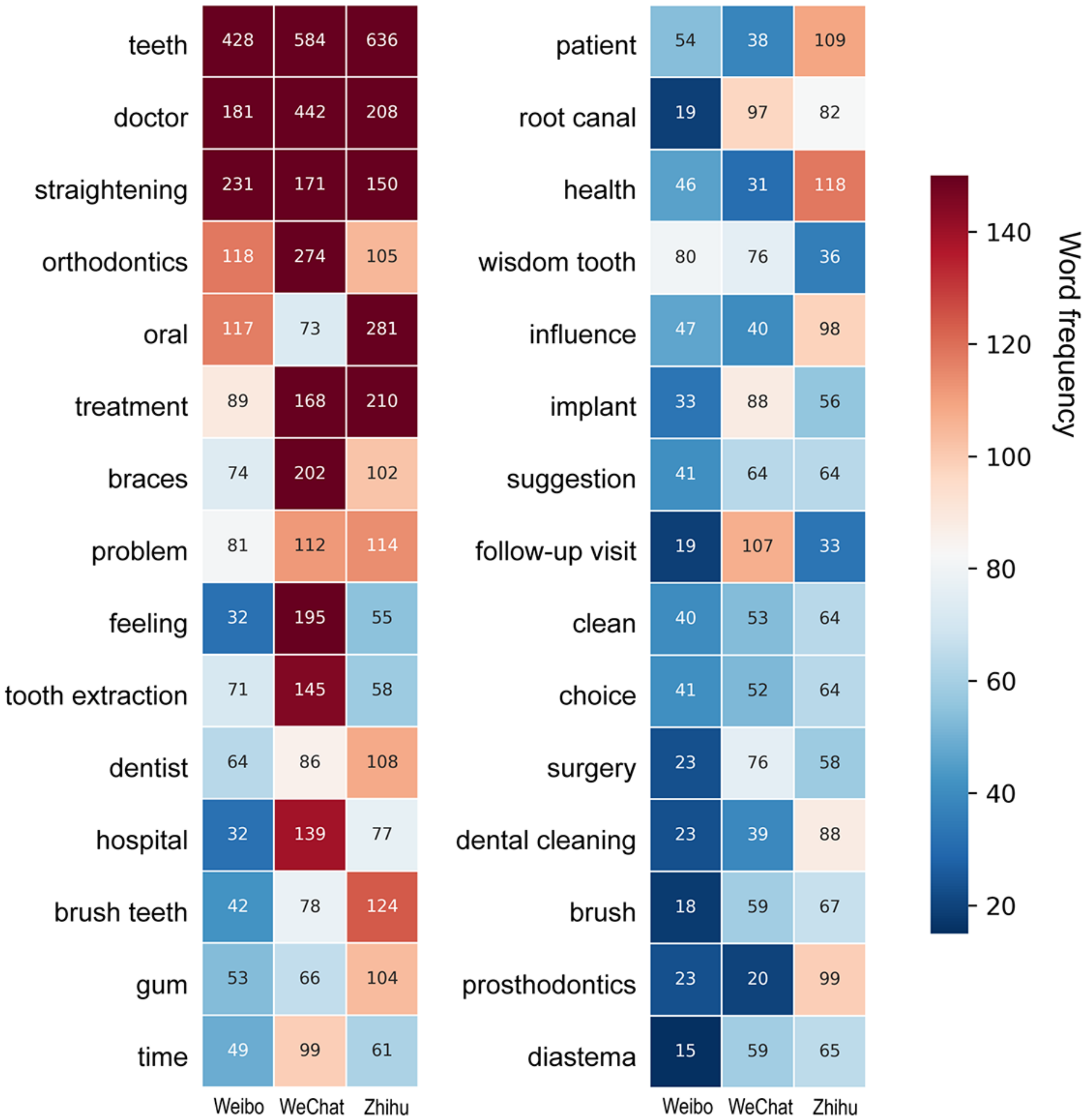
XSL•FO
RenderX

Figure 5. Heatmap of the frequencies of the top 30 words of dental health care information in the ADB in March 2018, 2020, and 2022. Orthodontics and tooth extraction have consistently been the most mentioned terms in Chinese SMPs in regard to dental health care. However, there was a shift from topics such as root canal treatments and metal crown restoration, which were discussed more often in 2018, to topics such as teeth brushing, gingival health, and dental cleaning by 2022. ADB: analysis database; SMP: social media platform.



XSL•FO
RenderX

Figure 6. Heatmap of the frequencies of the top 30 words of dental health care information across the 3 SMPs in the ADB. ADB: analysis database; SMP: social media platform.



Discussion

Principal Findings

This study extracted dental health care-related content from Weibo, WeChat, and Zhihu for March 2018, 2020, and 2022 and constructed an ODB. We also constructed an ADB consisting of the most popular long-text posts on each platform. By analyzing the nature, themes, public engagement, information quality, and word frequency, this study tracked the evolution of Chinese social media content related to dental health care. In the field of dentistry, previous studies have investigated SMPs, such as Facebook, Instagram, and YouTube, analyzing topic trends and evaluating the adequacy of these platforms as patient sources of information or education [31-33]. Graf et al [32] explored the nature and potential attitude differences in German orthodontic content on Twitter and Instagram, finding “getting braces” and “getting braces removed” to be the most crucial events for orthodontic patients, and Instagram contained more posts with positive emotions. Yahya et al [34] studied the 63 most viewed videos on YouTube related to miniscrew anchorage and found that the information quality of videos uploaded by dental professionals is not perfect, especially in terms of treatment duration, maintenance, and costs. Similarly, Samur et al [35] reported that the reliability and information quality of content related to facial trauma on SMPs are generally low, highlighting the need for caution when

XSL•FO
RenderX

recommending SMPs as a source of facial trauma–related information. To the best of our knowledge, similar systematic research on dental health care information on Chinese SMPs has not been reported in the literature.

Our findings revealed significant growth, by more than 4 times, in Weibo posts concerning dental health care from 2018 to 2022, with the fastest increase observed in the discussion of orthodontics, surpassing general dentistry content since 2020. The long-text posts with the highest engagement on Weibo and Zhihu platforms in the ADB also displayed an upward trend. SMPs plays a progressively significant role in people's lives [36], and our data confirmed that the amount of dental health care–related content on Chinese SMPs is also steadily increasing. Yang et al [37] reported that Instagram accounts created by oral and maxillofacial surgery residency programs increased exponentially from the period of the 7 months from June to December 2020 compared to the 18 months from December 2018 to May 2020. This indirectly reflected the increasing trend of dental health care information on US SMPs, which is consistent with our results.

In the ADB, which consisted of the most popular content on the 3 SMPs, 79.4% of the long-text posts were written by individuals without a health care background, and 58.3% shared personal medical experiences. This finding was consistent with the study by Samur et al [35], who found that personal experience–based content posted by laypersons receives more interactions. This phenomenon aligns with the theories of cognitive dissonance and selective exposure: people are inclined to consume content that is more similar and relatable to them [38,39]. It is worth noting that long-text posts authored by health care professionals received significantly higher DISCERN scores, indicating superior information quality. Nonetheless, these long-text posts were not rewarded with the same level of engagement as those written by nonprofessionals. Furthermore, there was a negative correlation between the DISCERN scores and engagement observed on Weibo and WeChat, suggesting that high-quality information may not generate a larger audience. Similar trends were observed in the study by Hegarty et al [40], who found that the most viewed YouTube videos are less helpful. However, a study by Kovalski et al [41] on oral leukoplakia–related content showed that more reliable videos of higher quality receive more likes and have higher viewing rates and interaction indices. Studies indicate that misinformation often features sensational headlines that are easy to understand without deep engagement or critical thinking. These posts quickly capture attention, eliciting strong emotional responses that prompt users to like, comment, and share [9,42]. The positive feedback loop of social media algorithms amplifies the spread of low-quality content, indirectly suppressing accurate, evidence-based, high-quality information and creating information silos [43].

Peek et al's [44] guidelines for mental health education and advocacy noted that using language that is more accessible to the public instead of medical jargon can make popular science even more popular. However, in the study by Yahya et al [34] on 31 videos uploaded by dental professionals about miniscrew anchorage on YouTube, only 2 videos avoided using technical terms. The remaining videos all used them, with 80.7% failing to provide explanations. The excessive use of technical terms, obscure and complicated principles, and stagnant formatting in health education materials may meet the requirements of the DISCERN questionnaire and thus can elicit high information quality scores. However, several studies have revealed that this may result in a loss of readers' interest and psychological resonance [35,39]. Conversely, using psychologically assisted writing techniques that foster affinity may prove more effective in attracting readers and achieving better health education outcomes. To effectively disseminate health care knowledge, dental professionals should improve writing methodologies, while ensuring the accuracy and high quality of the evidence-based information conveyed. For instance, incorporating patient-centered medical experiences and visual materials can attract more readers and stimulate discussions. This could ultimately make a greater impact and benefit a larger audience. In addition to health care professionals, governments and health departments should take measures to promote the dissemination of high-quality health information online, while enhancing the public's critical-thinking skills to discern true from false information and make informed decisions [9,13]. Technology platforms should transparentize and optimize recommendation algorithms and regulate the quality of health [9,42].

Oliveira et al [45] performed 2 searches on Twitter using the keywords "dentist" and "teeth" and generated a word cloud based on the collected tweets, finding that the most commonly used terms are "third molar" and "orthodontic appliance." On this basis, they determined that the most common dental needs during the COVID-19 pandemic were pain, urgencies, and orthodontic follow-ups. Graf et al [32] showcased word clouds of orthodontically posts with different sentiments. Positive posts revolved around the effectiveness of orthodontic treatment and the excitement of wearing or removing braces, while negative posts covered complaints about appointments, waiting times, pain, and side effects during orthodontic treatment. In this study, word clouds were generated for 3 SMPs during the observation period. We found that "teeth" and "doctors" are consistently the core subjects in the field of dental health care on Chinese SMPs. In line with the findings of Oliveira et al [45], orthodontics and tooth extraction have been the most discussed topics across different years and platforms, suggesting that the most prevalent keywords on Chinese SMPs are similar to those in other languages. Moreover, there has been a noticeable shift from topics such as endodontic treatment or dental crown restoration in 2018 to a stronger emphasis on topics such as periodontal maintenance and early prevention by 2022. This is indicated by the increased discussions around teeth brushing, gingival health, and dental cleaning. Additionally, topics such as dental implantation and orthognathic surgery have gradually gained popularity in the word clouds, suggesting that concepts in these areas are becoming more universally accepted by the public. Patients interest has gradually evolved from basic dental treatments to functional dentofacial aesthetics and preventative care.

This study demonstrates the increase in the quantity and engagement of dental health information on Chinese social media from 2018 to 2022, emphasizing the importance of

XSL•FO
RenderX

offering high-quality information online in the digital age. Additionally, we explored changes in public interest in dental topics, providing insights into the evolving awareness of patients. This highlights the need for dental health care providers to supplement evidence-based information, particularly on topics of interest to the general public. They should take measures to improve the popularity of online scientific materials and mitigate the impact of low-quality information.

Limitations

This study has a few limitations. First, there are constrained capabilities of NLP tools when evaluating complex indicators, such as DISCERN scores, across large datasets, as well as identifying meaningless content, such as bot-generated text. Consequently, this task was performed manually as a substitute, which inevitably restricted the scale of relevant data analysis, necessitating the use of sampling methods rather than analyzing the entire dataset. Second, platforms such as WeChat and Zhihu do not provide a straightforward way to access comprehensive data, prompting us to use alternative strategies for collecting a representative sample. We also note that including data from 2019 and 2021 could have resulted in more continuous and reliable sampling points. Third, this study was retrospective in design, while, given the dynamic nature of social media, some users may have hidden or deleted previously published posts, potentially introducing bias into the findings. Finally, expanding the scope of this study to include data from additional sources, such as government agencies or dental associations, would facilitate a comparative analysis. Future research could focus on these aspects.

Conclusion

During 2018-2022, despite the increase in the dissemination and evolution of public interest in dental health care information on Chinese social media, the quality of the most popular long-text posts was rated as moderate or low, which may mislead patients and the public. These findings could yield insights for dental practitioners, investigators, and educators into patients' evolving perceptions and interests in the era of social media. We also emphasize the importance of enhancing the provision of high-quality and popular health information on Chinese SMPs.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2022YFC2405904), the National Natural Science Foundation of China (11932012), and the Fundamental research program funding of Ninth People's Hospital affiliated to the Shanghai Jiao Tong Uuniversity School of Medicine (JYZZ085B).

Data Availability

The databases generated and analyzed in this work are available from the corresponding author upon reasonable request.

Authors' Contributions

ZZ and ZY contributed equally to this work. BF, LX, and ZL also contributed equally to this work. ZZ was responsible for conceptualization, methodology, formal analysis, investigation, and writing—original draft; ZY for conceptualization, methodology, and investigation; QW for software, investigation, data curation, and writing—review; RL for investigation, visualization, writing—review, and supervision; HL for conceptualization and methodology; WG for formal analysis and writing—review; ZL for writing—review and project administration; LX for methodology, writing—review, and supervision; and BF for conceptualization, writing—review and editing, supervision, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Chinese keywords for searching posts and their corresponding English translation; distribution of the number and proportion of Weibo and long-text posts on different themes in the ODB; engagement of long-text posts on Weibo, WeChat, and Zhihu in the ADB; DISCERN scores of long-text posts on the 3 SMPs in the ADB during the observation period; DISCERN scores of each question of long-text posts on the 3 SMPs in the ADB during the observation period; frequencies of the top 30 words in the ADB in March 2018, 2020, and 2022 and on Weibo, WeChat, and Zhihu. ADB: analysis database; ODB: original database; SMP: social media platform.

[DOCX File , 35 KB - infodemiology_v5i1e55065_app1.docx]

Multimedia Appendix 2

DISCERN health information quality assessment questionnaire.

[DOCX File , 19 KB - infodemiology_v5i1e55065_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Word clouds in English.
[PDF File (Adobe PDF File), 24432 KB - infodemiology_v5i1e55065_app3.pdf]

Multimedia Appendix 4
Word clouds in Chinese.
[PDF File (Adobe PDF File), 12340 KB - infodemiology_v5i1e55065_app4.pdf]

References

1. Global social media statistics. DataReportal. URL: https://datareportal.com/social-media-users [accessed 2025-04-02]
2. Simon K. Digital 2024: China. DataReportal. 2024 Feb 21. URL: https://datareportal.com/reports/digital-2024-china [accessed 2025-04-02]
3. Farsi D. Social media and health care, part I: literature review of social media use by health care providers. J Med Internet Res 2021 Apr 05;23(4):e23205 [FREE Full text] [doi: 10.2196/23205] [Medline: 33664014]
4. Farsi D, Martinez-Menchaca HR, Ahmed M, Farsi N. Social media and health care (part ii): narrative review of social media use by patients. J Med Internet Res 2022 Jan 07;24(1):e30379 [FREE Full text] [doi: 10.2196/30379] [Medline: 34994706]
5. Zhang X, Wen D, Liang J, Lei J. How the public uses social media wechat to obtain health information in China: a survey study. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
6. Braun LA, Zomorodbakhsch B, Keinki C, Huebner J. Information needs, communication and usage of social media by cancer patients and their relatives. J Cancer Res Clin Oncol 2019 Jul;145(7):1865-1875 [FREE Full text] [doi: 10.1007/s00432-019-02929-9] [Medline: 31123824]
7. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 09;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
8. Sylvia Chou W, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. Am J Public Health 2020 Oct;110(S3):S273-S275. [doi: 10.2105/AJPH.2020.305905] [Medline: 33001722]
9. Office of the Surgeon General (OSG). Confronting Health Misinformation: The U.S. Surgeon General's Advisory on Building a Healthy Information Environment. Washington, DC: US Department of Health and Human Services; 2021.
10. Bari A, Heymann M, Cohen RJ, Zhao R, Szabo L, Apas Vasandani S, et al. Exploring coronavirus disease 2019 vaccine hesitancy on Twitter using sentiment analysis and natural language processing algorithms. Clin Infect Dis 2022 May 15;74(Suppl_3):e4-e9. [doi: 10.1093/cid/ciac141] [Medline: 35568473]
11. Smirnov O, Hsieh P. COVID-19, climate change, and the finite pool of worry in 2019 to 2021 Twitter discussions. Proc Natl Acad Sci U S A 2022 Oct 25;119(43):e2210988119 [FREE Full text] [doi: 10.1073/pnas.2210988119] [Medline: 36251993]
12. Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. Annu Rev Public Health 2020 Apr 02;41:101-118 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094402] [Medline: 31905322]
13. Sage WM, Yang YT. Reducing "COVID-19 misinformation" while preserving free speech. JAMA 2022 Apr 19;327(15):1443-1444. [doi: 10.1001/jama.2022.4231] [Medline: 35357403]
14. Burki T. Social media and misinformation in diabetes and obesity. Lancet Diabetes Endocrinol 2022 Dec;10(12):845. [doi: 10.1016/S2213-8587(22)00318-7] [Medline: 36335973]
15. Yeung AWK, Tosevska A, Klager E, Eibensteiner F, Tsagkaris C, Parvanov ED, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. J Med Internet Res 2022 Jan 25;24(1):e28152 [FREE Full text] [doi: 10.2196/28152] [Medline: 34951864]
16. Tao Z, Chu G, McGrath C, Hua F, Leung YY, Yang W, et al. Nature and diffusion of COVID-19-related oral health information on Chinese social media: analysis of tweets on Weibo. J Med Internet Res 2020 Jun 15;22(6):e19981 [FREE Full text] [doi: 10.2196/19981] [Medline: 32501808]
17. Guo F, Tang B, Qin D, Zhao T, Su Y, McGrath C, et al. The impact of the COVID-19 epidemic on orthodontic patients in China: an analysis of posts on Weibo. Front Med (Lausanne) 2020;7:577468 [FREE Full text] [doi: 10.3389/fmed.2020.577468] [Medline: 33364245]
18. Somani S, van Buchem MM, Sarraju A, Hernandez-Boussard T, Rodriguez F. Artificial intelligence-enabled analysis of statin-related topics and sentiments on social media. JAMA Netw Open 2023 Apr 03;6(4):e239747 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.9747] [Medline: 37093597]
19. Hirschberg J, Manning CD. Advances in natural language processing. Science 2015 Jul 17;349(6245):261-266. [doi: 10.1126/science.aaa8685] [Medline: 26185244]
20. Samaras L, García-Barriocanal E, Sicilia M. Comparing social media and Google to detect and predict severe epidemics. Sci Rep 2020 Mar 16;10(1):4747 [FREE Full text] [doi: 10.1038/s41598-020-61686-9] [Medline: 32179780]
21. Garrett RK. Social media's contribution to political misperceptions in U.S. presidential elections. PLoS One 2019;14(3):e0213500 [FREE Full text] [doi: 10.1371/journal.pone.0213500] [Medline: 30917154]

XSL•FO
RenderX

22. Jafari Sadr M, Mirtaheri SL, Greco S, Borna K. Popular tag recommendation by neural network in social media. Comput Intell Neurosci 2023;2023:4300408 [FREE Full text] [doi: 10.1155/2023/4300408] [Medline: 37288169]
23. Carpenter KA, Altman RB. Using GPT-3 to build a lexicon of drugs of abuse synonyms for social media pharmacovigilance. Biomolecules 2023 Feb 18;13(2):387 [FREE Full text] [doi: 10.3390/biom13020387] [Medline: 36830756]
24. Zhang Z, Zhou X, Guo C, Cheng B. Stomatology (9th Edition). China: People's Medical Publishing House; 2018.
25. Proffit W, Fields H, Larson B, Sarver D. Contemporary Orthodontics. Amsterdam, the Netherlands: Elsevier Health Sciences; 2018.
26. Berman L, Hargreaves K, Cohen S. Cohen's Pathways of the Pulp Expert Consult. Amsterdam, the Netherlands: Elsevier Health Sciences; 2010.
27. Gkikas DC, Tzafilkou K, Theodoridis PK, Garmpis A, Gkikas MC. How do text characteristics impact user engagement in social media posts: modeling content readability, length, and hashtags number in Facebook. Int J Inf Manag Data Insights 2022 Apr;2(1):100067. [doi: 10.1016/j.jjimei.2022.100067]
28. Loeb S, Reines K, Abu-Salha Y, French W, Butaney M, Macaluso JN, et al. Quality of bladder cancer information on YouTube. Eur Urol 2021 Jan;79(1):56-59. [doi: 10.1016/j.eururo.2020.09.014] [Medline: 33010986]
29. Gorrepati PL, Smith GP. DISCERN scores of YouTube information on eczema treatments. J Am Acad Dermatol 2021 Nov;85(5):1354-1355. [doi: 10.1016/j.jaad.2020.11.007] [Medline: 33181240]
30. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
31. Giese H, Neth H, Moussaïd M, Betsch C, Gaissmaier W. The echo in flu-vaccination echo chambers: selective attention trumps social influence. Vaccine 2020 Feb 18;38(8):2070-2076. [doi: 10.1016/j.vaccine.2019.11.038] [Medline: 31864854]
32. Graf I, Gerwing H, Hoefer K, Ehlebracht D, Christ H, Braumann B. Social media and orthodontics: a mixed-methods analysis of orthodontic-related posts on Twitter and Instagram. Am J Orthod Dentofacial Orthop 2020 Aug;158(2):221-228. [doi: 10.1016/j.ajodo.2019.08.012] [Medline: 32576427]
33. Meade MJ, Dreyer CW. Orthodontic temporary anchorage devices: a qualitative evaluation of Internet information available to the general public. Am J Orthod Dentofacial Orthop 2020 Oct;158(4):612-620. [doi: 10.1016/j.ajodo.2020.02.008] [Medline: 32988569]
34. Yahya M, Wang GM, Nimbalkar S. Orthodontic treatment with miniscrew anchorage: analysis of quality of information on YouTube. Am J Orthod Dentofacial Orthop 2023 Jul;164(1):97-105. [doi: 10.1016/j.ajodo.2022.11.013] [Medline: 36890012]
35. Samur Erguven S, Topsakal KG. Are social media platforms appropriate sources of information for patients regarding the topic of facial trauma? J Oral Maxillofac Surg 2023 Oct;81(10):1270-1278. [doi: 10.1016/j.joms.2023.06.004] [Medline: 37391172]
36. Chiang AL. Social media and medicine. Nat Rev Gastroenterol Hepatol 2020 May;17(5):256-257. [doi: 10.1038/s41575-020-0289-5] [Medline: 32203401]
37. Yang SC, Kwon WM, Saghezchi S. An update on Instagram utilization by oral and maxillofacial surgery residency programs: the social media big bang. J Oral Maxillofac Surg 2021 Jul;79(7):1394-1395. [doi: 10.1016/j.joms.2021.02.029] [Medline: 33775607]
38. Hart W, Albarracín D, Eagly AH, Brechan I, Lindberg MJ, Merrill L. Feeling validated versus being correct: a meta-analysis of selective exposure to information. Psychol Bull 2009 Jul;135(4):555-588 [FREE Full text] [doi: 10.1037/a0015701] [Medline: 19586162]
39. Yeung A, Ng E, Abi-Jaoude E. TikTok and attention-deficit/hyperactivity disorder: a cross-sectional study of social media content quality. Can J Psychiatry 2022 Dec;67(12):899-906 [FREE Full text] [doi: 10.1177/07067437221082854] [Medline: 35196157]
40. Hegarty E, Campbell C, Grammatopoulos E, DiBiase AT, Sherriff M, Cobourne MT. YouTube™ as an information resource for orthognathic surgery. J Orthod 2017 Jun;44(2):90-96. [doi: 10.1080/14653125.2017.1319010] [Medline: 28463076]
41. Kovalski LNS, Cardoso FB, D'Avila OP, Corrêa APB, Martins MAT, Martins MD, et al. Is the YouTube™ an useful source of information on oral leukoplakia? Oral Dis 2019 Nov;25(8):1897-1905. [doi: 10.1111/odi.13161] [Medline: 31332880]
42. Tang L, Fujimoto K, Amith MT, Cunningham R, Costantini RA, York F, et al. "Down the rabbit hole" of vaccine misinformation on YouTube: network exposure study. J Med Internet Res 2021 Jan 05;23(1):e23262 [FREE Full text] [doi: 10.2196/23262] [Medline: 33399543]
43. Lindström B, Bellander M, Schultner DT, Chang A, Tobler PN, Amodio DM. A computational reward learning account of social media engagement. Nat Commun 2021 Feb 26;12(1):1311 [FREE Full text] [doi: 10.1038/s41467-020-19607-x] [Medline: 33637702]
44. Peek HS, Richards M, Muir O, Chan SR, Caton M, MacMillan C. Blogging and social media for mental health education and advocacy: a review for psychiatrists. Curr Psychiatry Rep 2015 Nov;17(11):88. [doi: 10.1007/s11920-015-0629-2] [Medline: 26377948]

XSL•FO
RenderX

45. Oliveira LM, Zanatta FB. Self-reported dental treatment needs during the COVID-19 outbreak in Brazil: an infodemiological study. Braz Oral Res 2020 Sep 04;34:e114 [FREE Full text] [doi: 10.1590/1807-3107bor-2020.vol34.0114] [Medline: 32901729]

Abbreviations

ADB: analysis database
AI: artificial intelligence
NLP: natural language processing
ODB: social media original database
SMP: social media platform

Edited by T Mackey; submitted 04.12.23; peer-reviewed by A Geiken, D Chrimes, N Kaur; comments to author 10.04.24; revised version received 04.06.24; accepted 19.03.25; published 10.04.25.

Please cite as:
Zhu Z, Ye Z, Wang Q, Li R, Li H, Guo W, Li Z, Xia L, Fang B
Evolutionary Trend of Dental Health Care Information on Chinese Social Media Platforms During 2018-2022: Retrospective Observational Study
JMIR Infodemiology 2025;5:e55065
URL: https://infodemiology.jmir.org/2025/1/e55065
doi:10.2196/55065
PMID:

©Zhiyu Zhu, Zhiyun Ye, Qian Wang, Ruomei Li, Hairui Li, Weiming Guo, Zhenxia Li, Lunguo Xia, Bing Fang. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 10.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Large-Scale Deep Learning–Enabled Infodemiological Analysis of Substance Use Patterns on Social Media: Insights From the COVID-19 Pandemic

Julina Maharjan[1], PhD; Jianfeng Zhu[1], PhD; Jennifer King[2], Prof Dr; NhatHai Phan[3], Prof Dr; Deric Kenne[2], Prof Dr; Ruoming Jin[1], Prof Dr

[1]Department of Computer Science, Kent State University, Kent, OH, United States

[2]Department of Public Health, Kent State University, Kent, OH, United States

[3]Data Science Department, New Jersey Institute of Technology, Newark, NJ, United States

Corresponding Author:
Julina Maharjan, PhD
Department of Computer Science
Kent State University
800 E Summit St
Kent, OH, 44242
United States
Phone: 1 3305931365
Email: jmaharja@kent.edu

Abstract

Background: The COVID-19 pandemic intensified the challenges associated with mental health and substance use (SU), with societal and economic upheavals leading to heightened stress and increased reliance on drugs as a coping mechanism. Centers for Disease Control and Prevention data from June 2020 showed that 13% of Americans used substances more frequently due to pandemic-related stress, accompanied by an 18% rise in drug overdoses early in the year. Simultaneously, a significant increase in social media engagement provided unique insights into these trends. Our study analyzed social media data from January 2019 to December 2021 to identify changes in SU patterns across the pandemic timeline, aiming to inform effective public health interventions.

Objective: This study aims to analyze SU from large-scale social media data during the COVID-19 pandemic, including the prepandemic and postpandemic periods as baseline and consequence periods. The objective was to examine the patterns related to a broader spectrum of drug types with underlying themes, aiming to provide a more comprehensive understanding of SU trends during the COVID-19 pandemic.

Methods: We leveraged a deep learning model, Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach (RoBERTa), to analyze 1.13 billion Twitter (subsequently rebranded X) posts from January 2019 to December 2021, aiming to identify SU posts. The model's performance was enhanced by a human-in-the-loop strategy that subsequently enriched the annotated data used during the fine-tuning phase. To gain insights into SU trends over the study period, we applied a range of statistical techniques, including trend analysis, k-means clustering, topic modeling, and thematic analysis. In addition, we integrated the system into a real-time application designed for monitoring and preventing SU within specific geographic locations.

Results: Our research identified 9 million SU posts in the studied period. Compared to 2019 and 2021, the most substantial display of SU-related posts occurred in 2020, with a sharp 21% increase within 3 days of the global COVID-19 pandemic declaration. Alcohol and cannabinoids remained the most discussed substances throughout the research period. The pandemic particularly influenced the rise in nonillicit substances, such as alcohol, prescription medication, and cannabinoids. In addition, thematic analysis highlighted COVID-19, mental health, and economic stress as the leading issues that contributed to the influx of substance-related posts during the study period.

Conclusions: This study demonstrates the potential of leveraging social media data for real-time detection of SU trends during global crises. By uncovering how factors such as mental health and economic stress drive SU spikes, particularly in alcohol and prescription medication, we offer crucial insights for public health strategies. Our approach paves the way for proactive, data-driven interventions that will help mitigate the impact of future crises on vulnerable populations.

XSL•FO
RenderX

(JMIR Infodemiology 2025;5:e59076) doi:10.2196/59076

KEYWORDS

substance use; social media; deep learning; Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach; human-in-the-loop; COVID-19

Introduction

Overview

Substance use (SU) is a pressing public health issue in the United States, with 58.7% of Americans aged ≥12 years using tobacco, alcohol, or illicit drugs in 2020, with an annual increase of 3.8% [1]. This includes 50% alcohol users, 18.7% tobacco users, and 13.5% illicit drug users [1]. The consequences of SU, such as deteriorating health and increased crime, have led to a significant rise in drug overdose deaths, reaching >91,000 in 2020 and >106,000 in 2021 [2]. The economic cost is substantial, with an estimated US $249 billion for alcohol misuse and >US $193 billion for illicit drug use annually [3]. The financial and health repercussions of SU demand a strategic focus on prevention research. Allocating resources to explore and counteract the causes of drug use can lead us toward a healthier and more economically resilient society.

Background

The year 2020, commonly referred to as the COVID-19 year, holds historical significance for health care researchers due to the emergence of the deadly coronavirus. The COVID-19 pandemic exhibited a profound connection with preexisting SU and mental health issues [4-6]. Various consequences, such as economic instability, social isolation, bereavement, and restricted access to health care services, escalated anxiety and stress levels among the population [7-10]. According to the Centers for Disease Control and Prevention, data as of June 2020 revealed that 13% of Americans reported initiating or intensifying SU as a means of coping with stress or emotions related to COVID-19 [11]. The Overdose Detection Mapping Application Program reports indicated an 18% rise in drug overdoses in the early months of the pandemic compared to the same period in 2019 [12]. Several other studies [13-15] also highlighted that changes in drug availability contributed to a rise in deaths related to illicit opioid use; for instance, if heroin became less accessible, individuals might resort to the more potent fentanyl.

Simultaneously, the COVID-19 pandemic led to internet use of up to 70% [16], leading to a record 11.1% growth in Twitter's (subsequently rebranded X) user base in 2020. This surge in social media engagement, while providing a vital connection for many, has also been directly linked to an increase in SU [17]. Research studies [18-22] have indicated the negative impacts of social media on mental health, including increased anxiety, depressive symptoms, and psychological burdens related to COVID-19, which have been correspondingly linked to an increase in SU as individuals seek coping mechanisms. Notably, previous studies [23-27] have also shown a strong correlation between social media use and SU, with evidence of users being influenced to use substances by their peers' behavior, such as tagging their social connections in their posts [26,27]. A few research studies [23,25] provide evidence that higher levels of exposure to substance-related content tend to develop positive norms and attitudes toward alcohol and drug use. Likewise, a study also showed that adolescents who are regularly active on social media have a greater likelihood of subsequent tobacco or cannabis use initiation [24]. In our research, we aim to identify these gaps in the knowledge of SU during the COVID-19 pandemic by analyzing social media content and making a comparison with pre- and postpandemic years. We achieve this through a deep learning model alongside various statistical methods. By comprehending the findings, the ultimate goal is to support public health sectors to develop more effective prevention and intervention strategies to control and prevent SU during global crises.

Related Studies

The onset of the COVID-19 pandemic has notably intensified global research on drug crises. Numerous studies [6,7,28-50] have examined the intersection of drug use and the pandemic's societal impacts. These investigations commonly revealed a significant correlation between the pandemic and shifts in SU patterns, impacting both people with or without SU disorder (SUD). Various studies [6,28,36,37] evidenced that the disruption in health care services during the COVID-19 pandemic period primarily impacted people with SUD and was thus linked to higher abuse of substances. However, many of those research studies relied on data from small cohorts [18,22,29,30] that predominantly used methodologies such as surveys or interviews for data collection. Few studies [30,35,39-41,47,48,51] have used social media data to explore SU during the pandemic. However, the scope of such studies often remains limited to peak pandemic periods and typically focuses on specific types of drugs, such as alcohol, tobacco, or opioids. Only 2 of the studies [14,52] accounted for multiple drug types (that are mostly consumed) to study the correlation between COVID-19 and use of substances, but they still did not consider other drug types (that are less widely used) to check if the use was altered during the global crisis. Likewise, most research only accounted for the peak pandemic period to study the SU trend during COVID-19. Only the study by Omare et al [47] accounted for the prepandemic period (2016-2020) as the baseline to compare the SU trend before the COVID-19 pandemic. Essentially, it established 2 prepandemic baselines, that is, 2016 to 2018 and 2018 to 2019, and compared SU trends over the studied period. However, it did not account for the postpandemic period or whether the SU was altered due to the consequences of COVID-19. This highlights a gap in the literature, underscoring the need for more expansive research that covers various substance types and multiple time frames to better understand the long-term impacts of the pandemic on drug use patterns.

Prominent national agencies such as the National Survey on Drug Use and Health [1], the National Institute on Drug Abuse

XSL•FO
RenderX

(NIDA) [Centers for Disease Control and Prevention [2], and the Substance Abuse and Mental Health Services Administration [50] routinely perform national-level analyses of drug use. Traditionally, these reports are based on survey methodologies, which may involve a relatively limited participant pool. The COVID-19 pandemic further complicated these efforts, limiting face-to-face data collection and necessitating a shift toward online surveys. This change compromised the depth and reliability of data in 2020; for example, the 2020 National Survey on Drug Use and Health report only includes data from the first quarter and used web-based methods for the fourth quarter [1]. In addition, the transition from *The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* to *The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* during this period introduced challenges in comparing the new data with those from previous years due to methodological changes.

The existing studies were limited to fewer drug types and demographics of smaller cohorts that mainly focused on the peak pandemic period and did not account for trends before and after the pandemic. Thus, in our research, we have aimed to use large-scale social media data to examine a broader spectrum of drug types, aiming to provide a more comprehensive understanding of drug use trends during the COVID-19 pandemic.

Previous research on social media often used keyword-based and traditional machine learning approaches to analyze drug-related content. Notably, studies [52,53] have identified potential SU incidents using keyword-based methods, that is, by detecting specific drug names such as Adderall, oxycodone, quetiapine, metformin, cocaine, marijuana, weed, methamphetamine, tranquilizer, etc. However, these keyword-based methods are limited, as they often fail to discern the context in which terms are used, resulting in significant ambiguities [54]. Users frequently use slang and metaphorical language that these models cannot adequately interpret. In addition, other studies [53,55-57] have used traditional machine learning classifiers such as naive Bayes, support vector machines, and decision trees. While enhancements such as word2vec for word embedding have been applied, these methods typically struggle with the subtleties of language used in social media. Despite some advancements in sequence-based models, such as long short-term memory or convolutional neural networks [54,58], these approaches still fall short of fully understanding contextual meanings, a challenge effectively addressed by the attention mechanism [59]. Thus, in our research, we have adopted the Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach (RoBERTa) [60] model, which leverages an advanced attention mechanism to overcome these limitations. This implementation represents a novel application in the analysis of large-scale social media data for drug use studies. Despite the challenges posed by limited annotated data availability, we have incorporated an iterative learning process inspired by human-in-the-loop (HITL) [61] and active learning techniques [62] to further enhance the accuracy of our model. This approach not only refines the model with each iteration but also focuses on learning from the most informative data points, streamlining the data annotation process.

In summary, in this research, we sought to study a large amount of data from Twitter spanning a 3-year period, including the prepandemic (2019) and postpandemic (2020) periods as baseline and consequence periods, to identify the patterns of drug use using a deep learning model (RoBERTa) and various other statistical methods (trend analysis, k-means clustering, topic analysis, and thematic analysis), which are explained in the Methods section in detail. In addition to this, we also aim to analyze different types of drugs and themes in the SU discourse. Specifically, we aim to answer the research questions presented in Textbox 1.

Textbox 1. Research questions.

1. How did the discourse on substance use (SU) evolve on Twitter (subsequently rebranded X) from 2019 to 2021, and what variations existed in the distribution of different substances during this time?
2. Following the announcement of the pandemic, what were the primary substance types that garnered significant discussion, and what were the themes of these dialogues?
3. How did the prevalence of the studied theme influence various types of substances during the underlying study period?
4. How did the identified themes correlate with the substance types?
5. What primary discussion topics arise from k-means analysis, specifically during the study period?
6. To what degree does the classifier's effectiveness in pinpointing SU-related tweets during the pandemic align with or differ from GPT-3?
7. How has the overall system contributed to the real-time tracking of SU, as evidenced by the research?

Contributions

The study's main contributions are as follows:

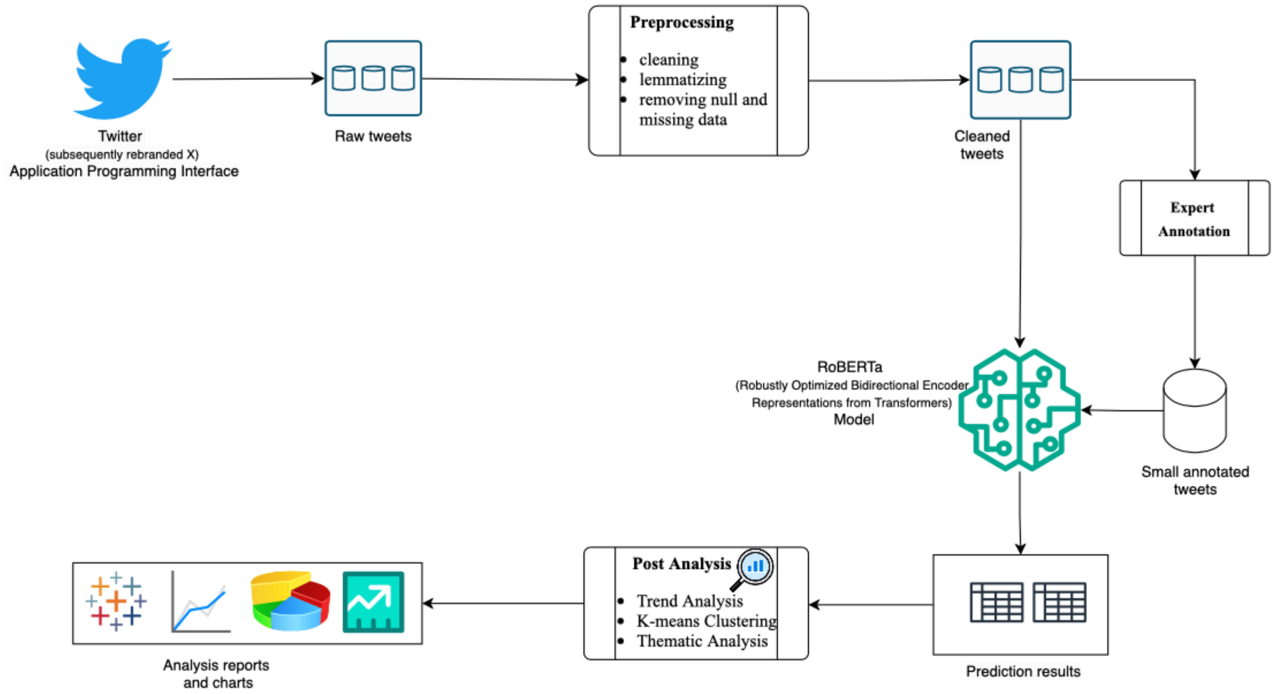
1. A large-scale SU behavior tweet collection system with expert-annotated tweets for supervised learning
2. A customized pretrained language model based on social media data (Twitter) and an iterative supervised deep learning algorithm for detecting SU posts
3. Insightful statistical analysis of the identified SU posts
4. A real-time search engine–based application for monitoring SU in temporal and spatial dimensions

Methods

Figure 1 shows the overall methodology used in the research. All the steps mentioned in the flow diagram are described subsequently.

XSL•FO
RenderX

Figure 1. Comprehensive research overview flowchart. API: application programming interface; BERT: Bidirectional Encoder Representations from Transformers; NA: not available; RoBERTa: Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach.



Data Collection

For this research, historical Twitter data were obtained from the Internet Archive [63], a digital library committed to providing free access to a wide array of digital information, including web pages, texts, audio, and videos. This nonprofit organization archives digital content to preserve it and make it accessible for future research. Among its many resources, the Internet Archive includes collections of Twitter data, which consist of tweets captured until July 2023. In our research, we downloaded the raw tweet data covering the period from January 2019 to December 2021. Initially, the data downloaded from this source were in compressed JSON formats, consisting of a large set of files for each day. A pipeline script was developed to extract these files and consolidate them into single-day JSON files. During the extraction process, we retrieved only the time stamp and the actual text of the posts for our analysis. It is important to note that the raw tweets for some days were missing in the data source, specifically in February 2020, January 2021, and April 2021. This absence resulted in skewed time series plots in these months, as discussed in the Results section.

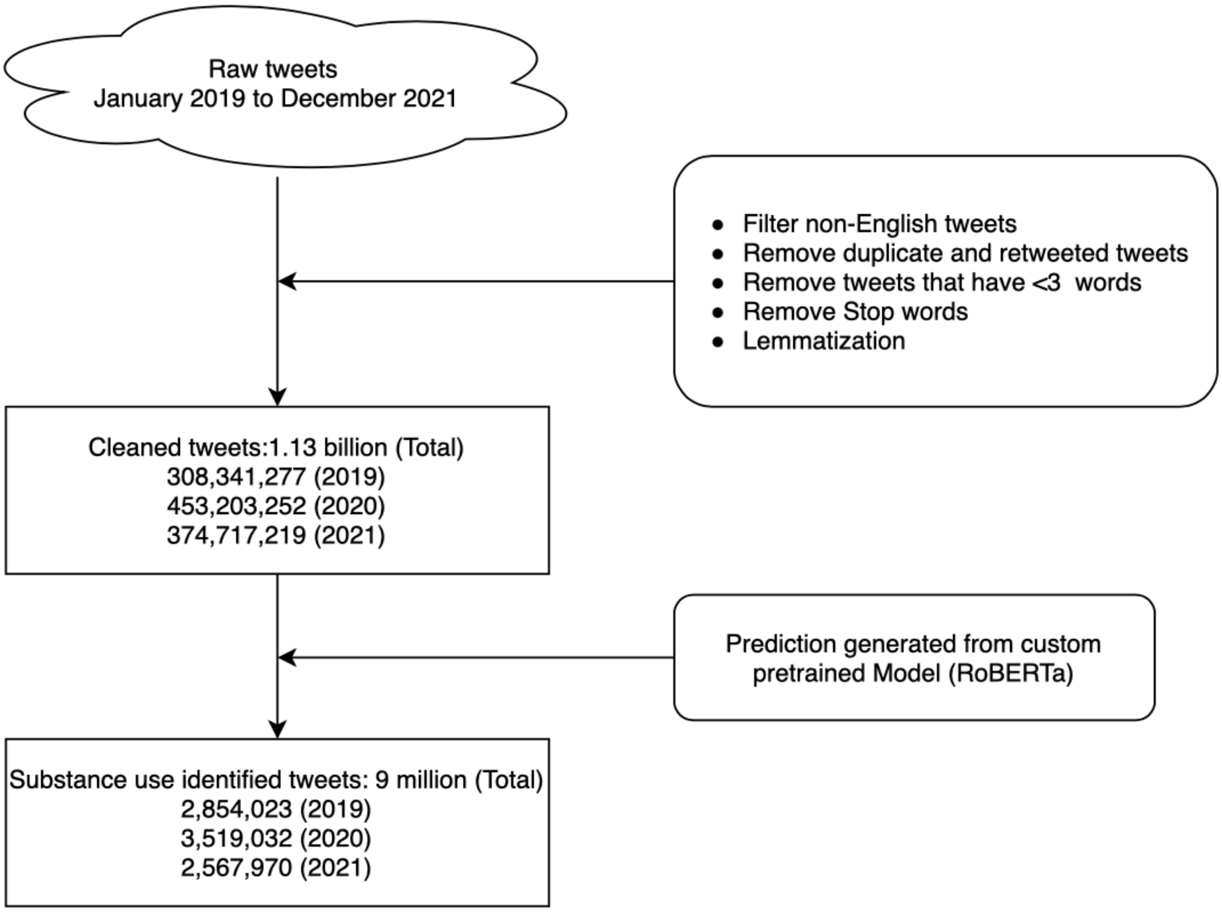
Data Preprocessing

The preprocessing of raw tweets was a crucial initial step to ensure the quality and relevance of the data for further analysis. To efficiently preprocess the large-size files, we divided each daily JSON file into smaller chunks, loaded them in memory, processed each chunk individually, and then merged them back into a single file. The preprocessing steps are described subsequently.

Initially, we filtered out all non-US tweets and duplicate or retweeted tweets to focus our research on English-language tweet posts and reduce redundancy, respectively. Then, we cleaned the text data by removing punctuation and stop words using the Nature Language ToolKit (*NLTK*) package and converted all characters to lowercase to maintain uniformity and prevent discrepancies caused by case sensitivity. Subsequently, we also replaced all the usernames, URLs, and hashtags in the post with the keywords “USER,” “HTTPURL,” and “HASHTAG” to hide the users’ identity and ease semantic understanding. Then, we performed lemmatization using the *NLTK* package to reduce words to their base form (eg, “drinking” to “drink”) to standardize text and improve consistency. Finally, we removed tweets containing <3 words, as these were deemed too brief to provide substantive insights. This comprehensive preprocessing approach resulted in a refined dataset of 1.13 billion cleaned tweets (n=308,341,277, 26.84% in 2019; n=453,203,252, 40.05% in 2020; and n=374,717,219, 33.11% in 2021) poised for further analysis, as depicted in Figure 2.

XSL•FO
RenderX

Figure 2. Flowchart of tweet processing. RoBERTa: Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach.



Feature Extraction and Data Annotation

Overview

Three specialized domain experts in mental health, SU, and public health performed the data annotation process. The main purpose of annotating data was to serve as a seeding dataset to train our deep learning RoBERTa model. Furthermore, as we intended to identify the SU posts in natural language social media data (where users might not clearly mention the drugs but still talk about SU), our goal was to collect and annotate data that were based on context rather than just keyword-based posts. Thus, we first outlined the context of SU based on 3 main criteria, namely *Types of Substance*, *Uses of Substances*, and *Intent to Use a Substance*.

Types of Substance

Substance type posts usually indicated either direct mention of drug names (that could be slang or street names) or described consuming them with or without actual drug names by specifying slang. The detailed list of such drug names, along with the street names and slang, are outlined in Table S1 in Multimedia Appendix 1. For instance, the tweet, “Man, just chill and smoke weed” had a direct mention of the substance “weed” with a clear meaning of SU. Likewise, “Just smoked a joint after work” had an indication of cannabis use hinted by the keyword *joint* even though the post had not specified the actual drug name. We acknowledge that Table S1 in Multimedia Appendix 1 contains a wide range of keywords or slang that might not have a direct association with SU. Hence, careful consideration has been made while annotating posts that contain slang but do not refer to SU. One counterexample of this would be “His joints and bones ache and his muscles seize up.” Here, the post does not have any context with SU even though it contains the keyword *joint*. Hence, it is labeled as a non-SU post.

Uses of Substances

SU posts were identified as posts that described the context of the use of substances, including experiences, effects, or consequences of consumption. The description usually covered personal anecdotes, stories, testimonials, promotions, advice, or recommendations about consumption, and information on obtaining substances. Examples included posts such as, “Feeling relaxed and happy after taking my meds—Xanax does wonders” and “Anyone needs advice on chilling out? I swear by CBD gummies.” In both examples, the post specified the consequence of consuming substances without mentioning the actual names of the substances.

Intent to Use a Substance

Substance intent posts were posts that exhibited actions or behaviors suggesting preparation for specific plans to engage in or a desire for SU and were classified as indicative of SU. Examples included “Planning to get some crystal tonight, can’t

XSL•FO
RenderX

wait" and "Thinking about getting high this weekend to unwind." These examples indicated the actual plan of consuming the substance without clearly mentioning the substance type.

Once the context of SU had been outlined, we proceeded with collecting tweets to annotate. We collected a subset of raw tweets (ie, without cleaning or preprocessing) from January 2020 through April 2020 and asked each domain expert to independently review and annotate a batch of collected tweets under the previously defined criteria. The annotation for each single post was confirmed only if at least 2 annotators voted the same. Upon discrepancies, the annotators further convened to discuss and repeated the process until a consensus was met. This iterative process ensured high reliability and validity in identifying instances of SU. The final annotation resulted in a corpus of 4011 posts. Sample examples of annotated SU and non-SU tweets are included in Table S2 in Multimedia Appendix 1. This thorough annotation process aided in creating a reliable training dataset for fine-tuning our SU classifier.

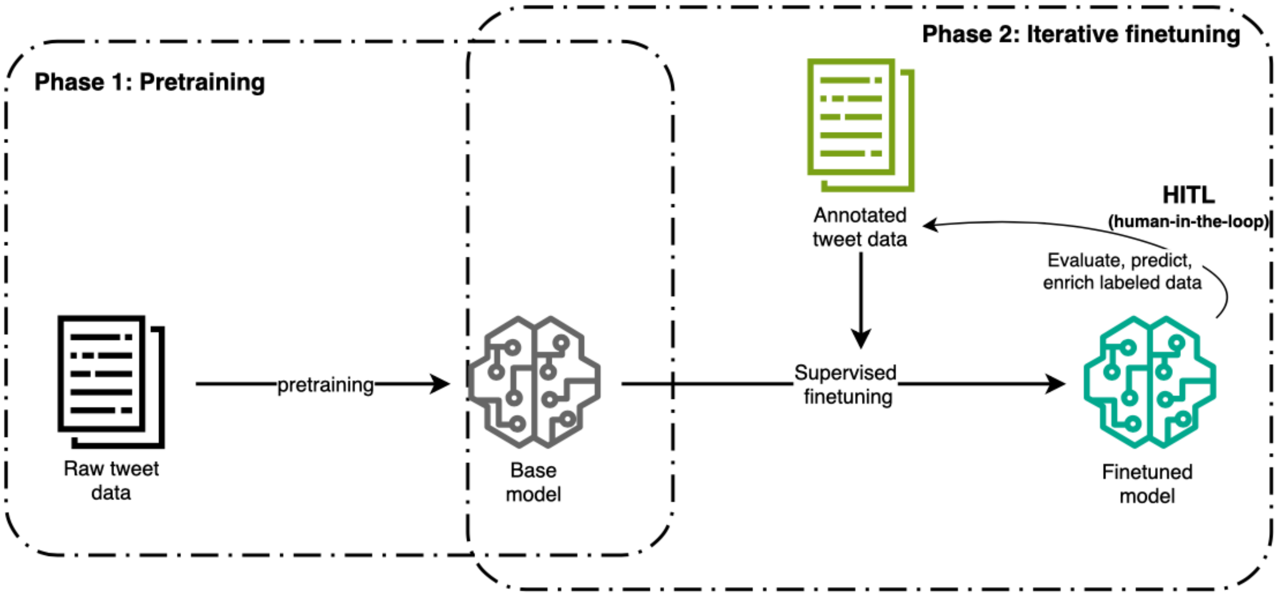
RoBERTa Model for Tweet Classification

Overview

In our research, we used the RoBERTa model [60], an advanced iteration of the Bidirectional Encoder Representations from Transformers (BERT) model [64], which itself marked a significant breakthrough in natural language processing. Developed by Google, BERT harnesses the power of the transformer architecture [59], notable for its innovative attention mechanism. This mechanism generates word embedding that captures deep contextual meanings within the text by enabling the model to consider each word in the context of all other words in a sentence rather than in isolation. The BERT model is structured to undergo 2 training phases, a pretraining phase and a fine-tuning phase, which is advantageous when adapting to specific tasks with limited available data. During the pretraining phase, the model learns general language patterns from a large text corpus through the masked language model (MLM) and next sentence prediction (NSP). MLM encourages the model to predict missing words based only on their context, enhancing its understanding of language nuances. NSP trains the model to understand the relationships between consecutive sentences, which is vital for tasks that require an appreciation of text flow. The fine-tuning phase then specifically adapts the pretrained model to nuanced tasks using smaller, specialized datasets, ensuring that the model maintains robust performance by refining the comprehensive linguistic capabilities developed during pretraining. Our objective could have been achieved by the BERT model; however, the elimination of the NSP task in the RoBERTa model simplifies the architecture, thereby making it the best fit for our use case. Unlike in BERT, RoBERTa only focuses on capturing contextual meaning (on the MLM task) rather than sentence relationships (on the NSP task), which is more relevant to tweet dataset context because tweet data are usually short sentences that do not require sentence relationship information. This modification makes RoBERTa more robust without compromising all the key features of BERT.

In the subsequent subsections, we explain the pretraining and fine-tuning phases carried out in our research, as depicted in Figure 3.

Figure 3. Illustration of training 2 phases Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach model for substance use tweet classification. HITL: human-in-the-loop.



Pretraining From Scratch

The specific linguistic challenges of our Twitter dataset necessitated a training-from-scratch approach to avoid biases from generic training datasets. In our method, we primarily adopted this customized approach to learn the language understanding of social media data from 33 million raw Tweet posts as shown in the phase 1 of Figure 3.

Initially, we performed tokenization using ByteLevelBPETokenizer. Essentially, this sub–word-level tokenizer broke down words into subword units, allowing it to handle out-of-vocabulary words and rare words more effectively than word-level tokenizers, thus enabling more coverages and

XSL•FO
RenderX

generalizations in domains with specialized terminology such as ours. We used 8192 vocab_size and min_frequency 2 as hyperparameters, along with [“<s>”, “<pad>”, “</s>,” “<unk>”, “mask>”] special tokens to indicate the start of the sentence token, padded token, end of the sentence token, unknown token, and masked token, respectively. Additional parameters are provided in Tables S3 and S4 in Multimedia Appendix 1.

After tokenization, we split the original dataset into 2 splits: as the training set (n=29.7 million, 90%), the testing set and (n=3.3 million, 10%). Then, we started training with the MLM objective, where a fraction of tokens (n=4.45 million, 0.15%) in each input sequence were masked, and the model learned to predict them based on contextual information. Training proceeded iteratively using stochastic gradient descent, with hyperparameters tuned based on validation performance. The model achieved a perplexity of 3.84 on the test data, which served as a baseline evaluation of the model. We ensured our language model was efficient by further evaluation after the fine-tuning step.

Iterative Fine-Tuning

Overview

After our model successfully deciphered language understanding in social media data, our next objective was to leverage this knowledge to distinguish posts related to SU. We achieved this by incorporating an additional binary classification layer into the existing model and retraining it with a newly labeled dataset, as depicted in the phase 2 of Figure 3. As with the unsupervised pretraining, we divided the dataset into training, validation, and test splits. We then retrained (fine-tuned) the newly annotated dataset, adjusting the model’s weights by calculating the error between the predictions and the actual labels using an optimization algorithm.

However, to prevent overfitting due to the limited size of the initially labeled data, we adopted an iterative fine-tuning approach inspired by HITL [61]. HITL is a collaborative technique that integrates human input at various stages of model development, such as training, testing, feedback, and decision-making [61]. In our case, we used HITL in only the training phase. We used human reviewers to only assess the model’s prediction results, which were then used to further train the model in successive rounds. Specifically, the process began with training the model on a seed-labeled dataset, followed by generating predictions for unseen data. Furthermore, these predictions were reviewed by human experts to refine and enrich the annotated dataset, which was used for subsequent training. This iterative cycle of training, prediction, and human review continuously improved the model’s performance by enhancing the quality of the training data.

The overall steps in our fine-tuning phase are detailed in the algorithm mentioned subsequently.

Step 1: Data Split

Before training the model, we split our 4011 annotated dataset into 3 sets: the training set (n=3208, 80%), the testing set (n=402, 10%), and the validation set (n=401, 10%).

Step 2: Initial Fine-Tuning and Cross-Validation

The initial parameters from the pretrained model were initialized. Then, the model was fine-tuned with the training dataset for 32 epochs on a batch size of 16 and a learning rate of 2e-5. A dropout layer was added to prevent overfitting, and the model was evaluated using a separate held-out dataset to ensure unbiased parameter tuning.

Step 3: HITL for Generating a New Labeled Dataset

Overview

Our ultimate objective at this step was to leverage human experts to pinpoint crucial data points that could enrich the annotated dataset and refine the model’s accuracy. Human experts reviewed the model’s predictions on unseen data and then identified and corrected errors. This feedback (corrected data) was then used to train the model in the next iteration.

Step 3.1: Prediction of Unseen Data

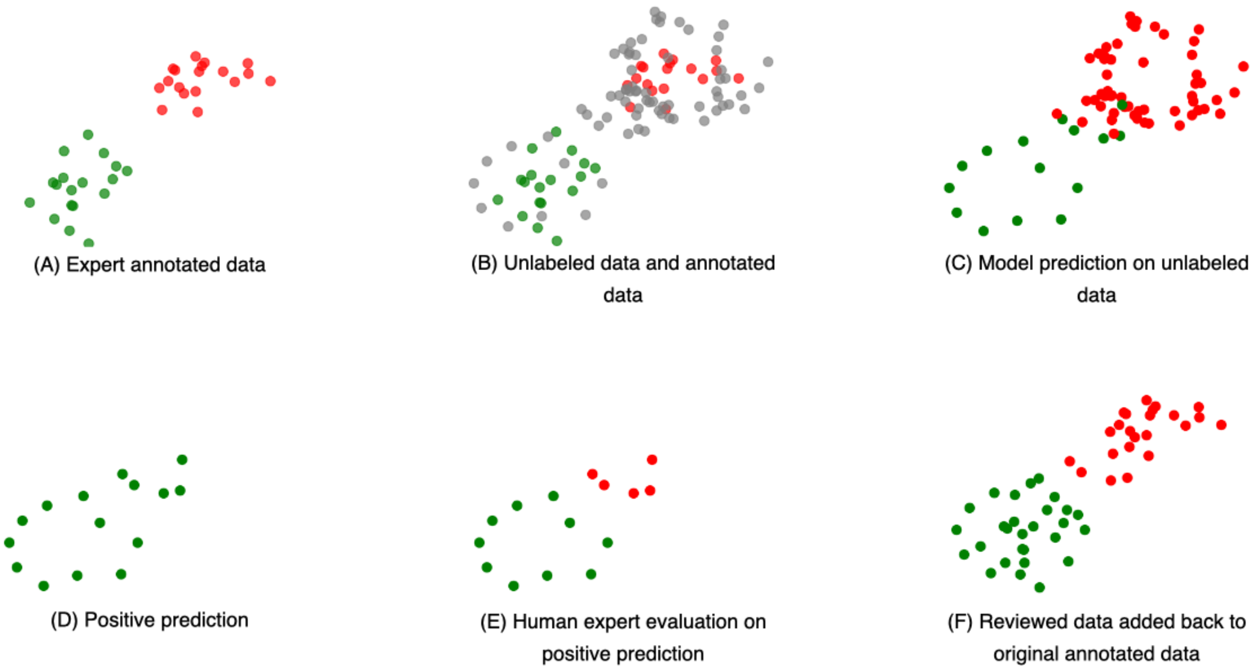
We used the refined model from step 2 to generate predictions of new, unseen data.

Step 3.2: Expert Review on Positive Predictions

Due to the severe imbalance in the dataset (0.05 positive, 0.95 negative), as shown in Figure 4, we focused exclusively on reviewing positive predictions. Concentrating on positives and making corrections to false negatives allowed us to directly improve the model’s sensitivity and precision. This approach ensured that the model better recognized the critical but infrequent true positive cases, thus enhancing overall accuracy and robustness.

XSL•FO
RenderX

Figure 4. Illustration of datasets in different stages used in the iterative fine-tuning phase. SU: substance use.



Step 3.3: Annotation

We leveraged our expert knowledge to annotate misclassified positives (false negatives) as negatives and correctly identified positives (true positives) as positives.

Step 3.4: Bias Reduction

To address potential bias by specifically focusing on positive predictions, we selectively reviewed a subset of complex positive predictions, which were likely to be misclassified as negatives. For example, posts using metaphoric language or slang, such as "riding the white horse," to indicate heroin use, required nuanced interpretation beyond simple keywords. By targeting these complex positives, we aimed to reduce the bias of not reviewing negative predictions. This careful attention to positive predictions ensured that we minimized the risk of failing to identify true instances of SU hidden within the data labeled as negative.

Step 3.5: Outcome

The final subset of true positives from Bias Reduction and misclassified positives from Annotation were considered as newly annotated data. In each iteration, we made sure the outcome contained 100 positive and 100 negative posts.

Step 4: Expansion of the Original Annotated Dataset

The outcomes from step 3 were added to the original annotated dataset.

Step 5: Evaluation and Iteration (Repeating Steps 1, 2, 3, and 4)

The fine-tuning was carried out for 20 iterations, expanding the annotated dataset in each iteration up to 6400 entries. At each round, we evaluated the model's accuracy on the test set and repeated the process until we achieved the desired accuracy of 80%.

In addition to achieving 80% accuracy, the model demonstrated strong performance across other metrics, with a recall of 79%, a precision of 85%, and an F_1-score of 81%. These scores indicated that the classifier effectively identified most SU instances while maintaining a low rate of false positives, ensuring balanced overall performance.

Substance Definitions and Their Types

A substance encompasses any psychoactive compound that can be legal, illegal, or medically prescribed, with potential impacts on health and society, including the risk of addiction. In our study, we classified substances into 10 primary categories based on their pharmacological and behavioral effects, following the categorization provided by the NIDA [65] and the Drug Enforcement Administration [66].

These categories are presented in Textbox 2.

The specifics for each substance category, including associated keywords, are detailed in Table S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Textbox 2. Classification of substances into 10 primary categories based on their pharmacological and behavioral effects.

1. Tobacco: includes cigarettes, vapor cigarettes, cigars, chewing tobacco, and snuff
2. Alcohol: covers all forms of beer, wine, and distilled spirits
3. Cannabinoids: encompasses marijuana, hashish, hash oil, and edibles containing cannabinoids
4. Opioids: includes drugs such as heroin, methadone, buprenorphine, oxycodone, Vicodin, and Lortab
5. Stimulants: includes cocaine, amphetamines, methamphetamine, methylphenidate (eg, Ritalin), and atomoxetine (eg, Strattera)
6. Club drugs: includes 3,4 methylenedioxymethamphetamine (MDMA) or ecstasy and gamma hydroxybutyrate (GHB)
7. Hallucinogens: lysergic acid diethylamide (LSD), psilocybin, mescaline, and dimethyltryptamine
8. Dissociative drugs: ketamine, phenyl cyclohexyl piperidine (PCP), and dextromethorphan
9. Prescription medications: a broad category that includes antibiotics, analgesics, statins, antidepressants, antihypertensives, hormonal contraceptives, and anticoagulants
10. Other compounds: features synthetic cannabinoids (eg, K2 or spice), anabolic steroids, inhalants, and synthetic cathinone (eg, mephedrone and methylenedioxypyrovalerone [MDPV])

Baseline Themes

The relationship between COVID-19 and SU patterns has garnered significant attention, with the COVID-19 pandemic serving as a critical case study. Previous research [4-6,14,15,31,36,37,67-70] highlighted various thematic areas that significantly influence SU, including stress and concerns related to COVID-19, economic instability, social dynamics, mental health issues, and disruptions in drug supply and health care services. Our study encompassed 6 key themes—COVID-19, economic factors, social influences, mental health, supply chain disruptions, and health care disruptions, as presented in Table 1. A short description of each theme and the impacted individuals (target population), along with study references, are listed in Table 1. To identify the themes in our dataset, we performed latent Dirichlet allocation (LDA) topic analysis to extract the tokens associated with each theme. Then, we refined the list of these tokens with the help of our experts. The complete list of tokens for each theme is detailed in Table S5 in Multimedia Appendix 1.

Table 1. Six major themes that impacted substance use during the global COVID-19 pandemic.

Themes	Description	Target	Studies
COVID-19	Worry or fear related to the virus and lockdown	All people with or without SUD[a]	[4,5,31,36,67]
Economic	Financial instability, job stress, housing, and food insecurity	All people with or without SUD	[5,68-70]
Social	Stress caused by the COVID-19 lockdown, social distancing policies, and change in daily routine	All people with or without SUD	[5,36,37,68,70]
Mental health	Anxiety and depression before COVID-19	Especially people with SUD	[4-6,36,37,68,70]
Supply disruption	Drug market disruptions	Especially people with SUD	[4,13-15]
Medical disruption	Decreased access to substance use treatment, harm reduction, and emergency services	Especially people with SUD	[4-6,68,70]

[a]SUD: substance use disorder.

Trend Analysis

Overview

Trend analysis is the most common technique for identifying patterns over time. In our study, trend analysis involved tracking and analyzing changes in types, discussion patterns, and themes associated with identified SU posts. We mainly used substance-type trend analysis, theme trend analysis, and k-means clustering analysis. At first, we identified the substance type, themes, and discussion pattern for each post by the keyword analysis based on Tables S1 and S5 in Multimedia Appendix 1, LDA topic analysis, and k-means clustering, respectively. Then, the subsequent trend analysis was performed.

Substance-Type Trend Analysis

To identify the substance type in the post, we first formulated a list of street names and slang words associated with the substance and labeled it with corresponding types, such as labeling post 1 for tobacco type if it contained any terms related to tobacco substance. The samples of posts, along with the identified substance type, are presented in Table S6 in Multimedia Appendix 1. Following identification, we aggregated the posts according to type and visualized the time series and histogram plot to identify and compare the growing trends in each substance type.

Theme Trend Analysis via LDA Topic Modeling

Theme trend analysis is a methodological approach that combines elements of theme analysis and trend analysis to

XSL•FO
RenderX

understand how specific themes or topics evolve within a dataset. In order to understand key topics of discussion, we used LDA topic modeling [71], a powerful unsupervised machine learning technique, to discover abstract topics within a collection of documents. We used this to answer question 2 (Textbox 1) specifically, where we generated the top 10 topics with the top 10 keywords and categorized the topics based on the identified baseline themes.

k-Means Clustering Analysis

k-means clustering is an unsupervised machine learning algorithm used to partition a dataset into k distinct, nonoverlapping clusters based on the similarity of data points by minimizing the variance within each cluster and maximizing the variance between different clusters [72]. The algorithm iteratively assigns data points to one of the k clusters based on the closest mean (centroid) of the cluster until the positions of the centroids stabilize. In our case, we used the scikit-learn library to perform the k-means clustering, where we used the term frequency-inverse document frequency scheme to create vectorization and considered the elbow method to identify the value of k for performing the clustering.

Thematic Analysis

Thematic analysis is used in qualitative research to analyze and interpret theme patterns within qualitative data. In our study, we used heat map analysis and factor analysis [73-75] to visually explore the relationship between identified themes and types of substances and to identify latent factors (patterns) from the observed themes, respectively.

Integration in Real-Time Application

We also integrated the trained model into a real-time application to monitor SU using the Elastic Logstash Kibana stack. We set up a search engine framework—using search database [76] and logstash [77]. Elasticsearch is an open-source, distributed, RESTful, JSON-based search engine originally based on Lucene (Solr) search that stores the document or the JSON object in an inverted index structure and allows the fastest full-text search. Logstash is a server-side data processing pipeline that usually sources or sinks data to and from multiple sources. In our work, we leveraged this pipeline to ingest the document and tweets from MongoDB [78], transformed the document by adding a custom call to generate a prediction result from the trained model, and finally wrote the document in Elasticsearch. The final document was a JSON comprising a tweet body with an additional prediction field from the trained model. Meanwhile, the Elastic Logstash Kibana stack had a built-in visualization tool to generate different trending charts based on real-time data during the development phase. We developed a full-fledged application in AngularJS and ReactJS frameworks for the client's real-time purposes. The snapshots demonstrating the chart showing the temporal and spatial analysis based on different filters are presented in the *Results* section.

Comparison With GPT-3

The advent of large language models, particularly GPT-3 [79], seemed to have raised questions regarding the efficiency and validity of custom models such as ours. Thus, we compared the reliability of our RoBERTa model and GPT-3 model in identifying SU posts. For this, we randomly sampled 3150 positive predictions from our customized model and queried GPT-3. Essentially, we designed a GPT-3 prompt, "Is this tweet <a real tweet post> related to substance use: Yes or No?" and queried for all sampled posts. Then, the predictions made by our model and GPT-3 were evaluated by our human experts.

Ethical Considerations

To ensure the privacy and confidentiality of individuals whose data were analyzed, all study data underwent a rigorous deidentification process before analysis. The data for this study were sourced from publicly available platforms [63] containing no personally identifiable information. In addition, all the sample posts were preprocessed, removing user IDs, emails, URLs, numbers, stop words, and lemmatizing, making the resulting tokens impossible to identify users' information. Thus, there was no personal information, including author names or any other private information, in the dataset. By addressing these ethical considerations, we conducted insightful research on SU patterns using social media content. In addition to this, our research was supported by the Substance Abuse and Mental Health Services Administration Strategic Prevention Framework-19 (grant 6H79SP081502), which was approved by the institutional review board at Kent State University (IRB20-182).

Results

Overview

Our primary objective was to comprehensively analyze the trends and patterns of SU over 3 years. To identify SU posts, we developed a self-trained deep learning model that achieved a precision rate of approximately 80%. This model was then used to detect SU tweets. The yearly breakdown of identified posts revealed 2,854,023 posts in 2019; 3,519,032 in 2020; and 2,567,970 in 2021. The identified data underwent various quantitative and qualitative analyses, such as trend analysis, k-means cluster analysis, topic and theme analysis, and factor analysis. To enhance the robustness of our findings, we validated our results by comparing them with those obtained from a GPT-3 model [79]. In the final section of our results, we present the outcomes of our integrated real-time application, showcasing the practical implications of our analyses.

Trend Analysis (Question 1: How Did the Discourse on SU Evolve on Twitter From 2019 to 2021, and What Variations Existed in the Distribution of Different Substances During This Time?)

We began our research by conducting a time series analysis to understand the SU trend in the following 3 periods: the prepandemic period, the pandemic period, and the postpandemic period. We aggregated identified SU posts monthly to plot in the chart. Figure 5 shows the average number of SU posts for the entire study period. The proportional representation of the same chart can be found in Figure S1 in Multimedia Appendix 1. While the trend in average number seems substantially high in 2020 in comparison to pre- and postpandemic periods (Figure 5), the proportion of posts for the same data is observed to decline from 2019 to 2021.

XSL•FO
RenderX

In addition, we also plotted a time series for 10 different substance categories to learn the trend of substances by categories. Thus, at first, we categorized each of the posts by applying a keyword-based method by referring to standard keywords from the NIDA [65], as outlined in Table S1 in Multimedia Appendix 1. For instance, we marked the post as alcohol if it contained any keywords associated with it and so forth. After the classification, we plotted the distribution for each substance type to visually understand the trend of each substance type in the study period, as shown in Figure 6.

Figure 5. Substance use distribution from 2019 to 2021.

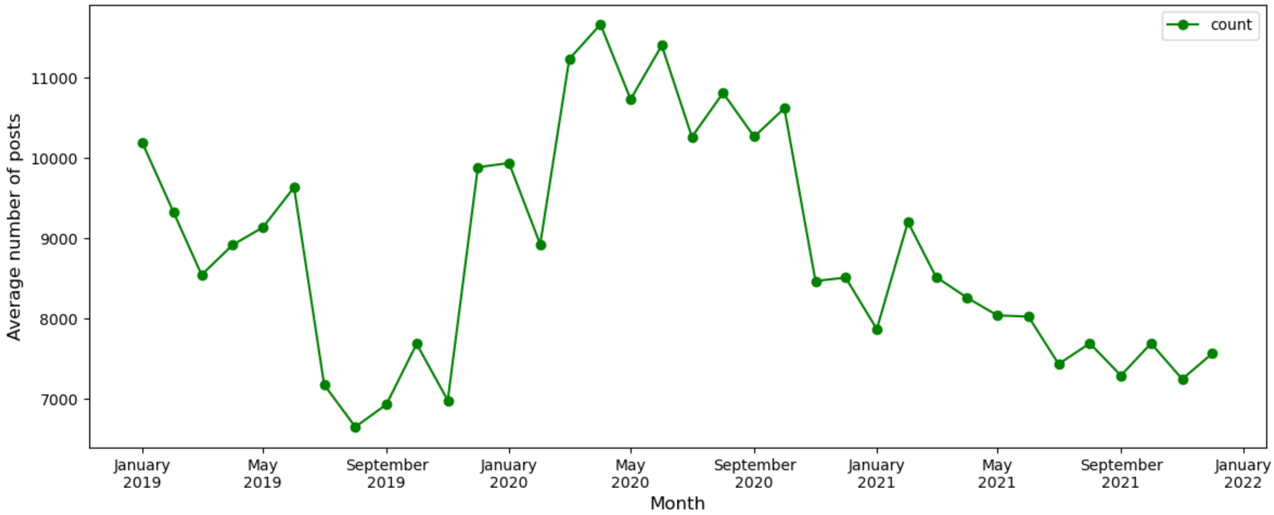
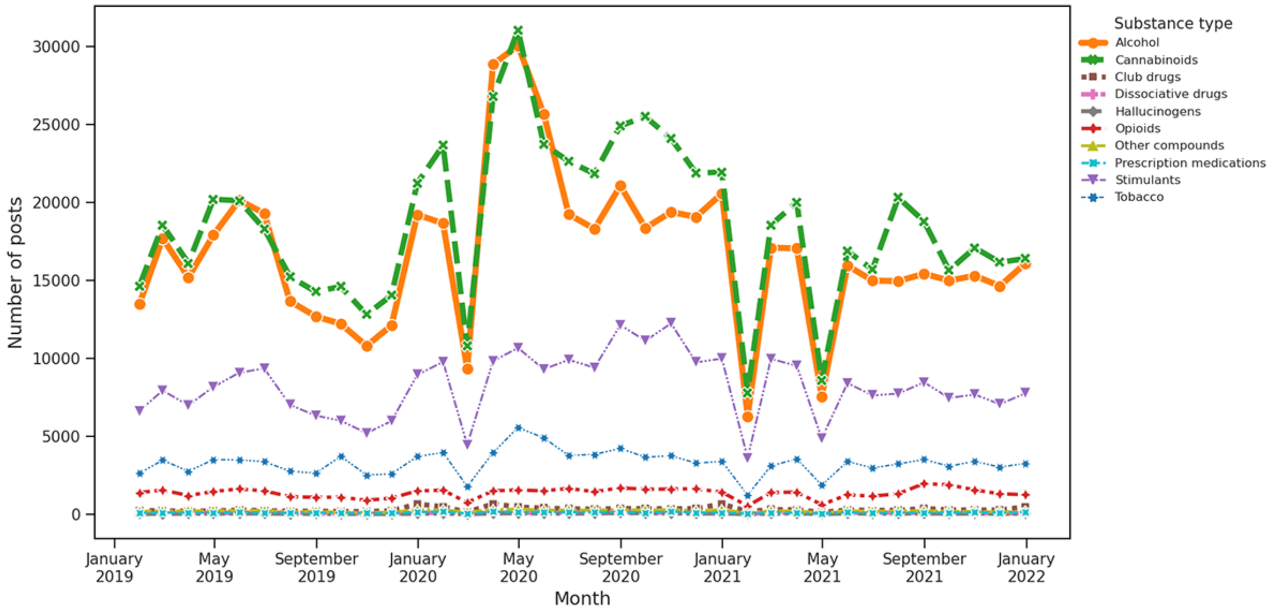


Figure 6. Substance type distribution from 2019 to 2021.



LDA Topic Analysis (Question 2: Following the Announcement of the Pandemic, What Were the Primary Substance Types That Garnered Significant Discussion, and What Were the Themes of These Dialogues?)

Drug Distribution 7 Days Before and After the Pandemic Declaration Day

Following the announcement of the pandemic on March 15, 2020, by Donald Trump, our result evidenced a significant 21% surge in the mentions of SU in tweets in just 3 days. Thus, understanding the change in pattern during that period was essential. We selected data from 7 days before and after March 15 to learn the impact of the pandemic declaration date on the trends. Thus, we aggregated the post count by each substance type 7 days before and after March 15, as shown in Table 2. The time series plot for the same period is also provided in Figure S3 in Multimedia Appendix 1.

XSL•FO
RenderX

Table 2. Substance type distribution 7 days before and after the pandemic declaration day.

Period and substance type	Proportion of posts, n (%)
Seven days before March 15, 2020 (n=54,671)	
Tobacco	2165 (3.96)
Alcohol	15,620 (28.57)
Cannabinoids	23,837 (43.6)
Opioids	1345 (2.46)
Stimulants	10,241 (18.74)
Club drugs	1000 (1.83)
Dissociative drugs	98 (0.18)
Hallucinogens	87 (0.16)
Other compounds	470 (0.86)
Prescription medications	98 (0.18)
Seven days after March 15, 2020 (n=56,773)	
Tobacco	1936 (3.41)
Alcohol	19,661 (34.63)
Cannabinoids	21,341 (37.59)
Opioids	835 (1.47)
Stimulants	7914 (13.94)
Club drugs	613 (1.08)
Dissociative drugs	131 (0.23)
Hallucinogens	57 (0.1)
Other compounds	199 (0.35)
Prescription medications	2606 (4.59)

LDA Topic Analysis

Furthermore, to comprehend the nuances of keywords and topics discussed following the declaration of the pandemic, we conducted an LDA topic analysis on these periods, 7 days before and after the official declaration. As shown in Tables 3 and 4, we highlighted the 10 main topics along with the distribution of the posts across each topic. Also, each topic consisted of the topmost terms that were extracted, excluding stop words.

Table 3. Top 10 terms of 10 latent Dirichlet allocation topics (7 days before the pandemic declaration day).

Topic	Top 10 terms	Distribution (n=54,671), n (%)
0	wine, buy, smoking, glass, everyone, water, red, drink, taste, beer	2733 (5)
1	alcohol, virus, corona, people, cigarette, leave, amp, covid, roll, hit	2733 (5)
2	beer, know, thing, fuck, try, man, cancel, cold, drink, problem	2733 (5)
3	drunk, bar, get drunk, blunt, pain, hold, kill, tonight, sick, coronavirus	2733 (5)
4	liquor, high, store, week, without, bitch, right, keep, always, low	2733 (5)
5	use, coke, would, call, drink, put, lmao, shot, really, enjoy, alcoholic	30,069 (55)
6	smoke, drink, drinking, sleep, drug, coffee, bro, work, fire, outside	2733 (5)
7	crack, night, love, last, stay, damn, smoke, cocaine, end, next	2733 (5)
8	weed, good, need, come, tequila, shit, smoke, drink, day, first	2733 (5)
9	nose, alcohol, bottle, year, please, beer, well, drink, hope, time	2733 (5)

XSL•FO
RenderX

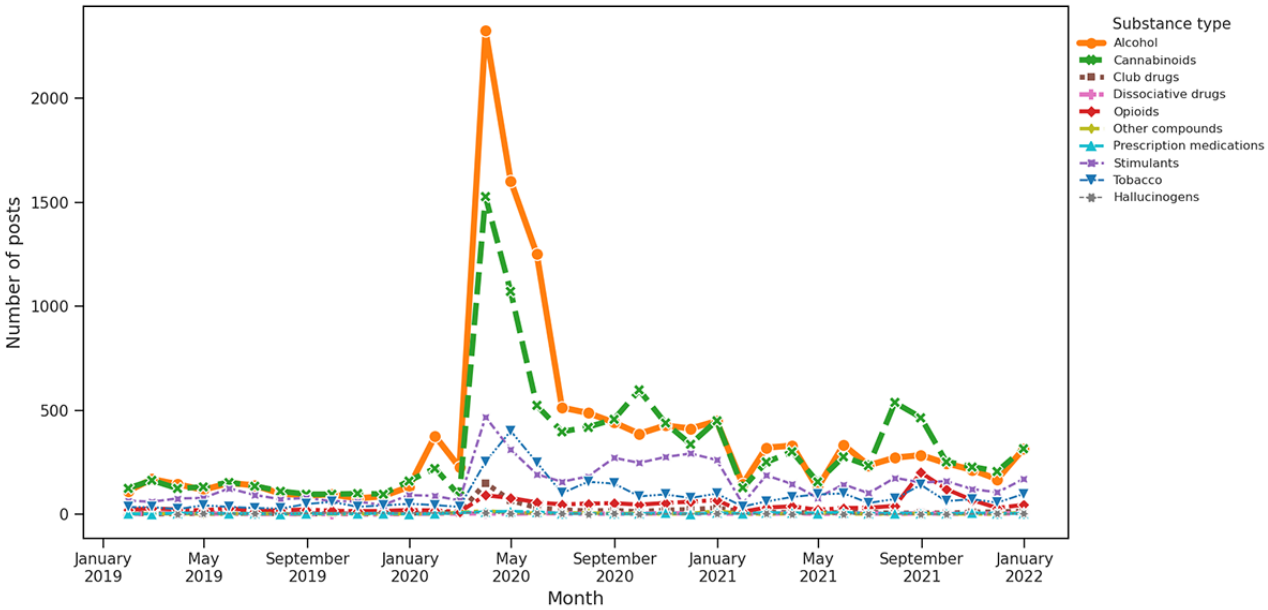
Table 4. Top 10 terms of 10 latent Dirichlet allocation topics (7 days after the pandemic declaration day).

Topic	Top 10 terms	Distribution (n=56,773), n (%)
0	come, smoking, back, man, alcoholic, street, eye, way, chinese, sell	1947 (3.43)
1	drink, drinking, beer, bottle, sleep, tequila, good, drive, tonight, nose	1947 (3.43)
2	liquor, stop, alcohol, store, order, close, bar, help, essential, turn	1947 (3.43)
3	drink, cigarette, smoke, eat, tell, weed, food, hold, even, talk	1947 (3.43)
4	fuck, virus, coke, year, high, shot, corona, covid, people, kill	1947 (3.43)
5	high, last, blunt, night, please, lit, amp, loudlycryingface, thought, die	1947 (3.43)
6	smoke, shit, start, open, feel, find, lmfao, woozyface, asf, miss, fire	1947 (3.43)
7	quarantine, pain, crack, really, use, cocaine, damn, bitch, drug, liquid	39,196 (69.04)
8	wine, would, weed, see, glass, need, someone, lmao, drunk	1947 (3.43)
9	drunk, get, love, friend, house, home, free, eat, wine, stay	1947 (3.43)

Theme Trend Analysis (Question 3: How Did the Prevalence of the Studied Theme Influence Various Types of Substances Used During the Studied Period?)

The theme in any subjective study is either a topic or a related subject name that best describes the group of the data. In our context, we wanted to identify such themes in the SU posts so that we could analyze the pattern and further investigate a correlation with different substances. Thus, we derived 6 major themes, namely COVID-19, economic, social, mental health, supply disruption, and medical disruption, as discussed in the Baseline Themes section in the Methods section. Subsequently, we plotted a time series for each substance type for all themes, as depicted in Figures 7-12. Our research yielded valuable insights through a trend analysis focused on the impact of prevalent themes on SU.

Figure 7. Substance distribution based on keywords associated with COVID-19.



XSL•FO
RenderX

Figure 8. Substance distribution based on keywords associated with economic stress.

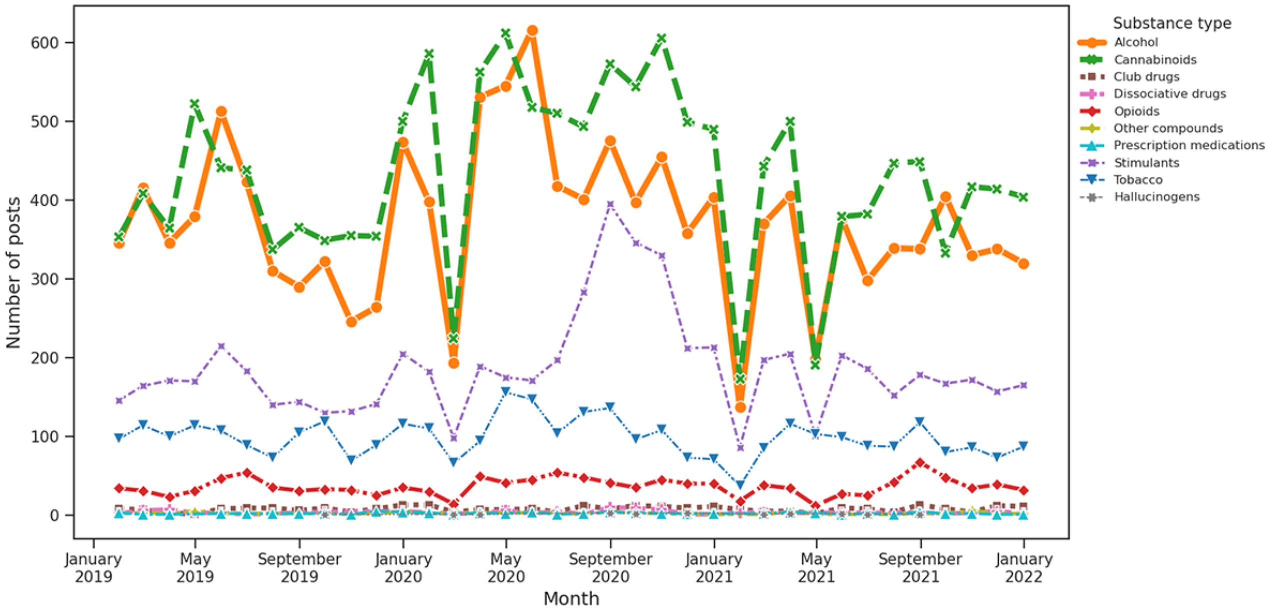
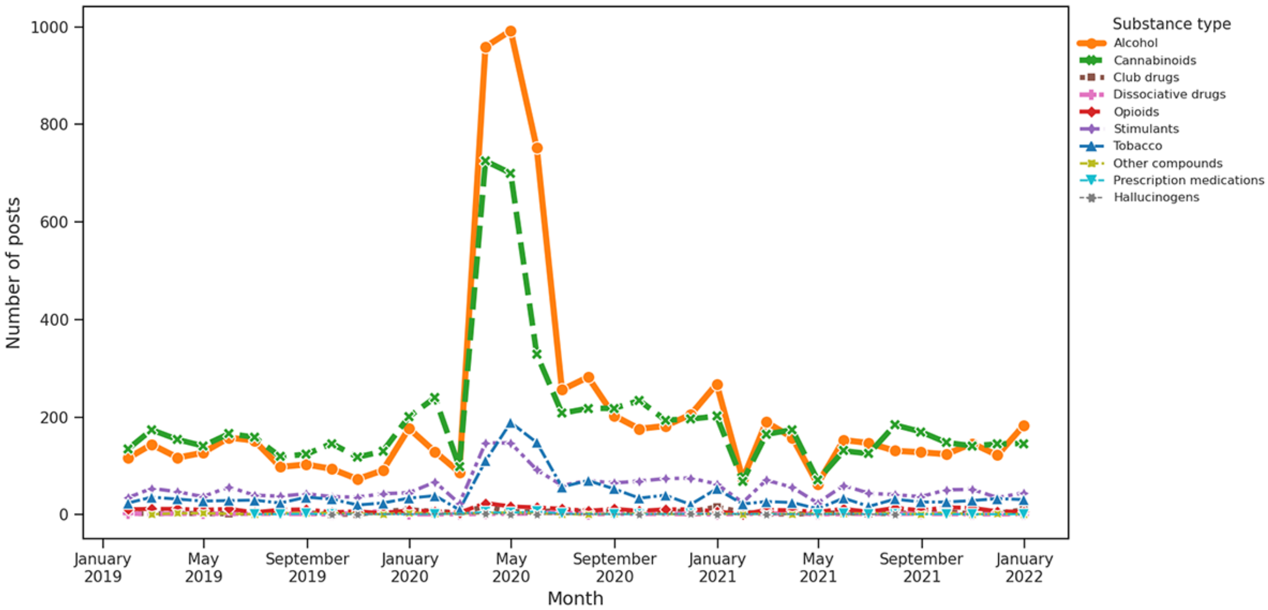


Figure 9. Substance distribution based on keywords associated with social stress.



XSL•FO
RenderX

Figure 10. Substance distribution based on keywords associated with mental health.

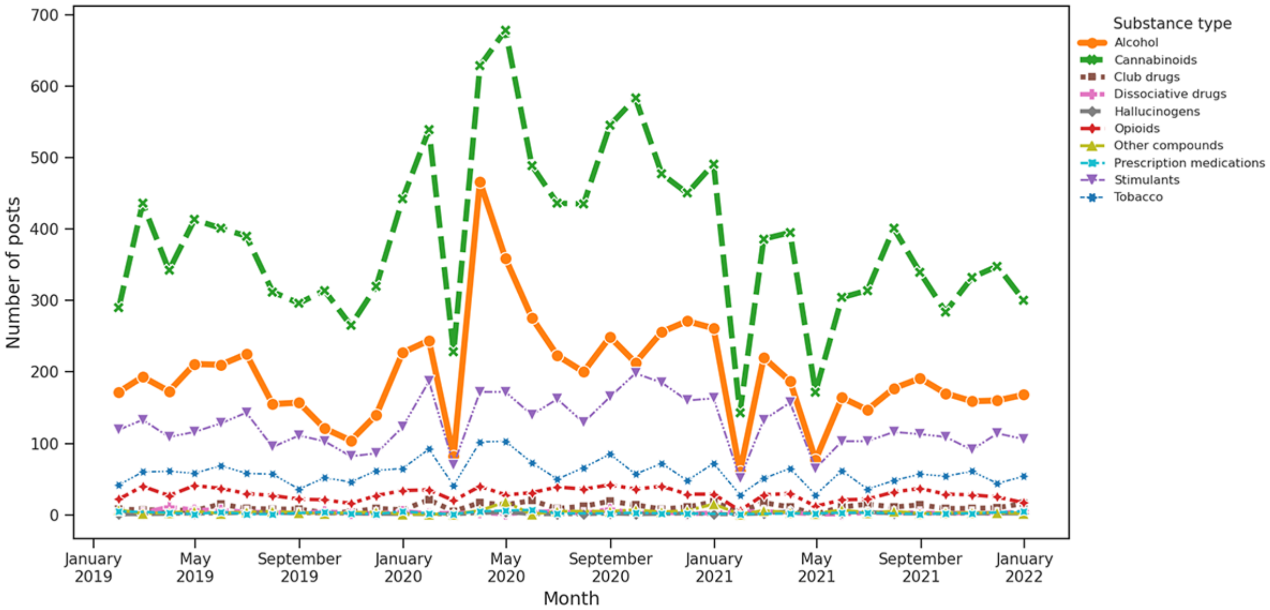
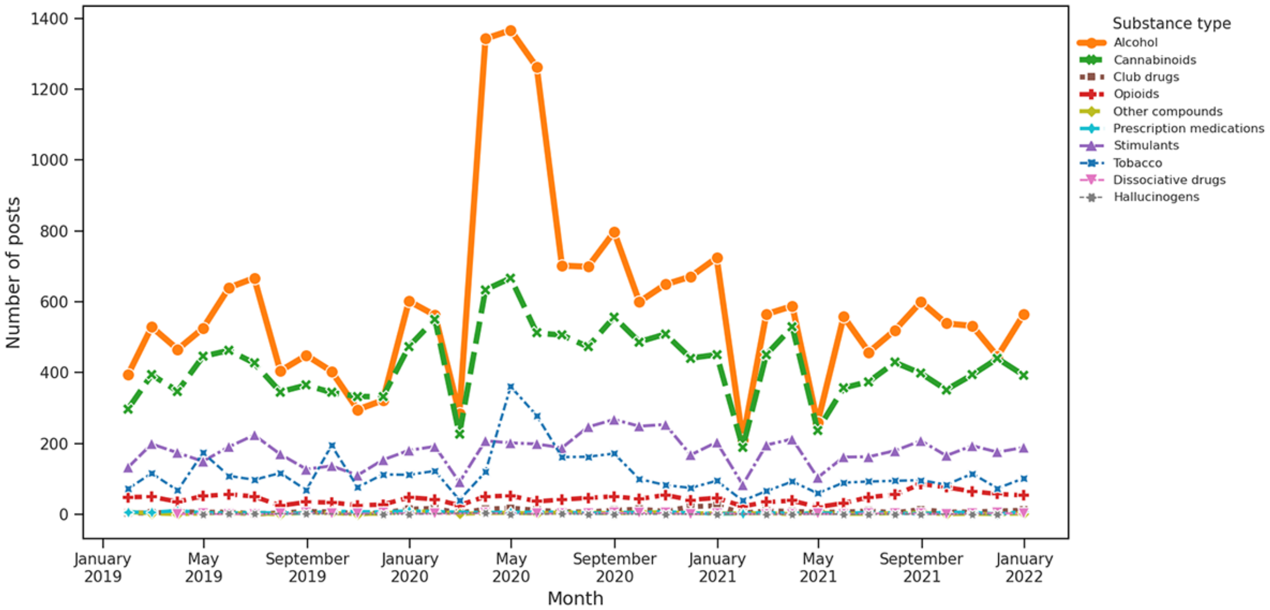
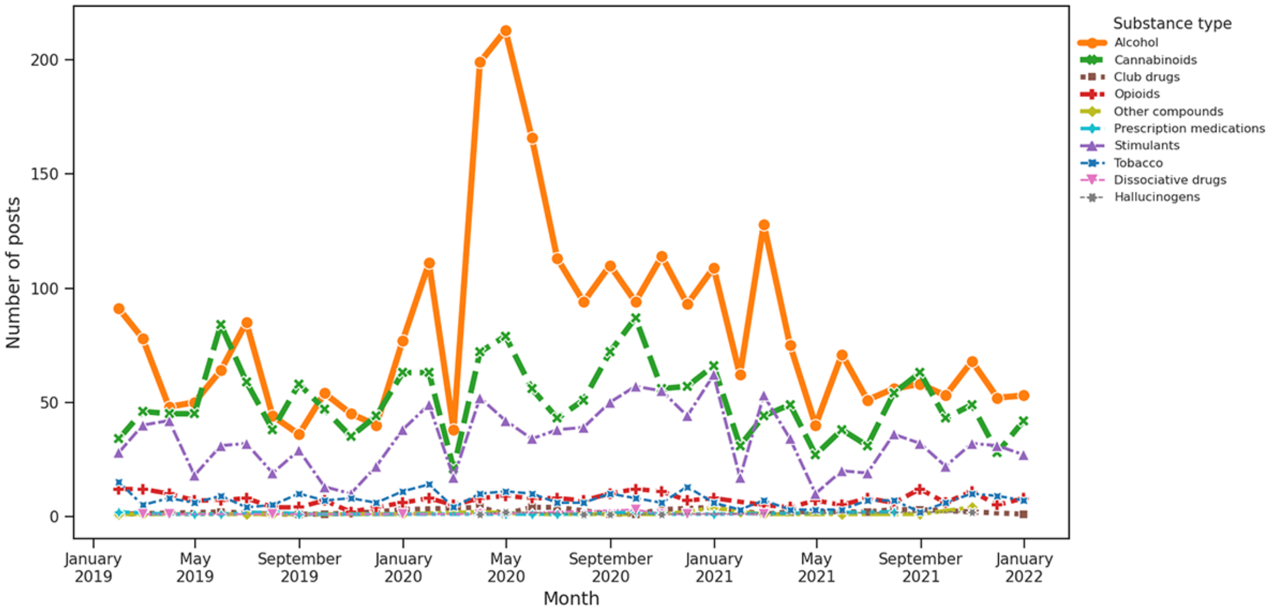


Figure 11. Substance distribution based on keywords associated with supply disruption.



XSL•FO
RenderX

Figure 12. Substance distribution based on keywords associated with medical disruption.



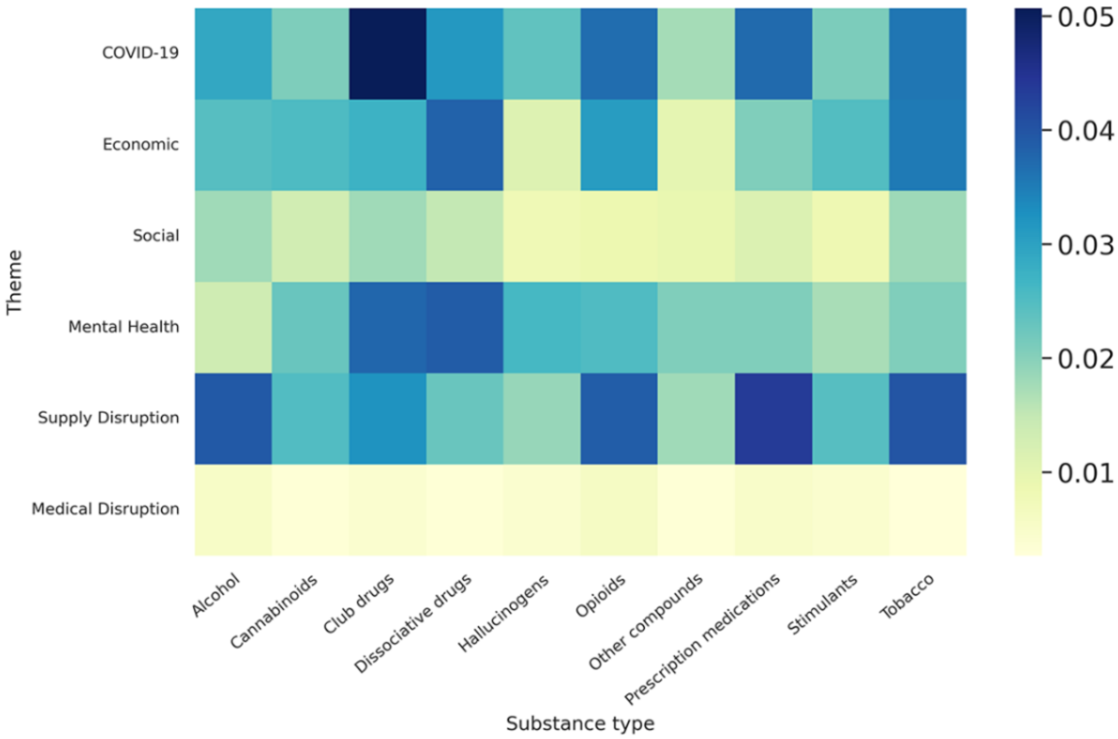
Thematic Analysis (Question 4: How Did the Identified Themes Correlate With the Substance Types?)

We performed heat map analysis and factor analysis to explore the correlation between identified themes.

Heat Map Analysis

In our study, we further used a heat map to visually analyze the relationships between identified themes (COVID-19, economic, social, mental health, supply disruption, and medical disruption) and substance types (alcohol, cannabinoids, club drugs, dissociative drugs, hallucinogens, opioids, other compounds, prescription medications, stimulants, and tobacco). The correlation plot is shown in Figure 13, where themes are represented on the y-axis and substances are represented on the x-axis. Each cell within the grid corresponds to a unique pairing of theme and substance type, with the color intensity indicating the strength of the association between them, and the color scale positioned along the right side of the vertical axis represents the intensity of association between these variables. Here, deeper shades of blue signify stronger associations, while lighter shades, reminiscent of lime, indicate weaker associations.

Figure 13. Heat map between themes and substance types.

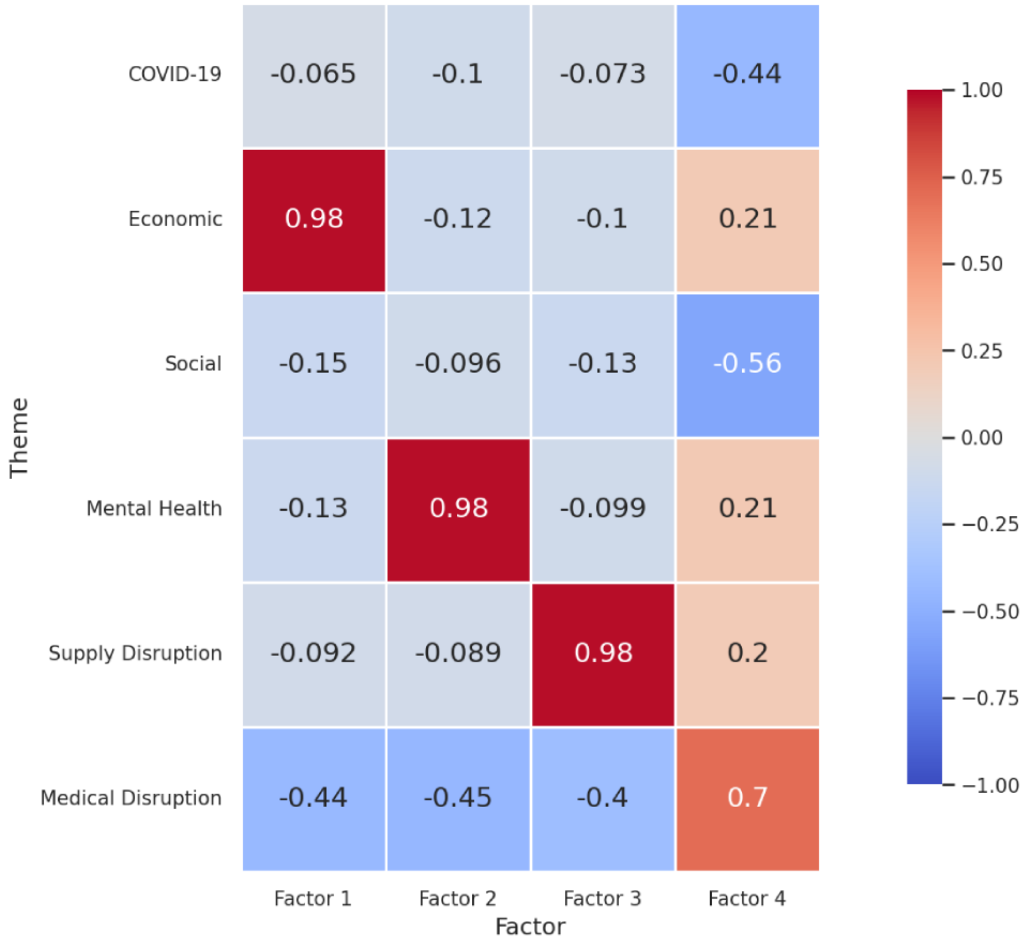


XSL•FO
RenderX

Factor Analysis

We performed factor analysis to examine the variability among the selected themes, aiming to distill these into a smaller set of unobserved, underlying variables known as factors. We determined the optimal number of factors to be 4 based on the Kaiser criterion, a decision further substantiated by the scree plot analysis, which revealed a distinct elbow point (Figure S5 in Multimedia Appendix 1). This analysis was facilitated by the *factor_analyzer* package within the Python application programming interface [45], which calculated the eigenvalues for each factor corresponding to the identified themes. The resultant factor loading heat map is shown in Figure 14. This heat map illustrates the relationships between factors and themes; negative values signify an inverse relationship, while positive values denote a direct relationship. The intensity of the relationship is indicated by values approaching 1 or –1 for strong relationships and values near 0 for weak ones. The heat map uses a color gradient where red shades indicate positive associations and blue shades indicate negative associations, providing a clear visual representation of these relationships.

Figure 14. Factor loading heat map.

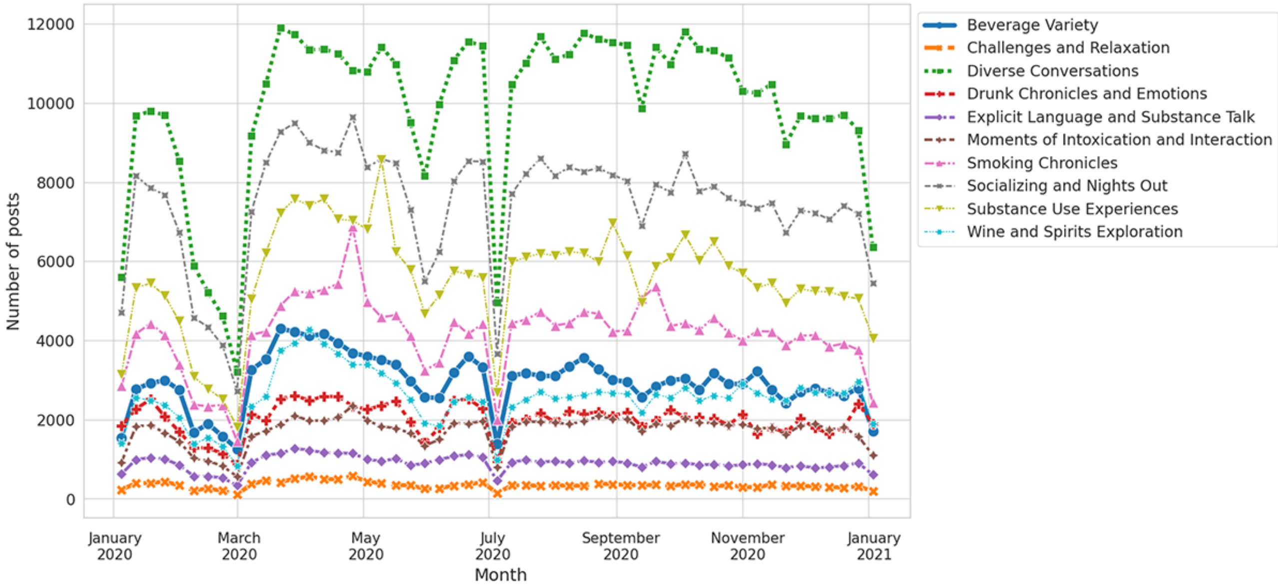


k-Means Clustering Analysis (Question 5: What Primary Discussion Topics Arise From k-Means Analysis, Specifically During the Pandemic Year?)

In addition, we also performed k-means clustering on SU posts from 2020 to identify the relevant groups or clusters by leveraging the similar distance algorithm inbuilt in the k-means clustering method [72]. Essentially, we started by applying the elbow method to determine an optimal cluster size, which turned out to be 19 for our data. The elbow diagram is depicted in Figure S4 in Multimedia Appendix 1. Furthermore, we applied k-means clustering to generate 19 clusters. However, due to redundant cluster keywords, we merged the relevant clusters, resulting in 10 main clusters, as shown in Figure 15. The cluster keywords and their respective details are presented in Table S8 and Figure S4 in Multimedia Appendix 1, respectively. Initially, information from each cluster was gathered and categorized into interaction, discussion, feelings, and perceptions. In addition, correlated clusters were amalgamated, such as cluster 0 and 13 labeled as raw conversations—explicit language and substance talk; cluster 1 and 6 as smoking chronicles—weed and cigarettes; cluster 2 and 12 as wine and spirits exploration; cluster 4 and 9 as social highs—moments of intoxication and interaction; cluster 5 and 16 as socializing and nights out—drinks and smokes; cluster 7 and 8 as beverage variety—beers, alcohol, and voting; cluster 10 and 14 as SU experiences—smoking, drugs, and liquor; and cluster 15 and 18 as challenges and relaxation—tiredness, and blunts. Cluster 3, encompassing diverse experiences and activities, was named diverse conversation.

XSL•FO
RenderX

Figure 15. K-means cluster analysis of substance use posts in 2020.



Comparison With GPT-3 (Question 6: To What Degree Does the Classifier's Effectiveness in Pinpointing SU-Related Tweets During the Pandemic Align With or Differ From GPT-3?)

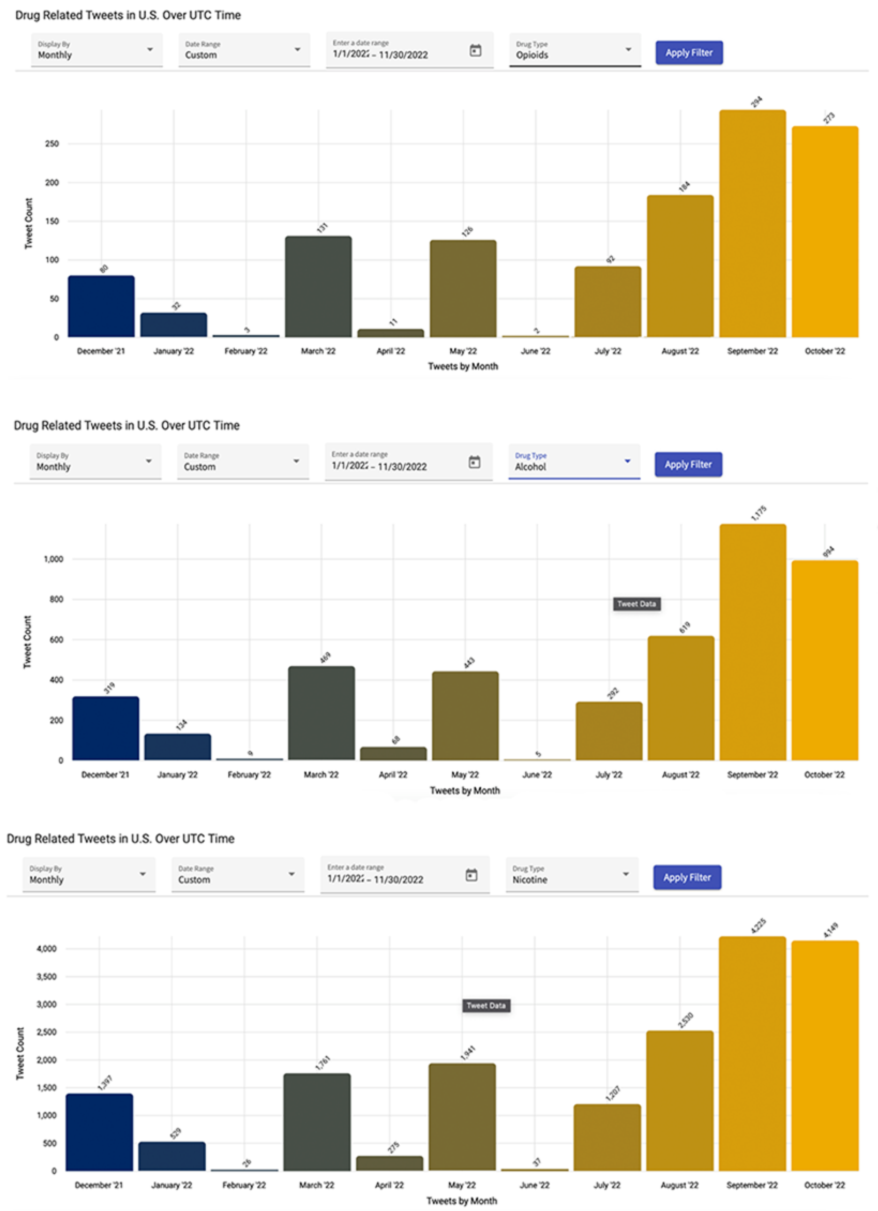
We also compared our results with the GPT-3 model by asking the model to classify tweets using GPT application programming interface [79]. We used a prompt, "Is this tweet '<a real tweet post>' related to substance use: Yes or No?" For this, we randomly sampled 3150 predicted positive results from our customized model and cross verified with the human (experts) and a machine (GPT-3). The human-verified 95.23% (3000/3150) of these predicted positive tweets were accurate, while GPT-3 only verified 53.73% (1693/3150) of the tweets as accurate. From this result, we concluded that generic powerful models such as GPT-3 do not necessarily generate true results when identifying hidden contexts in domain-specific data. This necessitates the need for domain-specific models for accurate results.

Real-Time Application (Question 7: How Has the Overall System Contributed to the Real-Time Tracking of SU, as Evidenced by Research?)

We further deployed our model to provide a real-time service in an application, Northeast Ohio Tri-County Prevention Infrastructure [80], specifically within the social media section designed for Ohio state. Primarily, the aim of the application was to serve as a monitoring and prevention dashboard for the state from static data. However, the real-time nature of social media data gave the application true power to monitor patterns of SU across areas of interest. Figure 16 provides snapshots of the application, illustrating how the stakeholders can dynamically monitor the SU segmented by time and substance type.

XSL•FO
RenderX

Figure 16. Snapshots of integrated real-time application. UTC: Universal Time Coordinated.



Discussion

We used our custom deep learning model and several statistical methods to perform this analysis to get insights into the trends and impacts of COVID-19 related to SU. The subsequent sections elaborate on the results in detail.

Trend Analysis

Time to Event Analysis and Substance Distribution for Question 1

The analysis of time to event reveals a significant increase in SU tweets in 2020, surpassing the counts for 2019 and 2021 by 17.6% and 22.35%, respectively (Figure 5). Notably, March 2020, April 2020, and June 2020 emerged as the focal months for SU discussions, with frequencies 16.55%, 21.18%, and 18.19% higher than other months in 2020. The elevated trend persisted until October 2020, likely coinciding with the availability of vaccines, highlighting a limitation in our study.

The examination of substance discussions over a 3-year time frame revealed a consistent focus on alcohol and cannabinoids, emerging as the predominant topics throughout the study. An intriguing observation during the pandemic period was the discernible surge in discussions surrounding alcohol, cannabinoids, and stimulant drugs, distinguishing them with an upward trend. In contrast, other substances did not exhibit substantial shifts in discourse.

It is crucial to exercise caution when interpreting the data for February 2020, January 2021, and April 2021, as the graph may be skewed due to limited available tweets in the Twitter source during those specific months. Despite this limitation, the broader insights gleaned from the study underscore the enduring prominence of alcohol and cannabinoids in public discourse.

XSL•FO
RenderX

The pandemic period, marked by unprecedented global challenges, evidently influenced a notable increase in discussions surrounding these substances and stimulant drugs, indicative of evolving societal dynamics and coping mechanisms. These findings prompt further exploration into the nuanced factors shaping substance-related discussions, offering valuable insights for public health considerations and policy implications.

Topic Analysis and Substance Distribution for Question 2

In order to observe the impact of the declaration of the global COVID-19 pandemic declaration day on March 15, we analyzed the posts by each substance type 7 days before and after March 15. The aggregated posts in these 2 weeks had distinct changes in each substance type. Notably, discourse in only 2 substance types, alcohol and prescription medication, were observed significantly increasing, while discourse in all other substance types were observed slightly declining. The trend can also be visualized in Figure S3 in Multimedia Appendix 1. The increased trend of alcohol discussion was likely due to the effect of COVID-19, particularly due to closed schools, social isolation, boredom, and various types of mental stress and anxiety, which is also supported by some studies [15,35,36]. A study by Farhoudian et al [15] that conducted a survey in May highlighted the increment in alcohol, prescription medication, and cannabinoids. However, in our study, cannabinoids showed a slight decrement in a 7-day period while it remained significantly discussed during the entire study period.

Moreover, our topic analysis for the same period indicated a shift in substance-related discourse. In general, during the prepandemic period, references to substances were casual in almost all the 10 topics, as depicted in Table 3. Although topic 5 had the highest proportion of keywords (30,069/54,671, 55%), the terms referred to casual keywords, insignificant to any particular substances or behavior. However, topics in the postpandemic period included keywords that concerned quarantine and SU as seen in topics 4 (1947/56,773, 3.43%) and 7 (39,196/56,773, 69.04%) in Table 4. The mention of keywords (such as *nose*, *coronavirus*, and *covid*) during the first period suggested that COVID-19 has been interlinked with few substance discussion; however, there were no negative words indicating stress or bad impact on mental health. By contrast, the topics in the second period included negative keywords (such as *pain*, *die*, *stress*, and *fuck*) along with SU keywords. This shift suggests a nuanced decline in mental health after the pandemic declaration day. Likewise, topics 1, 2, 8, and 9 in the second week contained more alcohol- and liquor-related keywords (such as *drink*, *beer*, *bottle*, *liquor*, *store*, and *drunk*), suggesting use of alcohol as the main substance during this period. Nevertheless, there were no major terms in the topic analysis that could support prescription medication use in the second week.

In conclusion, our detailed analysis on 7 days before and after the pandemic declaration day highlights the immediate impact on the use of substances, particularly alcohol and prescription medication.

Theme Trend Analysis and Substance Distribution for Question 3

As per our keyword-based theme analysis, COVID-19 had a notably significant impact on the discussion of SU. The early pandemic period showed a significant rise in alcohol and cannabinoids associated with 2 main themes as follows: COVID-19 and social isolation. This surge was most evident at the onset of the pandemic in early 2020, likely reflecting a response to the stress, uncertainty, and lifestyle changes imposed by the health crisis. The data indicated that these increases were particularly influenced by COVID-19–related factors, with social and economic aspects also playing a role. In contrast, factors related to supply and medical disruptions did not drastically affect use patterns. This concentrated spike in alcohol and cannabinoid use during challenging periods highlights the broader impact of the pandemic on SU behaviors.

k-Means Clustering Analysis for Question 5

From the k-means clustering, we identified 10 main clusters as an indication of what was discussed in the pandemic year as follows: beverage variety, challenges and relaxation, drunk chronicles and emotions, explicit language and substance talk, moments of intoxication and interaction, smoking chronicles, socializing and nights out, SU experiences, wine and spirits exploration, and diverse conversations. The diverse conversation cluster includes all the remaining tweets that do not belong to particular clusters. Hence, the number of posts in it has the highest counts. Excluding this cluster, SU-associated posts were mostly seen in socializing and nights out, followed by SU experiences and smoking chronicles as the 3 main top discussions.

Thematic Analysis for Question 4

Heat Map Analysis

The heat map analysis provided insightful revelations regarding the factors influencing SU discourse, highlighting COVID-19, economic stress, mental health concerns, and alterations in drug supply as the principal elements. Specifically, there is a stronger correlation between the "COVID-19" theme and cannabinoid use, possibly signifying an increase in this substance's consumption as a direct response to the pandemic's stressors. The "economic" theme shows a somewhat lower yet noticeable correlation with alcohol, which might reflect economic uncertainty's impact on alcohol consumption. The "social" theme has a less pronounced correlation across all substance types, implying that social factors had a milder influence on SU during this period. "mental health" has a moderate correlation with both cannabinoids and alcohol, highlighting these as coping mechanisms during mentally challenging times. "supply disruption" shows a varied correlation but is not significantly linked with any substance, suggesting that supply issues did not drastically alter consumption patterns. Finally, "medical disruption" seems to have the least correlation with SU, suggesting that medical service disruptions during the pandemic had minimal influence on the consumption of these substances. Overall, the heat map indicates that COVID-19–related factors had the most significant correlation with changes in SU, with

XSL•FO
RenderX

economic and mental health factors also being relevant but to a lesser extent.

Factor Analysis

The factor analysis gave insights into a combination of themes that had an impact on SU. Factor 1 indicated that mental health was the leading factor. Factor 2 was strongly and positively associated with the economic theme, suggesting that this factor could represent financial stress or economic consequences of the pandemic. The social theme had a moderate negative loading on factor 2, implying that social aspects may decrease in relevance as economic concerns increase or vice versa. Factor 3 showed a very strong negative loading with the medical disruption theme, indicating that this factor was significantly influenced by disruptions in medical services. This could represent the strain on health care systems and the impact of health care access on the population. Mental health and supply disruption themes had a strong positive loading on factor 4, implying that this factor may represent the psychological impact of the pandemic and its influence on drug supply chains.

In summary, the factor analysis suggested that economic and mental health themes were major dimensions of the pandemic's impact, with medical disruptions also playing a significant but negatively associated role.

Comparison With GPT-3 for Question 6

Our comparative analysis with the GPT-3 model yielded valuable insights into the effectiveness of powerful generic models in identifying hidden contexts in domain-specific data, particularly related to drug use in tweets. The experiment involved using a prompt to classify tweets as either related or unrelated to SU. The results demonstrated a substantial discrepancy in accuracy between human verification and GPT-3. When comparing the randomly sampled predicted positive tweets, human experts confirmed the accuracy of 95.23%, whereas GPT-3 verified only 53.73% of the tweets as accurate. This notable difference underscores the limitations of generic models such as GPT-3 in accurately discerning domain-specific nuances. Although we have not performed a detailed analysis to find out the reason behind this discrepancy, we anticipate the limitation of contextual awareness as a primary reason for this, as indicated in the studies by Ray [81] and Moradi et al [82]. For instance, Moradi et al [82] highlighted similar cases where GPT-3 underperformed in the biomedical corpora in comparison to domain-specific pretrained model BioBERT [82]. By contrast, generic pretrained models such as ours can provide rich contextual understanding as they are pretrained solely on social media data, making them powerful in understanding slang-like languages. Thus, the limitation in GPT-3 is well addressed by our custom model pretrained on domain-specific data.

Real-Time Integration for Question 7

The successful integration of our trained model into the practical application Tri County Prevention Infrastructure [80], particularly within the social media section tailored for Ohio, marks a significant achievement. This integration empowers real-time users by allowing them to visually explore the distribution of substance-use posts in both temporal and spatial dimensions. For instance, the users can explore and analyze the trend of any substance (eg, alcohol) in real time and take immediate actions to mitigate the use in the areas of interest. In addition, the applicability of our models' integration is promising during crisis periods such as the COVID-19 pandemic, when physical intervention is unfeasible.

Limitations

Our study has several limitations. Initially, data inconsistencies in certain months were due to incomplete datasets from the sources [63]. Moreover, our analysis was confined to English-language posts, potentially excluding non-English speaking users and thus not reflecting the full spectrum of users during the study period. The initial annotated data used for the training model were collected from a specific time frame (January 2020 through April 2020). The selection of data from this particular time frame could have introduced some bias in the SU identification process. In addition, the consideration of precision as our primary evaluation metric during iterative fine-tuning steps could have missed real SU posts, limiting to the small spectrum of patterns learned by the model and leading to overfitting. Also, the overall accuracy of the classifier reached 80%, which could have led to non-SU posts being identified as SU posts and vice versa. Consequently, this could actually deviate the count of SU posts identified in our study, thus deviating from trend studies. Although the choice of classifier, RoBERTa, seems to have performed better, the identification of SU tweet posts for multiple sequences could have been misclassified. Likewise, the limited labeled dataset during fine-tuning could have underfitted the performance in the initial rounds. While we used HITL [61] in our iterative fine-tuning approach to enrich the annotated data, the human reviewers involved in the process were only tasked with reviewing model predictions without providing feedback. This lack of active human feedback may have limited the model's capacity for improvement as corrections to errors and mispredictions or rewarding accurate predictions could have enhanced its performance further. In the future, incorporating a full HITL at different stages of model development could significantly improve accuracy and model refinement. Finally, the scope of keywords used in processing tweet data may have been too narrow, possibly leading to an overrepresentation of certain themes and factors in our results.

Future Work

This study only considered text data for the identification of SU. Future research could use multimedia, such as images and videos, to enhance the accuracy of the identification of SU. Furthermore, our iterative fine-tuning approach could be enhanced through active learning [62], where the most critical samples are selected for annotation in each iteration, optimizing model performance. Another potential improvement involves incorporating full HITL feedback [61], allowing human reviewers not only to review but also to correct errors or reward accurate predictions. This approach could significantly refine model accuracy. In addition, a user-level analysis could be conducted to investigate factors influencing the intention and purpose behind substance misuse. In addition to this, demographic factors such as age, gender, race, emotion, socioeconomic status, personality trait, and mental and physical

XSL•FO
RenderX

health status could be considered for investigation to understand the most impacted cohort during the pandemic. By understanding these cohorts and factors, we can develop strategies and interventions to prevent and control SU during global crises.

Conclusions

In this study, we conducted an extensive infodemiology analysis of Twitter posts from 2019 to 2021, focusing on SU patterns during the COVID-19 pandemic. Using a deep learning model (RoBERTa) alongside techniques with human involvement in iterative fine-tuning, our classifier achieved an optimal accuracy of 80%, even with limited resources. This performance is notable as even a powerful state-of-the-art model such as GPT-3 struggled with domain-specific data such as SU.

In summary, the results from our study showed the key patterns in SU trends during both the pandemic and overall study periods. The analysis of the pandemic period has shown that COVID-19 had a huge impact on the influx of SU. As indicated by trend analysis, the numbers were higher during the peak pandemic period, mainly between March and October 2020. Furthermore, the theme analysis showed a higher association of SU posts with COVID-19 and social themes in comparison to other themes during the pandemic period. In addition to this, the immediate declaration of the pandemic introduced stress and anxiety in public, as evidenced by our LDA topic analysis, causing a significant rise in SU (21% in just 3 days), primarily in readily available substances such as alcohol and prescription medication. These findings suggest that the authorities should pay attention to key factors such as social isolation, stress, and anxiety, and focus on strengthening regulations around the sale of accessible substances such as alcohol, prescription medications, and cannabinoids (though not legal in all areas) to have control the SU during the global COVID-19 pandemic crises. By contrast, economic, mental health, and supply disruptions seem to be the major contributing factors for SU throughout the study period, as indicated by our factor analysis, with cannabinoids, alcohol, and stimulants as dominating substances. Thus, public health agencies should focus on controlling the economic and mental health of global citizens as key actions, alongside surveilling drug supplies, in order to control global SU.

In summary, our study demonstrates the applicability of social media data used along with a deep learning model to analyze trends in global issues such as SU. The findings and methodology from this study can help public health sectors develop real-time strategies and prevent SU during future crises.

Acknowledgments

This work was supported by the Substance Abuse and Mental Health Services Administration Strategic Prevention Framework-19 (Grant No. 6H79SP081502). We gratefully acknowledge Megan Anderson, Sheryl Chatfield, Kaylie Kenne, Kayla Marker, Olivia Anderson, and Cassidy Shokles (Center for Public Policy & Health, College of Public Health, Kent State University) and Anthony Coetzer-Liversage (Social Sciences Research Center, University of Rhode Island) for their assistance with data annotation. We also sincerely thank the editors and reviewers for their valuable feedback, which greatly improved the manuscript.

Data Availability

The data and code supporting this study are publicly available on GitHub [83].

Authors' Contributions

JM contributed to the conceptualization, methodology, investigation, formal analysis, model development, and writing of the original draft. RJ assisted in conceptualization and supervised the study. JZ provided feedback on the analyses, while JK contributed to data curation and validation. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional dataset details, classified tweet samples, and analysis figures to support this paper's investigation.

[DOCX File , 3734 KB - infodemiology_v5i1e59076_app1.docx]

References

1. Drug abuse statistics. National Center of Drug Abuse Statistics. URL: https://drugabusestatistics.org/ [accessed 2023-10-20]
2. Drug overdose deaths in the United States, 2003-2023. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/nchs/products/databriefs/db522.htm [accessed 2025-03-31]
3. Addiction and substance misuse reports and publications. US Department of Health and Human Services. URL: https://www.hhs.gov/surgeongeneral/reports-and-publications/addiction-and-substance-misuse/index.html [accessed 2023-08-10]
4. Covid-19 and substance use. National Institute on Drug Abuse. URL: https://nida.nih.gov/research-topics/covid-19-substance-use [accessed 2023-08-14]

XSL•FO
RenderX

5. Ornell F, Moura HF, Scherer JN, Pechansky F, Kessler FH, von Diemen L. The COVID-19 pandemic and its impact on substance use: implications for prevention and treatment. Psychiatry Res 2020 Jul;289:113096 [FREE Full text] [doi: 10.1016/j.psychres.2020.113096] [Medline: 32405115]
6. Volkow ND. Collision of the COVID-19 and addiction epidemics. Ann Intern Med 2020 Apr 02;173:61-62 [FREE Full text] [doi: 10.7326/m20-1212]
7. Witteveen D, Velthorst E. Economic hardship and mental health complaints during COVID-19. Proc Natl Acad Sci U S A 2020 Nov 03;117(44):27277-27284 [FREE Full text] [doi: 10.1073/pnas.2009609117] [Medline: 33046648]
8. Godinic D, Obrenovic B, Khudaykulov A. Effects of economic uncertainty on mental health in the COVID-19 pandemic context: social identity disturbance, job uncertainty and psychological well-being model. Int J Innov Econ Dev 2020 Apr;6(1):61-74. [doi: 10.18775/ijied.1849-7551-7020.2015.61.2005]
9. Galea S, Merchant RM, Lurie N. The mental health consequences of COVID-19 and physical distancing: the need for prevention and early intervention. JAMA Intern Med 2020 Jun 01;180(6):817-818. [doi: 10.1001/jamainternmed.2020.1562] [Medline: 32275292]
10. Sheridan Rains L, Johnson S, Barnett P, Steare T, Needle JJ, Carr S, et al. Early impacts of the COVID-19 pandemic on mental health care and on people with mental health conditions: framework synthesis of international experiences and responses. Soc Psychiatry Psychiatr Epidemiol 2021 Jan;56(1):13-24 [FREE Full text] [doi: 10.1007/s00127-020-01924-7] [Medline: 32804258]
11. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
12. Alter A, Yeager C. COVID-19 impact on US national overdose crisis. Overdose Detection Mapping Application Program. URL: https://www.odmap.org/content/docs/news/2020/odmap-report-june-2020.pdf [accessed 2023-08-14]
13. COVID-19 and the drug supply chain: from production and trafficking to use. United Nations Office on Drugs and Crime. 2020. URL: https://www.unodc.org/documents/data-and-analysis/covid/Covid-19-and-drug-supply-chain-Mai2020.pdf [accessed 2025-03-27]
14. Bolinski RS, Walters S, Salisbury-Afshar E, Ouellet LJ, Jenkins WD, Almirol E, et al. The impact of the COVID-19 pandemic on drug use behaviors, fentanyl exposure, and harm reduction service support among people who use drugs in rural settings. Int J Environ Res Public Health 2022 Feb 16;19(4):2230 [FREE Full text] [doi: 10.3390/ijerph19042230] [Medline: 35206421]
15. Farhoudian A, Radfar SR, Mohaddes Ardabili H, Rafei P, Ebrahimi M, Khojasteh Zonoozi A, ISAM Global Survey Consortium (ISAM-GSC), et al. A global survey on changes in the supply, price, and use of illicit drugs and alcohol, and related complications during the 2020 COVID-19 pandemic. Front Psychiatry 2021 Aug 6;12:646206 [FREE Full text] [doi: 10.3389/fpsyt.2021.646206] [Medline: 34421664]
16. Beech M. COVID-19 pushes up internet use 70% and streaming more than 12%, first figures reveal. Forbes. 2020 Mar 25. URL: https://www.forbes.com/sites/markbeech/2020/03/25/covid-19-pushes-up-internet-use-70-streaming-more-than-12-first-figures-reveal/?sh=337c58783104 [accessed 2023-09-27]
17. Amram O, Borah P, Kubsad D, McPherson SM. Media exposure and substance use increase during COVID-19. Int J Environ Res Public Health 2021 Jun 11;18(12):6318 [FREE Full text] [doi: 10.3390/ijerph18126318] [Medline: 34207964]
18. Ahmad AR, Murad HR. The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. J Med Internet Res 2020 May 19;22(5):e19556 [FREE Full text] [doi: 10.2196/19556] [Medline: 32369026]
19. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020 Apr 16;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
20. Brailovskaia E, Swarlik VJ, Grethe GA, Schillack H, Margraf J. Experimental longitudinal evidence for causal role of social media use and physical activity in COVID-19 burden and mental health. Z Gesundh Wiss 2022 Sep 02:1-14 [FREE Full text] [doi: 10.1007/s10389-022-01751-x] [Medline: 36068852]
21. Cho H, Li P, Ngien A, Tan MG, Chen A, Nekmat E. The bright and dark sides of social media use during COVID-19 lockdown: contrasting social media effects through social liability vs. social support. Comput Human Behav 2023 Sep;146:107795 [FREE Full text] [doi: 10.1016/j.chb.2023.107795] [Medline: 37124630]
22. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Care Services, Forum on Mental Health and Substance Use Disorders, Graig L, Friedman K. Mental Health and Substance Use Disorders in the Era of COVID-19: The Impact of the Pandemic on Communities of Color: Proceedings of a Workshop. Washington, DC: National Academies Press; 2021.
23. Davis JP, Pedersen ER, Tucker JS, Dunbar MS, Seelam R, Shih R, et al. Long-term associations between substance use-related media exposure, descriptive norms, and alcohol use from adolescence to young adulthood. J Youth Adolesc 2019 Jul;48(7):1311-1326 [FREE Full text] [doi: 10.1007/s10964-019-01024-z] [Medline: 31025156]

XSL•FO
RenderX

24. Kelleghan AR, Leventhal AM, Cruz TB, Bello MS, Liu F, Unger JB, et al. Digital media use and subsequent cannabis and tobacco product use initiation among adolescents. Drug Alcohol Depend 2020 Jul 01;212:108017 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.108017] [Medline: 32408138]
25. Vogel EA, Ramo DE, Rubinstein ML, Delucchi KL, Darrow SM, Costello C, et al. Effects of social media on adolescents' willingness and intention to use e-cigarettes: an experimental investigation. Nicotine Tob Res 2021 Mar 19;23(4):694-701 [FREE Full text] [doi: 10.1093/ntr/ntaa003] [Medline: 31912147]
26. Shen L, Assanangkornchai S, Liu W, Cai L, Li F, Tang S, et al. Influence of social network on drug use among clients of methadone maintenance treatment centers in Kunming, China. PLoS One 2018 Jul 3;13(7):e0200105 [FREE Full text] [doi: 10.1371/journal.pone.0200105] [Medline: 29969481]
27. Cook SH, Bauermeister JA, Gordon-Messer D, Zimmerman MA. Online network influences on emerging adults' alcohol and drug use. J Youth Adolesc 2013 Nov;42(11):1674-1686 [FREE Full text] [doi: 10.1007/s10964-012-9869-1] [Medline: 23212348]
28. Dubey MJ, Ghosh R, Chatterjee S, Biswas P, Chatterjee S, Dubey S. COVID-19 and addiction. Diabetes Metab Syndr 2020;14(5):817-823 [FREE Full text] [doi: 10.1016/j.dsx.2020.06.008] [Medline: 32540735]
29. Brown P, Watts V, Hanna M, Rizk M, Tucker E, Saddlemire A, et al. Two epidemics and a pandemic: the collision of prescription drug misuse and racism during COVID-19. J Psychoactive Drugs 2021;53(5):413-421. [doi: 10.1080/02791072.2021.1992048] [Medline: 34694200]
30. Chacon NC, Walia N, Allen A, Sciancalepore A, Tiong J, Quick R, et al. Substance use during COVID-19 pandemic: impact on the underserved communities. Discoveries (Craiova) 2021 Dec 31;9(4):e141 [FREE Full text] [doi: 10.15190/d.2021.20] [Medline: 35261922]
31. Rogers AH, Shepherd JM, Garey L, Zvolensky MJ. Psychological factors associated with substance use initiation during the COVID-19 pandemic. Psychiatry Res 2020 Nov;293:113407 [FREE Full text] [doi: 10.1016/j.psychres.2020.113407] [Medline: 32827993]
32. Avena NM, Simkus J, Lewandowski A, Gold MS, Potenza MN. Substance use disorders and behavioral addictions during the COVID-19 pandemic and COVID-19-related restrictions. Front Psychiatry 2021 Apr 16;12:653674 [FREE Full text] [doi: 10.3389/fpsyt.2021.653674] [Medline: 33935838]
33. Ignaszewski MJ. The epidemiology of drug abuse. J Clin Pharmacol 2021 Aug;61 Suppl 2:S10-S17. [doi: 10.1002/jcph.1937] [Medline: 34396554]
34. Zaami S, Marinelli E, Varì MR. New trends of substance abuse during COVID-19 pandemic: an international perspective. Front Psychiatry 2020 Jul 16;11:700 [FREE Full text] [doi: 10.3389/fpsyt.2020.00700] [Medline: 32765328]
35. Capasso A, Jones AM, Ali SH, Foreman J, Tozan Y, DiClemente RJ. Increased alcohol use during the COVID-19 pandemic: the effect of mental health and age in a cross-sectional sample of social media users in the U.S. Prev Med 2021 Apr;145:106422 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106422] [Medline: 33422577]
36. Avery AR, Tsang S, Seto EY, Duncan GE. Stress, anxiety, and change in alcohol use during the COVID-19 pandemic: findings among adult twin pairs. Front Psychiatry 2020 Sep 25;11:571084 [FREE Full text] [doi: 10.3389/fpsyt.2020.571084] [Medline: 33088277]
37. Roberts A, Rogers J, Mason R, Siriwardena AN, Hogue T, Whitley GA, et al. Alcohol and other substance use during the COVID-19 pandemic: a systematic review. Drug Alcohol Depend 2021 Dec 01;229(Pt A):109150 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.109150] [Medline: 34749198]
38. Chodkiewicz J, Talarowska M, Miniszewska J, Nawrocka N, Bilinski P. Alcohol consumption reported during the COVID-19 pandemic: the initial stage. Int J Environ Res Public Health 2020 Jun 29;17(13):4677 [FREE Full text] [doi: 10.3390/ijerph17134677] [Medline: 32610613]
39. Sarker A, Nataraj N, Siu W, Li S, Jones CM, Sumner SA. Concerns among people who use opioids during the COVID-19 pandemic: a natural language processing analysis of social media posts. Subst Abuse Treat Prev Policy 2022 Mar 05;17(1):16 [FREE Full text] [doi: 10.1186/s13011-022-00442-w] [Medline: 35248103]
40. Yoon S, Odlum M, Broadwell P, Davis N, Cho H, Deng N, et al. Application of social network analysis of COVID-19 related tweets mentioning cannabis and opioids to gain insights for drug abuse research. Stud Health Technol Inform 2020 Jun 26;272:5-8 [FREE Full text] [doi: 10.3233/SHTI200479] [Medline: 32604586]
41. Chiappini S, Guirguis A, Corkery JM, Schifano F. Misuse of prescription and over-the-counter drugs to obtain illicit highs: how pharmacists can prevent abuse. Pharm J 2020 Nov 17;305(7943) [FREE Full text] [doi: 10.1211/PJ.2020.20208538]
42. Niles JK, Gudin J, Radcliff J, Kaufman HW. The opioid epidemic within the COVID-19 pandemic: drug testing in 2020. Popul Health Manag 2021 Feb;24(S1):S43-S51 [FREE Full text] [doi: 10.1089/pop.2020.0230] [Medline: 33031013]
43. Alexander GC, Stoller KB, Haffajee RL, Saloner B. An epidemic in the midst of a pandemic: opioid use disorder and COVID-19. Ann Intern Med 2020 Apr 02;173:57-58. [doi: 10.7326/m20-1141]
44. Fidancı İ, Aksoy H, Yengil Taci D, Ayhan Başer D, Cankurtaran M. Evaluation of the effect of the Covid-19 pandemic on smoking addiction levels. Int J Clin Pract 2021 May;75(5):e14012 [FREE Full text] [doi: 10.1111/ijcp.14012] [Medline: 33411384]

XSL•FO
RenderX

45. Yang H, Ma J. How the COVID-19 pandemic impacts tobacco addiction: changes in smoking behavior and associations with well-being. Addict Behav 2021 Aug;119:106917 [FREE Full text] [doi: 10.1016/j.addbeh.2021.106917] [Medline: 33862579]
46. Shankar R, Bhattacharya V. Trends in substance use and abuse during COVID-19 among the general population. Addict Health 2022 Oct;14(4):296-308 [FREE Full text] [doi: 10.34172/ahj.2022.1274] [Medline: 37559794]
47. Omare MO, Kibet JK, Cherutoi JK, Kengara FO. A review of tobacco abuse and its epidemiological consequences. Z Gesundh Wiss 2022;30(6):1485-1500 [FREE Full text] [doi: 10.1007/s10389-020-01443-4] [Medline: 33425659]
48. Compton WM, Flannagan KS, Silveira ML, Creamer MR, Kimmel HL, Kanel M, et al. Tobacco, alcohol, cannabis, and other drug use in the US before and during the early phase of the COVID-19 pandemic. JAMA Netw Open 2023 Jan 03;6(1):e2254566 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.54566] [Medline: 36719678]
49. Bolinski RS, Walters S, Salisbury-Afshar E, Ouellet LJ, Jenkins WD, Almirol E, et al. The Impact of the COVID-19 Pandemic on Drug Use Behaviors, Fentanyl Exposure, and Harm Reduction Service Support among People Who Use Drugs in Rural Settings. IJERPH 2022 Feb 16;19(4):2230. [doi: 10.3390/ijerph19042230]
50. Key substance and mental health indicators in the United States: results from the 2020 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration. URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt35325/NSDUHFFRPDFWHTMLFiles2020/2020NSDUHFFR1PDFW102121.pdf [accessed 2023-08-14]
51. Chancellor S, Nitzburg G, Hu A, Zampieri F, De Choudhury M. Discovering alternative treatments for opioid use recovery using social media. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300354]
52. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. J Med Internet Res 2013 Apr 17;15(4):e62 [FREE Full text] [doi: 10.2196/jmir.2503] [Medline: 23594933]
53. Katsuki T, Mackey TK, Cuomo R. Establishing a link between prescription drug abuse and illicit online pharmacies: analysis of Twitter data. J Med Internet Res 2015 Dec 16;17(12):e280 [FREE Full text] [doi: 10.2196/jmir.5144] [Medline: 26677966]
54. Hu H, Phan N, Chun SA, Geller J, Vo H, Ye X, et al. An insight analysis and detection of drug-abuse risk behavior on Twitter with self-taught deep learning. Comput Soc Netw 2019 Nov 06;6:10. [doi: 10.1186/s40649-019-0071-4]
55. Sarker A, O'Connor K, Ginn R, Scotch M, Smith K, Malone D, et al. Social media mining for toxicovigilance: automatic monitoring of prescription medication abuse from Twitter. Drug Saf 2016 Mar;39(3):231-240 [FREE Full text] [doi: 10.1007/s40264-015-0379-4] [Medline: 26748505]
56. Phan N, Chun SA, Bhole M, Geller J. Enabling real-time drug abuse detection in tweets. In: Proceedings of the IEEE 33rd International Conference on Data Engineering. 2017 Presented at: ICDE 2017; April 19-22, 2017; San Diego, CA. [doi: 10.1109/icde.2017.221]
57. Cheng Q, Li TM, Kwok CL, Zhu T, Yip PS. Assessing suicide risk and emotional distress in Chinese social media: a text mining and machine learning study. J Med Internet Res 2017 Jul 10;19(7):e243 [FREE Full text] [doi: 10.2196/jmir.7276] [Medline: 28694239]
58. Sarker A, Gonzalez-Hernandez G, Ruan Y, Perrone J. Machine learning and natural language processing for geolocation-centric monitoring and characterization of opioid-related social media chatter. JAMA Netw Open 2019 Nov 01;2(11):e1914672 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.14672] [Medline: 31693125]
59. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv. Preprint posted online on June 12, 2017 2025 [FREE Full text]
60. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. arXiv. Preprint posted online on July 26, 2019 2025 [FREE Full text]
61. Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal Á. Human-in-the-loop machine learning: a state of the art. Artif Intell Rev 2023;56:3005-3054. [doi: 10.1007/s10462-022-10246-w]
62. Settles B. Active Learning. Cham, Switzerland: Springer; 2012.
63. The Twitter stream grab. Internet Archive. 2021. URL: https://archive.org/details/twitterstream [accessed 2022-09-30]
64. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv. Preprint posted online on October 11, 2018 2025 [FREE Full text]
65. Commonly abused drugs. National Institute on Drug Abuse. URL: https://nida.nih.gov/sites/default/files/cadchart.pdf [accessed 2022-09-24]
66. Slang terms and code words: a reference for law enforcement personnel. Drug Enforcement Administration Intelligence Report. 2018. URL: https://ndews.umd.edu/sites/ndews.umd.edu/files/dea-drug-slang-terms-and-code-words-july2018.pdf [accessed 2025-03-27]
67. Kumar N, Janmohamed K, Nyhan K, Martins SS, Cerda M, Hasin D, et al. Substance, use in relation to COVID-19: a scoping review. Addict Behav 2022 Apr;127:107213 [FREE Full text] [doi: 10.1016/j.addbeh.2021.107213] [Medline: 34959077]
68. Walters SM, Bolinski RS, Almirol E, Grundy S, Fletcher S, Schneider J, et al. Structural and community changes during COVID-19 and their effects on overdose precursors among rural people who use drugs: a mixed-methods analysis. Addict Sci Clin Pract 2022 Apr 25;17(1):24 [FREE Full text] [doi: 10.1186/s13722-022-00303-8] [Medline: 35468860]

XSL•FO
RenderX

69. Mellos E, Paparrigopoulos T. Substance use during the COVID-19 pandemic: what is really happening? Psychiatriki 2022 Mar 28;33(1):17-20 [FREE Full text] [doi: 10.22365/jpsych.2022.072] [Medline: 35255473]
70. Cisneros IE, Cunningham KA. Covid-19 interface with drug misuse and substance use disorders. Neuropharmacology 2021 Oct 15;198:108766 [FREE Full text] [doi: 10.1016/j.neuropharm.2021.108766] [Medline: 34454912]
71. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022. [doi: 10.7551/mitpress/1120.003.0082]
72. Gurung P, Wagh R. A study on topic identification using K means clustering algorithm: big vs. small documents. Adv Comput Sci Technol 2017;10(2):221-233. [doi: 10.13140/RG.2.2.16409.98405]
73. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958 Jun;53:457-481. [doi: 10.2307/2281868]
74. Welcome to the FactorAnalyzer documentation!. FactorAnalyzer. URL: https://factor-analyzer.readthedocs.io/en/latest/ [accessed [accessed 2023-03-03]
75. Gorsuch RL. Factor Analysis Second Edition. Milton Park, UK: Routledge; 1983.
76. Elastic Stack homepage. Elastic Stack. URL: https://www.elastic.co/elastic-stack [accessed 2023-01-01]
77. Logstash 7.17.0 Release Notes. Elastic. URL: https://www.elastic.co/guide/en/logstash/7.17/logstash-7-17-0.html [accessed 2022-12-27]
78. MongoDB homepage. MongoDB. URL: https://www.mongodb.com/ [accessed 2022-12-01]
79. OpenAI homepage. OpenAI. URL: https://openai.com/api/ [accessed 2023-12-01]
80. Northeast Ohio tri-county prevention infrastructure. Kent State University. URL: https://www.kent.edu/mhsu/tcpi [accessed 2025-03-27]
81. Ray PP. ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber Phys Syst 2023;3:121-154. [doi: 10.1016/j.iotcps.2023.04.003]
82. Moradi M, Blagec K, Haberl F, Samwald M. GPT-3 models are poor few-shot learners in the biomedical domain. arXiv. Preprint posted online on September 6, 2021 2025 [FREE Full text]
83. JulinaM / robertaForTweetAnalysis. Github. URL: https://github.com/JulinaM/robertaForTweetAnalysis [accessed 2025-03-27]

Abbreviations

BERT: bidirectional encoder representations from transformers
HITL: human-in-the-loop
LDA: latent Dirichlet allocation
MLM: masked language model
NIDA: national institute on drug abuse
NLTK: Nature Language ToolKit
NSP: next sentence prediction
RoBERTa: robustly optimized bidirectional encoder representations from transformers pretraining approach
SU: substance use
SUD: substance use disorder

Edited by T Mackey; submitted 01.04.24; peer-reviewed by A Fisher, GK Gupta, M Bagewadi Ellur, S Mao, M Elbattah; comments to author 01.05.24; revised version received 31.05.24; accepted 02.03.25; published 17.04.25.

Please cite as:
Maharjan J, Zhu J, King J, Phan N, Kenne D, Jin R
Large-Scale Deep Learning–Enabled Infodemiological Analysis of Substance Use Patterns on Social Media: Insights From the COVID-19 Pandemic
JMIR Infodemiology 2025;5:e59076
URL: https://infodemiology.jmir.org/2025/1/e59076
doi:10.2196/59076
PMID:40244656

©Julina Maharjan, Jianfeng Zhu, Jennifer King, NhatHai Phan, Deric Kenne, Ruoming Jin. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 17.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Use of Social Media for Activism by Mexican Nongovernmental Organizations Using Posts From the 16 Days of Activism Against Gender-Based Violence Campaign: Thematic Content Analysis

Marian Marian[1,2], MS, PhD; Ramona L Pérez[2], MA, PhD; Elizabeth Reed[2], MPH, SCD; Samantha Hurst[1], MA, PhD; Rebecka Lundgren[1], MPH, PhD; Amanda C McClain[2], MS, PhD; Kathryn M Barker[1], MPH, SCD

[1]University of California San Diego, La Jolla, CA, United States

[2]San Diego State University, San Diego, CA, United States

Corresponding Author:
Marian Marian, MS, PhD
University of California San Diego
9500 Gilman Dr
La Jolla, CA, 92093
United States
Phone: 1 (619) 594 1103
Email: mmarian@health.ucsd.edu

Abstract

Background: In the past decade, *hashtag feminism* has emerged in Mexico as a prevalent strategy to build social movements against gender-based violence (GBV). For example, during the global "16 Days of Activism Against GBV" campaign held between November 25 and December 10 each year, Mexico-based nongovernmental organizations (NGOs) turn to X (formerly known as Twitter) to share messages. Despite this prevalence, there is limited research on the type of information shared by these NGO activists on social media and the public's engagement with these messages.

Objective: This study aims to explore the themes covered by Mexican NGOs on X and examine what types of messages related to GBV potentially resonated more with the public.

Methods: We collated and reviewed posts (commonly known as tweets) published in Spanish on the platform X by Mexico-based NGOs between November 25 and December 10 of 2020, 2021, and 2022, a period when digital interactions increased during the COVID-19 pandemic. We then extracted posts using the following 4 hashtags: #16días, #16DiasdeActivismo, or #16DíasdeActivismo; #25N or #25Noviembre; #DiaNaranja or #DíaNaranja; and #PintaElMundoDeNaranja. We subsequently assessed the number of likes each post had and retained the top 200 posts from each year with the highest number of likes. We used the iterative content analysis process and the inductive 6-step qualitative thematic analysis method in NVivo software to code and analyze the final 600 posts.

Results: Five themes emerged from the 16 Days of Activism Against GBV campaigns, covering both knowledge-sharing and activism-generating messages as follows: (1) activism and how to be an activist, (2) types of GBV most commonly highlighted in posts, (3) changing public discourse surrounding GBV, (4) GBV as a violation of human rights, and (5) the COVID-19 pandemic's impact on GBV. Most of the messages on these posts exclusively mentioned women and younger girls, while a few included adolescents. Gaps in the representation of vulnerable populations were also found.

Conclusions: The posts from this campaign that were highly liked by the public reflect some of the most significant societal issues currently present in the country. Our results could help guide further GBV campaigns. Still, further research related to hashtag feminism by Mexico-based NGOs on GBV is needed to understand the population that NGOs reach and how the messages shared on these campaigns translate into activism on online and offline social media platforms.

(JMIR Infodemiology 2025;5:e67368) doi:10.2196/67368

XSL•FO
RenderX

KEYWORDS

gender-based violence; Mexico; hashtag activism; feminist social activism; hashtag feminism; Twitter; X; nongovernmental organization; social media

Introduction

Background

Violence against women and girls, also commonly referred to as gender-based violence (GBV), is one of the world's most persistent forms of human rights violations and a major problem in the public health and clinical health arenas [1,2]. Specifically, GBV continues to be a serious problem in Mexico, as 70% of the women aged ≥15 years have experienced at least one incidence of GBV throughout their lives [3]. GBV takes many forms, with the most prevalent forms in Mexican society being psychological and sexual violence [3]. Disappearances and femicide, which is the intentional murder of a girl or woman because of her gender, are also common forms of violence in Mexico, with approximately 25% of women, adolescent girls, and young girls experiencing these forms of violence [4,5]. The COVID-19 pandemic increased the risk of violence against women, adolescents, and girls; women in North America, Central America, and South America were disproportionately affected by GBV during the pandemic [6]. Particularly in Mexico, domestic violence increased by more than 15% between 2020 and 2021 [7]. In 2020, the number of emergency calls related to violence against women increased by >30%, and the requests for support due to GBV increased by 40% [8,9]. Forms of structural GBV, such as violence at the social and economic levels, were also impacted during the COVID-19 pandemic in Mexico, for example, by the continuous increase in unpaid domestic and care work [10]. In addition, digital violence, which involves gender-based acts of violence through communication technology or digital media, such as image-based abuse, cyberstalking, blackmailing, and online harassment, increased in visibility in Mexico in the years surrounding the limitations of social interactions [10-13].

In Mexico, government institutions provide some forms of social support in the fight against the various forms of GBV, although important gaps in coverage are addressed by nongovernmental organizations (NGOs). Specifically, NGOs provide health care, psychosocial support, and legal aid for women, adolescents, and girls experiencing this type of violence [14]. As GBV and gender inequality became more visible during the COVID-19 pandemic, feminist movements, such as hashtag activism, feminist social activism, and hashtag feminism, in which NGOs actively collaborate, also became more visible [15]. Hashtag activism can facilitate social change, policy formation, and the provision of resources for the public [16]. Feminist social activism, popular since the late 2000s and present in Mexico since 2011, involves feminist individuals and organizations using social media for online civic engagement to denounce GBV and related aspects (such as sexism and gender discrimination); initiate social movements; and connect to share experiences, support, and resources, to name a few [17].

Online engagement to denounce GBV and other forms of gender-inequitable practices includes the initiation of social movements by connecting individuals to share their experiences, provide support, and offer sources of support. Hashtag feminism is a form of activism that combines hashtag activism with feminist social activism using specific hashtags across different digital platforms to call for action by sharing information and stories, connecting people, and organizing and mobilizing events against gender inequities [18-20]. NGOs have played a significant role in catalyzing hashtag feminist movements on GBV worldwide and in Mexico by engaging stakeholders and the general public through dialogue and community-building practices, mainly through Facebook and X (formerly known as Twitter; X Corp) [21-29]. An example of this is the 16 Days of Activism Against GBV, an annual international campaign led by civil society that runs from November 25, the International Day for the Elimination of Violence Against Women, to December 10, Human Rights Day [30,31]. The intention of this global campaign is to call for the prevention and elimination of violence against girls, adolescents, and women in all its different forms, including child marriage, sexual harassment, and intimate partner violence, to name a few [31,32].

Hashtag feminist movements stand to make a great impact in Mexican society, as social media use is widespread. In 2021, 89.2% of the internet users in Mexico connected to the internet daily, and 89.8% used it to access social media [33]. As of January 2022, Mexico ranked ninth in the world for most X users, with 13.9 million users [34]. Similar to what happened in other countries, digital interactions increased during the COVID-19 pandemic in Mexico, and the use of social media to promote strategies against violence, such as support groups, helplines, and screening for violence, became more popular [35-37]. Despite the prevalent use of X in the world and in Mexico, research is limited in regard to hashtag activism and hashtag feminism through X on GBV, specifically on interpersonal and sexual violence [22,38-40]. To our knowledge, only a few publications exist on feminist social activism in Mexico that are specific to GBV at the local level [17,41]. Furthermore, the X content from campaigns that address the umbrella of different types of GBV by Mexico-based NGOs has not previously been explored. Understanding the public's engagement on the GBV hashtag feminist campaigns is critical, as this can inform future GBV campaigns and provide valuable information to NGOs, governments, and researchers for tailored online and offline GBV prevention and intervention strategies.

This Study

To fill this gap, this research aimed to describe the range of GBV themes that Mexican NGOs posted on X during the annual 16 Days of Activism Against GBV campaign during the COVID-19 pandemic (from 2020 to 2022) and elucidate what types of messages related to GBV the public engaged with the most across and between the 3 different years. It is important to specify that while in Mexico the more specific term "violence against women" is commonly used, as it is the term described in the General Law on Women's Access to a Life Free from Violence first published in 2007 [3], this research followed the

XSL•FO
RenderX

term GBV, as it is an umbrella term commonly used interchangeably [42] with violence against women and is the term used in the global campaign that is the focus of our investigation. We sought to describe the GBV themes raised by Mexico-based NGOs on X during these global campaigns and, based on the public's engagement, how the main content themes of posts (commonly known as tweets) compare between the 3 different years. Our findings can provide information on how these campaigns and the public's engagement could modify the agenda against GBV at different societal levels that was weakened during the COVID-19 pandemic [43] by contributing information on the topics that are generating the attention of the public, providing voice to certain communities, and adapting these results into current or future political or legal initiatives.

Methods

Study Design and Data Collection

Our research methodology followed a thematic content analysis of posts posted by Mexico-based NGOs during the annual 16 Days of Activism Against GBV campaign from 2020 to 2022. The following four hashtags chosen for this analysis were based on informal exploration on X of relevant hashtags used by Mexico-based NGOs during the campaign in 2020, 2021, and 2022: (1) the overall name of the campaign and its variations, namely #16días, #16DiasdeActivismo, or #16DíasdeActivismo (#16daysofactivism); (2) the first day of the campaign and the International Day for the Elimination of Violence Against Women (November 25), namely #25N or #25deNoviembre; (3) Orange Day, which occurs on the 25th of each month to create awareness and prevent violence toward girls and women, namely #DiaNaranja and #DíaNaranja (#OrangeDay); and (4) a related but broader hashtag, namely #PintaElMundoDeNaranja (#PaintTheWorldOrange). Given that the hashtag #10D, referring to December 10, the last day of the campaign and Human Rights Day, is also used annually in Argentina for democratic campaigns, we excluded it from analysis as a strategy to ensure data focused on the GBV campaign.

During data collection, the advanced search option from X was used to filter data based on the following study inclusion criteria: (1) only original content (no reposts or retweets); (2) included one of the hashtags used by Mexico-based NGOs during the campaign that were selected for this analysis (#16días, #16DiasdeActivismo, #16DíasdeActivismo, #25N, #25deNoviembre, #DiaNaranja, #DíaNaranja, or #PintaElMundoDeNaranja); (3) published in Spanish; (4) posted by a public X user with "non-governmental and nonprofit organization" as the professional category in order to only collect posts from local, state, national, or global NGOs; (5) the location of the X user marked as Mexico or a city or state in it; and (6) posted between November 25 and December 10 of 2020, 2021, and 2022, to align with the dates the campaign occurred. We used the Zeeschuimer and the 4CAT research tools to capture and download data from X. Zeeschuimer is a browser extension that collects data from social media sites, such as Instagram, TikTok, and X, while 4CAT is a research tool that can be connected to Zeeschuimer to store and download the data collected by the browser extension [44].

Data Cleaning

Posts posted between November 25 and December 10 of 2020, 2021, and 2022 were downloaded from 4CAT to a Microsoft Excel spreadsheet in August 2023. Posts that included videos (<1 min) and informational graphics were also included in the analysis. A total of 1914 posts (n=662, 34.59% for 2020; n=654, 34.17% for 2021; and n=598, 31.24% for 2022) were collected for the hashtags and the selected time period. To avoid duplicates, identical posts published during the same campaign year were collapsed and considered as one individual post. The collected posts were reviewed to ensure they met the inclusion criteria (only posts in Spanish, X user being an NGO, and location of the X user in Mexico).

We used the number of likes received by a post as the metric to determine the public's interest in and engagement with a post [45]. The frequency of likes has been previously used in the literature, including studies on GBV and feminist activism, to measure the public's engagement [21,39,46-50]. Reposts, which have been considered another form of public engagement in which the public contributes and creates content through reposting or forwarding a post [22], were not included in the analysis in order to focus on an original and representative set of posts posted by Mexico-based NGOs that participated in the 16 Days of Activism Against GBV campaigns.

We limited the analysis to the top 200 posts per year with the most likes (N=600 total posts) to assess the messages that generated the most engagement from the public. This analytic decision was based on previously published literature that has demonstrated that at least 500 posts are sufficient to identify themes and understand how individuals engage in health behaviors on X [51,52]. In the event of identical posts from different NGOs, we collapsed the number of likes.

Data Coding and Analysis

After obtaining and cleaning the data, the final list of 600 posts was uploaded into NVivo (version 14; Lumivero) for coding and analysis in the original language. We retained the original language of the posts to more accurately capture the cultural connotations and local meaning of the messages [53]. We completed an iterative content analysis process and a 6-step inductive thematic content analysis to identify and categorize the content of the posts into themes and explore socially produced and reproduced experiences [21,22,54]. These systematic approaches to interpreting qualitative research data have been used by other researchers when conducting thematic content analysis of social media data, including X hashtags in GBV campaigns [38,55-59]. The thematic process involved familiarization with the data, selection of keywords, coding, theme development, interpretation of themes, and development of a conceptual model [55]. An iterative inductive content analysis approach [60] was followed for the creation of the codebook by the first author and 2 research assistants (Yxchel Tejeda and Julia Godinez), as code categorization and subcategorization were extracted from the posts. Both research assistants were bilingual and were trained in thematic content analysis. Regular debriefing meetings between the primary author and the 2 research assistants took place during the coding and analysis process to refine the working codebook, discuss

XSL•FO
RenderX

findings and emergent themes, resolve any coding discrepancies, and address any potential questions. The research team used NVivo (version 14) for coding and analysis. After the initial analysis, data were translated from Spanish to English by the first author (MM), who is bilingual.

Ethical Considerations

This study used publicly available, deidentified posts and did not directly involve human participants. Per the US Department of Health and Human Services' regulations for the protection of human subjects in research (45 CFR 46), approval from a research ethics committee was therefore not required or obtained. However, we adhered to ethical guidelines for handling and analyzing publicly available data, ensuring user anonymity and data privacy throughout the research process, including findings reported in this study.

Results

Overview

The 600 posts included in the final sample came from 61 different Mexico-based organizations. The number of likes for each post ranged from 9 to 2095. As shown in Table 1, the hashtag using the campaign's name (16 Days of Activism Against GBV) was the most frequently used hashtag in all 3 campaign years.

Table 1. Frequency of hashtags mentioned on most liked posts per campaign year by Mexico-based nongovernmental organizations.

Hashtag	2020, n (%)	2021, n (%)	2022, n (%)	Total, n (%)
#25N or #25Noviembre (n=216)	56 (25.9)	58 (26.9)	102 (47.2)	216 (100)
#DiaNaranja or #DíaNaranja (n=283)	113 (39.9)	111 (39.2)	59 (20.8)	283 (100)
#16días, #16DiasdeActivismo, or #16DíasdeActivismo (n=528)	167 (31.6)	204 (38.6)	157 (29.7)	528 (100)
#PintaElMundoDeNaranja (n=76)	49 (64.5)	22 (28.9)	5 (6.6)	76 (100)

The following five themes emerged across the 3 campaign years: (1) activism and how to be an activist, (2) types of GBV most commonly highlighted in posts, (3) changing public discourse surrounding GBV, (4) GBV as a violation of human rights, and (5) the impact of the COVID-19 pandemic on GBV. Within some of these themes, we also identified several cross-cutting subthemes, as described subsequently.

Theme 1: Activism and How to Be an Activist

The most common theme among the highly liked posts covered descriptions of what the 16 Days of Activism Against GBV campaign is, along with what activism is and how to be an activist (Table 2). Many of these posts were invitations to join the campaign or be activists. Other posts stressed that activism to support the rights of women and girls does not stop when the campaign stops but should be practiced on a regular basis. Invitations, live updates, and links to different types of events, such as marches, podcasts, and webinars organized during the campaigns, were also shared in these posts.

Table 2. Representative quotes from theme 1 (activism and how to be an activist).

Subthemes	Examples (posts translated into English)	Examples (posts in original language)
Description of the campaign	• "During these #16Días and all year long, stand in solidarity with women's rights activists and support feminist movements to resist the rollback of women's and girls' rights. #Únete #25N" [2022 campaign] • "What are the #16Días de activismo? It is a call to action and a reminder that violence against women and girls is the most widespread human rights violation around the world. #DíaNaranja Take action to end violence against women and girls." [2022 campaign]	• "Durante estos #16Días y todo el año, solidarízate con las activistas de los derechos de las mujeres y apoya a los movimientos feministas para resistir el retroceso de los derechos de las mujeres y las niñas. #Únete #25N" [Campaña 2022] • "¿Qué son los #16Días de activismo? Es un llamado a la acción y un recordatorio de que la violencia contra las mujeres y las niñas es la violación de derechos humanos más extendida en todo el mundo. #DíaNaranja Actúa para poner fin a la violencia contra las mujeres y las niñas." [Campaña 2022]
Campaign events	• "Women musicians, visual artists, activists, journalists, photographers and filmmakers join their voices to create #25NMás16, a sound, visual and informative piece promoted by the Initiative #SpotlightMX #DíaNaranja #Únete #16Días" [2020 campaign]	• "Mujeres músicas, artistas visuales, activistas, periodistas, fotógrafas y cineastas unen sus voces para crear #25NMás16, una pieza sonora, visual e informativa impulsada por la Iniciativa #SpotlightMX #DíaNaranja #Únete #16Días" [Campaña 2020]
Geographical examples	• "#25N| This is a fight for all women 'Neither in Tijuana, nor in Chiapas, nor in this city, NOT ONE MORE MURDERED' #NiUnaMenos #VivasNosQueremos" [2020 campaign]	• #25N| Esta es una lucha de todas las mujeres "Ni en Tijuana, ni en Chiapas, ni en esta ciudad, NI UNA ASESINADA MÁS" #NiUnaMenos #VivasNosQueremos" [Campaña 2020]

XSL•FO
RenderX

Different NGOs used different types of communication materials to explain the campaign's purpose and why and how to be an activist, including songs, short videos, and celebrity involvement. Particularly in the 2021 campaign, posts included images and quotes related to the campaign from musicians, politicians, ambassadors, NGO directors and representatives, and leaders in industry and sport. Figure 1 shows examples of visuals used in highly liked posts for the 16 Days of Activism Against GBV campaigns in 2020, 2021, and 2022.

Figure 1. Examples of visuals used by Mexico-based nongovernmental organizations in the frequently liked X (formerly known as Twitter) posts for the 16 Days of Activism Against Gender-Based Violence campaigns in 2020, 2021, and 2022.

Posts' messages on activism and how to be an activist were mainly generalized to the national level or provided data at a country level, although some included specific activities and examples of GBV cases from different municipalities, states, and, in 3 instances, other countries.

Theme 2: Types of GBV Most Commonly Highlighted in Posts

Many posts consisted of information about a specific type of violence (Table 3). Sexual violence was the most common type of violence found in these posts, followed by messages related to trafficking, disappearances, and femicide. These types of posts included data on the number of femicides or women and girls who have disappeared in the country and advice on what needs to change to prevent trafficking. Specific examples of women who have experienced GBV, particularly sexual violence and torture, were also included. Violence at home and physical violence were other forms of GBV that were also commonly present in these posts. Posts about digital violence were present in the data as well, particularly in 2020, at the start of the COVID-19 pandemic.

XSL•FO
RenderX

Table 3. Representative quotes from theme 2 (types of gender-based violence most commonly highlighted in posts).

Subthemes	Examples (posts translated into English)	Examples (posts in original language)
Sexual violence	• "Let us join forces against those who attempt to blur the boundaries of sexual consent, blame victims, and excuse perpetrators. During the #16Days of Activism, let's say it loud and clear: NO IS NO. #Únete #DíaNaranja" [2020 campaign]	• "Unamos nuestros esfuerzos contra quienes intenten desdibujar los límites del consentimiento sexual, culpar a las víctimas y excusar a los agresores. Durante los #16Días de Activismo, digámoslo alto y claro: NO ES NO. #Únete #DíaNaranja" [Campaña 2020]
Femicides and trafficking	• "Mexico closes the year with the terrible number of 27 thousand missing girls and women. Within the framework of #16Días of Activism, it is essential to demand differentiated strategies with a gender perspective to find them all" [2022 campaign] • "In Mexico, 11 femicides occur DAILY. Today #25N International Day for the Elimination of Violence against Women, we raise our voices for freedom, security, justice and sisterhood for all girls and women in Mexico. #NiUnaMas #VivasNosQueremos #DiaNaranja" [2020 campaign]	• "México cierra el año con la terrible cifra de 27 mil niñas y mujeres desaparecidas. En el marco de #16Días de activismo es fundamental exigir estrategias diferenciadas con perspectiva de género para encontrarlas a todas" [Campaña 2022] • "En México ocurren 11 feminicidios DIARIOS. Hoy #25N Día Internacional de la Eliminación de la Violencia contra la Mujer, alzamos la voz por la libertad, seguridad, justicia y sororidad para con todas las niñas y mujeres en México. #NiUnaMas #VivasNosQueremos #DiaNaranja" [Campaña 2020]
Violence at home	• "Women, girls and adolescents need to have peace of mind in their home and feel that their home is a safe and reliable space. Their home is NOT a space of #violence. Let's make families and homes #EspaciosSeguros! #25N #PintaElMundoDeNaranja #16Días" [2021 campaign]	• "Las mujeres, niñas y adolescentes necesitan tener tranquilidad en su hogar y sentir que su casa es un espacio seguro y confiable. Su hogar NO es un espacio de #violencia. ¡Hagamos de las familias y los hogares #EspaciosSeguros! #25N #PintaElMundoDeNaranja #16Días" [Campaña 2021]
Digital violence	• "The increase in internet use during #COVID19 has made women and girls targets of online violence. In these #16Días of Activism, #pintaelmundodenaranja using your networks to raise awareness and promote solidarity. #DíaNaranja #Únete" [2020 campaign]	• "El incremento del uso de internet durante #COVID19 ha convertido a mujeres y niñas en blanco de violencias en línea. En estos #16Días de Activismo, #pintaelmundodenaranja usando tus redes para concientizar y fomentar la solidaridad. #DíaNaranja #Únete" [Campaña 2020]
Physical violence	• " Did you know that 1 in 3 women in the world has suffered physical or sexual violence throughout their lives? Share this post and #Únete the #16días of activism to end this global pandemic. #25N" [2021 campaign]	• " ¿Sabías que 1 de cada 3 mujeres en el mundo ha sufrido violencia física o sexual a lo largo de su vida? Comparte este post y #Únete a los #16días de activismo para acabar con esta pandemia mundial. #25N" [Campaña 2021]

Other forms of violence, such as the lack of access to sexual and reproductive resources, obstetric violence, vicarious violence, and precarious work, were present in posts, but they were not as common as sexual violence, femicide, and digital violence. Violence against women in the form of harassment in specific locations, such as school and at work, was also highlighted by the campaign but to a lesser extent.

Theme 3: Changing Public Discourse Surrounding GBV

Posts in this theme centered on changing current narratives in Mexican society surrounding GBV that blame and silence survivors and normalize violence (Table 4). The posts provided examples of how to change the narrative, such as believing the victim, not excusing the perpetrator, and reporting any form of violence. The actors responsible for making these changes included both individuals and institutions. At the individual level, most posts were not targeted at a specific gender; however, some did aim toward male individuals and their role in eradicating GBV. Some highly liked posts were focused on female individuals, while other posts were specific to inviting both men and women to become activists and eradicate GBV. Throughout the 3 campaign years, the highly liked posts provided specific information about GBV toward women or women and younger girls. A few highly liked posts included adolescents in the messages, and, when included, these posts were about violence against younger girls and adolescents or against younger girls, adolescents, and women. At the institutional level, posts provided reasons for the importance of these groups in combating GBV, such as health and government personnel providing help to women after experiencing some form of GBV, while also providing examples of what these institutions can do, such as treating survivors of GBV with dignity and defending the rights of women, adolescents, and girls.

XSL•FO
RenderX

Table 4. Representative quotes from theme 3 (changing public discourse surrounding gender-based violence).

Subthemes	Examples (posts translated into English)	Examples (posts in original language)
Believe survivors	• "Every time a woman talks about her experience of sexual violence and we don't believe her, rape culture grows stronger. Every time you hear a survivor's story: 1-Listen. 2-Believe. 3-Support. #DíaNaranja #16Días #Únete" [2020 campaign]	• "Cada vez que una mujer habla de su experiencia de violencia sexual y no le creemos, la cultura de la violación se hace más fuerte. Cada vez que escuches la historia de una sobreviviente: 1-Escucha. 2-Cree. 3-Apoya. #DíaNaranja #16Días #Únete" [Campaña 2020]
End the silence	• "If you face violence, don't be silent, raise your voice. Call 911 for help immediately and tell your trusted people. #16Días #DíaNaranja" [2021 campaign] • "Report when you see: Abuse Harassment on the street Sexist jokes Unwanted behavior Inappropriate sexual comments Sexual harassment is never acceptable. #16Días #Únete #DíaNaranja" [2020 campaign]	• "Si enfrentas violencia, no te calles, alza la voz. Pide ayuda inmediatamente al 911 y cuéntalo a tus personas de confianza. #16Días #DíaNaranja" [Campaña 2021] • "Denuncia cuando veas: Abuso Acoso en la calle Bromas sexistas Comportamiento no deseado Comentarios sexuales inapropiados El acoso sexual nunca es aceptable. #16Días #Únete #DíaNaranja" [Campaña 2020]
Call out perpetrators' actions	• "Don't be an accomplice to those who commit violence, don't look the other way. Speak up, intervene and show your support for survivors during the #16Días, and every day. #Únete #DíaNaranja" [2020 campaign]	• "No seas cómplice de quien ejerce violencia, no mires a otro lado. Habla, interviene y muestra tu apoyo a las sobrevivientes durante los #16Días, y todos los días. #Únete #DíaNaranja" [Campaña 2020]
Engage men	• "As men we have to reflect, question ourselves, inform ourselves and act. #25N Let's build relationships of respect, peace and equality! " [2022 campaign]	• "Como hombres nos toca reflexionar, cuestionarnos, informarnos y actuar. #25N ¡Construyamos relaciones de respeto, paz e igualdad! " [Campaña 2022]
Focus on women and younger girls	• "Eliminating violence against girls and adolescents is everyone's task. Get informed and learn about different ways in which you can prevent gender violence. #16Días #DíaNaranja" [2021 campaign] • "Violence will affect 1 in 3 girls and women throughout their lives. This must stop! #PintaElMundoDeNaranja and unite against violence." [2020 campaign]	• "Eliminar la violencia contra las niñas y las adolescentes es tarea de todas y todos. Infórmate y conoce diversas formas en que tú puedes prevenir la violencia de género. #16Días #DíaNaranja" [Campaña 2021] • "La violencia afectará a 1 de cada 3 niñas y mujeres a lo largo de su vida. ¡Esto debe parar! #PintaElMundoDeNaranja y únete contra la violencia." [Campaña 2020]
Responsibility of societal institutions	• "People in leadership must implement prevention measures that address inequalities in power relations between genders, which are at the root of violence against women and girls #16Días #DíaNaranja" [2021 campaign] • "Health providers can also defend the life and health of girls who have been forced to become mothers. Taking care of the health of survivors also involves having the ability to treat them humanely and with dignity. They are #GirlsNotMothers. #25N" [2021 campaign]	• "Las personas en liderazgo, deben poner en marcha medidas de prevención que hagan frente a las desigualdades en las relaciones de poder entre los géneros, que se encuentran en la raíz de la violencia contra las mujeres y las niñas. #16Días #DíaNaranja" [Campaña 2021] • "Las personas proveedoras de salud también pueden defender la vida y la salud de las niñas que han sido forzadas a ser madres. Cuidar la salud de las sobrevivientes también pasa por tener la capacidad de darles un trato humano y digno. Son #NiñasNoMadres. #25N" [Campaña 2021]

Theme 4: GBV as a Violation of Human Rights

Posts related to GBV as a human rights concern rooted in gender inequalities and power imbalances were also common in the 3 campaign years (Table 5). These posts that stated the importance of challenging inequitable gender norms to promote equity and justice for girls and women were highly liked during the 3 campaign years. Particularly, for posts from 2020 related to justice and equity, many highly liked posts described real-life events of women in Mexico who had experienced sexual violence, wrongful convictions, and unjust incarcerations.

The focus on inequities did not stop at gender inequities. Posts also focused on violence toward specific vulnerable groups, such as people with disabilities, migrants, and sexual minority groups, and they were also found in the highly liked posts for the 3 campaign years, especially in 2021 and 2022. The two most common messages among posts related to vulnerable groups were (1) events to visualize the GBV that migrants experience in Mexico and (2) information on vulnerable populations being at a higher risk of experiencing GBV.

XSL•FO
RenderX

Table 5. Representative quotes from theme 4 (gender-based violence as a violation of human rights).

Subthemes	Examples (posts translated into English)	Examples (posts in original language)
Human rights violation	" Violence against women is the most widespread violation of human rights in the world and the latest UN data show that the greatest threat to women and girls lies in the people in their immediate environment. #16Días #YaEsYa #DíaNaranja #Únete" [2022 campaign]	" La violencia contras las mujeres es la violación más generalizada de los derechos humanos en el mundo y los últimos datos de ONU demuestran que la mayor amenaza para las mujeres y las niñas reside en las personas de su entorno más cercano. #16Días #YaEsYa #DíaNaranja #Únete" [Campaña 2022]
Origins in gender inequity	"The origin of violence is discrimination and gender inequality, as well as gender stereotypes and harmful masculinities still in force in our societies. #16Días #DíaNaranja #Únete" [2021 campaign]	"El origen de la violencia es la discriminación y la desigualdad de género, así como los estereotipos de género y las masculinidades nocivas aún vigentes en nuestras sociedades. #16Días #DíaNaranja #Únete" [Campaña 2021]
Violence toward vulnerable groups	"Inequality, poverty, ethnic origin, disability, immigration status, among others, increase the vulnerability of women and girls. Let's end violence against women and girls NOW! #16Días #DíaNaranja" [2021 campaign]	"La desigualdad, la pobreza, el origen étnico, la discapacidad, el estatus #migratorio, entre otros, aumentan la vulnerabilidad de las mujeres y las niñas. ¡Pongamos fin a la violencia contra las mujeres y las niñas YA! #16Días #DíaNaranja" [Campaña 2021]

Theme 5: The Impact of the COVID-19 Pandemic on GBV

Finally, a common theme among posts from 2020 (but not common in the 2021 and 2022 campaign posts) was the impact of the COVID-19 pandemic on GBV in Mexico. Posts in this theme highlighted the importance of the resources for GBV victims during this time and shared data or examples of how the pandemic increased GBV during the lockdown period in Mexico. Posts also cautioned that inaction could lead to setbacks in progress made toward gender equity. Table 6 presents selected quotations representing this theme.

Table 6. Representative quotes from theme 5 (the impact of the COVID-19 pandemic on gender-based violence).

Examples (posts translated into English)	Examples (posts in original language)
"We have stayed at home to avoid risks of contagion by #COVID19 but what happens if the risk is at home? Violence against women in the home has increased during the pandemic, but we can all help victims. #DíaNaranja #16Días #Únete" [2020 campaign]	"Nos hemos quedado en casa para evitar riesgos de contagio por #COVID19 pero ¿qué pasa si el riesgo está en casa? La violencia contra las mujeres en el hogar ha aumentado durante la pandemia, pero toda/os podemos ayudar a las víctimas. #DíaNaranja #16Días #Únete" [Campaña 2020]
" Shelters Helplines Advice Services for survivors are essential. Any type of support for victims of violence must be available to anyone who needs it, even during the #COVID19 pandemic. #DíaNaranja #16Días #Únete" [2020 campaign]	" Refugios Líneas telefónicas de ayuda Asesoramiento Los servicios para las sobrevivientes son esenciales. Cualquier tipo de apoyo a las víctimas de violencia debe estar disponible para quien lo necesite, incluso durante la pandemia de #COVID19. #DíaNaranja #16Días #Únete" [Campaña 2020]
"The pandemic is being particularly painful for women. If no action is taken, the progress in gender equality made over the last 25 years will be lost. #16Días #DíaNaranja #Únete" [2020 campaign]	"La pandemia está siendo particularmente dolorosa para las mujeres. Si no se adoptan medidas, los avances en la igualdad de género logrados en los últimos 25 años se perderán. #16Días #DíaNaranja #Únete" [Campaña 2020]

Discussion

Principal Findings

Using data from 3 years of social media activism by Mexican NGOs on X, we found that messaging surrounding the 16 Days of Activism Against GBV centered on the following: (1) activism and how to be an activist, (2) types of GBV, (3) changing the public discourse surrounding GBV, (4) GBV as a human rights violation, and (5) the impact of the COVID-19 pandemic on GBV. The posts from this campaign that were highly liked by the public reflect some of the biggest societal issues currently present in Mexico. In total, 3 of the 5 main themes identified—the types of GBV, activism and how to be an activist, and changing public discourse surrounding GBV—were present across the 3 campaign years. The impact of the COVID-19 pandemic on GBV and GBV as a violation of human rights were the more prevalent themes in 2020 and 2021 and 2022, respectively. Subsequent research on future campaigns could help better understand how these themes evolve in Mexican society over time.

Activism and how to be an activist was the most common theme among the posts analyzed during the 16 Days of Activism against GBV campaigns from 2020 to 2022. The activism in this annual campaign educates the public on GBV through posts that show the information in different ways, such as videos, songs, data, infographics, and citing influential people, to name a few. Similar to what other studies have found related to digital activism, the posts from this campaign provide information that makes the problem of GBV visible, which helps destigmatize GBV in Mexico [61,62]. A mixed methods study published in 2018 analyzed whether GBV activism through Facebook creates sisterhood among followers and found that the sisterhood created

XSL•FO
RenderX

on Facebook through the interaction and creation of groups and communities was translated into activism online and offline; this study noted that the lack of groups and communities on X does not allow this sisterhood to be created among X users [28]. However, we suggest that further analysis on X is needed to understand the potential of activism developing a sisterhood on this social media platform through reposts, replies, and other communication tools present on this platform.

Related to the different forms of GBV, sexual violence was the main form of violence highlighted in the analyzed posts. This resonates with the prevalence of sexual violence among the Mexican population, as this type of violence is the second most common form of GBV in Mexico [3]. Sexual violence prevalence among women aged ≥15 years increased from 41.3% in 2016 to 49.7% in 2021 [3]. Digital violence was another form of violence featured in the campaign posts. Digital violence soared in 2020 as a consequence of the COVID-19 pandemic [63]. This type of violence has continued to increase after the pandemic and affects girls and adolescent girls more than their male counterparts. Between 2021 and 2022, more than 33% of girls and adolescent girls in Mexico with access to a phone or internet experienced some form of digital violence, compared to 12% to 18% of Mexican boys and male adolescents [64]. As both forms of GBV continue to increase, the 16 Days of Activism against GBV campaign as well as other NGO- and non–NGO-based campaigns and prevention and intervention strategies should focus on sexual and digital violence in girls, adolescents, and women in Mexico.

We found that less commonly discussed forms of violence, such as economic violence, vicarious violence, lack of access to sexual and reproductive resources, and obstetric violence [65-69] were covered in the posts for these campaigns. Although other more commonly known forms of violence, such as sexual violence and intimate partner violence, were more commonly highlighted in these posts, having information on less commonly discussed forms of GBV within the campaign educated the public and provided tools to identify more types of violence. Indeed, although these forms of GBV are less commonly known by the public, it does not necessarily mean that these forms of violence are less prevalent. For example, Mexican national data have found that economic violence, which is not a well-known form of GBV in the country compared to physical, psychological, and sexual violence, is the most common form of GBV in the workplace [70].

A notable finding from theme 3, changing the narrative surrounding GBV, is that those called upon for change were not only individuals and the general public but also those in the public sector, including people in government leadership positions and health care providers. For the latter, the highly liked posts from the 3 campaign years acknowledged health care professionals as pillars of society for eradicating GBV in Mexico, one of the few countries with laws for the health care sector to address GBV at different prevention levels [71]. However, limited research has examined whether the Mexican health care sector is addressing GBV with their patients or working in other ways toward prevention. For example, a study completed among health care professionals in the states of Quintana Roo, Coahuila, and Mexico City found that while health care professionals were willing to address the issue of domestic violence with their patients, the care and attention needed for this specific type of violence was insufficient [72]. The awareness provided by the NGOs' posts and the engagement of the public on the vital role of these medical professionals in preventing and eliminating GBV could lead to the health care sector and other types of leadership organizations finally taking ownership of their role in reducing GBV.

A current social issue represented as a subtheme under theme 4, GBV as a human rights violation, was the violence experienced by vulnerable groups in Mexico, including migrant girls, adolescents, and women. Since 2018, caravans of migrants from Central and South America have traveled to Mexico in order to reach the United States, with girls, adolescents, and women being particularly vulnerable to different forms of GBV through their travels in Mexico [73,74]. According to the literature, migrants are more likely to be the victims of human trafficking [73], another form of GBV also present in highly liked posts for the 3 campaign years. However, the posts analyzed did not provide resources that could help prevent human trafficking, as these posts only had data on the prevalence of this GBV. Specific resources, such as phone numbers, may be more instrumental, particularly in states such as Chiapas, where migrants tend to stay longer and where it has been shown that GBV is more prevalent for this group [73]. A vulnerable group not mentioned in the highly liked posts that were analyzed was of Indigenous girls, adolescents, and women, even though they have consistently experienced GBV, structural violence, and institutional violence [75,76]. It is unknown if posts with fewer likes and thus not included in these analyses specifically mentioned this vulnerable group in posts by NGOs during the 3 campaign years. Future research on this topic is suggested to not only better understand the public's interest or awareness but also to identify different vulnerable groups that are not represented or addressed in nation-based campaigns or are not receiving broader support through recognition (likes).

Our results found that most of the messages on these posts exclusively mentioned women and younger girls, while only a few included adolescents. However, in Mexico, 60% of the adolescents aged between 15 and 17 years have experienced some type of GBV, and 40% have experienced sexual violence, while 80% of the minors aged <18 years who have disappeared in this country were adolescent girls aged between 12 and 17 years [4,77]. In addition, different studies have found a high prevalence of GBV in Mexican adolescent girls. For example, a study in the northern city of Tijuana found that most adolescent girls have experienced some form of intimate partner violence, more specifically, emotional, sexual, or physical abuse [78]. It is imperative for Mexico-based NGOs involved in future campaigns for 16 Days of Activism against GBV to be more inclusive of adolescent girls in their posts and have specific messages toward this age group, as they are vulnerable and in need of information to combat GBV, and to take account of the influence of social media toward adolescents. It is also important to state that the messages from most highly liked posts followed a collective voice, as these were written from a collective point of view in order to express ideas about GBV and activism against GBV through group unity, shared responsibility, and

XSL•FO
RenderX

changes as a group. This characteristic of the highly liked posts from Mexican NGOs is in line with existing literature, which has previously identified that Latin American feminist movements are distinguished by a collective voice compared to non-Latin American campaigns that have been framed through an individualized lens, such as the #MeToo movement [79].

Research using social media data to learn more about hashtag activism and hashtag feminism has grown in the last decade. Particularly, published literature on X data to analyze activism against GBV in the world has focused on hashtags in English [38,80], and most of it has been related to sexual violence [22,39,40,81]. Previous examples of hashtag feminism in Mexico include the #MeToo movement in 2017; #SiMeMatan (#IfTheyKillMe), used by women in response to the media and authority's secondary victimization of femicide victims; and #MiPrimerAcoso (#MyFirstAssault) in 2016, related to childhood sexual abuse, to name a few [82]. Some of these hashtags have been previously analyzed; however, to the best of our knowledge, we believe this research is one of the first studies to focus on Mexico-based NGOs, different campaign years, and more than 1 hashtag. This study provides important information on the types of messages that Mexico-based NGOs focused on in the prevention and elimination of GBV as well as what types of messages are potentially of more interest to the public. Future studies could further expand the limited knowledge on hashtag feminism against GBV in Mexico. For example, analyzing the demographics of the X users who liked these posts would provide context for understanding the population that follows this campaign. Analyzing other campaigns on GBV as well as the messages published and liked on other popular social media platforms in Mexico, such as Facebook, TikTok, and Instagram, is worth pursuing in future studies to further explore and better understand topics related to the prevention and elimination of GBV. Moreover, future research that quantifies post volume and interactive users through in-depth social network and semantic network analysis could provide more information on our findings by further exploring the level of user engagement and the evolution of discussions on GBV and activism over time. This tone of voice emerged across the 5 themes.

Finally, due to the inductive nature of this research, our study did not follow a theoretical framework. Nonetheless, our results lay the groundwork for future research that could benefit from the use of theoretical frameworks, such as the intersectionality framework [83], to further understand the intersecting identities of the public engaging in GBV campaigns and the feminist theory [84] to further study one of the main themes for this research, GBV as a violation of human rights, including gender inequities.

Strengths and Limitations

As noted earlier, to the best of our knowledge, this study is the first of its kind to conduct a comprehensive thematic content analysis using X data to examine GBV activism across Mexico at a national level. It delivers a groundbreaking perspective on the powerful impact of hashtag activism and feminist movements, offering an unprecedented understanding of social activism in the country. The research brings valuable insights into the topics related to GBV and the messages from this activist campaign that the public engages with the most. This high-level synthesis provides an overview of priorities and topics that can help NGOs and other types of organizations be more strategic in their social media campaigns and messages to combat GBV. For example, we note a gap in the provision of specific resources for those facing GBV; this may be a key area for future campaigns to build on. Despite these strengths, we recognize the inherent limitations in working with X data. First, the text content of posts is limited to up to 280 characters. This particular characteristic of X could hinder the amount of content that can be shared with the public. While some posts analyzed in this study included videos or infographics that could further expand the message shared by the NGO, most posts were messages of <280 characters with concise and specific yet limited information. First, we chose to use X data for analysis instead of another platform such as Facebook, as X offers a diverse dataset and has been one of the main channels used for hashtag feminism [39]. Second, while we collected and analyzed all the text present in the highly liked posts, we only analyzed the visual elements from the posts in the form of pictures and short videos (<1 min), losing the potential for more emerging themes to come from longer videos. Third, our study is limited to posts posted by NGOs that participated in the 16 Days of Activism against GBV campaigns. The results are not generalizable to other types of users (such as government agencies and celebrities) or to other hashtags related to the campaign that were not analyzed in this research. Fourth, this analysis was specific to NGOs based in Mexico and posts published in Spanish, so our results might not be generalizable to the campaign in other countries. Fifth, we focused our research on 4 hashtags (and their variations) specific to this campaign; therefore, it is possible that our research missed capturing other themes from other hashtags (such as #10D, #niunamenos [#notonemore], #niunaasesinadamás [#notonemoremurdered]) related to the 2020, 2021, and 2022 campaigns. Sixth, because of the limited number of posts (N=600) analyzed in this research, themes could have been missed; however, a sample size of 500 posts has been demonstrated to be sufficient to identify themes and understand the public's engagement [52,85]. Finally, the posts chosen for this analysis were based on the number of likes at the time of data collection. The number of likes on a post changes through time; therefore, the messages from these posts that generated the most attention and engagement from the public at the time of the campaign could have changed by the time the data were collected and might not reflect the engagement of the public with the different topics related to GBV. However, this is unlikely, as research has found that most likes on a post are given within 24 hours after the post is posted on X [86].

Conclusions

Our findings reveal that the highly liked posts for the 16 Days of Activism against GBV campaigns posted by Mexico-based NGOs in 2020, 2021, and 2022 are a reflection of the forms of violence and social issues that currently occur in the country. According to our results, informational posts about the types of GBV, the role of activism, and changing public discourse

XSL•FO
RenderX

surrounding GBV are the types of messages that seem to engage the public the most. Our results could help inform and guide future GBV campaigns while also informing the NGOs of the lack of representation in their messages about GBV toward important vulnerable populations, such as adolescents and Indigenous groups. Nevertheless, further research related to hashtag feminism by Mexico-based NGOs on GBV, such as the demographics of X users who engage on these posts and the types of activism interactions among X users on other social media platforms, is vital to understand the population that NGOs reach and how the messages shared on these campaigns translate into activism online and offline.

Acknowledgments

The authors would like to thank Yxchel Tejeda and Julia Godinez for their support in the coding and analysis of the data for this project.

Conflicts of Interest

None declared.

References

1. Violence against women. Pan American Health Organization & World Health Organization. URL: https://www.paho.org/en/topics/violence-against-women [accessed 2024-02-08]
2. Violence against women. World Health Organization. URL: https://www.who.int/health-topics/violence-against-women [accessed 2024-02-08]
3. Violencia contra las mujeres en México. INEGI. 2022. URL: https://www.inegi.org.mx/tablerosestadisticos/vcmm/ [accessed 2023-01-27]
4. UNICEF. Proteger a las niñas y a las adolescentes de cualquier forma de violencia. United Nations Children's Fund. URL: https://www.unicef.org/mexico/protecci%C3%B3n-la-ni%C3%B1ez-y-adolescencia/proteger-las-ni%C3%B1as-y-las-adolescentes-de-cualquier-forma-de [accessed 2024-04-05]
5. Strickland RD. Violencia contra niñas, niños y adolescentes en México. Psicol Iberoam. URL: https://www.redalyc.org/journal/1339/133972771005/html/ [accessed 2024-04-05]
6. ONU México: poner fin a la violencia contra las mujeres y las niñas no está en pausa. UNICEF México. 2020. URL: https://www.unicef.org/mexico/comunicados-prensa/onu-m%C3%A9xico-poner-fin-la-violencia-contra-las-mujeres-y-las-ni%C3%B1as-no-est%C3%A1-en [accessed 2023-02-08]
7. Felipe de Alba. Violencia Familiar en México (2015-2022). Centro de Estudios Sociales y de Opinión Pública. 2023. URL: https://portalhcd.diputados.gob.mx/PortalWeb/Micrositios/8657bfd2-a41c-46bc-837d-9a1d492c2c39.pdf [accessed 2024-04-29]
8. Murray C. México: las llamadas de emergencia por violencia contra las mujeres aumentaron en 2020. Glob Citiz. 2021. URL: https://www.globalcitizen.org/es/content/mexico-violence-women-emergency-calls/ [accessed 2023-01-30]
9. Galván M. La otra pandemia,631 mujeres huyeron por violencia de enero a mayo de 2021. Expans Política Mexico. URL: https://politica.expansion.mx/voces/2021/07/24/pandemia-historica-mexico-13631-mujeres-huyen-por-violencia [accessed 2023-01-30]
10. PAHO. Gendered health analysis COVID-19 in the Americas. Pan American Health Organization. URL: https://iris.paho.org/bitstream/handle/10665.2/55432/PAHOEGCCOVID-19210006_eng.pdf?sequence=4&isAllowed=y [accessed 2024-04-29]
11. Care in Latin America and the Caribbean during the COVID-19: towards comprehensive systems to strengthen response and recovery. UNWOMEN & ECLAC. URL: https://tinyurl.com/3mjcb5sf [accessed 2024-04-29]
12. Rosas Villicaña RM. Orfandad y violencia a niñas, niños y adolescentes en la pandemia de Covid-19. El caso de México en el contexto latinoamericano. Relaciones 2022 Jun 01;42(166):1. [doi: 10.24901/rehs.v42i166.861]
13. Reporting tip sheet on digital violence: a practical reference guide for journalists and media. United Nations Population Fund. URL: https://www.unfpa.org/resources/digital-violence-tip-sheet-for-journalists [accessed 2024-06-25]
14. Cantú-Martínez PC. Organizaciones No Gubernamentales y Violencia de Género: Caso Nuevo León, México. Univ Rev Cienc Soc Humanas 2021 Aug 04(35):151-165. [doi: 10.17163/uni.n35.2021.07]
15. González-Malabet MA, Sanandres Campis E, May R, Molinares Guerrero IS, Durán-Oviedo S. The hybrid political role of feminism on Twitter during COVID-19: SISMA Mujer in Colombia. Womens Stud Int Forum 2023 Jul;99:102778 [FREE Full text] [doi: 10.1016/j.wsif.2023.102778] [Medline: 37332898]
16. Yang G. Narrative agency in hashtag activism: the case of #BlackLivesMatter. Media Commun 2016 Aug 11;4(4):13-17. [doi: 10.17645/mac.v4i4.692]
17. Ramírez Plascencia D, Carvalho Gurgel B, Plaw A. The Politics of Technology in Latin America. Volume 2. London, UK: Routledge; 2022.

XSL•FO
RenderX

18. Núñez Puente S, D'Antonio Maceiras S, Romero DF. Twitter activism and ethical witnessing: possibilities and challenges of feminist politics against gender-based violence. Soc Sci Comput Rev 2019 Jul 24;39(2):295-311 [FREE Full text] [doi: 10.1177/0894439319864898]
19. Kermani H, Hooman N. Hashtag feminism in a blocked context: the mechanisms of unfolding and disrupting #rape on Persian Twitter. New Media Soc 2022 Oct 19;26(8):4750-4784 [FREE Full text] [doi: 10.1177/14614448221128827]
20. Chen G, Pain P, Barner B. “Hashtag feminism”: activism or slacktivism? In: Harp D, Loke J, Bachmann I, editors. Feminist Approaches to Media Theory and Research. Cham, Switzerland: Springer; 2018:197-218.
21. Li M, Turki N, Izaguirre CR, DeMahy C, Thibodeaux BL, Gage T. Twitter as a tool for social movement: an analysis of feminist activism on social media communities. J Community Psychol 2021 Apr;49(3):854-868. [doi: 10.1002/jcop.22324] [Medline: 32032443]
22. Xiong Y, Cho M, Boatwright B. Hashtag activism and message frames among social movement organizations: Semantic network analysis and thematic analysis of Twitter during the #MeToo movement. Public Relat Rev 2019 Mar;45(1):10-23. [doi: 10.1016/j.pubrev.2018.10.014]
23. Nau C, Quan-Haase A, McLaughlin R. Women-focused nonprofit organizations and their use of twitter during the COVID-19 pandemic: characterizing a gendered pandemic through information, community, and action. Soc Media Soc 2023;9(1):20563051221146489 [FREE Full text] [doi: 10.1177/20563051221146489] [Medline: 36686064]
24. Matos C. NGOs and advocacy communications on sexual and reproductive health and rights: from the North to the South. Fem. Media Stud 2020 Nov 10;22(2):183-204. [doi: 10.1080/14680777.2020.1841813]
25. Bermúdez G, Bonino A. Ni una más, ni una menos, manifestaciones de mujeres como fuente del derecho. Inventio 2017;13(29):12 [FREE Full text]
26. #NoEstásSola Material Informativo Elaborado por ONU Mujeres México. ONU Mujeres México. 2020. URL: https://mexico.unwomen.org/sites/default/files/2022-12/onu%20mujeres%20mexico%20folleto%20NoEstasSola%20201222.pdf [accessed 2024-06-01]
27. #NoEstásSola: Contra la violencia hacia las mujeres en el hogar. Spotlight Intiative. URL: https://www.spotlightinitiative.org/es/news/noestassola-contra-la-violencia-hacia-las-mujeres-en-el-hogar [accessed 2024-07-01]
28. Navarro C, Martínez García L, Vialás Fernández S. Usos y valoración de Facebook como herramienta para la lucha contra la violencia de género en México. Teknokultura 2018 Sep 25;15(2):285-299. [doi: 10.5209/tekn.59546]
29. Lovejoy K, Saxton GD. Information, community, and action: how nonprofit organizations use social media. J Comput Mediat Commun 2012;17(3):353. [doi: 10.1111/j.1083-6101.2012.01576.x]
30. En la mira: 16 Días de activismo contra la violencia de género. ONU Mujeres. URL: https://lac.unwomen.org/es/stories/noticia/2022/11/en-la-mira-16-dias-de-activismo-contra-la-violencia-de-genero-2022 [accessed 2023-01-20]
31. 16 days of activism against gender-based violence. UN Women. URL: https://www.unwomen.org/en/what-we-do/ending-violence-against-women/unite/16-days-of-activism [accessed 2023-01-20]
32. 16 days of activism against gender-based violence. United Nations Population Fund. URL: https://www.unfpa.org/16days [accessed 2023-01-20]
33. Encuesta Nacional Sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2021. INEGI. URL: https://www.ift.org.mx/sites/default/files/comunicacion-y-medios/comunicados-ift/comunicadoendutih_1.pdf [accessed 2024-06-30]
34. Countries with most Twitter users 2022. Statista. URL: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ [accessed 2023-01-30]
35. Rivera Rivera L, Séris Martínez M, Reynales Shigematsu LM, Gómez García JA, Austria Corrales F, Toledano-Toledano F, et al. Violence against women during the COVID-19 pandemic in Mexico. Healthcare (Basel) 2023 Feb 01;11(3):419 [FREE Full text] [doi: 10.3390/healthcare11030419] [Medline: 36766994]
36. de la Garza Montemayor DJ, Ibáñez D, Brosig Rodríguez ME. Digital habits of users in the post-pandemic context: a study on the transition of Mexican internet and media users from the Monterrey metropolitan area. Societies 2023 Mar 16;13(3):72. [doi: 10.3390/soc13030072]
37. Lara MA, Navarrete L, Medina E, Patiño P, Tiburcio M. Impact of Facebook on social support and emotional wellbeing in perinatal women during three waves of the COVID-19 pandemic in Mexico: a descriptive qualitative study. Int J Environ Res Public Health 2023 Jan 30;20(3):2472 [FREE Full text] [doi: 10.3390/ijerph20032472] [Medline: 36767838]
38. Bogen KW, Bhuptani PH, Haikalis M, Orchowski LM. A qualitative analysis of how individuals utilized the Twitter hashtags #NotOkay and #MeToo to comment on the perpetration of interpersonal violence. Soc Media Soc 2022 Mar 22;8(1):20563051221086229. [doi: 10.1177/20563051221086229]
39. Bogen KW, Bleiweiss KK, Leach NR, Orchowski LM. #MeToo: disclosure and response to sexual victimization on Twitter. J Interpers Violence 2021 Sep;36(17-18):8257-8288. [doi: 10.1177/0886260519851211] [Medline: 31117851]
40. Bogen KW, Bleiweiss K, Orchowski LM. Sexual violence is #NotOkay: social reactions to disclosures of sexual victimization on Twitter. Psychol Violence 2019 Jan;9(1):127-137. [doi: 10.1037/vio0000192]
41. Kräh LM. Social media and the Mexican feminist movement: leading or impeding the fight against gender-based violence? University of Twente. URL: https://essay.utwente.nl/85806/1/Kraeh_BA_BMS_2.pdf [accessed 2024-04-29]

XSL•FO
RenderX

42. Violence against women & girls -- resource guide. World Bank. URL: https://www.worldbank.org/en/programs/violence-against-women-and-girls/terminology [accessed 2025-02-21]
43. Boeff DF. Agenda 2030 en México. Pandemia de COVID-19 asevera desigualdades de género. Int Soc Sci Rev 2022 Aug 04;11(1):25-45. [doi: 10.37467/gkarevsocial.v11.3285]
44. Peeters S, Hagen S. The 4CAT capture and analysis toolkit: a modular tool for transparent and traceable social media research. Comput Commun Res 2022;4(2):571-589. [doi: 10.5117/ccr2022.2.007.hage]
45. Chen J. Twitter metrics: how & why you should track them. sproutsocial. 2021 Sep 14. URL: https://sproutsocial.com/insights/twitter-metrics/ [accessed 2023-02-17]
46. Charbonneau E, Mellouli S, Chouikh A, Couture L, Desroches S. The information sharing behaviors of dietitians and Twitter users in the nutrition and COVID-19 infodemic: content analysis study of tweets. JMIR Infodemiology 2022;2(2):e38573 [FREE Full text] [doi: 10.2196/38573] [Medline: 36188421]
47. Ripberger JT, Jenkins-Smith HC, Silva CL, Carlson DE, Henderson M. Social media and severe weather: do tweets provide a valid indicator of public attention to severe weather risk communication? Weather Clim Soc 2014;6(4):520-530. [doi: 10.1175/wcas-d-13-00028.1]
48. Silver A, Andrey J. Public attention to extreme weather as reflected by social media activity. J Contingencies Crisis Manag 2019 May 03;27(4):346-358. [doi: 10.1111/1468-5973.12265]
49. Tahamtan I, Potnis D, Mohammadi E, Miller L, Singh V. Framing of and attention to COVID-19 on Twitter: thematic analysis of hashtags. J Med Internet Res 2021 Sep 10;23(9):e30800 [FREE Full text] [doi: 10.2196/30800] [Medline: 34406961]
50. Viguria I, Alvarez-Mon MA, Llavero-Valero M, Asunsolo Del Barco A, Ortuño F, Alvarez-Mon M. Eating disorder awareness campaigns: thematic and quantitative analysis using Twitter. J Med Internet Res 2020 Jul 14;22(7):e17626 [FREE Full text] [doi: 10.2196/17626] [Medline: 32673225]
51. Sharma G. Deaths prompt Nepal to offer cash to women who shun 'menstrual huts'. Reuters. URL: https://www.reuters.com/article/us-nepal-women-menstruation/deaths-prompt-nepal-to-offer-cash-to-women-who-shun-menstrual-huts-idUSKBN1Y71VG [accessed 2020-11-10]
52. Chan B, Lopez A, Sarkar U. The canary in the coal mine tweets: social media reveals public perceptions of non-medical use of opioids. PLoS One 2015;10(8):e0135072 [FREE Full text] [doi: 10.1371/journal.pone.0135072] [Medline: 26252774]
53. Smith HJ, Chen J, Liu X. Language and rigour in qualitative research: problems and principles in analyzing data collected in Mandarin. BMC Med Res Methodol 2008 Jul 10;8(1):44 [FREE Full text] [doi: 10.1186/1471-2288-8-44] [Medline: 18616812]
54. Rufai SR, Bunce C. World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. J Public Health (Oxf) 2020 Aug 18;42(3):510-516 [FREE Full text] [doi: 10.1093/pubmed/fdaa049] [Medline: 32309854]
55. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
56. Bacsu JD, O'Connell ME, Cammer A, Ahmadi S, Berger C, Azizi M, et al. Examining the impact of COVID-19 on people with dementia from the perspective of family and friends: thematic analysis of tweets. JMIR Aging 2022 Jun 27;5(2):e38363 [FREE Full text] [doi: 10.2196/38363] [Medline: 35667087]
57. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023 Nov 08;22:16094069231205789. [doi: 10.1177/16094069231205789]
58. Supthanasup A, Banwell C, Kelly M, Yiengprugsawan VS, Davis JL. Child feeding practices and concerns: thematic content analysis of Thai virtual communities. Matern Child Nutr 2021 Apr;17(2):e13095 [FREE Full text] [doi: 10.1111/mcn.13095] [Medline: 33258296]
59. Dehghani A, Kojuri J, Dehghani MR, Keshavarzi A, Najafipour S. Experiences of students and faculty members about using virtual social networks in education: a qualitative content analysis. J Adv Med Educ Prof 2019 Apr;7(2):86-94 [FREE Full text] [doi: 10.30476/JAMP.2019.44712] [Medline: 31086800]
60. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
61. Rodríguez AP, González-García R. La incidencia de la acción colectiva feminista virtual en las respuestas del Estado a la violencia de género en México. Sci Bull Soc Sci Humanit 2018 Dec 05;7(13):33. [doi: 10.29057/icshu.v7i13.3490]
62. Di Napoli PN. Jóvenes, activismos feministas y violencia de género en la Unam: genealogía de un conflicto. Rev Latinoam Cienc Soc Niñez Juv 2021 Apr 29;19(2):1-27. [doi: 10.11600/rlcsnj.19.2.4567]
63. Violencia digital contra las mujeres y las niñas. ONU Mujeres. URL: https://mexico.unwomen.org/sites/default/files/2023-03/Brief_ViolenciaDigital.pdf [accessed 2024-05-14]
64. Estadísticas a propósito del Día Internacional de la Eliminación de la Violencia contra la Mujer. INEGI. URL: https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2024/EAP_VCM_24.pdf [accessed 2024-04-29]
65. Castro R, Frías SM. Obstetric violence in Mexico: results from a 2016 national household survey. Violence Against Women 2020 May;26(6-7):555-572. [doi: 10.1177/1077801219836732] [Medline: 30957706]
66. Arévalo A. 25N | economic violence, a silent aggression. Friedrich Naumann Found Freedom. URL: https://www.freiheit.org/mexico/25n-economic-violence-silent-aggression [accessed 2024-09-29]

XSL•FO
RenderX

67. Aguilar Peláez E. Vicarious violence: the insidious crime that affects thousands of families in Mexico. Global Press Journal. URL: https://globalpressjournal.com/americas/mexico/vicarious-violence-insidious-crime-affects-thousands-families-mexico/ [accessed 2024-09-29]
68. Mitchell A, Gutmann-Gonzalez A, Brindis CD, Decker MJ. Contraceptive access experiences and perspectives of Mexican-origin youth: a binational qualitative study. Sex Reprod Health Matters 2023 Dec;31(1):2216527 [FREE Full text] [doi: 10.1080/26410397.2023.2216527] [Medline: 37335382]
69. Villalobos A, Estrada F, Hubert C, Torres-Ibarra L, Rodríguez A, Romero I, et al. Sexual and reproductive health among adolescents in vulnerable contexts in Mexico: needs, knowledge, and rights. PLOS Glob Public Health 2023;3(11):e0002396 [FREE Full text] [doi: 10.1371/journal.pgph.0002396] [Medline: 37910453]
70. Rodríguez Pérez RE, Castro Lugo D, Ibarra Gutiérrez S. Violencia laboral de género como factor que inhibe la participación de las mujeres en el mercado de trabajo de México. Rev Econ 2022 Aug 25;39(99):110-133. [doi: 10.33937/reveco.2022.277]
71. Ortiz-Barreda G, Vives-Cases C. Violencia contra las mujeres: el papel del sector salud en la legislación internacional. Gac Sanit 2012 Sep;26(5):483-489 [FREE Full text] [doi: 10.1016/j.gaceta.2011.09.027] [Medline: 22265648]
72. Herrera C, Rajsbaum A, Agoff C, Franco A. Entre la negación y la impotencia: prestadores de servicios de salud ante la violencia contra las mujeres en México. Salud Publica Mex 2006;48 Suppl 2:S259-S267. [doi: 10.1590/s0036-36342006000800006] [Medline: 16884164]
73. Fuentes Alcalá ML, Hernández Engrandes C, Alcay Méndez SG. Situations of vulnerability to human trafficking in mobile and resident populations in the municipality of Tapachula in the context of the 2018-2019 migrant caravans. Frontera Norte 2022 Jan 01;34:1-23. [doi: 10.33679/rfn.v1i1.2191]
74. Torre Cantalapiedra E. El estudio de las caravanas migrantes en México. Norteamérica 2022 May 27;17(2):30. [doi: 10.22201/cisan.24487228e.2022.2.525]
75. Torres Beltrán XK, Tena Guerrero O. Violencia estructural e institucional hacia mujeres rurales mixtecas: el caso del Programa de Inclusión Social PROSPERA en el estado de Guerrero, México. Culturales 2017 Jan 01;7(1):1-30. [doi: 10.22234/recu.20190701.e392]
76. Raby D. Nuestra Madre sufre y llora. Violencia intrafamiliar y Buen Vivir desde la relación con el maíz en una comunidad nahua de Guerrero, México. Relac Estud Hist Soc 2018 Aug 17;39(155):43-75. [doi: 10.24901/rehs.v39i155.288]
77. Principales Resultados (Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH). INEGI. 2017. URL: https://www.inegi.org.mx/contenidos/programas/endireh/2016/doc/endireh2016_presentacion_ejecutiva.pdf [accessed 2024-04-29]
78. Uysal J, Stockman JK, Miller E, Rocha-Jimenez T, Rangel GM, Mercado AP, et al. "At least I didn't get raped": a qualitative exploration of IPV and reproductive coercion among adolescent girls seeking family planning in Mexico. J Interpers Violence 2022 Apr;37(7-8):NP4740-NP4761 [FREE Full text] [doi: 10.1177/0886260520959571] [Medline: 33183147]
79. Torres M, O'Connor E. Framing the final issue on women's movements and the shape of feminist theory and praxis in Latin America. J Int Women Stud 2019;20(8):1-4 [FREE Full text]
80. Silva M, Anaba U, Jani Tulsani N, Sripad P, Walker J, Aisiri A. Gender-based violence narratives in internet-based conversations in Nigeria: social listening study. J Med Internet Res 2023 Sep 15;25:e46814. [doi: 10.2196/46814] [Medline: 37713260]
81. Xue J, Zhang B, Zhang Q, Hu R, Jiang J, Liu N, et al. Using Twitter-based data for sexual violence research: scoping review. J Med Internet Res 2023 May 15;25:e46084 [FREE Full text] [doi: 10.2196/46084] [Medline: 37184899]
82. Rovira-Sancho G. Activism and affective labor for digital direct action: the Mexican #MeToo campaign. Soc Mov Stud 2021 Dec 06;22(2):145-162. [doi: 10.1080/14742837.2021.2010530]
83. Essue BM, Chadambuka C, Perez-Brumer A, Arruda-Caycho I, Tocallino D, Balasa R, et al. Women's experiences of gender-based violence supports through an intersectional lens: a global scoping review. BMJ Public Health 2025 Jan 1;3(1):e001405-e001441. [doi: 10.1136/bmjph-2024-001405] [Medline: 40017981]
84. Yodanis CL. Gender inequality, violence against women, and fear: a cross-national test of the feminist theory of violence against women. J Interpers Violence 2004 Jun;19(6):655-675. [doi: 10.1177/0886260504263868] [Medline: 15140317]
85. Sharma AE, Mann Z, Cherian R, Del Rosario JB, Yang J, Sarkar U. Recommendations from the Twitter hashtag #DoctorsAreDickheads: qualitative analysis. J Med Internet Res 2020 Oct 28;22(10):e17595 [FREE Full text] [doi: 10.2196/17595] [Medline: 33112246]
86. Pfeffer J, Matter D, Sargsyan A. The half-life of a Tweet. Proc Int AAAI Conf Web Soc Media 2023 Jun 02;17:1163-1167. [doi: 10.1609/icwsm.v17i1.22228]

Abbreviations

GBV: gender-based violence
NGO: nongovernmental organization

XSL•FO
RenderX

Edited by I Brooks; submitted 09.10.24; peer-reviewed by C Nau, AA Villa-Rueda; comments to author 16.02.25; revised version received 03.03.25; accepted 20.03.25; published 17.04.25.

Please cite as:
Marian M, Pérez RL, Reed E, Hurst S, Lundgren R, McClain AC, Barker KM
Exploring the Use of Social Media for Activism by Mexican Nongovernmental Organizations Using Posts From the 16 Days of Activism Against Gender-Based Violence Campaign: Thematic Content Analysis
JMIR Infodemiology 2025;5:e67368
URL: https://infodemiology.jmir.org/2025/1/e67368
doi:10.2196/67368
PMID:

©Marian Marian, Ramona L Pérez, Elizabeth Reed, Samantha Hurst, Rebecka Lundgren, Amanda C McClain, Kathryn M Barker. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 17.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring Topics, Emotions, and Sentiments in Health Organization Posts and Public Responses on Instagram: Content Analysis

Abigail Paradise Vit[1], PhD; Avi Magid[2,3], PhD

[1]Department of Information Systems, The Max Stern Emek Yezreel College, Jezreel Valley Regional Council, Israel

[2]Management, Rambam Healthcare Campus, Haifa, Israel

[3]Department of International Health, Maastricht University, Maastricht, The Netherlands

Corresponding Author:
Abigail Paradise Vit, PhD
Department of Information Systems
The Max Stern Emek Yezreel College
D.N. Emek Yezreel
Jezreel Valley Regional Council, 1930600
Israel
Phone: 972 509903930
Email: abigailp@yvc.ac.il

Abstract

Background: Social media is a vital tool for health organizations, enabling them to share evidence-based information, educate the public, correct misinformation, and support a more informed and healthier society.

Objective: This study aimed to categorize health organizations' content on social media into topics; examine public engagement, sentiment, and emotional responses to these topics; and identify gaps in fear between health organizations' messages and the public response.

Methods: Real data were collected from the official Instagram accounts of health organizations worldwide. The BERTopic algorithm for topic modeling was used to categorize health organizations' posts into distinct topics. For each identified topic, we analyzed the engagement metrics (number of comments and likes) of posts categorized under the same topic, calculating the average engagement received. We examined the sentiment and emotional content of both posts and responses within the same topic, providing insights into the distributions of sentiment and emotions for each topic. Special attention was given to identifying emotions, such as fear, expressed in the posts and responses. In addition, a linguistic analysis and an analysis of sentiments and emotions over time were conducted.

Results: A total of 6082 posts and 82,982 comments were collected from the official Instagram accounts of 8 health organizations. The study revealed that topics related to COVID-19, vaccines, and humanitarian crises (such as the Ukraine conflict and the war in Gaza) generated the highest engagement. Our sentiment analysis of the responses to health organizations' posts showed that topics related to vaccines and monkeypox generated the highest percentage of negative responses. Fear was the dominant emotion expressed in the posts' text, while the public's responses showed more varied emotions, with anger notably high in discussions around vaccines. Gaps were observed between the level of fear conveyed in posts published by health organizations and in the fear conveyed in the public's responses to such posts, especially regarding mask wearing during COVID-19 and the influenza vaccine.

Conclusions: This study underscores the importance of transparent communication that considers the emotional and sentiment-driven responses of the public on social media, particularly regarding vaccines. Understanding the psychological and social dynamics associated with public interaction with health information online can help health organizations achieve public health goals, fostering trust, countering misinformation, and promoting informed health behavior.

(JMIR Infodemiology 2025;5:e70576) doi:10.2196/70576

XSL•FO
RenderX

KEYWORDS

emotion analysis; fear; health communication; health care; Instagram; official health organizations; sentiment analysis; social media; vaccines

Introduction

Background

Social networks are used by billions of people around the world, making them an effective platform for reaching a wide range of audiences. Health organizations, such as the World Health Organization (WHO), use social networks to disseminate important health-related information [1-5]; provide real-time updates, news, and emergency guidelines [6]; and promote awareness of diseases [7] and mental health [8]. Health organizations also promote vaccination compliance by disseminating information regarding vaccination importance, safety, and disease severity [9].

In contrast to the content published by social network users unaffiliated with health organizations, who can spread a large amount of inaccurate, false, and misleading information about health-related issues [10,11], the health-related content published on social networks by official health organizations enables the public to have access to reliable and useful information [12-14]. Therefore, health organizations are important actors in social media as reliable sources, providing evidence-based and authoritative information [15]. Through their efforts, these organizations educate the public, dispel myths, and help the community become healthier and more informed [16].

In addition to serving as a means of disseminating health-related information, social networks provide the public with the opportunity to participate in discussions and conversations with health organizations [17] by posting comments, sharing posts, and liking posts. Members of the public can also ask questions about the health issue being discussed, and the organizations that disseminate the information can respond. Active public participation can enhance individuals' understanding of health-related content [18,19] and can foster public trust in and appreciation for science [20]. In addition, health organizations are able to pinpoint concerns related to specific health topics, gain insight into public opinion [17-25], and identify topics that result in misinformation [26].

However, health-related messages disseminated by health organizations can provoke negative public reactions [27]. When combined with contradictory messages from unaffiliated social media users, this can lead to undesired health behaviors, such as vaccine hesitancy, noncompliance with health directives, and diminished trust in the reliability of health organizations [28-30].

Understanding public emotions in response to information disseminated by health organizations on social media is important for assessing the effectiveness of health communication strategies [31]. Sentiment and emotion analysis are widely used for determining sentiment polarity and detecting specific emotions expressed in textual data [32]. Sentiment analysis categorizes a text as positive, negative, or neutral, while emotion analysis identifies specific emotions expressed within a text [32]. Fundamental emotions, such as happiness, sadness, anger, disgust, surprise, and fear, can be detected, along with more nuanced emotions such as confusion and trust [33,34]. Among these emotions, fear is an important factor in health communication, as it can influence public perception, engagement, and behavioral responses [35].

Various theoretical models, such as the extended parallel process model [36], have explored how fear is used in public health messaging to encourage protective behaviors. This model suggests that individuals respond to fear-based messages depending on their perceived threat level and their sense of efficacy in managing the threat. According to this model, when perceived risk is high and the message also offers a clear solution, individuals are more likely to engage positively and adopt protective behaviors. However, if either condition is not met, responses may be negative, leading to outcomes such as avoidance or denial [36]. Therefore, fear is a crucial factor to consider in public health communication.

Despite the need to examine how the public emotionally responds to information shared by health organizations on social media, studies examining the topics communicated by these organizations and the corresponding responses from the public are sparse. The aim of this study is to analyze and characterize the content disseminated by health organizations on social media into topics, as well as social media users' responses to this content and the engagement, sentiment, and emotions induced by this content. In addition, we aim to identify gaps in fear between health organizations' messages and public responses.

Related Work

In this section, we provide an overview of various studies associated with health-related content and the topic modeling and sentiment analysis of such content.

Sentiment Analysis of Health-Related Social Media Content

Many studies examined the sentiment and public opinion surrounding the COVID-19 vaccine on Twitter [21,23,24,37-53]. For example, an analysis of Twitter data was conducted by Niu et al [24] to examine public opinion and sentiment before and during the administration of the COVID-19 vaccines in Japan. They found that negative sentiment toward the vaccines dominated positive sentiment in Japan, and concerns about side effects may have outweighed fears of infection at the beginning of the vaccination process.

Numerous studies leveraged machine learning techniques to classify tweets as positive, negative, or neutral sentiment toward vaccines, enabling the identification of vaccine hesitancy among communities and social media users [35,54-68]. Most of these studies collected data from Twitter using keywords or hashtags related to vaccinations. Chakraborty et al [56] used deep learning to analyze 226,668 COVID-19 tweets from December 2019 to May 2020, achieving 81% accuracy. Most tweets showed positive or neutral sentiment, while highly retweeted posts were predominantly neutral or negative.

XSL•FO
RenderX

Topic Modeling for Health-Related Social Media Content

Several studies have used topic modeling to examine health-related discussions on social media, focusing on topics such as blood donation [69], cancer-related content [70], and vaccine-related conversations [71]. Paul and Dredze [72] proposed the ailment topic aspect model to identify health topics on Twitter. Analyzing 144 million tweets, they identified 13 topics linked to seasonal influenza, allergies, temporal surveillance, and obesity-related geographic data in the United States.

Seltzer et al [73] analyzed 500 Instagram (Meta Platforms, Inc) images tagged #zika from May to August 2016, analyzing them by sentiment, content, and engagement. A total of 299 images were related to health, while 193 focused on topics of public interest. Sentiments and emotion analysis revealed that fear and negative emotions were linked to Zika transmission and response uncertainty. The study highlighted Instagram's value in understanding public sentiment and addressing gaps in health communication. Furthermore, Muralidhara and Paul [74] analyzed 96,426 Instagram posts collected between September and October 2016, using 269 health-related hashtags. Polylingual topic modeling approach was used to identify 47 health-related topics spanning 10 broad categories: acute illness, alternative medicine, chronic illness and pain, diet, exercise, health care and medicine, mental health, musculoskeletal health and dermatology, sleep, and substance abuse. Kim et al [75] analyzed 96,302 Instagram photos and 513,694 comments with antivaccination hashtags, focusing on photo features, engagement, and sentiment. Most photos (52.24%) were categorized as "text." "Food" and "plant" photos received the most positive comments, while "text" photos, despite high engagement, received fewer positive responses.

Other studies focused on dividing vaccine content on social networks into topics [40,41,43,45,52,53,76]. The study by Kwok et al [76] examined tweets of Australian users regarding COVID-19 vaccination on Twitter. Using a latent Dirichlet allocation topic model, they identified 3 commonly discussed topics: attitudes toward COVID-19 and vaccination, advocacy for infection control measures against COVID-19, and misconceptions and complaints regarding COVID-19. Similarly, Lyu et al [40] used latent Dirichlet allocation to analyze COVID-19 vaccine discussions on Twitter, identifying 16 topics grouped into 5 themes. Vaccination opinions were the most discussed topic. Emotion analysis showed trust as the dominant emotion, followed by anticipation, fear, and sadness. In addition, Chandrasekaran et al [43] used the correlation explanation topic modeling algorithm to examine COVID-19 vaccine–related tweets. The authors identified 16 topics in the COVID-19 vaccination tweets, which were grouped into 6 broader themes. Most tweets regarding COVID-19 vaccination centered on vaccine policy, vaccine hesitancy, and postvaccination symptoms and side effects.

Analysis of Health Care Providers' Content and Public Responses on Social Media

Several studies have examined content published on social networks by health care providers. Among them, Kim and Kim [77] analyzed 1545 Instagram photos published by the US Centers for Disease Control and Prevention (CDC) and public comments using Microsoft Azure Cognitive Services. Their findings showed that most images featured text or people, but those with larger faces or flashy elements tended to receive less engagement. Happiness and neutral emotions in comments were negatively correlated with interaction levels. Pinto et al [78] analyzed 632 Instagram posts from Portugal's National Health Service (NHS) and Brazil's Ministry of Health (MH) in 2019, mapping 53 topics for the NHS and 63 for the MH. The NHS emphasized healthy eating and blood donation, while the MH focused on vaccination campaigns, dengue prevention, and HIV awareness.

Mello et al [79] analyzed 726 Instagram posts from the WHO and CDC in 2020 to explore how these organizations communicated COVID-19 risks. Their study focused on messaging related to threat and efficacy, cues to action, and indicators of credibility. According to the findings, efficacy messages, such as those promoting preventive behaviors, were more prevalent, while threat messages addressing the susceptibility and severity of COVID-19 were less common. The study concluded that improving credibility cues, using compelling visuals, tailoring content for diverse audiences, and leveraging Instagram's interactive features could enhance public health communication, boosting engagement, trust, and impact.

Vaghefi et al [80] analyzed health care providers' messages on Twitter from May 2018 to May 2019 using machine learning, including Bidirectional Encoder Representations from Transformers (BERT)–based models, to classify tweets as professional communications or health-related information, further categorizing them as fear based or hope based. Results showed that fear-based messages were widely shared by the public but were less effective at motivating constructive health actions, while hope-based messages resonated more with health care providers. While this study [80] examined health care providers' content and the fear evoked by their messages on Twitter, our study focuses on categorizing health care providers' posts into distinct topics; analyzing engagement metrics, sentiment, and emotional content within each topic; and identifying the gaps between the fear expressed in messages from health organizations and the fear observed in public responses to such content. In contrast, Vaghefi et al [80] did not analyze the public comments to measure fear in relation to health care providers' messages; instead, they focused on public interactions, such as retweets and replies, to study information propagation.

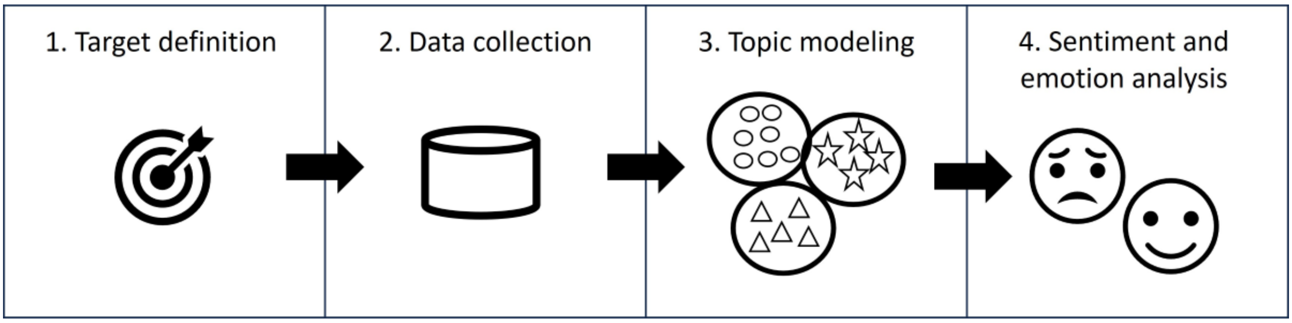
Methods

Overview

This section outlines the proposed methodology, which consists of four main phases, as illustrated in Figure 1: (1) defining the targeted health organizations, (2) gathering data from the selected health organization accounts, (3) performing topic modeling on health organization content, and (4) analyzing the sentiment and emotion associated with each topic and the public's response to it.

XSL•FO
RenderX

Figure 1. An overview of the methodology's main phases.



Detailed descriptions of each phase are provided in the subsections that follow.

Target Definition

The first phase of our methodology was to define the group of health organizations to be examined. These organizations were selected based on their status as official government entities or government-supported health agencies responsible for public health policies, research, and regulations, ensuring that the study focused on authoritative sources [81-88]. In addition, a minimum threshold of 10,000 followers was set for health organization accounts to ensure sufficient activity and public engagement.

A brief description of each organization is provided in Textbox 1.

Textbox 1. Description of the health organizations included in this study.

Description

1. The World Health Organization (WHO): the WHO is the United Nations agency tasked with connecting nations around the world to promote health, keep the world safe, and serve populations considered vulnerable [81]. The organization's official Instagram account is named WHO.
2. Department of Health and Human Services (HHS): the HHS aims to enhance the health and well-being of the residents or citizens of the United States by providing effective health and human services and fostering sound, sustained advances in the sciences underlying medicine, public health, and social services [82]. HHSgov is the official Instagram account of the HHS.
3. The Office of Minority Health (OMH): the OMH is part of the HHS dedicated to improving the health of racial and ethnic minority groups. The OMH fulfills its commitment to improving the health of racial and ethnic minority groups in large part by developing health policies and programs that help eliminate health disparities [83]. MINORITYHEALTH is the official Instagram account of the OMH.
4. National Institutes of Health (NIH): the NIH is the leading federal agency in the United States responsible for conducting and supporting medical research. The NIH, which is part of the HHS, is one of the world's most prominent centers for medical research. Its mission is to enhance human health by advancing research across a wide range of scientific disciplines [84]. The organization's official Instagram account is named NIHgov.
5. The National Institute of Mental Health (NIMH): the NIMH is the US agency responsible for research on mental health. Its primary objective is to understand, treat, and prevent mental illness by conducting basic and clinical research. The NIMH is one of the 27 institutes and centers that comprise the NIH, which is part of the HHS [85]. NIMHgov is the official Instagram account of the NIMH.
6. The US Centers for Disease Control and Prevention (CDC): the CDC is a science-based, data-driven organization that leads the United States' efforts to protect the public's health. The CDC is one of the major components of the HHS, and it aims to protect the residents of the United States from health, safety, and security threats, both foreign and domestic [86]. CDCgov is the official Instagram account of the CDC.
7. The UK National Health Service (NHS): the NHS was established as the public health care system of the United Kingdom. It is one of the largest and most comprehensive health care systems in the world. The NHS provides various types of services, including mental health services, general practitioners, hospitals, and treatment facilities. The NHS is dedicated to improving the overall health of the population by promoting public health initiatives, health education, vaccination programs, and disease prevention campaigns [87]. The organization's official Instagram account is named NHS.
8. The US Food and Drug Administration (FDA): the FDA is a governmental regulatory agency responsible for protecting public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices [88]. The organization's official Instagram account is named FDA.

It is important to note that, because we selected health organizations with a high number of followers, the resulting sample skewed toward US-based organizations. Most accounts originated from the United States, with 1 from the United Kingdom and 1 global organization (WHO). This concentration may affect the generalizability of the findings, as both the published content and the public responses are likely influenced by US-specific health priorities and cultural context.

Data Collection

We chose Instagram as the social media platform for its visual and interactive nature, which allows health organizations to share information and engage with the public effectively [89]. We searched for the official Instagram accounts of health care organizations using Instagram's search box. Table 1 provides the name of each official Instagram account for the organizations, along with the total number of posts published since the account's creation and the number of followers, as recorded in August 2024. As can be seen in Table 1, the CDC,

XSL•FO
RenderX

WHO, and Office of Minority Health are the organizations that publish the most posts. The WHO's account has the largest number of followers.

To collect the health organizations' Instagram posts, we connected with the Instagram application programming interface (API) using the RapidAPI website. This website is a large API hub that allows to connect with tens of thousands of public Representational State Transfer APIs over the internet.

The posts were collected between April 7, 2017, and November 17, 2023. Each Instagram post included the publication date, the number of likes, the number of comments, text, and a photo. For each post, we collected the comments, including the publication date, the comment text, and the number of likes. A total of 6082 posts and 82,982 comments were collected using the Instagram API. All retrieved posts and comments were included in the analysis. Table 2 presents the relevant statistics for the collected posts and comments for each health organization's Instagram account. We analyzed all the posts that the API allowed us to retrieve, without applying selection criteria or filtering specific posts. Comments were collected only from the original posts, excluding replies to comments.

Table 1. Statistics for the health organizations' Instagram accounts, ranked by number of followers (highest to lowest).

Organization name	Instagram account name	Published posts, n	Followers, n
Centers for Disease Control	CDCGOV	6359	2.5 million
World Health Organization	WHO	3893	12 million
National Health Service	NHS	695	564,000
Department of Health and Human Services	HHSGOV	3269	202,000
Food and Drug Administration	FDA	801	122,000
National Institute of Mental Health	NIMHGOV	674	64,100
National Institutes of Health	NIHGOV	1867	279,000
Office of Minority Health	MINORITYHEALTH	3909	15,400

Table 2. Statistics for the collected posts and comments of each health organization's Instagram account.

Name of health organization	Collected posts, n	Collected comments, n	Number of likes, mean (SD)
Centers for Disease Control	4298	58,471	2725.341 (3870.086)
World Health Organization	527	15,448	18349.774 (28197.828)
National Health Service	277	3826	1443.018 (1386.701)
Department of Health and Human Services	116	896	207.836 (270.549)
Food and Drug Administration	85	565	203.259 (183.758)
National Institute of Mental Health	301	1096	185.920 (142.871)
National Institutes of Health	253	1987	723.447 (709.083)
Office of Minority Health	225	603	34.124 (31.448)

Ethical Considerations

The data collection process and analysis were approved by the Emek Yezreel College Ethical Review Board (2023-81 YVC EMEK).

As the research relied solely on publicly available social media data and did not involve direct interaction with individuals, informed consent was not applicable. No compensation was offered or provided, as the study did not involve direct participation of human participants.

No identifiable private user information was collected or analyzed. All data used in the analysis were publicly available and did not contain personally identifiable information.

Topic Modeling

Text from the posts published by health organization accounts was analyzed using BERTopic [90]. Note that we removed the names of the health organizations and their abbreviations from the text to avoid creating topics around each organization. Stop words were also removed from the text.

We used the C_V coherence score to evaluate the quality and interpretability of the topics produced by the model. This measure combines the indirect cosine measure with normalized pointwise mutual information and a Boolean sliding window [91]. The coherence score indicates how closely related and coherent the words are in a topic. Using the average coherence metric, which is the average of the coherence metrics within each topic, we measured the model's ability to generate coherent and meaningful topics. This score enabled us to evaluate the overall effectiveness of the model in producing topics with a high degree of semantic similarity.

We performed 6 steps in BERTopic to analyze the posts (Textbox 2).

XSL•FO
RenderX

Textbox 2. Six-step BERTopic analysis of posts.

1. Embedding tweets: in this step, the text in the posts was converted into numerical representations using a sentence-transformers model named All-MiniLM-L6-v2 [92].
2. Dimensionality reduction: uniform manifold approximation and projection [93] was used to reduce the dimensionality of the embedded text, with the following parameters: _neighbors=15, n_components=3, min_dist=0, and metric = 'cosine.'
3. Cluster tweets: the text was grouped into clusters using the hierarchical density-based spatial clustering of applications with noise density-based clustering technique [94], with the following parameters: min_cluster_size=40, metric="Euclidean," and cluster_selection_method='eom.'
4. Word frequency analysis in clusters: the frequency of each word in each cluster was determined at the cluster level.
5. Topic representation: to represent the topics in the Instagram posts and responses, term frequency–inverse document frequency (TF-IDF) was adapted to work on a cluster or topic level instead of a tweet level. A new TF-IDF representation was used called class-based TF-IDF (c-TF-IDF).
6. Outlier reduction: hierarchical density-based spatial clustering of applications with noise identified texts that were outliers, meaning they did not belong to any of the established topics. To address this, we calculated the c-TF-IDF representation for each outlier text and compared its cosine similarity with the c-TF-IDF representations of the existing topics. By associating outlier texts with the closest matching topic based on similarity, we minimized the number of texts classified as outliers.

Sentiment and Emotion Analysis

Having categorized the health organizations' posts into topics, we analyzed the emotions and sentiment of each post and comment associated with a particular topic. Sentiment analysis was conducted using *distilbert - base - multilingual - cased - sentiments - student*, which achieved an average accuracy of 0.808 on the test set [95]. This is a distilled version of a zero-shot classification pipeline trained on the multilingual sentiment dataset. Zero-shot classification is a machine learning technique that allows models to classify data into categories they have never encountered during training without requiring labeled examples to be provided. It accomplishes this by leveraging contextual understanding and semantic relationships between seen and unseen classes, often using embeddings or natural language models [96]. In this case, a larger "teacher" model, *MoritzLaurer/mDeBERTa-v3-base-mnli-xnli*, was used to train a smaller "student" model, *distilbert-base-multilingual-cased*. Using this distillation process, the student model maintains high classification performance while being more efficient and lightweight. According to the training log, the student model achieved an impressive agreement rate of 88.29% with its teacher model.

Emotion analysis of 6 basic emotions (fear, anger, disgust, joy, sadness, and surprise) was conducted using a fine-tuned checkpoint of the DistilRoBERTa-base model called *j-hartmann/emotion-english-distilroberta-base*. The model was trained on 6 diverse datasets [97]. The model was trained on a balanced subset from several datasets of nearly 20,000 observations in total. In total, 80% of this balanced subset was used for training and 20% for evaluation. The evaluation accuracy was 66%.

To ensure the accuracy of the models in identifying sentiments and emotions, we randomly selected 100 posts and comments. We manually classified them based on their dominant sentiment and dominant emotion. The results showed that the sentiment model achieved 92% accuracy, while the emotion model demonstrated 84% accuracy. Considering sentiment models predict binary classification and emotion models face greater complexity due to emotions' multidimensional nature, the results are logical. Therefore, applying the models to the data was expected to provide sufficiently reliable outcomes.

Each post and posts' comments received a sentiment score of positive, negative, or neutral, as well as an emotion score of fear, anger, disgust, joy, sadness, or surprise.

For each topic, we calculated the following:

- Average number of comments (the average number of comments for all posts in the topic).
- Average number of likes (the average number of likes for all posts in the topic).
- Average post sentiment scores (the average positive, negative, and neutral sentiment scores for all posts in the topic).
- Average post emotion scores (the average fear, anger, disgust, joy, sadness, and surprise scores for all posts in the topic).
- Average comment sentiment scores (the average positive, negative, and neutral sentiment scores for all post comments in the topic).
- Average comment emotion scores (average fear, anger, disgust, joy, sadness, and surprise scores were calculated for all post comments in the topic).
- Gap (the difference between the average fear score of the posts and the average fear score of the comments).

Linguistic Analysis

The objective of this analysis was to identify the most significant phrases used by health organizations in posts that resonated more positively with the public (ie, associated with higher positive sentiment in comments) compared to those that elicited more negative reactions (higher negative sentiment in comments).

For this purpose, we calculated the average sentiment score (positive and negative) of all comments related to each post. Posts were then categorized based on the dominant sentiment (positive or negative) derived from these average sentiment scores in comments, resulting in 2 groups: posts with predominantly positive comments and posts with predominantly negative comments.

XSL•FO
RenderX

For each group of posts, we extracted the top 50 most significant phrases, including single words (unigrams), 2-word phrases (bigrams), 3-word phrases (trigrams), and 4-word phrases (four-grams). The preprocessing involved removing hashtags from the text, eliminating stopwords, and calculating term importance using TF-IDF. TF-IDF estimates the significance of terms within a group of posts based on their frequency within each post as compared to the frequency across all posts in the group. The cumulative TF-IDF scores for each term were calculated by summing across all posts in the group, enabling the identification of the most significant phrases for each group of posts.

Time Analysis

In the time analysis section, we examined the average positive and negative sentiment in both comments and posts over the years. In addition, we analyzed the average levels of various emotions, including fear and anger, expressed in both posts and comments throughout the data collection period. Furthermore, we decided to calculate additional emotions of trust, disappointment, and confusion. The purpose of including these emotions was to extend the analysis beyond the 6 basic emotions, providing a broader range of emotional insights.

To calculate trust, we used the *ayoubkirouane/BERT-Emotions-Classifier* model [98], a fine-tuned BERT-based model designed for multilabel emotion classification. This model was trained on the sem_eval_2018_task_1 dataset. This model includes a wide range of emotions, such as anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, and trust.

For confusion and disappointment emotions, we used the *SamLowe/roberta-base-go_emotions* model [99], which is trained on the go_emotions dataset for multilabel classification. The go_emotions dataset, based on Reddit (Reddit, Inc) data, contains 28 emotion labels. The model achieved high performance in identifying confusion (with an accuracy of 0.972) and disappointment (with an accuracy of 0.974).

As part of the temporal analysis, we incorporated weekly COVID-19 mortality data in the United States using a publicly available dataset published by the National Center for Health Statistics [100]. The analysis included calculating a "COVID-19 death ratio," which represents the proportion of US COVID-19–related deaths to the total number of US deaths. We applied a rolling average smoothing technique across a 10-week window to reduce noise and variability in the data, allowing us to identify patterns over time. This approach allowed us to combine the mortality data alongside our sentiment and emotional analyses, providing valuable insights into the alignment between public responses and real-world outcomes.

Results

Topic Modeling

Using BERTopic, 36 topics were identified for the posts of the 8 health organization Instagram accounts. The average coherence score was 0.7374. A health expert reviewed the list of topics and suggested that we combine related topics.

Therefore, the following topics were combined: 7 topics related to COVID-19 were grouped together, 2 topics related to vaccines for children were grouped together, 2 topics related to monkeypox were grouped together, and 2 topics related to booster vaccines were combined. Combining the topics resulted in 25 topics, with an average coherence score of 0.7298. The health expert assigned a representative name to each topic. The topic names were assigned by analyzing the 10 most significant words for each topic, as determined by the c-TF-IDF scores (refer to the Methods section for more information about c-TF-IDF) from the BERTopic algorithm and cross-referencing them with example posts related to that topic. Each topic was automatically assigned a unique number starting from 0, following the default numbering convention used by the BERTopic algorithm. Table 3 lists the topics, indicating their number, name, and number of posts. Refer to Multimedia Appendix 1 for the 5 words with the highest c-TF-IDF for each topic.

XSL•FO
RenderX

Table 3. The number, name, and size of each topic.

Topic number	Topic name	Topic size, n
0	COVID-19	1446
1	Mental health	752
2	Children and vaccines	496
3	COVID-19 booster vaccine	288
4	Pregnant	246
5	Foodborne	235
6	Research	226
7	Community health	223
8	Flu	216
9	Cancer	198
10	Monkeypox	197
11	Cardiovascular diseases	193
12	Public health	183
13	Climate	148
14	Sepsis	141
15	Masks	127
16	Vector	122
17	Health equity	108
18	Antibiotics	108
19	Ebola	92
20	Humanitarian	91
21	Smoking	85
22	RSV[a]	59
23	Sun damage	52
24	Noise damage	50

[a]RSV: respiratory syncytial virus.

Engagement Analysis

Our dataset includes engagement metrics for each post, specifically the number of likes and user comments.

We calculated the average number of comments and likes for all posts in each topic. Multimedia Appendix 2 presents the average number of comments received per topic, while Multimedia Appendix 3 displays the average number of likes. Among the topics, those with the highest user engagement based on the average number of post comments were humanitarian issues, masks, and COVID-19. In terms of engagement measured by the average number of post likes, the leading topics were humanitarian issues, masks, and cancer. Multimedia Appendices 2 and 3 also show increased engagement for topics related to vaccines and COVID-19.

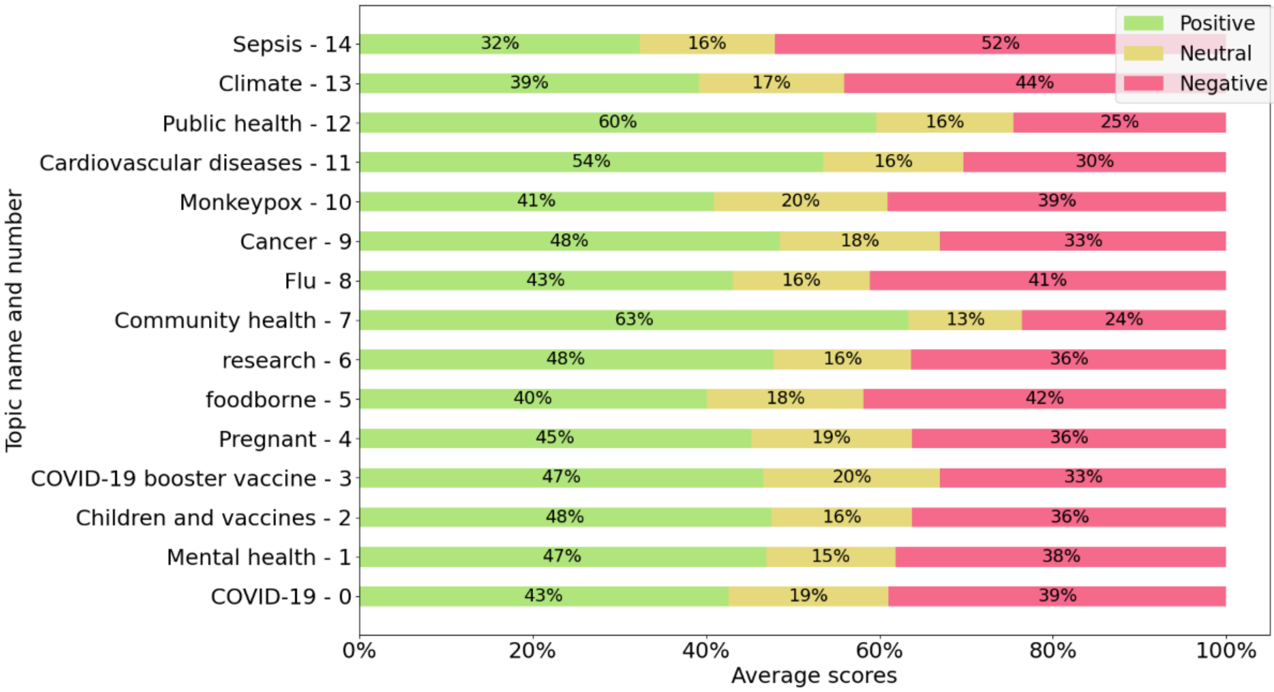
Sentiment Analysis

We analyzed the sentiment in the text of the posts and comments and calculated the sentiment scores of the posts and comments, as described in the Methods section.

Figures 2 and 3 present the average sentiment scores for health organizations' posts and comments, respectively. In Figures 2 and 3, we selected only the 15 largest topics (those containing the greatest number of posts). Figure 2 shows that in certain topics, such as sepsis, climate, and foodborne illnesses, negative sentiment is the most prominent in the health organizations' posts. However, in Figure 3, which presents the scores for the comments to the health organizations' posts, we see that there are more topics where negative sentiment is dominant than in Figure 2, with the greatest negative sentiment found in the comments for posts about vaccines, specifically for posts in the COVID-19, children and vaccines, booster vaccines, and monkeypox topics.

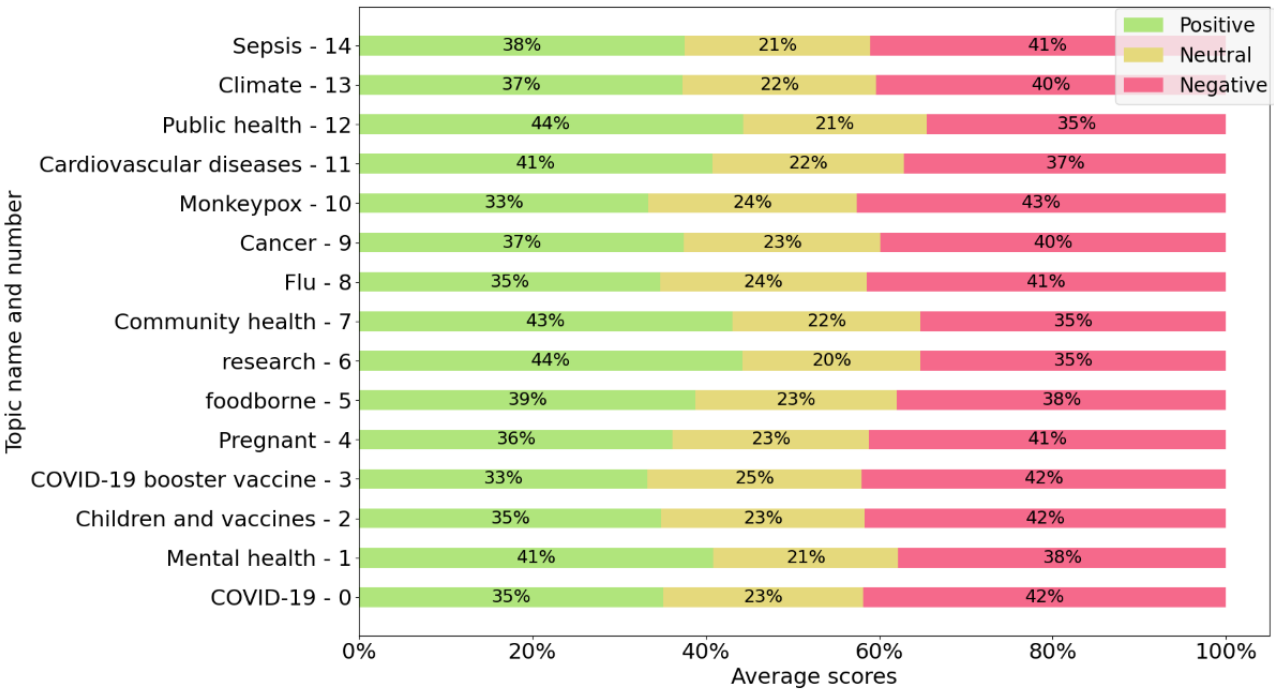
XSL•FO
RenderX

Figure 2. Sentiment scores of posts (average positive, neutral, and negative) for each topic.



While certain topics showed a dominant positive sentiment in the posts published by health organizations (as illustrated in Figure 2), this does not necessarily reflect how the public responds to those posts. The dominance of positive sentiment in these topics was based on the average sentiment scores, where the positive score was higher than both the negative and neutral scores. This indicated that these subjects were presented mostly positively by the organizations. However, when examining the sentiment expressed in user comments (Figure 3), we observed a contrasting response. In topics such as COVID-19, children and vaccines, and the COVID-19 booster vaccine, the comments exhibited predominantly negative sentiment, even though the organizations framed these topics positively. In these cases, the average negative score in the comments was higher than both the positive and neutral scores. In other words, while the organizations tried to communicate these topics in a positive light, the public's reaction to them was largely negative.

Figure 3. Sentiment scores of comments (average positive, neutral, and negative) for each topic.



XSL•FO
RenderX

Emotion Analysis

We analyzed the emotion in the text of the posts and comments and calculated the emotion scores of the posts and comments, as described in the Methods section.

The emotions present were anger, disgust, fear, joy, neutral, sadness, and surprise. Figure 4 presents the average emotion scores for the posts, and Figure 5 displays the average emotion scores for the comments. In Figures 4 and 5, we selected only the 15 largest topics (those containing the greatest number of posts).

As seen in Figure 4, the dominant emotion in all topics was fear, and its scores were higher than those of all the other emotions in the posts. In other words, the text in the health organization posts was characterized by a very high level of fear. Among the topics with the highest fear emotion were sepsis, monkeypox, cancer, and flu.

Figure 4. Emotion scores of posts (average anger, disgust, fear, joy, neutral, sadness, and surprise) for each topic.

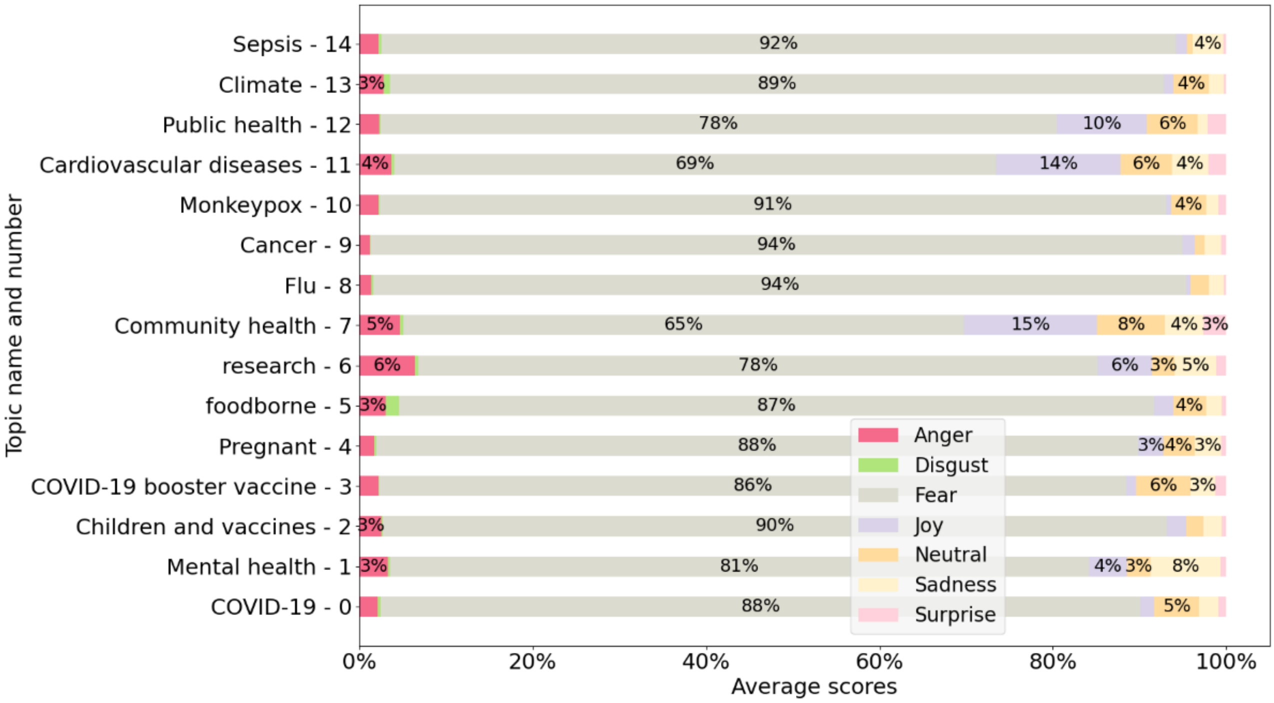
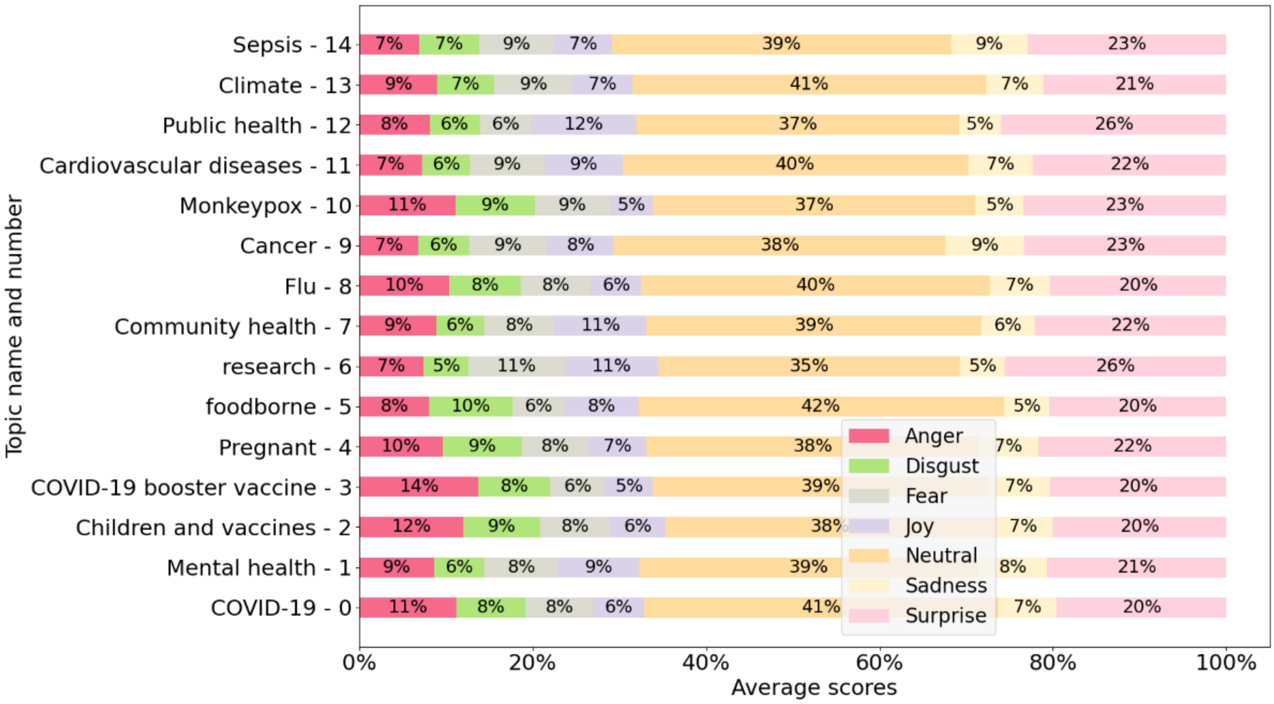


Figure 5. Emotion scores of comments (average anger, disgust, fear, joy, neutral, sadness, and surprise) for each topic.



XSL•FO
RenderX

However, as shown in Figure 5, fear is not the dominant emotion; instead, neutral responses prevail. This suggests a disconnect between what health organizations consider fear-inducing issues and how the public actually responds, displaying less fear. When examining other emotions in Figure 5, we see that the topic generating the most anger is vaccines, with the COVID-19 booster vaccine specifically eliciting the highest levels of anger. In terms of sadness, cancer and sepsis were the topics that evoked the strongest feelings of sadness.

Gap Analysis

Given the difference observed between the high average fear in the health organizations' posts compared to the very low fear in the public's responses to the posts, we examined the topics with the largest gap between the 2 average fears as revealed in the emotion analysis. For each topic, Table 4 presents the average fear in the posts of health organizations, the average fear in the comments, and the difference between the two. As can be seen in Table 4, the topics with the highest gap are masks and flu.

Table 4. Each topic's average scores for fear in post and comments and the gap.

Topic name	Topic number	Fear in posts, mean (SD)	Fear in comments, mean (SD)	Gap
Masks	15	0.958 (0.072)	0.092 (0.166)	0.881
Flu	8	0.938 (0.156)	0.066 (0.180)	0.857
Vector	16	0.931 (0.090)	0.101 (0.196)	0.846
Cancer	9	0.936 (0.126)	0.062 (0.191)	0.846
Sepsis	14	0.917 (0.124)	0.106 (0.191)	0.831
Children and vaccines	2	0.904 (0.152)	0.086 (0.181)	0.823
Noise damage	24	0.913 (0.150)	0.084 (0.198)	0.821
Monkeypox	10	0.906 (0.138)	0.061 (0.197)	0.819
Ebola	19	0.898 (0.170)	0.085 (0.193)	0.812
Foodborne	5	0.870 (0.198)	0.077 (0.155)	0.810
Pregnant	4	0.879 (0.192)	0.086 (0.172)	0.803
Climate	13	0.893 (0.208)	0.090 (0.197)	0.803
COVID-19	0	0.876 (0.201)	0.060 (0.181)	0.799
COVID-19 booster vaccine	3	0.861 (0.178)	0.086 (0.155)	0.798
RSV[a]	22	0.897 (0.161)	0.087 (0.222)	0.796
Health equity	17	0.815 (0.223)	0.091 (0.147)	0.754
Antibiotics	18	0.835 (0.274)	0.081 (0.192)	0.751
Sun damage	23	0.800 (0.272)	0.081 (0.169)	0.734
Mental health	1	0.806 (0.283)	0.111 (0.188)	0.722
Public health	12	0.780 (0.281)	0.060 (0.152)	0.720
Smoking	21	0.777 (0.254)	0.076 (0.143)	0.715
Research	6	0.783 (0.218)	0.064 (0.240)	0.672
Humanitarian	20	0.724 (0.341)	0.081 (0.207)	0.618
Cardiovascular diseases	11	0.694 (0.332)	0.084 (0.192)	0.608
Community health	7	0.646 (0.331)	0.077 (0.178)	0.565

[a]RSV: respiratory syncytial virus.

Linguistic Analysis

The linguistic analysis results are presented in Multimedia Appendix 4. It contains the top 50 phrases for unigrams, bigrams, trigrams, and four-grams in posts receiving predominantly positive or negative responses, along with their TF-IDF scores.

Results revealed that phrases associated with positive public responses promote public health awareness, vaccination benefits, and preventive measures. As an example, the phrases that were only included in the top 50 phrases of the health organizations' posts that received positive responses included "stay healthy" "awareness month," "signs symptoms," "raise awareness," "save lives," "help slow spread," "health care provider," and "better health better understanding." In contrast, phrases associated with negative sentiment related to vaccination efforts, outcomes, policy mandates, and health monitoring include terms such as "data tracker," "vaccinated covid," "dose vaccine," "severe

XSL•FO
RenderX

illness hospitalization," "illness hospitalization death," and "covid 19 vaccine booster."

Time Analysis

Figures 6-8 illustrate sentiment and emotional shifts in posts and comments from 2018 to 2023 along with the COVID-19 death ratio. The numbers 1 to 5 in Figures 6-8 represent 5 significant milestones during the COVID-19 pandemic. These include the following: (1) first report of COVID-19 in late 2019, (2) the declaration of COVID-19 as a global pandemic by the WHO in March 2020, (3) the administration of the first COVID-19 vaccine in December 2020 in the United Kingdom, (4) the peak of the Omicron wave in early 2022, and (5) the WHO's declaration in May 2023 that COVID-19 was no longer a global health emergency.

Figure 6. Evolution of sentiment (average positive and negative scores) in posts and comments over the years, combined with the COVID-19 death ratio.

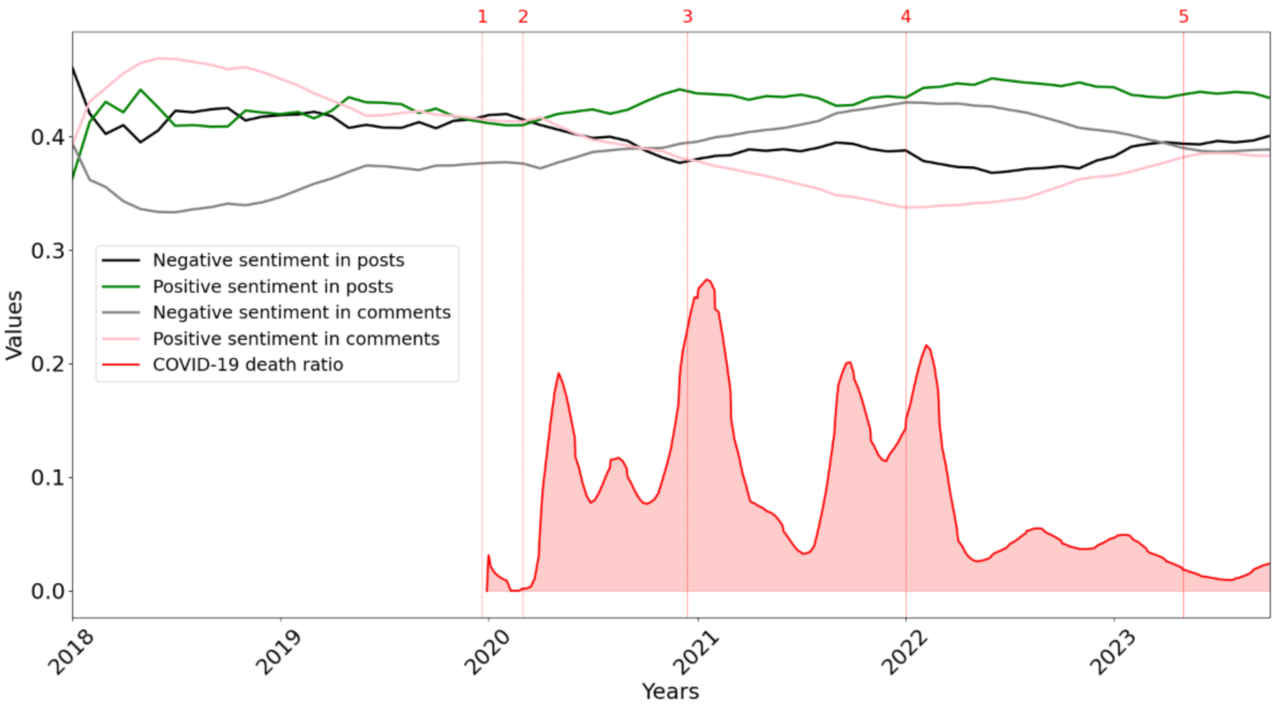


Figure 6 presents the average positive and negative sentiment expressed in posts and comments over time. The negative sentiment in comments increased over time and then began to decrease after the peak of the Omicron wave. Positive sentiment in comments showed a similar but opposite pattern, as it decreased and then increased after the peak of the Omicron wave. Regarding the sentiment of posts, it appeared that positivity in posts increased slightly over the years, while negativity decreased. The peaks of the COVID-19 death ratio aligned with increased negative sentiment in comments.

Figure 7 explores the average levels of emotions—fear, trust, disappointment, anger, and confusion—in posts. Fear consistently dominated posts and remained relatively steady until it declined, coinciding with a significant reduction in the death ratio. Trust, anger, and disappointment remained relatively steady throughout the years. Confusion gradually increased following the declaration of the pandemic.

XSL•FO
RenderX

Figure 7. Evolution of emotion (average scores for trust, fear, anger, confusion, and disappointment) in posts over the years, combined with the COVID-19 death ratio.

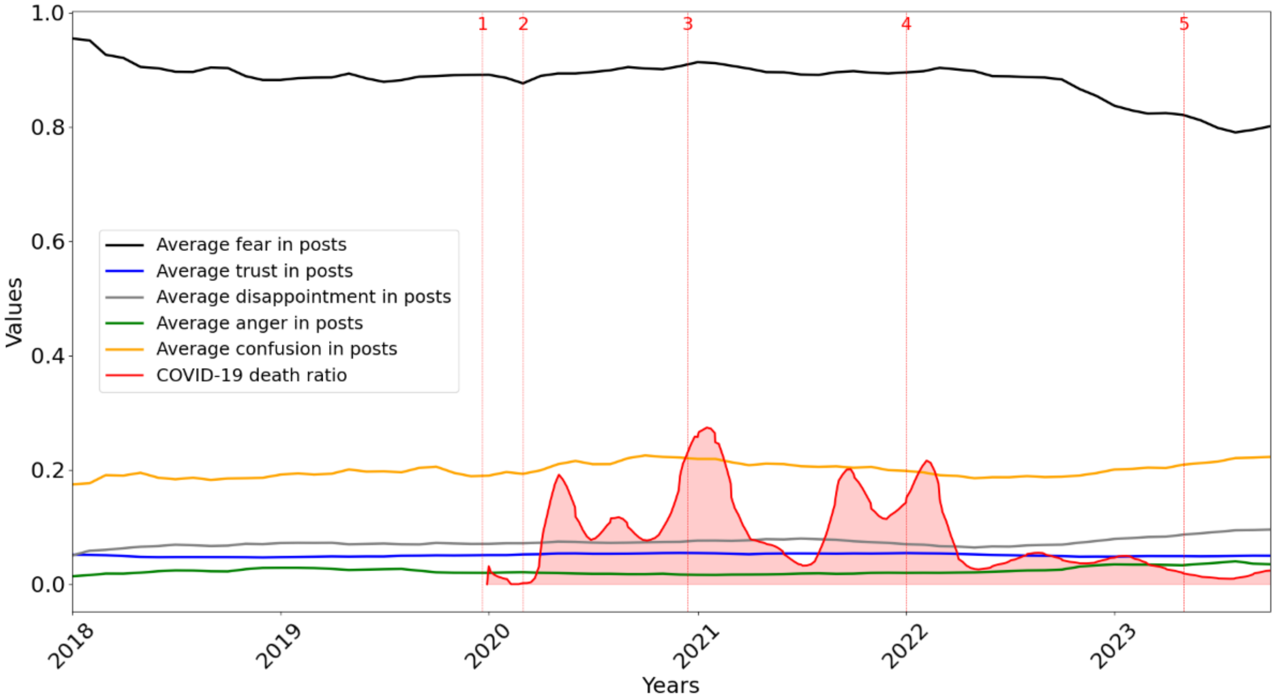
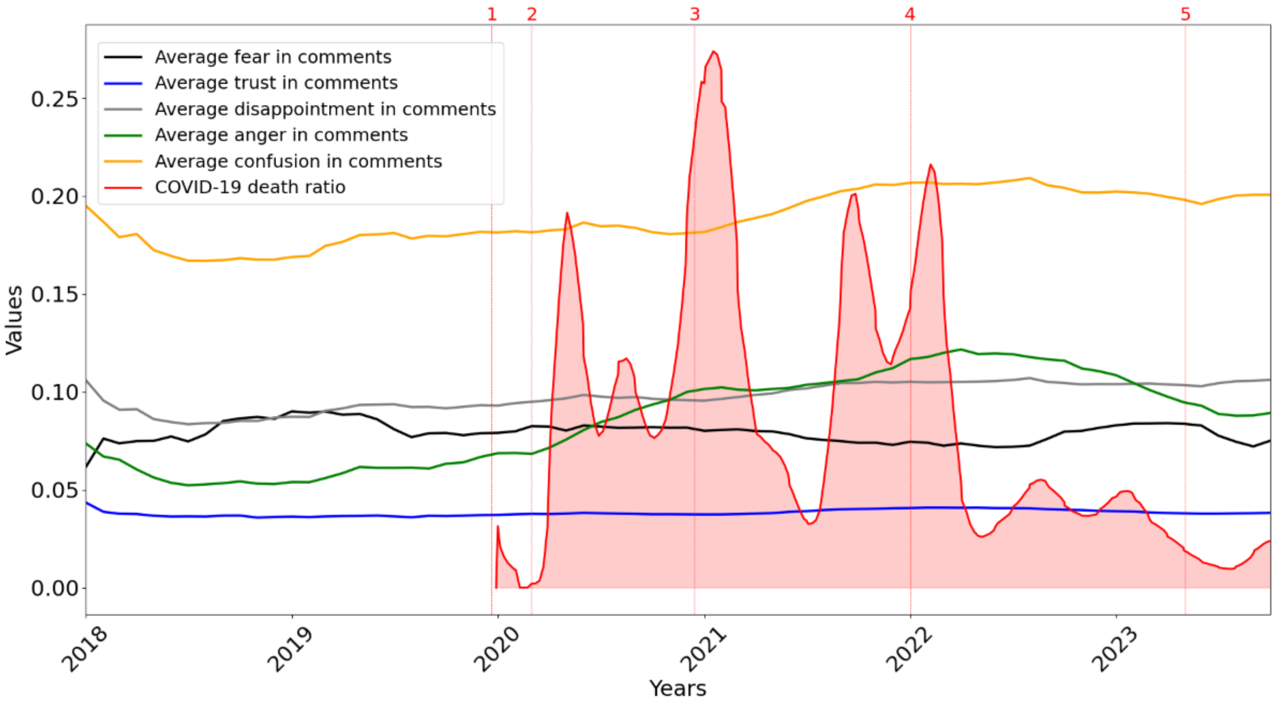


Figure 8 shows the average levels of emotions—fear, trust, disappointment, anger, and confusion—in comments over the same period. Emotions of fear and trust remained relatively steady throughout the years. Anger and disappointment increased around death ratio peaks and, overall, showed a general upward pattern over the years. However, anger began to decline slightly after the Omicron peak and the subsequent decrease in the death ratio.

Figure 8. Evolution of emotion (average scores for trust, fear, anger, confusion, and disappointment) in comments over the years, combined with the COVID-19 death ratio.



XSL•FO
RenderX

Discussion

Principal Findings

The findings of this study should be interpreted with consideration of the dataset's geographic composition. As most of the health organizations analyzed are based in the United States, the topics emphasized may reflect US-specific health priorities, and the patterns of public engagement, sentiment, and emotional responses to these topics may also be shaped by US cultural and social contexts. The selected health organizations have a high number of followers on the platform, hold substantial influence, and reach audiences beyond national boundaries. Still, the framing and intensity of reactions observed in this study may not be fully generalizable to health organizations in other countries or regions, particularly because the characteristics of the users engaging with the content (such as their geographic location or demographics) are unknown.

Our findings on engagement analysis show that humanitarian issues, including the Ukraine war and the war in Gaza, and the COVID-19 pandemic received the greatest response from the public. This supports prior research, which has shown that those disasters, including humanitarian crises, such as nature-related crises and wars, are a major concern of the public and policy makers [101]. In recent years, conflicts, such as the ongoing Ukraine war and the war in Gaza, have had a significant impact on international stability, affecting economies, migration patterns, and security measures, particularly in Europe [102]. These conflicts have not only disrupted local societies but have also prompted reevaluation of humanitarian aid and intervention strategies globally. The connection between humanitarian conflicts and adverse health outcomes is well documented. Studies have shown increased incidence of mental health disorders, infectious disease outbreaks, and chronic health conditions in conflict-affected areas [103]. As a result, health organizations worldwide have recognized the urgent need to develop comprehensive strategies for emergency preparedness. This includes enhancing public health response capabilities, improving prehospital care, and integrating disaster medicine in public health frameworks to better address humanitarian crises [104,105].

The COVID-19 pandemic has fundamentally altered global perceptions and practices across various sectors, revealing both strengths and weaknesses in systems and policies worldwide [106]. During the COVID-19 pandemic, the public turned to health organization websites and social media channels to receive timely updates on infection rates, government guidelines, and evolving safety protocols [107,108]. Trust in organizations such as the WHO and CDC was crucial, as these bodies communicated essential information, adapting advice as new data emerged about the virus's spread and impact. The results of this study further highlight that a significant portion of posts shared by health organizations focused on the topic of COVID-19. In addition, these posts generated substantial public engagement, demonstrating the public's heightened interest and concern regarding COVID-19–related information. Given the profound impact of COVID-19, the public remains extremely interested in these insights, recognizing that understanding what was done well and what failed can better prepare societies for future health crises. Consequently, the ongoing dialogue among scientists, policy makers, and the public about these findings continues to shape postpandemic recovery strategies and fortify global health preparedness [109].

Our sentiment and emotion analysis of responses to health organizations' posts showed that the highest percentage of negative scores was in topics related to vaccines and monkeypox. Moreover, our results show that the highest level of anger was observed in topics related to the COVID-19 booster and children's vaccination in general. The high anger and negative sentiment can be attributed to vaccine hesitancy.

Vaccine hesitancy and antivaccine movements have long posed significant challenges to public health [110], particularly in communities with historically low trust in government and pharmaceutical institutions [111]. The analysis of topics associated with high anger emotions, such as COVID-19 vaccines, can be enriched by examining the emotional characteristics and motivations of specific user groups with vaccine hesitancy. For instance, women are generally more likely to exhibit vaccine hesitancy than men [112]. Alternatively, older individuals tend to display lower levels of hesitancy, likely due to their awareness of their vulnerability to COVID-19 complications, which may alleviate their fear or anger [112]. Vaccine hesitancy is also more prevalent among individuals with lower economic security [112]. Trust also shapes emotional and cognitive characteristics. Higher trust in health care providers, scientists, and global health organizations such as the WHO correlates with reduced vaccine hesitancy [112]. Conversely, in some contexts, higher trust in religious leaders is linked to increased hesitancy [112].

The COVID-19 pandemic has increased skepticism toward routine immunizations and increased hesitancy even among those previously compliant with vaccination schedules. The rapid development of COVID-19 vaccines, coupled with widely publicized adverse effects, heightened fears and deepened mistrust [113]. In addition, misinformation and conflicting narratives on social media further fueled fear, anger, and uncertainty, particularly in communities already skeptical of public health authorities. Our findings reflect these trends, showing increases in public negativity, anger, and disappointment over time.

Our linguistic analysis highlights that messages emphasizing safety, prevention, and public health benefits tend to elicit positive sentiments, while messages emphasizing vaccination efforts, outcomes, policy mandates, and health monitoring tend to elicit negative sentiments. This underscores the need for health organizations to refine communication strategies, focusing on clear, trust-building messaging to address concerns and counteract negative sentiments effectively.

Our analysis also reveals a gap between the fear conveyed in posts by health organizations and the public's responses to these posts, particularly regarding the influenza vaccine and face masks during the COVID-19 pandemic. While health organizations intended to convey urgency, public comments did not reflect the same fear. Addressing skepticism and ambivalence about influenza vaccine and mask use is critical.

XSL•FO
RenderX

Face masks are a critical tool in pandemic response [114], and enhancing public understanding of their effectiveness is essential for improving adherence and compliance during future outbreaks.

The difference in fear responses between the influenza vaccine and face masks can be understood by examining the interplay of social, psychological, and cultural factors [114]. One factor is psychological reactance, which occurs when individuals perceive mandates, such as mask requirements, as threats to their autonomy. This perception, combined with beliefs that masks are ineffective and an aversion to being forced to wear them, can trigger anger and counterarguments. Individuals with strong psychological reactance are particularly likely to exhibit these responses, reinforcing and intensifying their antimask attitudes [115,116]. Such resistance is further exacerbated by personality traits, which not only strengthen antimask sentiments but also link to broader vaccine skepticism, exaggeration of COVID-19 risks, and resistance to social distancing, often influenced by political conservatism [117].

Social and cultural dynamics also shape mask perceptions. For example, in many Western societies, masks are seen as extraordinary artifacts associated with emergency, while in Asian cultures, they are normalized as part of daily life [118]. In collectivist cultures, mask wearing aligns with a sense of duty to protect the community [116]. In addition, stigma, appearance concerns, and fears of being perceived as overly cautious further hinder acceptance [119].

Policy decisions and public health campaigns significantly influence perceptions. During the COVID-19 pandemic, inconsistent messaging from health authorities and varying guidance on mask use between countries and organizations contributed to public confusion and skepticism regarding their efficacy [120], despite substantial scientific evidence supporting their role in reducing transmission rates [120,121]. Therefore, consistent messaging and targeted communication strategies should focus on populations with antimask perceptions, aiming to reduce stigmas and address the underlying factors driving these attitudes.

Historically, the influenza vaccine has engendered significant hesitancy and objection from segments of the general public, a trend particularly evident among health care workers, who are critical in promoting vaccination [122]. Factors contributing to this hesitancy include misconceptions about vaccine efficacy, fears of adverse effects, and a perceived lack of urgency surrounding influenza, particularly when compared to more severe illnesses such as COVID-19 [123,124]. The emergence of the COVID-19 pandemic exacerbated this situation. During the COVID-19 pandemic, public focus shifted intensely toward COVID-19 vaccination campaigns, leading to a diversion of attention and resources from influenza vaccination efforts [125]. The flood of information regarding COVID-19 vaccines overshadowed the long-standing influenza vaccine campaigns. As a result, many individuals prioritized the COVID-19 vaccine over the seasonal influenza vaccine [125]. Data indicating that compliance with the influenza vaccine plummeted in several countries during 2023 and 2024 compared to prepandemic years highlight the ongoing challenges faced by public health authorities [125].

Despite the availability of the influenza vaccine, social media misinformation and evolving health narratives regarding influenza and COVID-19 have led to an atmosphere of uncertainty. For instance, some individuals may mistakenly believe that if COVID-19 variants pose health risks, the influenza virus may be less significant or that acquiring one vaccine negates the need for others. This shift necessitates a reassessment of public health strategies to reengage communities with the importance of influenza vaccination. Health organizations must develop targeted messaging that addresses misconceptions, enhances understanding of the influenza virus's potential impact, and reinforces the protective benefits of vaccination, even during times when attention is focused on other diseases.

To summarize, in this study, we analyzed the quality of disseminated messages along 2 dimensions: public responses to topics and the emotions and sentiments expressed in those responses. Mapping these reactions to messages posted by health organizations is important for evaluating public engagement with specific health-related issues. Identifying emotional gaps can also help assess the effectiveness of health-related messages, revealing potential discrepancies between the importance health organizations assign to certain topics and the public's perceived importance.

Our findings highlight gaps in fear responses regarding the influenza vaccine and wearing face masks during the COVID-19 pandemic, underscoring the need for transparent communication from health authorities and comprehensive education campaigns to address misconceptions and reassure the public. Understanding the psychological and social factors driving vaccine hesitancy after the pandemic is essential for tailoring effective public health strategies. Efforts must focus on fostering trust through consistent and clear messaging, transparency about vaccine development processes, and open forums for addressing public concerns. By doing so, it will be possible to mitigate misinformation, reduce fear, enhance public compliance, and improve vaccine uptake across populations.

Limitations

Our study may have some limitations. We collected data from Instagram, and therefore, our results and conclusions are based on the posts and interactions on this social media platform. Our data did not include specific information about the users who commented on the health organizations' posts. However, according to general information about Instagram users by Statista [126], Instagram had 2 billion monthly active users in 2024, with India leading the platform's user base at approximately 360 million, followed by the United States with 169 million and Brazil with 134 million users. Younger users dominate the platform, with the 18- to 24-year age group being the largest demographic, followed by the 25- to 34-year age group. Participation declines significantly among older age groups, particularly those aged ≥55 years, who represent only a small fraction of the audience. Moreover, most users in the 18- to 34-year age group are men, while most users >34 years are women.

XSL•FO
RenderX

The absence of detailed user characteristics in our dataset may limit the generalizability of our findings, as biases could arise due to age, gender, educational level, or other demographic factors that influence engagement and sentiment patterns. Certain user groups may be overrepresented or underrepresented in the data, potentially shaping the interpretation of public responses. Future research should aim to systematically characterize respondents' profiles, leveraging available metadata or incorporating external surveys to gain a clearer understanding of the audience engaging with health-related content.

The study also has limitations regarding the influence of social media algorithms on post visibility and engagement. The social media algorithms prioritize content based on user interactions, interests, and platform-specific ranking mechanisms, potentially introducing bias in public sentiment patterns. Nevertheless, the decision as to whether to engage and how to respond is left to the users. Therefore, although engagement data may not fully represent the broader public, it still offers insights into those actively participating in discussions. To mitigate bias, we ensured a diverse dataset by including a substantial number of posts from multiple health organizations.

It is important to acknowledge that the models used in this study may have limitations. Although BERTopic is an effective topic modeling technique, it assumes that each document (in our case, each post) is associated with a single dominant topic, which does not always reflect reality. As posts can discuss a variety of interconnected topics, it is often difficult to classify them accurately under a single theme. Furthermore, topic separation relies on clustering techniques, which may not always produce clear or optimal topic divisions, resulting in the merging of distinct topics or the fragmentation of related discussions, reducing interpretability.

Similarly, the sentiment and emotion models may inherit biases from their training datasets, influencing detection. These models are not always fully accurate and often struggle to capture context-dependent sentiment shifts, irony, and implicit emotional expressions, which may result in misinterpretations.

In addition, as previously mentioned, the health organizations in this study are predominantly US based, with limited representation from other countries. This concentration may influence the topics represented in the dataset, reflecting US-specific health priorities and cultural dynamics, as well as the public sentiment, emotional responses, and engagement observed in relation to those topics. Future research could incorporate data from a more geographically diverse set of health organizations, enabling cross-cultural comparisons and providing a broader understanding of global public health communication and audience responses.

Conclusions

This study demonstrates the value of our methodology in assessing public responses and emotions expressed regarding health-related messages. By identifying emotional gaps, particularly fear, we were able to uncover discrepancies between the fear health organizations assign to issues and the fear that the public expresses in response. The greatest gaps were revealed with regard to influenza vaccines and face masks during COVID-19.

These findings emphasize the need for transparent communication and trust-building strategies that consider the emotional and sentiment-driven responses of the public on social media. By understanding the psychological and social dynamics of public interaction with health information, particularly regarding vaccine resistance, organizations can support public health goals, foster trust and efficient engagement, counter misinformation, and encourage informed health behaviors.

In future research, we plan to apply our topic modeling and sentiment and emotion analysis approach on other social media platforms to gain a more comprehensive view of the public's response to the posts of health organizations across various digital platforms. In addition, we aim to extend our analysis beyond social media to offline settings, such as newspapers and television campaigns. Because public responses are essential to our method, we will also incorporate surveys to assess audience reactions and engagement with these campaigns. Investigating how fear-based messaging and public emotional responses manifest in both online and offline environments will provide deeper insights into the broader impact of health communication strategies. In addition, we plan to investigate misinformation within the comments on health organization posts and identify the topics where inaccuracies are prevalent.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

APV contributed to the study's conceptualization, software, methodology, data curation, formal analysis, visualization, and writing. AM contributed to the study design, methodology, and writing. All authors approved the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The number, name, size, and top-5 most significant words of each topic.

[XLSX File (Microsoft Excel File), 11 KB - infodemiology_v5i1e70576_app1.xlsx]

XSL•FO
RenderX

Multimedia Appendix 2
Average number of post comments per topic. RSV: respiratory syncytial virus.
[PNG File , 92 KB - infodemiology_v5i1e70576_app2.png]

Multimedia Appendix 3
Average number of post likes per topic. RSV: respiratory syncytial virus.
[PNG File , 101 KB - infodemiology_v5i1e70576_app3.png]

Multimedia Appendix 4
Top-50 phrases for unigrams, bigrams, trigrams, and four-grams in posts receiving predominantly positive or negative responses with corresponding term frequency–inverse document frequency scores.
[XLSX File (Microsoft Excel File), 23 KB - infodemiology_v5i1e70576_app4.xlsx]

References

1. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. Lancet 2010 Oct 09;376(9748):1261-1271 [FREE Full text] [doi: 10.1016/S0140-6736(10)60809-4] [Medline: 20933263]
2. Gatewood J, Monks SL, Singletary CR, Vidrascu E, Moore JB. Social media in public health: strategies to distill, package, and disseminate public health research. J Public Health Manag Pract 2020;26(5):489-492 [FREE Full text] [doi: 10.1097/PHH.0000000000001096] [Medline: 32732723]
3. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917 [FREE Full text] [doi: 10.2196/17917] [Medline: 33978589]
4. Tseng TS, Gonzalez G. Social media and types with their current applications in public health and healthcare. In: Batra K, Sharma M, editors. Effective Use of Social Media in Public Health. New York, NY: Academic Press; 2023:3-22.
5. Jeyaraman M, Ramasubramanian S, Kumar S, Jeyaraman N, Selvaraj P, Nallakumarasamy A, et al. Multifaceted role of social media in healthcare: opportunities, challenges, and the need for quality control. Cureus 2023 May;15(5):e39111 [FREE Full text] [doi: 10.7759/cureus.39111] [Medline: 37332420]
6. Ramanadhan S, Mendez SR, Rao M, Viswanath K. Social media use by community-based organizations conducting health promotion: a content analysis. BMC Public Health 2013 Dec 05;13:1129 [FREE Full text] [doi: 10.1186/1471-2458-13-1129] [Medline: 24313999]
7. Xu S, Markson C, Costello KL, Xing CY, Demissie K, Llanos AA. Leveraging social media to promote public health knowledge: example of cancer awareness via Twitter. JMIR Public Health Surveill 2016;2(1):e17 [FREE Full text] [doi: 10.2196/publichealth.5205] [Medline: 27227152]
8. Latha K, Meena KS, Pravitha MR, Dasgupta M, Chaturvedi SK. Effective use of social media platforms for promotion of mental health awareness. J Educ Health Promot 2020;9(1):124 [FREE Full text] [doi: 10.4103/jehp.jehp_90_20] [Medline: 32642480]
9. Anwar A, Malik M, Raees V, Anwar A. Role of mass media and public health communications in the COVID-19 pandemic. Cureus 2020 Sep 14;12(9):e10453 [FREE Full text] [doi: 10.7759/cureus.10453] [Medline: 33072461]
10. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
11. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
12. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]
13. Edney S, Bogomolova S, Ryan J, Olds T, Sanders I, Maher C. Creating engaging health promotion campaigns on social media: observations and lessons from Fitbit and Garmin. J Med Internet Res 2018 Dec 10;20(12):e10911 [FREE Full text] [doi: 10.2196/10911] [Medline: 30530449]
14. Afful-Dadzie E, Afful-Dadzie A, Egala SB. Social media in health communication: a literature review of information quality. Health Inf Manag 2023 Jan 04;52(1):3-17. [doi: 10.1177/1833358321992683] [Medline: 33818176]
15. Kington RS, Arnesen S, Chou WY, Curry SJ, Lazer D, Villarruel AM. Identifying credible sources of health information in social media: principles and attributes. NAM Perspect 2021 Jul 16;2021:31478 [FREE Full text] [doi: 10.31478/202107a] [Medline: 34611600]
16. Eysenbach G. Credibility of health information and digital media: new perspectives and implications for youth. Digit Media Learn 2007;9(2):123 [FREE Full text] [doi: 10.1162/dmal.9780262562324.123]
17. Heldman AB, Schindelar J, Weaver JB. Social media engagement and public health communication: implications for public health organizations being truly “social”. Public Health Rev 2013 Jun 3;35(1):1-18. [doi: 10.1007/BF03391698]

XSL•FO
RenderX

18. Aydin G. Improving health and efficiency with strategic social media use in health organizations: a critical review of the status quo. In: Dinçer H, Yüksel S, editors. Multidimensional Perspectives and Global Analysis of Universal Health Coverage. New York, NY: IGI Global; 2020:309-336.
19. Ghahramani A, Prokofieva M, de Courten MP. Health promotion campaigns using social media: association rules mining and co-occurrence network analysis of Twitter hashtags. BMC Public Health 2025 Jan 07;25(1):67 [FREE Full text] [doi: 10.1186/s12889-024-21255-5] [Medline: 39773431]
20. De FreitasqC FC, Silva S, Maia T, Amorim M. Public participation in health data governance: a scoping review. Eur J Public Health 2019;29(Suppl 4):185-647. [doi: 10.1093/eurpub/ckz185.647]
21. Tavoschi L, Quattrone F, D'Andrea E, Ducange P, Vabanesi M, Marcelloni F, et al. Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. Hum Vaccin Immunother 2020 May 03;16(5):1062-1069 [FREE Full text] [doi: 10.1080/21645515.2020.1714311] [Medline: 32118519]
22. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. J Med Internet Res 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: 10.2196/26627] [Medline: 33724919]
23. Hu T, Wang S, Luo W, Zhang M, Huang X, Yan Y, et al. Revealing public opinion towards COVID-19 vaccines with twitter data in the United States: spatiotemporal perspective. J Med Internet Res 2021 Sep 10;23(9):e30854 [FREE Full text] [doi: 10.2196/30854] [Medline: 34346888]
24. Niu Q, Liu J, Kato M, Shinohara Y, Matsumura N, Aoyama T, et al. Public opinion and sentiment before and at the beginning of COVID-19 vaccinations in Japan: Twitter analysis. JMIR Infodemiology 2022 May 9;2(1):e32335 [FREE Full text] [doi: 10.2196/32335] [Medline: 35578643]
25. Saleh SN, McDonald SA, Basit MA, Kumar S, Arasaratnam RJ, Perl TM, et al. Public perception of COVID-19 vaccines through analysis of Twitter content and users. Vaccine 2023 Jul 25;41(33):4844-4853 [FREE Full text] [doi: 10.1016/j.vaccine.2023.06.058] [Medline: 37385887]
26. Finset A, Bosworth H, Butow P, Gulbrandsen P, Hulsman RL, Pieterse AH, et al. Effective health communication - a key factor in fighting the COVID-19 pandemic. Patient Educ Couns 2020 May;103(5):873-876 [FREE Full text] [doi: 10.1016/j.pec.2020.03.027] [Medline: 32336348]
27. Rovetta A. World Health Organization is losing online credibility towards health-sensitive topics: infodemiological analysis of Facebook users' reactions. Health Promot Perspect 2022;12(4):367-371 [FREE Full text] [doi: 10.34172/hpp.2022.48] [Medline: 36852199]
28. Gemenis K. COVID-19 conspiracy beliefs and vaccination intentions among social media users. Stat Polit Policy 2022;13(3):279-296. [doi: 10.1515/spp-2022-0005]
29. Lee J, Bissell K. Correcting vaccine misinformation on social media: the inadvertent effects of repeating misinformation within such corrections on COVID-19 vaccine misperceptions. Curr Psychol 2024 Jan 25;43(26):22754-22766. [doi: 10.1007/s12144-024-05651-z]
30. Ayaz H, Celik MH, Koytak HZ, Yanik IE. Exploring vaccine hesitancy in digital public discourse: from tribal polarization to socio-economic disparities. PLoS One 2024 Nov 5;19(11):e0308122 [FREE Full text] [doi: 10.1371/journal.pone.0308122] [Medline: 39499705]
31. Kite J, Chan L, MacKay K, Corbett L, Reyes-Marcelino G, Nguyen B, et al. A model of social media effects in public health communication campaigns: systematic review. J Med Internet Res 2023 Jul 14;25:e46345 [FREE Full text] [doi: 10.2196/46345] [Medline: 37450325]
32. Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. Soc Netw Anal Min 2021 Aug 28;11(1):81 [FREE Full text] [doi: 10.1007/s13278-021-00776-6] [Medline: 34484462]
33. Demszky D, Movshovitz-Attias D, Ko J, Cowen A, Nemade G, Ravi S. GoEmotions: a dataset of fine-grained emotions. arXiv Preprint posted online May 1, 2020 [FREE Full text] [doi: 10.18653/v1/2020.acl-main.372]
34. Chutia T, Baruah N. A review on emotion detection by using deep learning techniques. Artif Intell Rev 2024 Jul 11;57(8):203. [doi: 10.1007/s10462-024-10831-1]
35. Chen L, Yang X, Fu L, Liu X, Yuan C. Using the extended parallel process model to examine the nature and impact of breast cancer prevention information on mobile-based social media: content analysis. JMIR Mhealth Uhealth 2019 Jun 24;7(6):e13987 [FREE Full text] [doi: 10.2196/13987] [Medline: 31237239]
36. Witte K. Putting the fear back into fear appeals: the extended parallel process model. Commun Monogr 1992 Dec;59(4):329-349. [doi: 10.1080/03637759209376276]
37. Raghupathi V, Ren J, Raghupathi W. Studying public perception about vaccination: a sentiment analysis of tweets. Int J Environ Res Public Health 2020 May 15;17(10):3464 [FREE Full text] [doi: 10.3390/ijerph17103464] [Medline: 32429223]
38. Ansari MT, Khan NA. Worldwide COVID-19 vaccines sentiment analysis through Twitter content. Electron J Gen Med 2021;18(6):em329. [doi: 10.29333/ejgm/11316]
39. Liu S, Liu J. Public attitudes toward COVID-19 vaccines on English-language Twitter: a sentiment analysis. Vaccine 2021 Sep 15;39(39):5499-5505 [FREE Full text] [doi: 10.1016/j.vaccine.2021.08.058] [Medline: 34452774]
40. Lyu JC, Han EL, Luli GK. COVID-19 vaccine-related discussion on Twitter: topic modeling and sentiment analysis. J Med Internet Res 2021 Jun 29;23(6):e24435 [FREE Full text] [doi: 10.2196/24435] [Medline: 34115608]

XSL•FO
RenderX

41. Monselise M, Chang CH, Ferreira G, Yang R, Yang CC. Topics and sentiments of public concerns regarding COVID-19 vaccines: social media trend analysis. J Med Internet Res 2021 Oct 21;23(10):e30765 [FREE Full text] [doi: 10.2196/30765] [Medline: 34581682]
42. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis 2021 Jul;108:256-262 [FREE Full text] [doi: 10.1016/j.ijid.2021.05.059] [Medline: 34052407]
43. Chandrasekaran R, Desai R, Shah H, Kumar V, Moustakas E. Examining public sentiments and attitudes toward COVID-19 vaccination: infoveillance study using Twitter posts. JMIR Infodemiology 2022 Apr 15;2(1):e33909 [FREE Full text] [doi: 10.2196/33909] [Medline: 35462735]
44. Chen N, Chen X, Pang J. A multilingual dataset of COVID-19 vaccination attitudes on Twitter. Data Brief 2022 Oct;44:108503 [FREE Full text] [doi: 10.1016/j.dib.2022.108503] [Medline: 35935093]
45. Huangfu L, Mo Y, Zhang P, Zeng DD, He S. COVID-19 vaccine tweets after vaccine rollout: sentiment-based topic modeling. J Med Internet Res 2022 Feb 08;24(2):e31726 [FREE Full text] [doi: 10.2196/31726] [Medline: 34783665]
46. Nguyen A, Catalan-Matamoros D. Anti-vaccine discourse on social media: an exploratory audit of negative tweets about vaccines and their posters. Vaccines (Basel) 2022 Dec 01;10(12):2067 [FREE Full text] [doi: 10.3390/vaccines10122067] [Medline: 36560477]
47. Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J 2022 Jul 09;98(1161):544-550. [doi: 10.1136/postgradmedj-2021-140685] [Medline: 34373343]
48. Mir AA, Sevukan R. Sentiment analysis of Indian Tweets about COVID-19 vaccines. J Inf Sci 2022 Sep 16;50(5):1308-1320. [doi: 10.1177/01655515221118049]
49. Reshi AA, Rustam F, Aljedaani W, Shafi S, Alhossan A, Alrabiah Z, et al. COVID-19 vaccination-related sentiments analysis: a case study using worldwide twitter dataset. Healthcare (Basel) 2022 Feb 22;10(3):411 [FREE Full text] [doi: 10.3390/healthcare10030411] [Medline: 35326889]
50. Stracqualursi L, Agati P. COVID-19 vaccines in Italian public opinion: identifying key issues using Twitter and natural language processing. PLoS One 2022 Nov 17;17(11):e0277394 [FREE Full text] [doi: 10.1371/journal.pone.0277394] [Medline: 36395254]
51. Yiannakoulias N, Darlington JC, Slavik CE, Benjamin G. Negative COVID-19 vaccine information on Twitter: content analysis. JMIR Infodemiology 2022 Aug 29;2(2):e38485 [FREE Full text] [doi: 10.2196/38485] [Medline: 36348980]
52. Zulfiker MS, Kabir N, Biswas AA, Zulfiker S, Uddin MS. Analyzing the public sentiment on COVID-19 vaccination in social media: Bangladesh context. Array (N Y) 2022 Sep;15:100204 [FREE Full text] [doi: 10.1016/j.array.2022.100204] [Medline: 35722449]
53. Xu H, Liu R, Luo Z, Xu M. COVID-19 vaccine sensing: sentiment analysis and subject distillation from Twitter data. Telemat Inform Rep 2022 Dec;8:100016. [doi: 10.1016/j.teler.2022.100016]
54. Rustam F, Khalid M, Aslam W, Rupapara V, Mehmood A, Choi GS. A performance comparison of supervised machine learning models for COVID-19 tweets sentiment analysis. PLoS One 2021 Feb 25;16(2):e0245909 [FREE Full text] [doi: 10.1371/journal.pone.0245909] [Medline: 33630869]
55. Blankenship E, Goff ME, Yin J, Tse ZT, Fu K, Liang H, et al. Sentiment, contents, and retweets: a study of two vaccine-related twitter datasets. Perm J 2018;22:17-138 [FREE Full text] [doi: 10.7812/TPP/17-138] [Medline: 29911966]
56. Chakraborty K, Bhatia S, Bhattacharyya S, Platos J, Bag R, Hassanien AE. Sentiment analysis of COVID-19 tweets by deep learning classifiers-a study to show how popularity is affecting accuracy in social media. Appl Soft Comput 2020 Dec;97:106754 [FREE Full text] [doi: 10.1016/j.asoc.2020.106754] [Medline: 33013254]
57. Argyris YA, Monu K, Tan PN, Aarts C, Jiang F, Wiseley KA. Using machine learning to compare provaccine and antivaccine discourse among the public on social media: algorithm development study. JMIR Public Health Surveill 2021 Jun 24;7(6):e23105 [FREE Full text] [doi: 10.2196/23105] [Medline: 34185004]
58. Aygun I, Kaya B, Kaya M. Aspect based twitter sentiment analysis on vaccination and vaccine types in COVID-19 pandemic with deep learning. IEEE J Biomed Health Inform 2022 May;26(5):2360-2369. [doi: 10.1109/jbhi.2021.3133103]
59. Basiri ME, Nemati S, Abdar M, Asadi S, Acharrya UR. A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets. Knowl Based Syst 2021 Sep 27;228:107242 [FREE Full text] [doi: 10.1016/j.knosys.2021.107242] [Medline: 36570870]
60. Kaur H, Ahsaan SU, Alankar B, Chang V. A Proposed sentiment analysis deep learning algorithm for analyzing COVID-19 tweets. Inf Syst Front 2021 Apr 20;23(6):1417-1429 [FREE Full text] [doi: 10.1007/s10796-021-10135-7] [Medline: 33897274]
61. Villavicencio C, Macrohon JJ, Inbaraj XA, Jeng JH, Hsieh JG. Twitter sentiment analysis towards COVID-19 vaccines in the Philippines using naïve bayes. Information 2021 May 11;12(5):204. [doi: 10.3390/info12050204]
62. Bari A, Heymann M, Cohen RJ, Zhao R, Szabo L, Apas Vasandani S, et al. Exploring coronavirus disease 2019 vaccine hesitancy on twitter using sentiment analysis and natural language processing algorithms. Clin Infect Dis 2022 May 15;74(Suppl_3):e4-e9. [doi: 10.1093/cid/ciac141] [Medline: 35568473]

XSL•FO
RenderX

63. Islam MB, Hasibunnahar S, Shukla PK, Rawat P, Dange J. Twitter opinion mining on COVID-19 vaccinations by machine learning presence. In: Proceedings of 3rd International Conference Doctoral Symposium on Computational Intelligence. 2022 Presented at: DoSCI '22; September 22-28, 2022; Lucknow, India p. 37-55 URL: https://link.springer.com/chapter/10.1007/978-981-19-3148-2_4 [doi: 10.1007/978-981-19-3148-2_4]
64. Maulana FI, Adi PD, Lestari D, Purnomo A, Prihatin SY. Twitter data sentiment analysis of COVID-19 vaccination using machine learning. In: Proceedings of the 5th International Seminar on Research of Information Technology and Intelligent Systems. 2022 Presented at: ISRITI '22; December 8-9, 2022; Yogyakarta, Indonesia p. 582-587 URL: https://ieeexplore.ieee.org/document/10053035 [doi: 10.1109/isriti56927.2022.10053035]
65. Nyawa S, Tchuente D, Fosso-Wamba S. COVID-19 vaccine hesitancy: a social media analysis using deep learning. Ann Oper Res 2022 Jun 16;339(1-2):1-39 [FREE Full text] [doi: 10.1007/s10479-022-04792-3] [Medline: 35729983]
66. Shahzad A, Zafar B, Ali N, Jamil U, Alghadhban AJ, Assam M, et al. COVID-19 vaccines related user's response categorization using machine learning techniques. Computation 2022 Aug 18;10(8):141. [doi: 10.3390/computation10080141]
67. Albahli S, Nawaz M. TSM-CV: Twitter sentiment analysis for COVID-19 vaccines using deep learning. Electronics 2023 Aug 07;12(15):3372. [doi: 10.3390/electronics12153372]
68. Biri I, Kucuktas UT, Uysal F, Hardalac F. Forecasting the future popularity of the anti-vax narrative on Twitter with machine learning. J Supercomput 2023 Aug 23;80(3):2917-2947. [doi: 10.1007/s11227-023-05567-8]
69. Ramondt S, Kerkhof P, Merz EM. Blood donation narratives on social media: a topic modeling study. Transfus Med Rev 2022 Jan;36(1):58-65 [FREE Full text] [doi: 10.1016/j.tmrv.2021.10.001] [Medline: 34810071]
70. Chen L, Wang P, Ma X, Wang X. Cancer communication and user engagement on Chinese social media: content analysis and topic modeling study. J Med Internet Res 2021 Nov 10;23(11):e26310 [FREE Full text] [doi: 10.2196/26310] [Medline: 34757320]
71. Jiang S, Wang P, Liu PL, Ngien A, Wu X. Social media communication about HPV vaccine in China: a study using topic modeling and survey. Health Commun 2023 May 24;38(5):935-946. [doi: 10.1080/10410236.2021.1983338] [Medline: 34555993]
72. Paul MJ, Dredze M. Discovering health topics in social media using topic models. PLoS One 2014;9(8):e103408 [FREE Full text] [doi: 10.1371/journal.pone.0103408] [Medline: 25084530]
73. Seltzer EK, Horst-Martz E, Lu M, Merchant RM. Public sentiment and discourse about Zika virus on Instagram. Public Health 2017 Sep;150:170-175. [doi: 10.1016/j.puhe.2017.07.015] [Medline: 28806618]
74. Muralidhara S, Paul MJ. #Healthy selfies: exploration of health topics on Instagram. JMIR Public Health Surveill 2018 Jun 29;4(2):e10150 [FREE Full text] [doi: 10.2196/10150] [Medline: 29959106]
75. Kim Y, Song D, Lee YJ. #Antivaccination on Instagram: a computational analysis of hashtag activism through photos and public responses. Int J Environ Res Public Health 2020 Oct 17;17(20):7550 [FREE Full text] [doi: 10.3390/ijerph17207550] [Medline: 33080782]
76. Kwok SW, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian twitter users: machine learning analysis. J Med Internet Res 2021 May 19;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: 33886492]
77. Kim Y, Kim JH. Using photos for public health communication: a computational analysis of the Centers for Disease Control and Prevention Instagram photos and public responses. Health Informatics J 2020 Sep 23;26(3):2159-2180 [FREE Full text] [doi: 10.1177/1460458219896673] [Medline: 31969051]
78. Pinto PA, Lopes Antunes MJ, Pisco Almeida AM. Public Health on Instagram: an analysis of health promotion strategies of Portugal and Brazil. Procedia Comput Sci 2021;181:231-238. [doi: 10.1016/j.procs.2021.01.142]
79. Mello S, Glowacki E, Fuentes I, Seabolt J. Communicating COVID-19 risk on Instagram: a content analysis of official public health messaging during the first year of the pandemic. J Health Commun 2023 Jan 02;28(1):38-52 [FREE Full text] [doi: 10.1080/10810730.2023.2175278] [Medline: 36760119]
80. Vaghefi MS, Beheshti N, Jain H. Dissemination of health messages in online social network: a study of healthcare providers' content generation and dissemination on Twitter. Inf Manag 2024 Mar;61(2):103925. [doi: 10.1016/j.im.2024.103925]
81. We champion health and a better future for all. World Health Organization. URL: https://www.who.int/about/ [accessed 2024-12-05]
82. About HHS. U.S. Department of Health and Human Services. URL: https://www.hhs.gov/about/index.html [accessed 2024-12-05]
83. About the Office of Minority Health. The Office of Minority Health. URL: https://minorityhealth.hhs.gov/about-office-minority-health [accessed 2024-12-05]
84. About NIH. The National Institutes of Health. URL: https://www.nih.gov/about-nih [accessed 2024-12-05]
85. About NIMH. The National Institute of Mental Health. URL: https://www.nimh.nih.gov/about [accessed 2024-12-05]
86. Featured topics. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/ [accessed 2024-12-05]
87. NHS services. National Health Service. URL: https://www.nhs.uk/nhs-services/ [accessed 2024-12-05]
88. About FDA. The U.S Food and Drug Administration. URL: https://www.fda.gov/ [accessed 2024-12-05]

XSL•FO
RenderX

89. Pinto PA, Antunes MJ, Almeida AM. Instagram as a communication tool in public health: a systematic review. In: Proceedings of the 15th Iberian Conference on Information Systems and Technologies. 2020 Presented at: CISTI '20; June 24-27, 2020; Seville, Spain p. 1-6 URL: https://ieeexplore.ieee.org/document/9140809 [doi: 10.23919/cisti49556.2020.9140809]
90. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online March 11, 2022 [FREE Full text]
91. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. 2015 Presented at: WSDM '15; February 2-6, 2015; Shanghai, China p. 399-408 URL: https://dl.acm.org/doi/10.1145/2684822.2685324 [doi: 10.1145/2684822.2685324]
92. Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. arXiv Preprint posted online August 27, 2019 [FREE Full text] [doi: 10.18653/v1/d19-1410]
93. McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. arXiv Preprint posted online February 9, 2018 [FREE Full text] [doi: 10.21105/joss.00861]
94. McInnes L, Healy J, Astels S. hdbscan: Hierarchical density based clustering. J Open Source Softw 2017 Mar;2(11):205. [doi: 10.21105/joss.00205]
95. Yuan LX. distilbert-base-multilingual-cased-sentiments-student. Hugging Face. URL: https://huggingface.co/lxyuan/distilbert-base-multilingual-cased-sentiments-student [accessed 2024-04-29]
96. Xian Y, Akata Z, Sharma G, Nguyen Q, Hein M, Schiele B. Latent embeddings for zero-shot classification. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016 Presented at: CVPR '16; June 27-30, 2016; Las Vegas, NV p. 69-77 URL: https://ieeexplore.ieee.org/document/7780384 [doi: 10.1109/cvpr.2016.15]
97. Hartmann J. Emotion english distilroberta-base. Hugging Face. URL: https://huggingface.co/j-hartmann/emotion-english-distilroberta-base [accessed 2024-04-29]
98. Kirouane A. Emotion english distilroberta-base. Hugging Face. URL: https://huggingface.co/ayoubkirouane/BERT-Emotions-Classifier [accessed 2024-04-29]
99. Lowe S. Emotion english distilroberta-base. Hugging Face. URL: https://huggingface.co/SamLowe/roberta-base-go_emotions [accessed 2024-04-29]
100. Provisional COVID-19 death counts by week ending date and state. National Center for Health. URL: https://data.cdc.gov/d/r8kw-7aab [accessed 2025-01-22]
101. Sandifer PA, Walker AH. Enhancing disaster resilience by reducing stress-associated health impacts. Front Public Health 2018 Dec 21;6:373 [FREE Full text] [doi: 10.3389/fpubh.2018.00373] [Medline: 30627528]
102. Colombo A, Magri P. Europe in the age of insecurity. Italian Institute for International Political Studies. URL: https://www.ispionline.it/en/publication/europe-in-the-age-of-insecurity-163702 [accessed 2024-04-29]
103. Garry S, Checchi F. Armed conflict and public health: into the 21st century. J Public Health (Oxf) 2020 Aug 18;42(3):e287-e298. [doi: 10.1093/pubmed/fdz095] [Medline: 31822891]
104. Sofuoglu Z, Jaeger B, Sofuoglu T, Kamyabi L, Bianchi S. Determining the emergency medical services innovation needs through a coordination and support action. Open Res Europe 2024 Dec 16;4:264. [doi: 10.12688/openreseurope.18449.1]
105. Lee JM, Jansen R, Sanderson KE, Guerra F, Keller-Olaman S, Murti M, et al. Public health emergency preparedness for infectious disease emergencies: a scoping review of recent evidence. BMC Public Health 2023 Mar 02;23(1):420 [FREE Full text] [doi: 10.1186/s12889-023-15313-7] [Medline: 36864415]
106. Lombatti S, Magid A, Davidovitch N, Middleton J, Sheek-Hussein M, Lopes H, et al. The COVID-19 pandemic as a public health teacher-the lessons we must learn. Public Health Rev 2024 May 7;45:1607232 [FREE Full text] [doi: 10.3389/phrs.2024.1607232] [Medline: 38774497]
107. Al-Dmour H, Masa'deh R, Salman A, Abuhashesh M, Al-Dmour R. Influence of social media platforms on public health protection against the COVID-19 pandemic via the mediating effects of public health awareness and behavioral changes: integrated model. J Med Internet Res 2020 Aug 19;22(8):e19996 [FREE Full text] [doi: 10.2196/19996] [Medline: 32750004]
108. Abuhashesh MY, Al-Dmour H, Masa'deh RE, Salman A, Al-Dmour R, Boguszewicz-Kreft M, et al. The role of social media in raising public health awareness during the pandemic COVID-19: an international comparative study. Informatics 2021 Nov 18;8(4):80. [doi: 10.3390/informatics8040080]
109. Codd M, Barros H, Davidovitch N, Razum O, Mabhala M, Kostoulas P, et al. ASPHER statement: a new public health curriculum for a "new normal". Public Health Rev 2023 Nov 15;44:1606539 [FREE Full text] [doi: 10.3389/phrs.2023.1606539] [Medline: 38034595]
110. Mir EA. The lingering challenge: addressing vaccine hesitancy in a post-pandemic world. Cent Eur J Public Health 2023 Dec 30;31(4):296-299 [FREE Full text] [doi: 10.21101/cejph.a7593] [Medline: 38309708]
111. Chen X, Lee W, Lin F. Infodemic, institutional trust, and COVID-19 vaccine hesitancy: a cross-national survey. Int J Environ Res Public Health 2022 Jun 30;19(13):8033 [FREE Full text] [doi: 10.3390/ijerph19138033] [Medline: 35805691]
112. Rozek LS, Jones P, Menon A, Hicken A, Apsley S, King EJ. Understanding vaccine hesitancy in the context of COVID-19: the role of trust and confidence in a seventeen-country survey. Int J Public Health 2021 May 14;66:636255. [doi: 10.3389/ijph.2021.636255] [Medline: 34744589]

XSL•FO
RenderX

113. Carpiano RM, Callaghan T, DiResta R, Brewer NT, Clinton C, Galvani AP, et al. Confronting the evolution and expansion of anti-vaccine activism in the USA in the COVID-19 era. Lancet 2023 Mar;401(10380):967-970. [doi: 10.1016/s0140-6736(23)00136-8]
114. Schönweitz F, Eichinger J, Kuiper JM, Ongolly F, Spahl W, Prainsack B, et al. The social meanings of artifacts: face masks in the COVID-19 pandemic. Front Public Health 2022 Apr 14;10:829904 [FREE Full text] [doi: 10.3389/fpubh.2022.829904] [Medline: 35493352]
115. Taylor S, Asmundson GJ. Negative attitudes about facemasks during the COVID-19 pandemic: the dual importance of perceived ineffectiveness and psychological reactance. PLoS One 2021 Feb 17;16(2):e0246317 [FREE Full text] [doi: 10.1371/journal.pone.0246317] [Medline: 33596207]
116. Shim JM. Mask sociology as a way of action theory: voices of the face mask among social individuals in the COVID-19 masquerade in South Korea. PLoS One 2023 Nov 2;18(11):e0293758 [FREE Full text] [doi: 10.1371/journal.pone.0293758] [Medline: 37917775]
117. Rosenfeld DL, Balcetis E, Bastian B, Berkman ET, Bosson JK, Brannon TN, et al. Psychological science in the wake of COVID-19: social, methodological, and metascientific considerations. Perspect Psychol Sci 2022 Mar;17(2):311-333 [FREE Full text] [doi: 10.1177/1745691621999374] [Medline: 34597198]
118. Syed Q, Sopwith W, Regan M, Bellis MA. Behind the mask. Journey through an epidemic: some observations of contrasting public health responses to SARS. J Epidemiol Community Health 2003 Nov 01;57(11):855-856 [FREE Full text] [doi: 10.1136/jech.57.11.855] [Medline: 14600109]
119. Buregyeya E, Mitchell EM, Rutebemberwa E, Colebunders R, Criel B, Kiguli J, et al. Acceptability of masking and patient separation to control nosocomial tuberculosis in Uganda: a qualitative study. J Public Health 2012 Mar 15;20(6):599-606. [doi: 10.1007/S10389-012-0503-1]
120. Unlu A, Truong S, Tammi T, Lohiniva AL. Exploring political mistrust in pandemic risk communication: mixed-method study using social media data analysis. J Med Internet Res 2023 Oct 20;25:e50199 [FREE Full text] [doi: 10.2196/50199] [Medline: 37862088]
121. Boulos L, Curran JA, Gallant A, Wong H, Johnson C, Delahunty-Pike A, et al. Effectiveness of face masks for reducing transmission of SARS-CoV-2: a rapid systematic review. Philos Trans A Math Phys Eng Sci 2023 Oct 09;381(2257):20230133 [FREE Full text] [doi: 10.1098/rsta.2023.0133] [Medline: 37611625]
122. Kisa S, Kisa A. A comprehensive analysis of COVID-19 misinformation, public health impacts, and communication strategies: scoping review. J Med Internet Res 2024 Aug 21;26:e56931 [FREE Full text] [doi: 10.2196/56931] [Medline: 39167790]
123. Kaur M, Coppeta L, Olesen OF. Vaccine hesitancy among healthcare workers in Europe: a systematic review. Vaccines (Basel) 2023 Oct 28;11(11):1657 [FREE Full text] [doi: 10.3390/vaccines11111657] [Medline: 38005989]
124. Magid A, Hussein K, Dabaja-Younis H, Szwarcwort-Cohen M, Almog R, Mekel M, et al. Knowledge of the serological response to the third BNT162b2 vaccination may influence compliance of healthcare workers to booster dose. Antibodies (Basel) 2024 Aug 01;13(3):63 [FREE Full text] [doi: 10.3390/antib13030063] [Medline: 39189234]
125. Walkowiak MP, Domaradzki J, Walkowiak D. Are we facing a tsunami of vaccine hesitancy or outdated pandemic policy in times of omicron? Analyzing changes of COVID-19 vaccination trends in Poland. Vaccines (Basel) 2023 Jun 05;11(6):1065 [FREE Full text] [doi: 10.3390/vaccines11061065] [Medline: 37376454]
126. Instagram - statistics and facts. Statista. URL: https://www.statista.com/topics/1882/instagram/ [accessed 2025-01-22]

Abbreviations

API: application programming interface
BERT: Bidirectional Encoder Representations from Transformers
CDC: Centers for Disease Control and Prevention
c-TF-IDF: class-based term frequency–inverse document frequency
MH: Ministry of Health
NHS: National Health Service
TF-IDF: term frequency–inverse document frequency
WHO: World Health Organization

XSL•FO
RenderX

Edited by T Mackey; submitted 25.12.24; peer-reviewed by N Hu, P Chidipudi, M Raimi, S Kath, PK Taisuwan; comments to author 13.01.25; revised version received 26.01.25; accepted 13.04.25; published 02.05.25.

Please cite as:
Paradise Vit A, Magid A
Exploring Topics, Emotions, and Sentiments in Health Organization Posts and Public Responses on Instagram: Content Analysis
JMIR Infodemiology 2025;5:e70576
URL: https://infodemiology.jmir.org/2025/1/e70576
doi:10.2196/70576
PMID:

©Abigail Paradise Vit, Avi Magid. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 02.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Shifting Narratives in Media Coverage Across a Decade of Drug Discourse in the Philadelphia Inquirer: Qualitative Sentiment Analysis

Layla Bouzoubaa[1], MSPH; Ramtin Ehsani[1], BS; Preetha Chatterjee[1], PhD; Rezvaneh Rezapour[1], PhD

College of Computing & Informatics, Drexel University, Philadelphia, PA, United States

Corresponding Author:
Layla Bouzoubaa, MSPH
College of Computing & Informatics
Drexel University
3675 Market St., Rm. 1120
Philadelphia, PA, 19104
United States
Phone: 1 215 895 2487
Email: lb3338@drexel.edu

Abstract

Background: The media has immense power in shaping public narratives surrounding sensitive topics such as substance use. Its portrayals can unintentionally fuel harmful stereotypes and stigma, negatively impacting individuals struggling with addiction, influencing policy decisions, and hindering broader public health efforts.

Objective: This study aimed to examine how the regional newspaper, The Philadelphia Inquirer, covered events related to illicit drug use between 2013 and 2022, focusing on linguistic patterns and themes associated with specific types of substances.

Methods: We collected a dataset of 157,476 articles published in The Philadelphia Inquirer between 2013 and 2022 and categorized mentioned substances into 8 classes: stimulants, narcotics, cannabis, hallucinogens, depressants, designer drugs, drugs of concern, and treatment medications. From these 157,476 articles, we identified 3661 (2.32%) that mentioned at least 1 substance with potential for misuse. Using dynamic topic modeling, we analyzed thematic evolution in coverage across different drug classes. We then applied aspect-based sentiment analysis to extract the most significant phrases mentioned in each distinct drug class annually and examined the sentiments around these aspects to understand shifting discourse patterns.

Results: Cannabis (1575/3661, 43.02%) and narcotics (1361/3661, 37.17%) dominated the coverage, with 2018 showing peak drug-related reporting (666/3661, 18.19%). Our substance co-occurrence analysis revealed that heroin was most frequently discussed alongside treatment medications (methadone, naloxone, and buprenorphine), reflecting evolving approaches to opioid use disorder. Topic modeling revealed distinct themes across drug classes: legislative and medical aspects dominated cannabis coverage, while narcotics coverage focused heavily on overdose deaths and safe injection sites, particularly during 2017 to 2018. Stimulant coverage centered on feature news and crime-related reporting, while treatment coverage showed an increasing focus on overdose prevention by 2021. The aspect-based sentiment analysis showed that 74.3% (165/222) of extracted aspects were portrayed negatively across all drug classes, with narcotics maintaining consistently negative sentiment throughout the period. However, some drug classes showed notable evolution: hallucinogens demonstrated a marked shift in sentiment score (SS) from negative coverage in 2013 (–0.79 SS) to positive coverage of therapeutic applications by 2021 (+0.47 SS), while cannabis coverage reflected complex societal debates, with industry and business aspects showing strong positive sentiment score peaks (0.64 SS in 2019) even as legislation and policy aspects remained volatile (–0.76 SS in 2013 to 0.61 SS in 2019 and declining to –0.31 SS by 2022).

Conclusions: Our analysis revealed a predominance of negative and punitive language in drug-related news coverage, with limited representation of harm reduction principles. While some drug classes, particularly cannabis and hallucinogens, saw evolving narratives toward medical applications and policy reform, coverage of narcotics remained primarily focused on crime and overdose. These findings suggest a need for more balanced reporting that incorporates harm reduction perspectives and avoids potentially stigmatizing language when covering substance use disorders.

(JMIR Infodemiology 2025;5:e56004) doi:10.2196/56004

XSL•FO
RenderX

KEYWORDS

natural language processing; drug use; media analysis; sentiment analysis; infodemiology; news

Introduction

Background

The relationship between media representation and public perception is a complex interplay that has profound implications for societal attitudes and policies [1-3]. Media, as a primary source of information for many, has the power to shape, reinforce, or challenge societal norms and beliefs [4,5]. When it comes to issues of public health, such as substance use disorder (SUD), the media's portrayal can either support evidence-based interventions or perpetuate misconceptions and stigmas [6].

Previous research underscores the media's powerful role in shaping perceptions about drugs and issues related to SUD [7-10]. For instance, in the late 1990s in Australia, a program aimed at controlling heroin-related issues faced political defeat, largely due to the media's negative portrayal of heroin users as "deviants" [11]. Similarly, exaggerated media narratives around bath salts overshadowed clinical studies, leading to their prohibition [12,13]. Research by Denham [14] further highlights the media's propensity to motivate moral panics about drugs, even when actual use rates remain stable [14].

Moreover, the media's portrayal of stigmatized subjects, such as illicit drug use, can lead to stigma, discrimination, and reluctance to seek treatment [15,16]. For example, the media often emphasizes punitive measures against users and dealers, potentially influencing public attitudes and behaviors [17-19]. Caburnay et al [16] also found that media coverage can sway individual health behaviors, suggesting its potential impact on attitudes toward drug use.

Harm reduction, a public health strategy, seeks to mitigate the adverse effects of drug use without mandating complete cessation [20]. However, media coverage often lacks a comprehensive view of this approach. A study on Canadian news outlets revealed a tendency to focus on singular, controversial harm reduction strategies [21]. However, when conveyed appropriately, harm reduction messages can significantly diminish the societal repercussions of drug use, especially when they resonate with the audience's values and present a holistic solution [22]. These examples show the importance of critically assessing media representations of substance use, ensuring that they are both accurate and holistic.

Study Rationale and Aims

The pervasive issue of illicit drug use and its associated consequences has been a topic of concern in many urban areas across the United States. The city of Philadelphia has been at the forefront of this crisis, dealing with the devastating effects of drug addiction and its ripple effects on the community. Philadelphia has the second highest rate of overdose deaths (71 per 100,000) in the country [23] and is commonly known as one of the epicenters of the "opioid epidemic" in the United States [24]. In 2021, there were >1250 unintentional drug overdose deaths in Philadelphia [25], the highest number on record. Most of these deaths were caused by opioids, such as heroin and fentanyl [26]. Despite the city's efforts to address the problem, such as expanding access to social services in the hardest-hit neighborhoods, several challenges remain, including the lack of affordable housing and the stigma associated with drug use and SUDs [27]. These challenges can make it difficult for people to seek help [28].

Given Philadelphia's position at the epicenter of the US drug crisis, examining local media coverage becomes crucial for understanding how drug-related issues are presented to the public. The Philadelphia Inquirer, established in 1829, is the city's largest newspaper, with a daily circulation of >100,000 readers, and serves as a primary source of local news coverage. As the third oldest surviving daily newspaper in the United States, it has historically provided extensive coverage of the city's social issues, including public health challenges such as substance use. This study analyzed a decade of coverage (2013 to 2022) from The Philadelphia Inquirer, a period that encompasses significant shifts in drug policy and public health approaches, including the rise of the opioid crisis, the emergence of fentanyl as a dominant threat, and evolving attitudes toward cannabis legalization. This time frame allows us to capture both the intensification of the opioid crisis and the evolution of public health responses to substance use. Our study aimed to understand how The Philadelphia Inquirer's coverage of illicit substances has evolved over time, with particular attention to their framing and contextual portrayal. Specifically, our study addresses 3 research questions (RQs):

- RQ1: How have the characteristics and nuances of mentions related to commonly abused substances evolved over the decade?
- RQ2: In the context of these substances, which overarching themes and subtopics have dominated the discourse, and how have they shifted over time?
- RQ3: What are the multifaceted dimensions, sentiments, and narratives associated with the portrayal of commonly abused substances?

Through this analysis, we seek to understand how media coverage may influence public perception and policy responses to substance use in a major urban center particularly affected by the ongoing drug crisis.

Prior Work

Few studies have investigated the media coverage of illicit drugs and their change over time, particularly in the United States [15,16,29]. A recent study [7] investigating the coverage of illicit drug use by a Canadian newspaper found that it focused on basic social representations, such as attribution of the responsibility for the opioid crisis to a few collectives, including pharmaceutical companies and physicians, or the overall drug supply, and that a shift to less stigmatizing language could positively influence public perception. The study suggests that closer collaboration between the media and the research community is needed to achieve a better understanding of the issue. Ketamine, a psychoactive drug that has attained positive

XSL•FO
RenderX

outcomes in treating severe, treatment-resistant depression, was found to be most associated with themes of abuse, legality, and clinical utility as an anesthetic when it was reported in North American news outlets between 2000 and 2015 [30]. This finding suggests that changes in news media reporting could influence how substances, such as ketamine, are received as viable treatment options, and that guidance is required for journalists on objective reporting of medical research findings.

Google News Archives and cause-of-death records published by the National Center for Health Statistics between 1999 and 2005 were explored for patterns in mentions of the opioid epidemic, and a noteworthy correlation was discovered between the number of news articles and the rates of opioid-related overdose deaths over time [31].

To comprehend the dissimilarities and similarities of media coverage based on drug types, Hayden Griffin et al [19] analyzed 487 news articles published in a national media outlet in Malaysia over a 2-year period. They found that amphetamines, opiates, and cannabis received most of the media coverage, and the discussion of these drugs was primarily in relation to criminal justice. Similarly, Hughes et al [18] conducted a framing analysis of Australian news media coverage of illicit drugs between 2003 and 2008. The study found that criminal justice topics are dominant, but nonlegal issues are also highlighted. The media frames can differ between drugs, with amphetamines portrayed most negatively and cocaine most neutrally. Their findings suggest that sensationalized reporting of drug events is more prevalent during specific episodes and may not be representative of the norm. Our findings do not indicate that The Philadelphia Inquirer uses tactics that sensationalize the current state of drug use in Philadelphia; however, there is a lack of messaging framed around harm reductionist principles or that normalizes evidence-based approaches to substance use.

Methods

Data Collection and Substance Classification

We collected 157,476 news articles from The Philadelphia Inquirer from January 1, 2013, to December 31, 2022, using the *ProQuest* database platform. Each article in our dataset comprises the complete text along with metadata attributes, such as publication date, author, title, links, and subject keywords.

Drug Class Assignment

We compiled our list of substances using the Commonly Used Drugs Charts from the National Institute on Drug Abuse [32], which provides an overview of frequently used legal and illegal drugs. We then classified these substances according to the categories established by the Drug Enforcement Administration [33]. To ensure accuracy in our classifications, we manually reviewed and removed ambiguous terms with multiple meanings (eg, the term “pot” could refer to “flowering pot” or “cooking pot” in addition to marijuana). The Drug Enforcement Administration classifies drugs into 9 categories:

1. Cannabis: marijuana is a mind-altering (psychoactive) drug produced by the cannabis sativa plant. “Cannabis” and “marijuana” are often used interchangeably, resulting in the appearance of cannabis as both the drug class and the drug name.
2. Depressants: these are known to induce sleep, relieve anxiety and muscle spasms, and prevent seizures (eg, barbiturates and sedative-hypnotic substances, such as gamma-hydroxybutyrate).
3. Designer drugs: these are produced illicitly with a slightly altered chemical structure to mimic the pharmacological effects of controlled substances (eg, synthetic marijuana or synthetic cathinones).
4. Drugs of concern: these are unregulated drugs that can be harmful if abused (eg, kratom and xylazine).
5. Hallucinogens: these are derived from plants and fungi and renowned for their capacity to modify human perception and mood (eg, lysergic acid diethylamide, mushrooms, and ecstasy).
6. Narcotics: these refer to opium, opium derivatives, and their semisynthetic substitutes. “Opioid” is a more current and precise term to describe these drugs (eg, heroin, OxyContin, codeine, morphine, and fentanyl).
7. Stimulants: these are drugs that accelerate the body’s functions (eg, methamphetamine, cocaine, and amphetamines).
8. Treatment: these are substances aiding the treatment of opioid addiction (eg, methadone, Suboxone, and naloxone).
9. Miscellaneous: these are substances that can be abused but do not belong to any class (eg, steroids).

Using our mapped list of drug names, we identified news articles that mention any of these drugs in either their title or text. Subsequently, we grouped these articles into clusters based on the respective drug classes. If an article mentioned drugs from multiple classes, we assigned it to all relevant drug classes.

Topic Modeling

To investigate how drug-related topics evolved over time in news coverage, we used dynamic topic modeling [34] using bidirectional encoder representations from transformers topic modeling technique (BERTopic) [35], an advanced method that identifies themes in a corpus of texts. Unlike traditional approaches, such as latent Dirichlet allocation, BERTopic better captures the context and meaning of words, producing more coherent and interpretable results [35]. We analyzed each drug class separately to understand how the coverage of different substances changed over time. Before topic modeling, we implemented several text preprocessing steps to improve the quality of our analysis. We standardized the text by converting all characters to lowercase and expanding contractions to their full form (eg, “don't” to “do not”). We then removed elements that could interfere with the analysis, such as URLs and common stop words (eg, “the,” “and,” as well as “or”) that do not carry significant meaning. These preprocessing steps reduced noise in the data and improved the model’s ability to identify meaningful patterns and topics. A technical limitation of our analysis with bidirectional encoder representations from transformers (BERT) is its maximum token limit of 512 tokens, while the median length of our articles was 749 (SD 573; IQR 511-1014) tokens. However, this limitation aligned well with journalistic writing conventions. News articles typically follow

XSL•FO
RenderX

the inverted pyramid style [36], in which the most newsworthy information appears at the beginning of an article, followed by supporting details. This structure allowed us to confidently use the first 512 tokens of each article, as they typically contained the most salient information and key themes of the full text. This approach allowed us to work within the token limit and retain the essential content for our analysis.

To ensure high-quality results, we refined our analysis through multiple iterations following established topic modeling validation approaches [37,38]. We set specific parameters to optimize topic coherence, including capping the top number of words per topic at 10, and looked for words and phrases that commonly appear together (ie, n-grams) to provide richer context. We evaluated the quality of our results using 3 criteria: topic coherence (measured by a coherence score with a threshold >0.5), topic distinctiveness (minimal overlap in the top 10 keywords between topics), and human interpretability (clear thematic focus discernible from top keywords) [39]. Through iterative refinement, we systematically reduced the number of topics to 4 per drug class based on achieving maximum coherence scores, minimal thematic overlap, and clear interpretability [40].

To characterize and validate the topics generated by BERTopic, we conducted a detailed analysis of the representative articles. For each topic identified, 2 researchers independently examined the top 20 articles with the highest topic representation scores, noting common themes and key narratives. This human analysis helped us understand how the algorithmically derived topics appeared in actual news coverage and enabled us to develop more meaningful topic labels. For instance, when examining articles highly associated with what BERTopic initially labeled as “law enforcement” within the cannabis class, our review revealed a more nuanced focus on legislative processes and policy reform, leading us to refine this topic’s characterization. Disagreements in topic interpretation were resolved through discussion with a third researcher until a consensus was reached. This process of combining computational topic modeling with human analysis of representative articles enhanced our understanding of the thematic evolution in drug-related coverage.

Aspect-Based Sentiment Analysis

We used aspect-based sentiment analysis rather than traditional sentiment analysis because drug-related articles often contain multiple viewpoints and sentiments within the same text. While traditional sentiment analysis assigns a single sentiment score to an entire document, aspect-based sentiment analysis allows us to capture nuanced sentiments associated with specific aspects of drug discourse, for instance, distinguishing between sentiments about law enforcement approaches versus public health initiatives, even within the same article. To extract these aspects from the articles, we used key phrase extraction to identify words and phrases with the highest significance and relevance within each news article. As key phrases encapsulate the essence of an entire document, they serve as essential tools for retrieving critical information from large and diverse datasets [41].

While various tools and techniques are available for key phrase extraction from documents, it is notable that most of these models generally focus on the statistical properties of text rather than semantic similarity [42]. We selected KeyBERT [42] for phrase extraction because of its ability to capture context and identify meaningful phrases. The model combines 3 key components: the TextRank algorithm [43] (which ranks phrases based on their importance within the text), BERT embeddings [44] (numerical representations that capture the meaning of words and phrases), and cosine similarity [42] (a measure of how closely related 2 pieces of text are). To improve accuracy, we used vectorizers (tools that convert text into structured patterns) that analyze grammatical patterns to ensure that extracted phrases were linguistically meaningful. We analyzed articles within each drug class year by year from 2013 to 2022, allowing us to track how key phrases and their associated topics evolved over time. For each article, we compared the extracted phrases with the overall article content using cosine similarity scores, retaining only those phrases that reached a similarity threshold of 40%. We determined this threshold through careful testing, finding that it effectively balanced between capturing relevant phrases and excluding irrelevant or redundant information.

To analyze sentiment, we first located all paragraphs containing each identified aspect, allowing us to focus on relevant content and reduce noise. We then used the NewSentiment model [45], which is specifically designed for analyzing news content, to determine whether the discussion of each aspect was positive, negative, or neutral. This model was pretrained on a large dataset of political news articles [45], making it well suited for analyzing journalistic writing. For each aspect, we calculated a normalized sentiment score (SS) by subtracting the negative sentiment probability from the positive sentiment probability and dividing by their sum, resulting in scores ranging from –1 (most negative) to 1 (most positive). To facilitate interpretation and identify broader patterns, we manually grouped semantically similar aspects into thematic categories for each drug class (eg, “Legislation/Policy,” “Medical Use” for cannabis; “Overdose Prevention,” “Opioid Treatment” for treatment drugs). We then calculated mean sentiment scores and CIs for each aspect group per year, allowing us to track how sentiment toward different themes evolved over time.

Ethical Considerations

This study analyzed publicly available news articles from The Philadelphia Inquirer accessed through the ProQuest database platform. No data involving human participants were collected, analyzed, or stored during this research. All article data were handled in aggregate form, and no personally identifiable information was extracted or analyzed.

Results

Substance Distribution and Co-Occurrence Patterns

Of the 157,476 articles analyzed, 3661 (2.32%) referenced at least one drug class. The distribution of these articles across drug classes and years is shown in Table 1. We observed a steady increase in explicit substance mentions, followed by a decline (COVID-19 pandemic effect, presumably around 2020)

XSL•FO
RenderX

in The Philadelphia Inquirer over the past decade. Cannabis was the most frequently discussed drug class, followed by narcotics (ie, opioids). The annual count of news articles indicates that 2018 saw the highest frequency of drug-related mentions. Narcotics were the second most discussed drug class and the most prevalent in 2017 and 2018, suggesting that the opioid crisis was a major focus of news coverage during those years.

Further analysis of drug mentions by class shows that cannabis is primarily associated with marijuana (Figure 1). Despite the controversial origins of the term "marijuana" [46], it has been consistently used in news articles over the years. Among the depressant drugs, Xanax and Ambien were the most frequently mentioned, likely due to their widespread use for anxiety and sleep disorders and their potential for abuse. Similarly, cocaine and ecstasy were the most frequently mentioned drugs in the stimulant and hallucinogen drug classes, respectively, indicating their enduring popularity.

Table 1. Distribution of articles per drug classes by year.

Year	Cannabis (n=1575, 43.02%)	Depressants (n=106, 2.89%)	DD[a] (n=16, 0.43%)	DC[b] (n=11, 0.3%)	Hallucinogens (n=137, 3.74%)	Narcotics (n=1361, 37.17%)	Stimulants (n=627, 17.13%)	Treatment (n=449, 12.26%)	Total[c] (n=3661, 100%)
2013	138 (8.76)	23 (21.7)	3 (18.75)	2 (18.18)	20 (14.6)	*142*[d] (10.43)	108 (17.22)	16 (3.56)	452 (100)
2014	*146 (9.27)*	10 (9.43)	0 (0)	1 (9.1)	10 (7.3)	117 (8.6)	64 (10.2)	21 (4.68)	369 (100)
2015	*122 (7.75)*	6 (5.66)	0 (0)	0 (0)	15 (10.95)	117 (8.6)	49 (7.8)	30 (6.68)	339 (100)
2016	*150 (9.52)*	12 (11.32)	0 (0)	0 (0)	19 (13.86)	132 (9.7)	63 (10.05)	43 (9.58)	419 (100)
2017	194 (12.32)	12 (11.32)	0 (0)	0 (0)	9 (6.57)	*223 (16.39)*	53 (8.45)	78 (17.37)	569 (100)
2018	221 (14.03)	14 (13.2)	8 (50)	6 (54.54)	18 (13.14)	*228 (16.75)*	73 (11.64)	98 (21.82)	666 (100)
2019	*231 (14.67)*	10 (9.43)	4 (25)	2 (18.18)	11 (8.03)	148 (10.87)	90 (14.35)	48 (10.69)	544 (100)
2020	*113 (7.17)*	7 (6.6)	0 (0)	0 (0)	7 (5.1)	69 (5.07)	48 (7.67)	31 (6.9)	275 (100)
2021	*118 (7.49)*	7 (6.6)	0 (0)	0 (0)	16 (11.68)	86 (6.32)	41 (6.54)	41 (9.13)	309 (100)
2022	*142 (9.02)*	5 (4.72)	1 (6.25)	0 (0)	12 (8.76)	99 (7.27)	38 (6.06)	43 (9.58)	340 (100)
Total	1575 (100)	106 (100)	16 (100)	11 (100)	137 (100)	1361 (100)	627 (100)	449 (100)	—[e]

[a]DD: designer drug.

[b]DC: drug of concern.

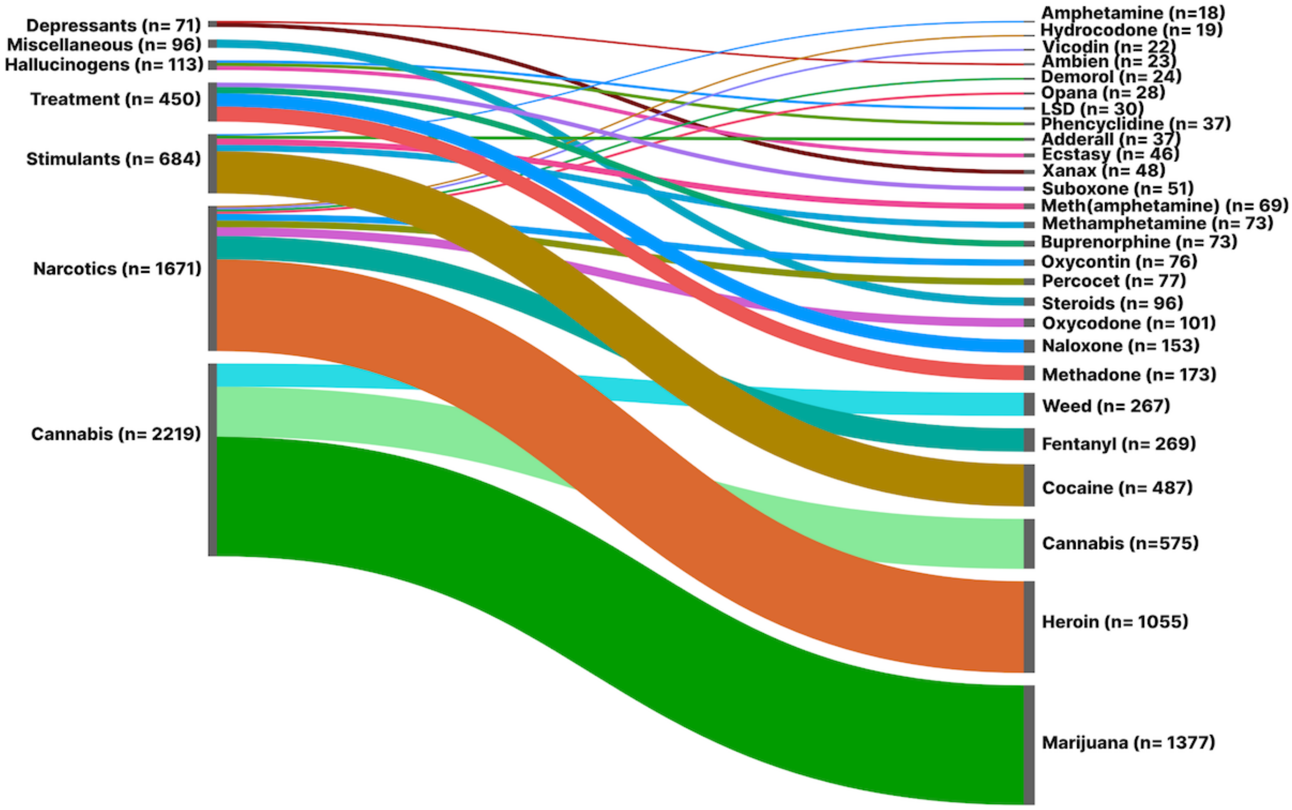
[c]The total sum of all drug class mentions exceeds the number of unique articles containing substance mentions (n=3661) because individual articles may mention substances from multiple drug classes.

[d]Italicization indicates the most frequent drug class.

[e]Not applicable.

XSL•FO
RenderX

Figure 1. Number of substances per drug class in our dataset. LSD: lysergic acid diethylamide.



Our co-occurrence analysis revealed an association between the frequency with which certain substances are discussed and provisional overdose and use patterns. Table 2 shows that heroin, one of the most prevalent substances in Philadelphia, was overtaken by fentanyl in 2018, buprenorphine in 2021, and buprenorphine and fentanyl in 2022. Furthermore, our analysis indicated that narcotics and treatment drugs, such as methadone, naloxone, and buprenorphine, were the second most commonly colocated drugs in articles. Treatment drugs such as methadone, naloxone, and buprenorphine are viewed as positive and proactive responses to opioid use disorder (OUD) [47]. Buprenorphine, a safe and effective treatment for OUD, emerged in 2020 alongside heroin, followed by cannabis in 2021 and 2022.

XSL•FO
RenderX

Table 2. Most frequently co-occurring drugs by year.

Year and drugs	Value, n (%)
2013 (n=452)	
Heroin and methadone	52 (11.5)
Heroin and cocaine	45 (10)
Barbiturates and cannabis	16 (3.5)
2014 (n=369)	
Heroin and naloxone	42 (11.4)
Marijuana and cocaine	36 (9.8)
Heroin and cocaine	34 (9.2)
2015 (n=339)	
Heroin and marijuana	40 (11.8)
Heroin and methadone	19 (5.6)
Heroin and methamphetamine	11 (3.2)
2016 (n=419)	
Heroin and marijuana	51 (12.2)
Heroin and cocaine	45 (10.7)
Heroin and naloxone	44 (10.5)
2017 (n=569)	
Heroin and methadone	157 (27.6)
Heroin and cocaine	90 (15.8)
Heroin and naloxone	80 (14.1)
2018 (n=666)	
Fentanyl and cocaine	270 (40.5)
Heroin and cocaine	194 (29.1)
Heroin and marijuana	127 (19.1)
2019 (n=544)	
Heroin and marijuana	115 (21.2)
Heroin and naloxone	77 (14.2)
Heroin and cocaine	71 (13.1)
2020 (n=275)	
Heroin and methadone	198 (72)
Heroin and buprenorphine	81 (29.5)
Fentanyl and cocaine	53 (19.3)
2021 (n=309)	
Buprenorphine and cannabis	108 (35)
Heroin and cannabis	78 (25.2)
Heroin and marijuana	41 (13.3)
2022 (n=340)	
Buprenorphine and cannabis	351 (103.2[a])
Buprenorphine and marijuana	153 (45)
Fentanyl and cocaine	114 (33.5)

[a]Percentages are calculated relative to the total number of drug-mentioning articles that year. Values may exceed 100% as substances can co-occur multiple times within individual articles.

XSL•FO
RenderX

Thematic Evolution by Drug Class

Our analysis revealed distinct thematic patterns across different drug classes (Figure 2). For hallucinogen-related coverage, we identified 4 main topics: phencyclidine (PCP) and criminal justice, music and film, feature news (eg, a human-interest story on how a local who experiences posttraumatic stress disorder [PTSD] uses lysergic acid diethylamide to treat their symptoms [48]), and decriminalization and treatment. Criminal justice coverage dominated early discussions in 2013 but showed a marked decline until a substantial spike in 2017. Feature news emerged as a consistent theme throughout the period, peaking in 2016, while coverage of hallucinogens in music and film maintained a relatively stable presence. Notably, discussions of decriminalization and treatment gained increasing prominence from 2019 onward, becoming a dominant theme by 2022.

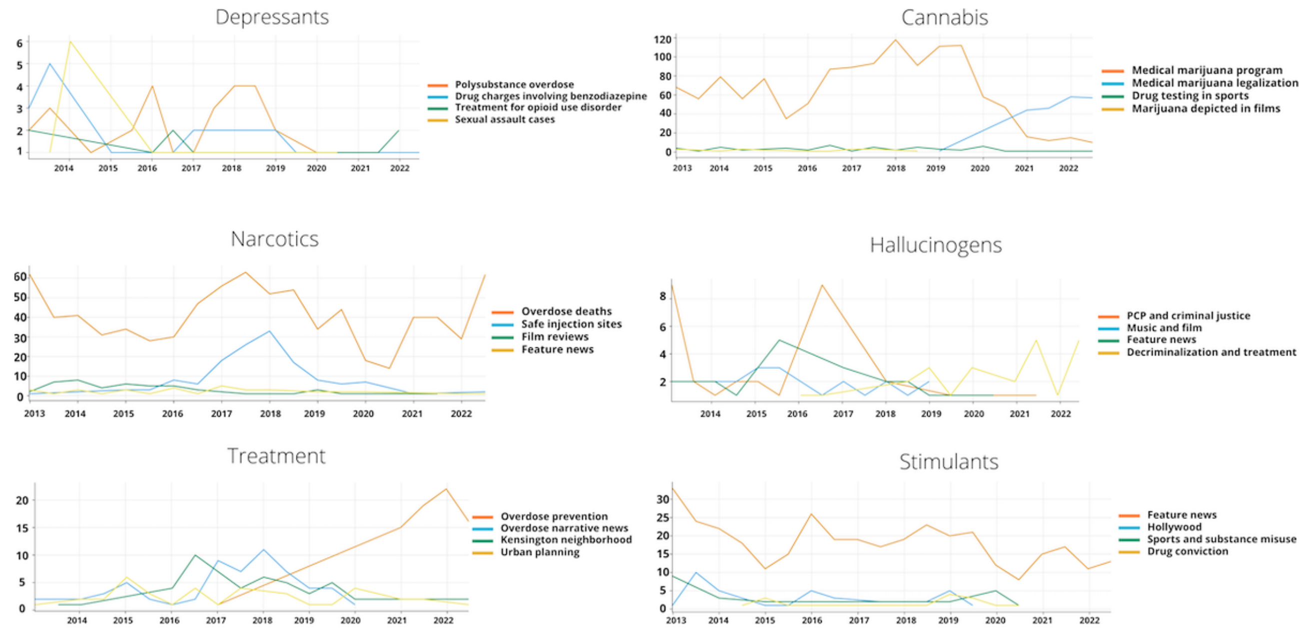


Figure 2. Topic evolution per class of drug per year. PCP: phencyclidine.

Articles discussing depressants centered on 4 key themes: polysubstance overdose, drug charges involving benzodiazepine, treatment for OUD, and sexual assault cases. Coverage of sexual assault cases saw a dramatic spike in 2014 before steadily declining, while reporting on polysubstance overdose showed cyclical patterns with peaks in 2016 and 2018 to 2019. Benzodiazepine-related charges received heightened attention in early years (peaking in 2014) but diminished over time, while treatment discussions remained relatively steady before showing renewed interest in 2022.

Cannabis coverage demonstrated the most dramatic thematic shifts over the decade. The 4 main topics identified were the medical marijuana program, medical marijuana legalization, drug testing in sports, and marijuana depicted in films. Coverage of the medical marijuana program dominated from 2013 to 2020, peaking in 2018 with nearly 96% (120/125) of articles. However, a significant transition occurred around 2020, when legalization discussions began to overtake program-related coverage. By 2022, legalization had become the primary focus, with about 75% (60/75) of articles, while program-related coverage declined to approximately 20% (15/75) of articles. Sport-related drug testing and film depictions maintained a minimal but consistent presence throughout the period, rarely exceeding >5 articles per year.

Among articles that mentioned narcotics, we observed the following top 4 themes: overdose deaths, safe injection sites, film reviews, and feature news. Coverage around overdose deaths within the city consistently received the most coverage throughout the decade, starting at around 85% (60/70) of the articles in 2013, declining to approximately 66% (30/45) of articles through the mid-2010s, before showing significant spikes in 2018 (>60 articles) and 2022 (>60 articles). Coverage of safe injection sites emerged as a prominent theme around 2017, peaking in 2018 with >30 articles, before gradually declining through 2022. Film reviews (articles discussing a film where narcotics are present) and feature news maintained relatively low but steady coverage, rarely exceeding 5 articles per year.

Stimulant-related coverage revealed 4 distinct themes: feature news, Hollywood, sports and substance misuse, and drug conviction. Feature news dominated the coverage, with higher frequency in earlier years (>30 articles in 2013) compared to later years (approximately 15/15, 100% in 2022), with notable fluctuations including a peak of 80% (25/30) of articles in 2016. Coverage related to Hollywood and sports showed higher prominence in earlier years (2013 to 2014) but maintained minimal presence afterward. Drug conviction coverage remained consistently low throughout the period, with slight increases in 2015 and 2019.

Treatment-related coverage demonstrated evolving priorities through overdose prevention, overdose narrative news, the Kensington neighborhood, and urban planning. Most notably, coverage of overdose prevention showed a dramatic increase from 2019 onward, reaching >20 articles by 2021. Overdose narrative news peaked around 2018 to 2019 with about 10 articles per year, while coverage of the Kensington

XSL•FO
RenderX

neighborhood showed heightened attention in 2017 (10/15, 66% of articles). Urban planning maintained relatively consistent but low coverage throughout the period.

These patterns highlight how coverage of different drug classes responded to evolving public health concerns and policy initiatives. While narcotics coverage maintained a consistent focus on overdose deaths, treatment coverage showed a clear shift toward prevention strategies in recent years. Meanwhile, stimulant coverage demonstrated a broader focus on social and cultural aspects rather than public health concerns.

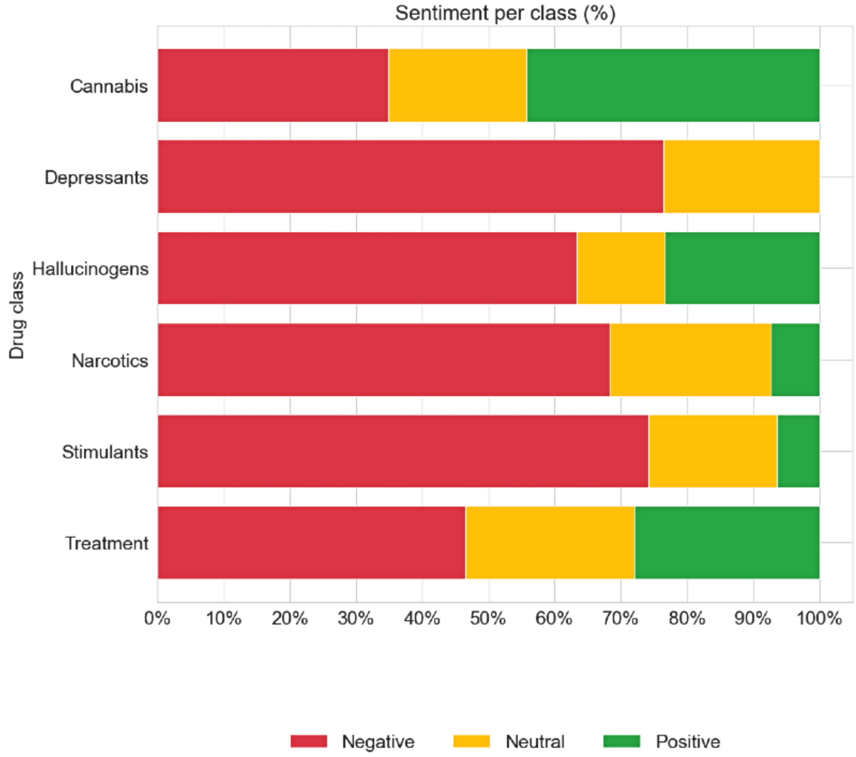
Sentiment Pattern and Aspect Analysis

Our sentiment analysis revealed distinct patterns across 3 dimensions: overall sentiment distribution by drug class, temporal changes in sentiment, and evolution of thematic aspect groups over time. The overall sentiment distribution (Figure 3) showed cannabis and hallucinogens having the highest proportion of positive aspects, while depressants and narcotics demonstrated predominantly negative coverage. Treatment-related coverage showed a more balanced distribution between positive and negative sentiments, reflecting the complex nature of intervention discussions.

Temporal sentiment patterns (Figure 4) revealed significant evolution in coverage tone across drug classes. Hallucinogens showed the most dramatic positive shift, moving from strongly negative coverage in 2013 to 2014 to increasingly positive coverage by 2022. Cannabis coverage demonstrated considerable volatility but maintained generally neutral to positive sentiment, while depressants consistently received the most negative coverage throughout the period.

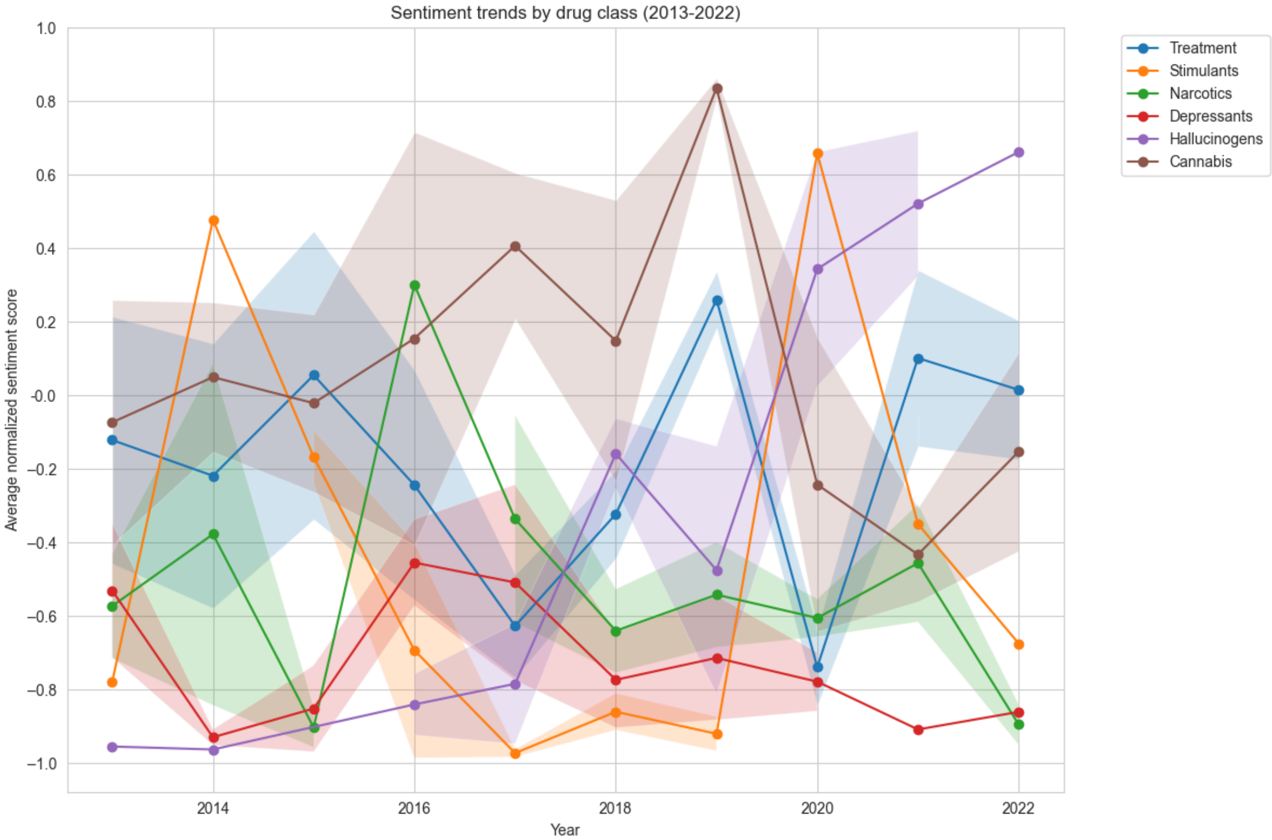
Analysis of thematic aspect groups over time (Table 3) provided deeper insight into these sentiment patterns (Figure 5). To facilitate interpretation and discussion, we grouped aspects into thematic categories within each drug class based on semantic similarity. As shown in Figure 5, for cannabis, industry and business aspects showed strong positive sentiment score (SS) peaks in 2016 (0.49) and 2019 (0.64), while legislation and policy aspects evolved from highly negative (–0.76 in 2013) to positive (0.61 in 2019) sentiment scores before declining again (–0.31 in 2022). Medical use maintained a relatively stable, slightly positive sentiment, while regulatory and enforcement aspects remained predominantly negative.

Figure 3. Proportion of aspects with positive, negative, and neutral sentiments per class.



XSL•FO
RenderX

Figure 4. Average sentiment scores for each drug class from 2013 to 2022. Sentiment scores are normalized to range from –1 (most negative) to 1 (most positive), with 0 representing neutral sentiment. Shaded areas represent CI.



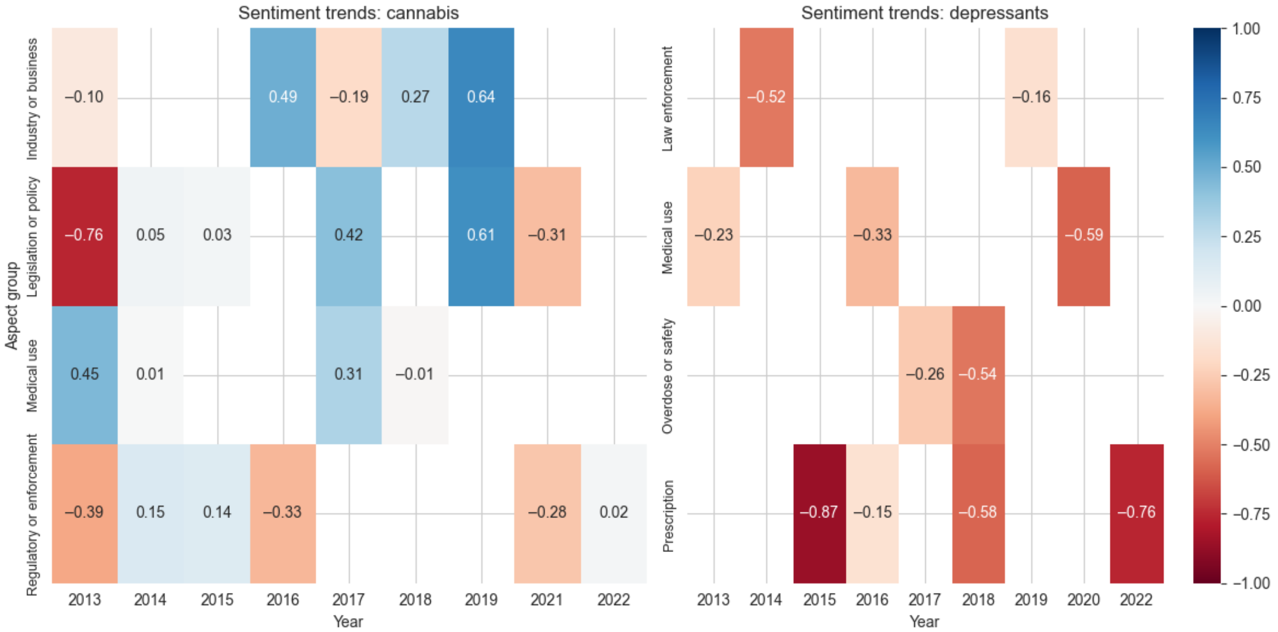
XSL•FO
RenderX

Table 3. Aspect groupings by drug class.

Drug class and aspect group	Representative aspects
Cannabis	
Legislation or policy	marijuana legalization, cannabis legislation, marijuana reform, marijuana laws
Medical use	medical marijuana, medical cannabis, medicinal marijuana, medical marijuana program
Industry or business	marijuana dispensaries, cannabis industry, cannabis investments, nascent marijuana industry
Regulatory or enforcement	marijuana convictions, marijuana offenses, recreational marijuana, cannabis impairment
Treatment	
Overdose prevention	naloxone spray, narcan, overdose drug, lifesaving drug, reversal medication
Opioid treatment	buprenorphine prescription, methadone programs, buprenorphine program, prescribing buprenorphine
Crisis response	overdose fatalities, drug costs, opioid crisis, opioid epidemic, drug response
Medical or clinical	opioid painkillers, opioid prescriptions, prescription drugs, opioid use disorder
Narcotics	
Overdose	overdose crisis, overdose victims, fatal overdose, overdose deaths, fentanyl overdoses
Law enforcement	drug possession, federal drug laws, drug convictions, border protection, drug trafficking
Public health	opioid epidemic, aids crisis, public drug use, neonatal abstinence syndrome
Policy or treatment	new jersey drug policy alliance, opioid litigation, drug laws ease, opioid drugmakers
Hallucinogens	
Medical or therapeutic	new psychedelic therapies, psychedelic therapeutics company, new psychiatric medicines
Law enforcement	drug charges, drug arrests, smuggled drugs, illegal psychedelic drugs
Substances	lsd, cannabis, marijuana, psychoactive compounds, crack cocaine
Social context	party drug, prevalent drug use, favorite drug conversation
Depressants	
Medical use	sleep drug zolpidem, insomnia drugs, medical marijuana, lorazepam, using ambien
Overdose or safety	xanax overdose, overdose deaths, overdose statistics, opioid overdoses
Prescription	prescription opioids, benzodiazepines, morphine, zaleplon prescription
Law enforcement	homicide case, county district attorney, montgomery county prosecutors
Stimulants	
Law enforcement	drug violations, cocaine possession, large cocaine distribution ring, drug trafficking
Health or treatment	addiction treatment program, few medication options, antidepressant, methamphetamine use
Crime related	crime spree, customs officials, probation violation
Substance specific	illicit fentanyl, cocaine use, lehtera cocaine, crack cocaine buyer, largest cocaine seizure

XSL•FO
RenderX

Figure 5. Sentiment scores across aspect groups for the drug classes cannabis and depressants. Values represent average sentiment scores ranging from –1 (most negative: red) to 1 (most positive: blue), with white indicating neutral sentiment.



The evolution of cannabis-related sentiment is particularly evident in key legislative and policy aspects (Table 4). Coverage of recreational marijuana use and decriminalization shifted from predominantly negative sentiment in 2013 (with quotes expressing opposition to decriminalization) to increasingly positive coverage by 2019, when discussions focused on potential economic and social benefits. Similarly, marijuana legalization aspects evolved from skeptical coverage in 2013 to 2014 to more supportive framing by 2019, although this positive sentiment declined in later years as implementation challenges emerged. This transformation in coverage tone reflects broader societal shifts in cannabis policy discourse over the decade. This evolution in cannabis coverage is further reflected in the terminology used, where aspects associated with "marijuana" often carried negative connotations, while "cannabis" was framed more positively. For example, the Marijuana Opportunity Reinvestment and Expungement Act's discussion in 2019 used the term "cannabis" when describing positive policy developments (Marijuana Opportunity Reinvestment and Expungement Act of 2019 [49]), while earlier negative coverage of decriminalization by New Jersey Governor Christie used "marijuana" (Inquirer Staff, June 18, 2015, NEW JERSEY: Sharp increase in backing for legalizing marijuana, the Philadelphia Inquirer; Multimedia Appendix 1). Within the cannabis class, "medicinal marijuana" consistently received more positive coverage than "recreational use."

Depressants showed consistently negative sentiment across all aspect groups, with prescription-related aspects showing particularly negative sentiment (–0.87 in 2015 and –0.76 in 2022). Law enforcement aspects improved slightly from –0.52 in 2014 to –0.16 in 2019, while medical use aspects grew increasingly negative over time, reaching –0.59 by 2020 (Figure 5). The overdose or safety category emerged with negative sentiment (–0.26 to –0.54) in later years, reflecting growing concerns about safety risks. The consistently negative coverage of depressants was particularly evident in discussions of combined drug use, especially regarding benzodiazepines with opioids and their associated overdose risks. The sentiment toward "insomnia drugs" notably deteriorated between 2013 and 2016, influenced by growing concerns about prescription policies and pharmaceutical companies' transparency in warning labels (Multimedia Appendix 1).

Hallucinogen coverage showed the most dramatic thematic evolution (Figure 6). While law enforcement aspects maintained strong negative sentiment (–0.79 in 2013 and –0.75 in 2016), medical or therapeutic aspects emerged in later years with notably positive sentiment (0.47 in 2021). As detailed in Multimedia Appendix 1, this shift is particularly evident in the coverage transition from negative associations with crack cocaine (2014) and cannabis (2018) to increasingly positive framing of medicinal cannabis (2019) and new psychedelic therapies (2021). This positive shift in medical or therapeutic aspects may be attributed to recent studies on psychedelic drugs for treating mental health disorders, such as treatment-resistant depression and PTSD. Substance-specific aspects showed gradual improvement from highly negative sentiment (–0.90 in 2014) to slightly negative sentiment (–0.09 in 2021), while social context aspects appeared sporadically with negative sentiment (–0.66 in 2019). Table 5 provides an exemplar excerpt from The Philadelphia Inquirer that reflects this evolution, particularly the positive sentiment regarding the therapeutic potential of hallucinogenic drugs for treating PTSD.

Building on these contrasting patterns, narcotics coverage maintained consistently negative sentiment across all aspect groups (Figure 6). Law enforcement aspects showed persistent negative sentiment, although improving slightly from –0.77 in 2015 to –0.54 in 2022. The overdose aspect group emerged in later years with moderately negative sentiment (–0.34 in 2020 and –0.36 in 2021), while public health aspects improved from –0.74 in 2018 to –0.38 in 2019. Policy or treatment aspects showed brief positive sentiment (0.24 in 2017) before turning

XSL•FO
RenderX

negative (–0.45 in 2018), reflecting the challenges in implementing treatment initiatives.

Similarly negative but with even stronger intensity, stimulant coverage demonstrated some of the most consistently negative sentiment patterns across all drug classes (Figure 7). Crime-related aspects maintained strongly negative sentiment throughout (–0.81 in 2013, –0.91 in 2017, and –0.63 in 2018), while law enforcement aspects showed similarly negative shifts (–0.93 to –0.66 between 2017 and 2019). Substance-specific coverage remained consistently negative (–0.84 to –0.85), although health or treatment aspects emerged in later years with slightly less negative sentiment (–0.22 in 2021).

Table 4. Changes of tone toward marijuana decriminalization.

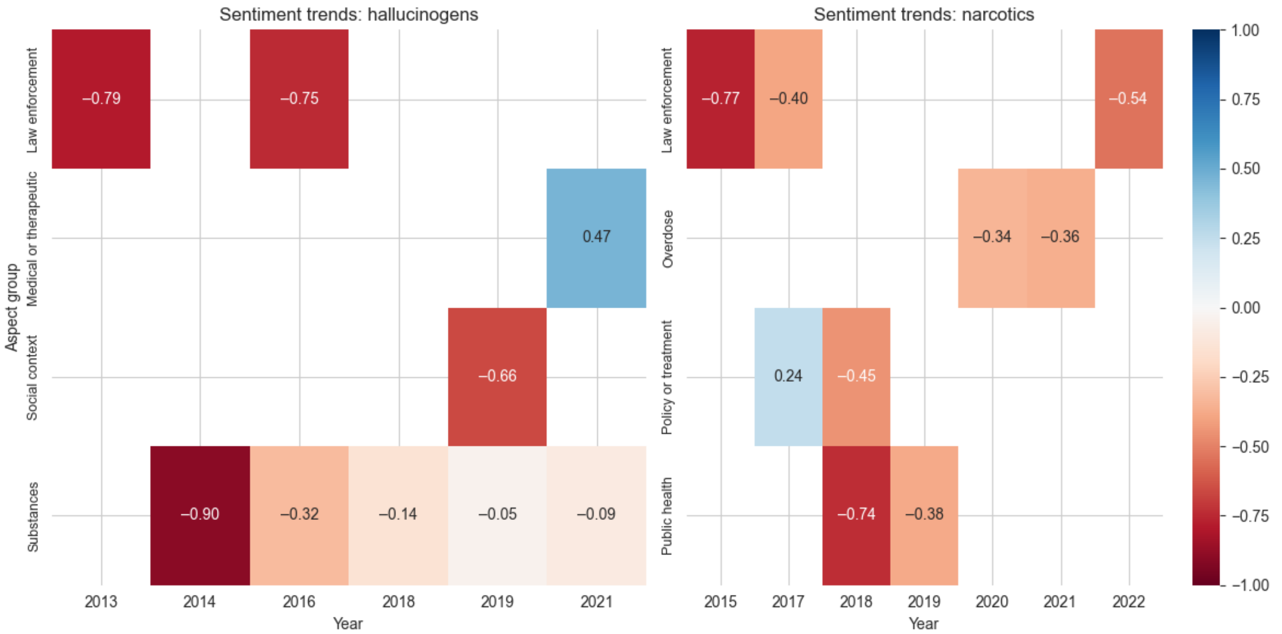
Article	Aspect	Year	Sentiment
"...Craig T. Steckler, president of the police officers' conference, led his introduction of holder by criticizing the justice department's decision this year not to challenge state laws in Colorado and Washington that allow recreational marijuana use[a]..."	Recreational marijuana use	2013	Negative
"...One question he answered: he said that he disagreed with U.S. Attorney General Eric Holder's memo Thursday saying the Obama administration would not challenge the new marijuana decriminalization[a] laws in Colorado and Washington. 'I think it's a mistake for him to turn his back and essentially by fiat legalize marijuana in Colorado and Washington,' Christie said..."	Marijuana decriminalization	2013	Negative
"...New Jersey is one of 23 states that have legalized medical marijuana despite a longtime federal prohibition against selling or using the drug for medical or recreational reasons...'this federal policy toward state-level marijuana legalization efforts[b] creates a situation in which the medical marijuana industry is in existence, integrating into local, state, and national economies, and employing thousands of people, some of whom are represented by labor unions or involved in labor organizing efforts despite the industry's illegality,' the opinion said..."	Marijuana legalization efforts	2015	Neutral
"...marijuana is a schedule 1 drug, which means the federal government treats it as if it were as dangerous as heroin or lsd and has no medical benefit. The act would also require authorities to remove federal cannabis convictions from millions of criminal records. More than two-thirds of American voters support full [marijuana] legalization[c], according to poll results released Nov. 14 by the Pew Research Center... 'It's the first piece of marijuana cannabis legislation[c] in Congress to move this far. And this could lead to more local reform as well,' said Goldstein, a South Jersey-based organizer for the National Organization for the Reform of Marijuana Laws."	Marijuana cannabis legislation	2019	Positive

[a]Negative aspects.

[b]Neutral aspects.

[c]Positive aspects.

Figure 6. Sentiment scores across aspect groups for the drug classes hallucinogens and narcotics. Values represent average sentiment scores ranging from –1 (most negative: red) to 1 (most positive: blue), with white indicating neutral sentiment.

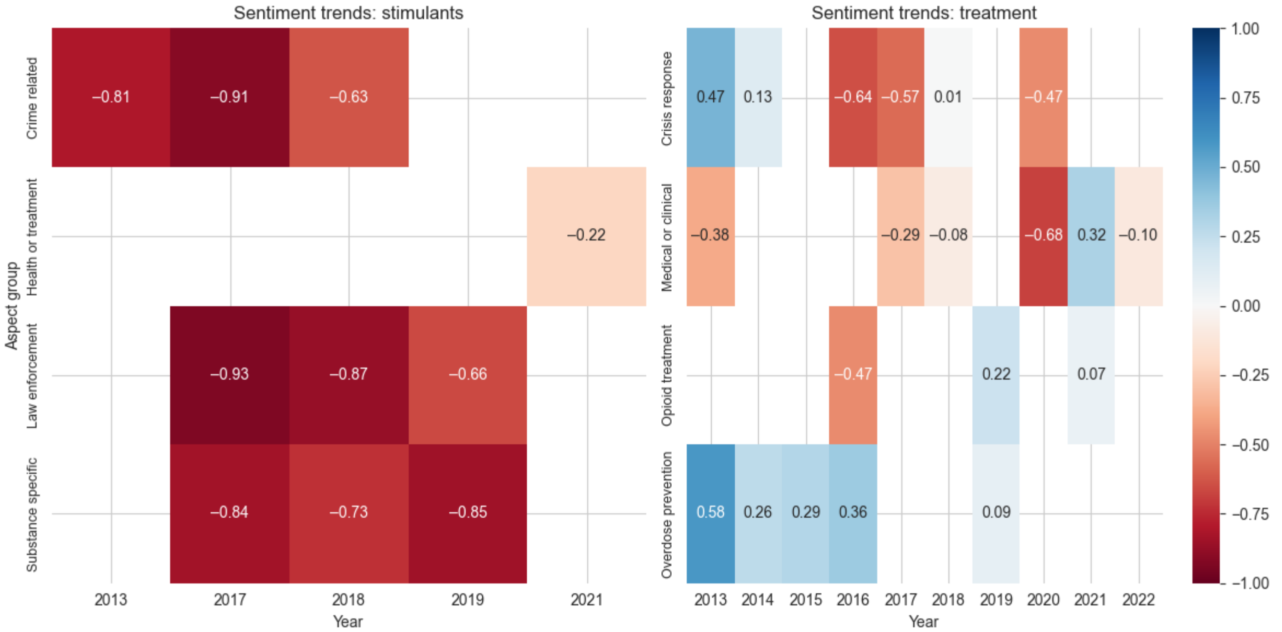


XSL•FO
RenderX

Table 5. Example aspects and excerpts of article text with sentiment: "+" denotes a positive sentiment and "–" denotes a negative sentiment.

Type of drugs and aspect; year	Sentiment	Excerpt
Hallucinogens		
"New psychiatric medicines"; 2021	+	"...last summer, the University of North Carolina received nearly $27 million from DARPA, the research arm of the U.S. Department of Defense, to develop new psychiatric medicines from psychedelics...Hopkins, along with New York University, recently completed phase 3 clinical trials with the nonprofit multidisciplinary association for psychedelic studies for MDMA-assisted therapy for PTSD..."
Narcotics		
"Medical marijuana"; 2017	+	"...he said. marijuana helped him with sobriety, and also helped him get off oxycodone and several other medications he was prescribed over the years for the pain caused by his many parachuting missions and the anxiety that PTSD triggered, he said. Karpowich said he participated in at least a dozen protests at the statehouse in Trenton, holding signs and 'ambushing lawmakers' to get them to consider expanding the medical marijuana program. 'veterans have served their country, and no one should tell them they can't use marijuana if it helps them,'..."
"Border protection"; 2022	−	"...he unleashed the worst border crisis in U.S. history. u.s. customs and border protection reported more than 1.7 million encounters with illegal migrants at the southern border, nearly four times the number the year before, the highest annual total on record—including 378,000 who were not from Mexico, Honduras, El Salvador, or Guatemala. seizures of deadly fentanyl more than doubled in 2021, and is closely connected to a surge in overdose deaths, which reached a historic high."
Treatment		
"Drug costs"; 2016	−	"...basic lifesaving medicines that emergency workers use every day are getting so costly, officials are scrambling to figure out how to pay for them. and as patients struggle with drug costs, EMS workers and emergency room doctors are seeing the impact. the price paid by Philadelphia emergency medical services for naloxone, which reverses opioid overdoses, has risen 150 percent since 2013...the opioid epidemic rages and record numbers of people die of overdoses, the cost of generic naloxone has more than doubled..."

Figure 7. Sentiment scores across aspect groups for the drug classes stimulants and treatment. Values represent average sentiment scores ranging from –1 (most negative: red) to 1 (most positive: blue), with white indicating neutral sentiment.



In contrast to these predominantly negative patterns, treatment coverage showed the most varied sentiment patterns, with distinct patterns across different aspect groups (Figure 6). Overdose prevention maintained a positive sentiment in the early years (0.58 in 2013, declining to 0.09 by 2019), while crisis response fluctuated dramatically, shifting from positive sentiment (0.47 in 2013) to negative (–0.64 and –0.57 in 2016 and 2017, respectively), back to neutral (0.01 in 2018), and then negative again (–0.47 in 2020). Medical or clinical aspects showed generally negative sentiment but with notable improvement in 2021 (0.32) before declining again (–0.10) in 2022. Opioid treatment aspects emerged with negative sentiment (–0.47 in 2016) but showed modest improvement in later years (0.22 in 2019 and 0.07 in 2022), suggesting growing acceptance of medication-assisted treatment approaches. The evolution of treatment coverage is particularly evident in the changing focus

XSL•FO
RenderX

on positive aspects over time. Naloxone-related aspects (described as a "life-saving drug") dominated positive coverage from 2013 to 2016, while buprenorphine prescription emerged as a central positive aspect from 2019 to 2021, before shifting to negative coverage due to concerns about misuse and street value (Multimedia Appendix 2). Cost-related aspects, particularly regarding naloxone prices and their impact on emergency medical service workers, contributed to negative sentiment spikes in 2016 (Table 5).

These temporal and thematic patterns reveal distinct evolutionary trajectories in media coverage of different substances and treatments. While some substances, particularly stimulants and depressants, maintained consistently negative coverage focused on crime and risk, others showed notable transitions. Hallucinogens evolved toward increasingly positive coverage centered on therapeutic potential, while cannabis coverage reflected complex societal debates around legalization and medical use. Treatment-related coverage, although variable, suggested growing acceptance of evidence-based interventions despite persistent challenges in implementation.

Discussion

Principal Findings

This study aimed to identify narrative shifts, including topics and sentiment, of drug news reporting within The Philadelphia Inquirer over a 10-year period. By grouping the articles based on drug classifications set by the National Institute on Drug Abuse and using co-occurrence analysis, dynamic topic modeling, and sentiment analysis, we uncovered several key findings. The co-occurrence analysis identified an expected connection between opiate narcotics (heroin and fentanyl) and substances such as cannabinoids or methadone, which are commonly used to alleviate opioid dependency symptoms, such as cravings and withdrawal severity. These findings align with provisional drug use and rise in overdose reported both nationwide [50] and within Philadelphia [51]. The emergence of "cocaine" comentions in 2018 matches the growing concern of opiate or stimulant co-use trends [52], as unintentional overdose deaths often involve the combined use of fentanyl or a stimulant such as cocaine. Buprenorphine, a safe and effective treatment for OUD, emerged in the coverage in 2020, with the most mentions (17 articles that mention buprenorphine in 2022 divided by 94 articles mentioning buprenorphine total, 18.1%) in 2022, likely correlated to policy changes making it more accessible [53].

The dynamic topic modeling revealed a strong focus on overdose deaths in the narcotics class. While we do not find evidence of sensationalism reported in other media studies [7,18,19,54], our analysis revealed a persistent focus on overdose deaths in narcotics coverage. Beyond direct event reporting, we observed substantial coverage of drug portrayals in entertainment media [55,56], suggesting news media's dual role in both reporting and reflecting on drug-related narratives.

Sentiment analysis revealed distinct trajectories for different substances. Cannabis coverage showed a marked evolution from negative to predominantly positive sentiment, particularly around medical applications and legalization. This shift coincided with significant local developments, including public figures' advocacy [57,58] and institutional initiatives [59]. In contrast, depressants and narcotics maintained consistently negative coverage, focusing on risks and criminal justice aspects. Hallucinogens showed the most dramatic positive shift, driven by emerging therapeutic applications, while treatment coverage reflected the complex challenges of implementing harm reduction approaches. These findings highlight the evolving nature of drug-related news coverage and its potential to shape public perceptions and policy discussions, as elaborated in the Implications section.

Implications

These findings have several important implications for both journalism practice and drug policy. Research has established that news media plays a significant role in shaping public perceptions and attitudes [1-5]. Through our analysis of The Philadelphia Inquirer's drug coverage, we track how these potentially influential media narratives have evolved over time. The evolution of cannabis coverage, characterized by a shift from negative to predominantly positive sentiment, particularly regarding medical applications and legalization, demonstrates how media reporting reflects shifting narratives around drug use.

The shift in cannabis coverage coincided with several significant local events in 2018: the admission of cannabis use by a former Philadelphia Eagles player [57], advocacy for recreational legalization by Mayor Jim Kenney [58], and the inaugural Cannabis Opportunity Conference in Philadelphia [59]. These events and the media's increasingly positive coverage of them exemplify the interplay between social change, policy developments, and media narratives.

Conversely, the persistent negative framing of certain substances, particularly in narcotics coverage (eg, "overdose crisis," "overdose victims," and "public drug use"), potentially reinforces stigma that can hinder public health approaches. The rare positive aspects (eg, the "New Jersey Drug Policy Alliance"; Multimedia Appendix 2) in narcotics coverage emerged only in relation to harm reduction initiatives, highlighting the potential challenge of incorporating harm reduction narratives into mainstream media discourse. While fear-based narratives can contribute to moral panic [60], our analysis suggests that balanced reporting can help normalize evidence-based interventions, as seen in the evolving coverage of cannabis and psychedelic therapies [61-63].

The pattern of coverage evolution raises important questions about the media's role in normalizing certain substances [64] while potentially stigmatizing others [65,66]. While cannabis coverage has grown increasingly nuanced and accepting, particularly regarding medical applications, coverage of narcotics remains predominantly negative, often emphasizing criminal justice aspects rather than public health approaches. News outlets have the potential to influence public discourse by providing more balanced coverage that includes harm reduction perspectives while maintaining responsible reporting on public health concerns.

XSL•FO
RenderX

Furthermore, our study found an increasing discussion of naloxone, an overdose prevention medication, and buprenorphine, a safe and effective treatment for OUD. Similar to previous work [67], the increasing coverage of these treatments suggests that news outlets can help to educate the public about evidence-based approaches to drug treatment and policy, which is important in the context of the ongoing opioid overdose crisis [67].

Finally, news outlets can impact policy by reporting on the work being done to reduce harm versus criminality aspects of drug use [17,66,68]. Our findings suggest that shifting news coverage from focusing on law enforcement to emphasizing prevention-oriented solutions may change the public's and nonexpert policy makers' views on drug issues, shifting from blame to more health-focused and harm reduction perspectives [22]. This shift mirrors a broader historical trajectory: while the stigmatization of drug users in the United States historically resulted in punitive policies [69], recent years have seen movement toward public health–oriented policies, such as expanding treatment and supporting overdose prevention laws [70].

However, our analyses reveal that discussions on harm reduction are still limited and have not been a prominent aspect of drug-related news in the past decade. As media coverage continues to evolve, it is hoped that future coverage will increasingly focus on compassionate and effective harm reduction strategies for affected communities and people who use drugs. By avoiding fear-based strategies and promoting balanced narratives, the media can greatly contribute to fostering a more empathetic and informed public discourse on drug use, ultimately influencing policy decisions and reducing harm in affected communities [65,66].

Limitations

Our findings should be interpreted in the context of several limitations. First, our analysis relies on articles from ProQuest's database of The Philadelphia Inquirer, which may not represent the complete set of articles published during this period. This limitation stems from potential gaps in digital archiving and database coverage. Future studies could incorporate multiple database sources or direct newspaper archives to ensure comprehensive coverage.

Second, our study focused solely on news articles published in The Philadelphia Inquirer, and therefore, our findings are specific to this source and its coverage of substance use within Philadelphia and the United States. While this focus provides valuable insight into local media coverage in a city significantly impacted by drug use, it limits generalizability to other regions or media outlets. Further research is needed to explore how the portrayal of substance use in the media varies across different regions and cultures.

Third, while previous research has established connections between news media coverage and public attitudes, our analysis of news did not enable us to evaluate the impact of this news exposure on public opinion regarding the issues. Future studies could combine media content analysis with public opinion surveys to directly measure these relationships.

Finally, the findings of our study do not correlate or imply popularity in use. As shown in Table 1, *designer drugs* and *drugs of concern* are mentioned in only 16 (0.43%) and 11 (0.3%) of the 3661 articles, respectively. Owing to the sparse data on these topics in our corpus, we excluded them from further analysis. Future work could specifically target coverage of emerging substances to better understand their media representation.

Conclusions

This study aimed to explore how The Philadelphia Inquirer, a prominent local newspaper covering Philadelphia, a city known nationwide for its ongoing drug crises, mentions commonly abused substances and how the narrativity and sentiment around these drugs change over time. To do so, we analyzed news articles published in The Philadelphia Inquirer between 2013 and 2022, using co-occurrence analysis, dynamic topic modeling, and aspect-based sentiment analysis, and found that cannabis was the most frequently discussed drug class, followed by narcotics. Aspects reported with hallucinogenic drugs tended to have a more positive tone compared to other categories of drugs, while articles on narcotics were the most negative. We also observed a significant focus on overdose and death-related aspects, but there was a noticeable lack of coverage related to harm reduction principles. This study highlights the linguistic shifts reported across various drug classes. It provides compelling evidence of the influence that news media outlets have in shaping discourse around drug use. This, in turn, contributes to creating a more informed and compassionate society, ultimately reducing the harm associated with drug use.

Acknowledgments

We gratefully acknowledge Dr Shahin Jabbari for his valuable support. This work was supported in part by the Institute of Museum and Library Services (IMLS) under grant number RE-246450-OLS-20. We also appreciate the reviewers for their thoughtful and constructive feedback.

Data Availability

The metadata of our dataset, codes, and results of this study are available in the Drugs in Inquirer repository [71].

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Representative aspects and article excerpts demonstrating sentiment shifts in media coverage of cannabis, hallucinogens, depressants, and treatment drugs (2013-2022). Color coding indicates negative (pink), neutral (yellow), or positive (green) sentiment, highlighting evolving media discourse around key drug-related topics.

[DOCX File , 25 KB - infodemiology_v5i1e56004_app1.docx]

Multimedia Appendix 2

Complete list of aspects (N=222) extracted from articles across all drug classes, showing neutral, positive, negative, and overall sentiment scores for each aspect (2013-2022).

[XLS File (Microsoft Excel File), 56 KB - infodemiology_v5i1e56004_app2.xls]

References

1. Iyengar S, Kinder DR. News That Matters: Television and American Opinion. Chicago, IL: University of Chicago Press; 1987.
2. Arias E. How does media influence social norms? Experimental evidence on the role of common knowledge. Political Sci Res Methods 2018 Feb 20;7(3):561-578. [doi: 10.1017/psrm.2018.1]
3. Olan F, Jayawickrama U, Arakpogun EO, Suklan J, Liu S. Fake news on social media: the impact on society. Inf Syst Front 2022 Jan 19;26(2):1-16 [FREE Full text] [doi: 10.1007/s10796-022-10242-z] [Medline: 35068999]
4. Schooler C, Flora JA. Pervasive media violence. Annu Rev Public Health 1996 Jan;17(1):275-298. [doi: 10.1146/annurev.pu.17.050196.001423] [Medline: 8724228]
5. Hjarvard S. The mediatization of society. Nordicom Rev 2008;29(2):102-131. [doi: 10.1515/nor-2017-0181]
6. Disrupting stigma: how understanding, empathy, and connection can improve outcomes for families affected by substance use and mental disorders. National Center on Substance Abuse and Child Welfare. URL: https://www.preventionnetwork.org/wp-content/uploads/disrupting-stigma-brief.pdf [accessed 2025-04-10]
7. Quan AM, Wilson LA, Mithani SS, Zhu DT, Bota AB, Wilson K. Reporting on the opioid crisis (2000-2018): role of The Globe and Mail, a Canadian English-language newspaper in influencing public opinion. Harm Reduct J 2020 Dec 30;17(1):93 [FREE Full text]
8. Harless JC, Hughes PM, Wilson C, Carpenter D, Ostrach B. Eligible prescriber experiences with substance use disorder treatment and perceptions of pharmacy barriers to buprenorphine. South Med J 2022 Aug;115(8):584-592. [doi: 10.14423/SMJ.0000000000001433] [Medline: 35922043]
9. Kuettel BT. Perceptions of substance use disorder and associated sanctions: a factorial vignette experiment. Crim Justice Rev 2022 May 31;48(2):203-231. [doi: 10.1177/07340168221105002]
10. Singh AK, Wu D. Sentiment analysis on Substance Use Disorder (SUD) tweets before and during COVID-19 pandemic. In: Proceedings of the 23rd HCI International Conference. 2021 Presented at: HCII 2021; July 24-29, 2021; Virtual Event. [doi: 10.1007/978-3-030-78645-8_77]
11. Lancaster K, Hughes CE, Spicer B, Matthew-Simmons F, Dillon P. Illicit drugs and the media: models of media effects for use in drug policy research. Drug Alcohol Rev 2011 Jul;30(4):397-402 [FREE Full text] [doi: 10.1111/j.1465-3362.2010.00239.x] [Medline: 21355898]
12. Swalve N, DeFoster R. Framing the danger of designer drugs: mass media, bath salts, and the “Miami zombie attack”. Contemp Drug Probl 2016 Apr 15;43(2):103-121. [doi: 10.1177/0091450916642515]
13. Miami-Dade Commission approves bath salt ban. CBS News. 2012 Jun 20. URL: https://www.cbsnews.com/miami/news/bath-salt-ban-approved-on-1st-reading-by-miami-dade-commission/ [accessed 2025-04-10]
14. Denham BE. Folk devils, news icons and the construction of moral panics. Journal Stud 2008 Dec;9(6):945-961 [FREE Full text] [doi: 10.1080/14616700802227811]
15. Johnston LD, O’Malley PM, Bachman JG, Schulenberg JE. Monitoring the future: national survey results on drug use, 1975-2003. National Institute on Drug Abuse. 2003. URL: https://files.eric.ed.gov/fulltext/ED518663.pdf [accessed 2023-09-26]
16. Caburnay CA, Kreuter MW, Luke DA, Logan RA, Jacobsen HA, Reddy VC, et al. The news on health behavior: coverage of diet, activity, and tobacco in local newspapers. Health Educ Behav 2003 Dec 01;30(6):709-722. [doi: 10.1177/1090198103255456] [Medline: 14655865]
17. McGinty EE, Kennedy-Hendricks A, Baller J, Niederdeppe J, Gollust S, Barry CL. Criminal activity or treatable health condition? News media framing of opioid analgesic abuse in the United States, 1998-2012. Psychiatr Serv 2016 May 01;67(4):405-411. [doi: 10.1176/appi.ps.201500065] [Medline: 26620290]

XSL•FO
RenderX

18. Hughes CE, Lancaster K, Spicer B. How do Australian news media depict illicit drug issues? An analysis of print media reporting across and between illicit drugs, 2003-2008. Int J Drug Policy 2011 Jul;22(4):285-291. [doi: 10.1016/j.drugpo.2011.05.008] [Medline: 21700441]
19. Hayden Griffin O3, Leban L, Singh D, Webb M, Narayanan S. Deconstructing differential drug coverage within a Malaysian media source. J Psychoactive Drugs 2024 Mar 02;56(2):272-278 [FREE Full text] [doi: 10.1080/02791072.2023.2186287] [Medline: 36864012]
20. Principles of Harm Reduction. National Harm Reduction Coalition. URL: https://harmreduction.org/about-us/principles-of-harm-reduction/ [accessed 2023-04-07]
21. Cameron Wild T, Koziel J, Anderson-Baron J, Hathaway J, McCurdy A, Xu X, et al. Media coverage of harm reduction, 2000-2016: a content analysis of tone, topics, and interventions in Canadian print news. Drug Alcohol Depend 2019 Dec 01;205:107599. [doi: 10.1016/j.drugalcdep.2019.107599] [Medline: 31610295]
22. White SA, Lee R, Kennedy-Hendricks A, Sherman SG, McGinty EE. Perspectives of U.S. harm reduction advocates on persuasive message strategies. Harm Reduct J 2023 Aug 18;20(1):112 [FREE Full text] [doi: 10.1186/s12954-023-00849-z] [Medline: 37596595]
23. Niamatullah S, Auchincloss A, Livengood K. Drug overdose deaths in big cities. Drexel University, Urban Health Collaborative. 2022. URL: https://drexel.edu/uhc/resources/briefs/BCHC%20Drug%20Overdose/ [accessed 2023-04-07]
24. Denis CM, Dominique T, Smith P, Fiore D, Ku YC, Culhane A, et al. HIV infection and depression among opiate users in a US epicenter of the opioid epidemic. AIDS Behav 2021 Jul 15;25(7):2230-2239 [FREE Full text] [doi: 10.1007/s10461-020-03151-2] [Medline: 33449236]
25. Philadelphia 2022: the state of the city. The Pew Charitable Trusts. 2022 Apr 20. URL: https://www.pewtrusts.org/en/research-and-analysis/reports/2022/04/philadelphia-2022-state-of-the-city [accessed 2025-04-10]
26. Opioid misuse and overdose data. City of Philadelphia Department of Public Health. 2020. URL: https://www.phila.gov/programs/combating-the-opioid-epidemic/reports-and-data/opioid-misuse-and-overdose-data/ [accessed 2023-09-22]
27. Project HOME to launch effort to tackle Philly's opioid epidemic. WHYY. URL: https://whyy.org/articles/project-home-philadelphia-housing-opioid-epidemic [accessed 2023-10-18]
28. Saldua M. Addressing social determinants of health among individuals experiencing homelessness. Substance Abuse and Mental Health Services Administration. 2024 Jan 22. URL: https://behavioralhealthnews.org/addressing-social-determinants-of-health-among-individuals-experiencing-homelessness/ [accessed 2024-01-02]
29. Blendon RJ, Young JT. The public and the war on illicit drugs. JAMA 1998 Mar 18;279(11):827-832. [doi: 10.1001/jama.279.11.827] [Medline: 9515986]
30. Zhang MWB, Hong YX, Husain SF, Harris KM, Ho RCM. Analysis of print news media framing of ketamine treatment in the United States and Canada from 2000 to 2015. PLoS ONE 2017 Mar 3;12(3):e0173202. [doi: 10.1371/journal.pone.0173202]
31. Dasgupta N, Mandl KD, Brownstein JS. Breaking the news or fueling the epidemic? Temporal association between news media report volume and opioid-related mortality. PLoS One 2009 Dec 18;4(11):e7758 [FREE Full text] [doi: 10.1371/journal.pone.0007758] [Medline: 19924221]
32. Commonly used drug chart. National Institutes of Health National Institute on Drug Abuse. URL: https://nida.nih.gov/research-topics/commonly-used-drugs-charts [accessed 2023-10-10]
33. Drugs of abuse: a DEA resource guide. Drug Enforcement Administration, U.S. Department of Justice. 2020. URL: http://drugsandalcohol.ie/33501/ [accessed 2023-04-07]
34. Blei DM, Lafferty JD. Dynamic topic models. In: Proceedings of the 23rd International Conference on Machine Learning. 2006 Presented at: ICML '06; June 25-29, 2006; Pittsburgh, PA. [doi: 10.1145/1143844.1143859]
35. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv. Preprint posted online on March 11, 2022 Preprint posted online on March 11, 2022 [FREE Full text]
36. Po¨ttker H. News and its communicative quality: the inverted pyramid—when and why did it appear? Journal Stud 2010 Jun 04;4(4):501-511. [doi: 10.1080/1461670032000136596]
37. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. 2015 Presented at: WSDM '15; February 2-6, 2015; Shanghai, China URL: https://doi.org/10.1145/2684822.2685324 [doi: 10.1145/2684822.2685324]
38. Newman D, Lau JH, Grieser K, Baldwin T. Automatic evaluation of topic coherence. In: Proceedings of the Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. 2010 Presented at: NAACL HLT 2010; June 2-4, 2010; Los Angeles, CA URL: https://aclanthology.org/N10-1012 [doi: 10.18653/v1/n16-1057]
39. Chang J, Boyd-Graber J, Gerrish S, Wang C, Blei DM. Reading tea leaves: how humans interpret topic models. In: Proceedings of the 23rd International Conference on Neural Information Processing Systems. 2009 Presented at: NIPS'09; December 7-10, 2009; Vancouver, BC URL: https://papers.nips.cc/paper_files/paper/2009/hash/f92586a25bb3145facd64ab20fd554ff-Abstract.html
40. O'Callaghan D, Greene D, Carthy J, Cunningham P. An analysis of the coherence of descriptors in topic modeling. Expert Syst Appl 2015 Aug;42(13):5645-5657 [FREE Full text] [doi: 10.1016/j.eswa.2015.02.055]

XSL•FO
RenderX

41. Alami Merrouni Z, Frikh B, Ouhbi B. Automatic keyphrase extraction: a survey and trends. J Intell Inf Syst 2019 May 2;54(2):391-424. [doi: 10.1007/s10844-019-00558-9]
42. Grootendorst M. KeyBERT: minimal keyword extraction with BERT. GitHub. 2020 Jul. URL: https://maartengr.github.io/KeyBERT/index.html [accessed 2025-03-31]
43. Mihalcea R, Tarau P. TextRank: bringing order into texts. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. 2004 Presented at: EMNLP 2004; July 25-26, 2004; Barcelona, Spain.
44. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv Preprint posted online on October 11, 2018 [FREE Full text] [doi: 10.18653/v1/n18-2]
45. Hamborg F, Donnay K. NewsMTSC: a dataset for (multi-)target-dependent sentiment classification in political news articles. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. 2021 Presented at: EACL 2021; April 19-23, 2021; Online. [doi: 10.18653/v1/2021.eacl-main.142]
46. Halperin A. Marijuana: is it time to stop using a word with racist roots? The Guardian. 2018 Jan 29. URL: https://tinyurl.com/mru4sj4e [accessed 2023-10-05]
47. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Committee on Medication-Assisted Treatment for Opioid Use Disorder, Mancher M, Leshner AI. The effectiveness of medication-based treatment for opioid use disorder. In: Medications for Opioid Use Disorder Save Lives. Washington, DC: National Academies Press; 2019.
48. Ao B. Can magic mushrooms and LSD treat depression and anxiety? Scientists are optimistic. The Philadelphia Inquirer. 2020 Feb 17. URL: https://www.inquirer.com/health/psychedelics-magic-mushrooms-acid-mental-health-20200217.html [accessed 2023-12-08]
49. Marijuana Opportunity Reinvestment and Expungement Act of 2019, S. 2227. Congress.Gov. 2019. URL: https://www.congress.gov/bill/116th-congress/senate-bill/2227 [accessed 2025-04-11]
50. Ahmad FB, Cisewski JA, Rossen LM, Sutton P. Provisional drug overdose death counts. National Center for Health Statistics. 2025. URL: https://dx.doi.org/10.15620/cdc/20250305008 [accessed 2025-04-10]
51. Public HP. Unintentional Drug Overdose Fatalities in Philadelphia, 2022. Department of public health. Philadelphia; 2023. URL: https://www.phila.gov/2023-10-02-philadelphia-records-more-than-1400-overdose-deaths-in-2022-deaths-among-black-residents-rose-nearly-20/ [accessed 2023-11-23]
52. Hoopsick RA, Andrew Yockey R. Methamphetamine-related mortality in the United States: co-involvement of heroin and fentanyl, 1999–2021. Am J Public Health 2023 Apr;113(4):416-419. [doi: 10.2105/ajph.2022.307212]
53. Buprenorphine treatment for opioid use disorder in Philadelphia. Pew Charitable Trusts. 2021 Sep 14. URL: https://www.pewtrusts.org/en/research-and-analysis/reports/2021/09/buprenorphine-treatment-for-opioid-use-disorder-in-philadelphia [accessed 2023-04-08]
54. Gussow L. Toxicology rounds: reports of new street drugs often sensationalized. Emergency Med News 2022;44(6):25. [doi: 10.1097/01.EEM.0000834160.23983.45]
55. Rezapour R, Diesner J. Classification and detection of micro-level impact of issue-focused documentary films based on reviews. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR. [doi: 10.1145/2998181.2998201]
56. Rezapour R, Diesner J, Jiang M. Assessing public awareness of social justice documentary films based on news coverage versus social media. In: Proceedings of the iConference 2016. 2016 Presented at: iConference 2016; March 20-23, 2016; Philadelphia, PA. [doi: 10.9776/16305]
57. Hermann A. Former Eagle Chris Long, now retired, says he regularly used marijuana during NFL career. Philly Voice. 2019 May 22. URL: https://www.phillyvoice.com/nfl-eagles-chris-long-marijuana-career-patriots-weed/ [accessed 2023-04-08]
58. Maas D. Mayor Kenney pushes marijuana legalization, again. AL DÍA. 2018 Oct 9. URL: https://aldianews.com/en/politics/policy/kenney-legalize-marijuana [accessed 2023-04-08]
59. Brown C. Senator sharif street. 2018. URL: https://temple-news.com/summit-brings-minorities-to-cannabis-industry/ [accessed 2023-04-08]
60. Baumgaertner J, Colvin P, Birch P. Moral panic and perceptions of community fear: a local community study examining drug use. Salus J 2018;6(2):21-36. [doi: 10.7208/chicago/9780226554228.003.0013]
61. Siegel JS, Daily JE, Perry DA, Nicol GE. Psychedelic drug legislative reform and legalization in the US. JAMA Psychiatry 2023 Jan 01;80(1):77-83 [FREE Full text] [doi: 10.1001/jamapsychiatry.2022.4101] [Medline: 36477830]
62. Weise E, della Cava M. Oakland in California decriminalizes magic mushrooms and peyote. USA Today. 2019 Jun 5. URL: https://www.usatoday.com/story/news/2019/06/05/oakland-california-legalizes-magic-mushrooms-and-peyote/1347888001/ [accessed 2023-04-08]
63. Center for Psychedelic and Consciousness Research homepage. Center for Psychedelic and Consciousness Research. URL: https://hopkinspsychedelic.org/index/#research [accessed 2023-11-25]

XSL•FO
RenderX

64. Egelhofer JL, Aaldering L, Eberl JM, Galyga S, Lecheler S. From novelty to normalization? How journalists use the term "fake news" in their reporting. Journal Stud 2020 Mar 30;21(10):1323-1343 [FREE Full text] [doi: 10.1080/1461670x.2020.1745667]
65. Salomon A, Bartlett M, Chenery M, Jauncey M, Roxburgh A. Outrage and algorithms: shifting drug-related stigma in a digital world. Int J Drug Policy 2023 Dec;122:104224. [doi: 10.1016/j.drugpo.2023.104224] [Medline: 37857181]
66. Mortensen TM, Moscowitz L, Wan A, Yang A. The marijuana user in US news media: an examination of visual stereotypes of race, culture, criminality and normification. Visual Commun 2019 Aug 22;19(2):231-255. [doi: 10.1177/1470357219864995]
67. Kennedy-Hendricks A, Levin J, Stone E, McGinty EE, Gollust SE, Barry CL. News media reporting on medication treatment for opioid use disorder amid the opioid epidemic. Health Aff (Millwood) 2019 Apr;38(4):643-651. [doi: 10.1377/hlthaff.2018.05075] [Medline: 30933576]
68. Iyengar S. Framing responsibility for political issues. Ann Am Acad Polit Soc Sci 1996 Jul 01;546(1):59-70. [doi: 10.1177/0002716296546001006]
69. Stangl AL, Earnshaw VA, Logie CH, van Brakel W, C Simbayi L, Barré I, et al. The Health Stigma and Discrimination Framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019 Feb 15;17(1):31 [FREE Full text] [doi: 10.1186/s12916-019-1271-3] [Medline: 30764826]
70. Drug and alcohol use. Office of Disease Prevention and Health Promotion, Office of the Assistant Secretary for Health, Office of the Secretary, U.S. Department of Health and Human Services. URL: https://odphp.health.gov/healthypeople/objectives-and-data/browse-objectives/drug-and-alcohol-use [accessed 2025-04-08]
71. social-nlp-lab / drugs_in_inquirer. GitHub. URL: https://github.com/social-nlp-lab/drugs_in_inquirer [accessed 2025-04-11]

Abbreviations

BERT: bidirectional encoder representations from transformers
BERTopic: bidirectional encoder representations from transformers topic modeling technique
OUD: opioid use disorder
PTSD: posttraumatic stress disorder
RQ: research question
SS: sentiment score
SUD: substance use disorder

Edited by T Mackey; submitted 04.01.24; peer-reviewed by S Rajput, M Haupt, F Lamy; comments to author 17.09.24; revised version received 12.11.24; accepted 25.01.25; published 13.05.25.

Please cite as:
Bouzoubaa L, Ehsani R, Chatterjee P, Rezapour R
Shifting Narratives in Media Coverage Across a Decade of Drug Discourse in the Philadelphia Inquirer: Qualitative Sentiment Analysis
JMIR Infodemiology 2025;5:e56004
URL: https://infodemiology.jmir.org/2025/1/e56004
doi:10.2196/56004
PMID:

©Layla Bouzoubaa, Ramtin Ehsani, Preetha Chatterjee, Rezvaneh Rezapour. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 13.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Natural Language Processing to Describe the Use of an Online Community for Abortion During 2022: Dynamic Topic Modeling Analysis of Reddit Posts

Elizabeth Pleasants[1,2], DrPH, MPH; Ndola Prata[3], MD, MSc; Ushma D Upadhyay[4], PhD, MPH; Cassondra Marshall[1], DrPH, MPH; Coye Cheshire[5], PhD, MA

[1]School of Public Health, University of California, Berkeley, Berkeley, CA, United States
[2]Department of Obstetrics and Gynecology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[3]Bixby Center for Population, Health, and Sustainability, School of Public Health, University of California, Berkeley, Berkeley, CA, United States
[4]Department of Obstetrics, Gynecology & Reproductive Sciences, University of California, San Francisco, San Francisco, CA, United States
[5]School of Information, University of California, Berkeley, Berkeley, CA, United States

Corresponding Author:
Elizabeth Pleasants, DrPH, MPH
School of Public Health
University of California, Berkeley
2121 Berkeley Way West
Berkeley, CA, 94704
United States
Phone: 1 919 966 2131
Email: b_pleasants@berkeley.edu

Abstract

Background: Abortion access in the United States has been in a state of rapid change and increasing restriction since the *Dobbs v Jackson Women's Health Organization* decision from the US Supreme Court in June 2022. With further constraints on access to abortion since *Dobbs*, the internet and online communities are playing an increasingly important role in people's abortion trajectories. There is a need for a broader understanding of how online resources are used for abortion and how they may reflect changes in the sociopolitical and legal context of abortion access. Research using online information and leveraging methods to work efficiently with large textual datasets has the potential to accelerate knowledge generation and provide novel insights into changing abortion-related experiences following *Dobbs*, helping address these knowledge gaps.

Objective: This project sought to use natural language processing techniques, specifically topic modeling, to explore the content of posts to 1 online community for abortion (r/abortion) in 2022 and assess how community use changed during that time.

Methods: This analysis described and explored posts shared throughout 2022 and for 3 subperiods of interest: *before the Dobbs leak* (December 24, 2021-May 1, 2022), *Dobbs leak to decision* (May 2, 2022-June 23, 2022), and *after the Dobbs decision* (June 24, 2022-December 23, 2022). We used topic modeling to obtain descriptive topics for the year and each subperiod and then classified posts. Topics were then aggregated into *conceptual groups* based on a combination of quantitative and qualitative assessments. The proportion of posts classified in each conceptual group was used to assess change in community interests across the 3 study subperiods.

Results: The 7273 posts shared in r/abortion in 2022 included in our analyses were categorized into 8 conceptual groups: abortion decision-making, navigating abortion access barriers, clinical abortion care, medication abortion processes, postabortion physical experiences, potential pregnancy, and self-managed abortion processes. Posts related to navigating access barriers were most common. The proportion of posts about abortion decision-making and self-management changed significantly across study periods (P=.006 and P<.001, respectively); abortion decision-making posts were more common before the *Dobbs* leak, whereas those related to self-management increased following the leak and decision.

Conclusions: This analysis provides a holistic view of r/abortion posts in 2022, highlighting the important role of online communities as abortion-supportive online resources and changing interests among posters with abortion policy changes. As policies and pathways to abortion access continue to change across the United States, approaches leveraging natural language

XSL•FO
RenderX

processing with sufficiently large samples of textual data present opportunities for timely monitoring, with the potential to reflect a broad range of abortion experiences, including those of people who have limited or no interaction with clinical abortion care.

(JMIR Infodemiol 2025;5:e72771) doi:10.2196/72771

KEYWORDS

abortion; online community; Reddit; Dobbs decision; natural language processing; topic modeling; BERTopic

Introduction

Overview

Abortion access in the United States has been in a state of rapid change and increasing restriction since the *Dobbs v Jackson Women's Health Organization* (*Dobbs*) decision in June 2022 [1]. While many people across the United States faced various challenges to accessing abortion before *Dobbs*, with this decision, the legal protection of abortion at the federal level in the United States was removed, returning control to individual states [2]. Since this change, >20 states have entirely or largely banned legal abortion [3], abortion clinics have closed across the United States, remaining and emerging clinics have struggled with the burden of demand resulting in long wait times, and many people have been unable to access in-clinic abortions [4-6]. Given the relative recency of *Dobbs*, the progressive and rolling impacts that have resulted from that decision, and the potential for a national ban on abortion, researchers seek to monitor the impacts of *Dobbs* to provide timely insights [5,7,8]. Research using online information and leveraging methods to work efficiently with large datasets of user-generated text has the potential to accelerate knowledge generation and offer meaningful insights into online behaviors and changing interests related to abortion. We present background describing the importance of online health communities related to abortion and research exploring people's use of these communities, with a specific interest in how natural language processing (NLP) techniques can further this area of research.

Use of the Internet and Online Health Communities for Abortion

People experienced various challenges to abortion access before *Dobbs*, and as such, they sought support to help navigate the processes of abortion decision-making and access [9]. Many people discover pregnancies and make decisions about abortion outside of clinical contexts [10]. In addition, there are many people considering, seeking, and having abortions in the United States who never interact with formal clinical abortion care [11-13]. The internet, particularly with increasing accessibility in recent decades, plays an important role in people's abortion decision-making and navigation. Research has shown that the internet is an important source of abortion information among adults in the United States, particularly for people living in contexts with more abortion restrictions [10,14-18]. With further constraints on abortion access, the internet is likely playing an increasingly important role in people's abortion trajectories as a source of information, peer interaction, social support, health services, and medication [10,14-17]—but more information is needed about the specifics of its use since *Dobbs*.

Online health communities, or online groups of people with a common health interest or purpose governed by a set of policies or norms, are one internet-based resource that people use for abortion-related support [19]. Reddit is a popular social networking site used by approximately one-quarter of adults in the United States in 2022 [20]. Within Reddit, user-generated content is aggregated and dispersed across millions of user-created and monitored message boards on specific subjects (or subreddits; r/subject) to which users can subscribe as members. Members can post material (text, images, videos, or links) with generous character count limits (40,000 characters, approximately 5000-10,000 words), which other members interact with using comments and an upvote and downvote system. Each subreddit is governed by a set of moderators elected from within the community who set subreddit-specific guidelines for posting content, review posts to ensure that those guidelines are being met, and have the option to delete noncompliant posts and ban deviant users. Some subreddits have developed around health topics and function as online health communities.

We chose Reddit as the focus for this research given its popularity and the uniquely valuable window that it offers into how people discuss health concerns outside of formal research settings or clinical encounters. Reddit provides anonymity with its use of pseudonyms and usernames, lack of requirement for users to share personal information in the creation of profiles, lack of collection of IP address information, and lack of restriction on the use of "throwaway" accounts (ie, temporary accounts created for a specific, single use or short-term purpose). These features allow users to engage in more open and honest discussions than they might on other social networking sites [21-23] or with family members or health care providers [24-26]. As a public, user-driven platform, Reddit enables peer-to-peer information exchange and emotional support, often among individuals who may face stigma or lack access to trusted clinical support. This makes it particularly appealing for communication about stigmatized and politicized health experiences such as abortion.

The r/abortion subreddit is a community that aims to "offer support and advice to people who are seeking or have had an abortion." In 2022, r/abortion had almost 45,000 members and was actively moderated by the Online Abortion Resource Squad (OARS), a group that trains volunteers and works to offer consistent, quality information and support to people coming to r/abortion [27]. The group of moderators working with OARS also actively enforces the r/abortion community rules. Previous research has established that people use online communities, particularly those on Reddit, to communicate about abortion decision-making; challenges to abortion access; taking action to end a pregnancy outside of clinical contexts, or self-managed

XSL•FO
RenderX

abortion (SMA); sharing experiences and advice; and needs for support [28-37]. These studies have used qualitative methods to analyze a defined subsample of content, with either a topical query (eg, mention of "abortion" [33]) or a narrow temporal focus (eg, posts from the previous 2 weeks [28,34]). While these qualitative analyses have allowed researchers to explore the nuances of abortion experience and communication in online communities, their scopes provide limited information about the range of uses for this community or any changes in use over time.

Abortion Research Using Machine Learning and NLP Techniques

Researchers with an interest in health behaviors and outcomes have quantitatively explored the use of online communities, including subreddits, for some topics related to sexual and reproductive health, including patient perceptions of prenatal diagnostic testing [38], disclosure and use of throwaway accounts when posting about parenting [39], perceptions of the human papillomavirus vaccine [40], discussion of sexually transmitted diseases [41], the use of message boards related to miscarriage and stillbirths [42], and discussions of contraception [43]. Some of these analyses have used machine learning tools to analyze the textual data. Similarly, other projects have used computational methods to explore discussions of abortion on social networking sites with an interest in assessing and improving research approaches, often using abortion as a test case for a polarizing and complex topic of communication in these spaces [44-47]. In addition, a recent study by Valdez et al [48] explored a subset of posts to r/abortion and r/AbortionDebate in early 2022 and found that these online communities were used differently, with r/abortion often serving as a place to seek and share abortion support and r/AbortionDebate as a platform to debate abortion attitudes. These quantitative analyses and the qualitative approaches described previously demonstrate that data from online communities can provide insights into population interests and needs, as well as into how these communities serve as sexual and reproductive health resources.

However, research using machine learning tools with an explicit interest in abortion as a health concern impacted by sociopolitical context is lacking. Machine learning tools present a particularly compelling opportunity to advance public health research, providing timely and holistic information about how people have used online communities for abortion. Machine learning has a range of applications, including speech recognition, medical diagnosis, and NLP [49]. NLP seeks to process and analyze large amounts of language data—both text and spoken—to understand, interpret, and generate human language meaningfully and with consideration of context. There are many techniques within NLP, including tokenizing text into smaller units, recognition and tagging of parts of speech, text classification, topic modeling, and word embeddings [49,50]. Similar to other machine learning techniques, NLP can be carried out with varying levels of human input. However, when human expertise is particularly crucial, a "human-in-the-loop approach"—or a collaborative approach in which humans and machines work together—can be used [51]. NLP techniques create the potential for accelerated analyses of large samples of language data compared to traditional qualitative analyses and have been leveraged by public health researchers to explore a range of topics [38,39,41,43,45,52-56]. However, to date, there are no applications of NLP techniques to analyze online communication about abortion focused on the *Dobbs* decision as a public health concern, leveraging content expertise with a human-in-the-loop approach.

Research Aims

With ongoing changes to abortion access in the United States and people's use of online communities for abortion, research exploring how people use these communities can provide important insights for researchers, care providers, advocates, and policy makers. Past research has explored relatively narrow aspects of online community use related to abortion or as a topic for a research case study. However, there is a need for a broader understanding of how this type of online community is used for abortion and how it may reflect changes in the sociopolitical and legal context of abortion access. Particularly given the rapidly changing landscape of abortion policy and access in 2022, with uncertainty and increasingly constrained pathways to care, online resources likely play a key role in providing access to information, support, and services. In addition, exploring people's use of online communities for abortion has the potential to provide timely insights from populations that may be hard to reach and not well represented by much of the research on abortion experiences [57-59].

Using NLP techniques to analyze text shared in online abortion communities can help address these gaps, allowing for efficient analysis and pattern detection in the large bodies of user-generated narratives shared in these digital spaces. As such, this research sought to use NLP techniques to explore the content of posts to 1 online community for abortion (r/abortion) in 2022 and assess whether the use of the community changed during that time. Specifically, we used unsupervised topic modeling on posts submitted during the year surrounding the *Dobbs* decision (6 months before and after June 24, 2022) to inductively discover the topical content of posts, integrating a human-in-the-loop approach in the quality review, naming, and aggregation of topics.

Methods

Overview

Reddit data have historically been accessible through free, publicly available application programming interfaces (APIs). Researchers have often used Pushshift's Reddit API and Reddit's official API to obtain compiled information about content shared on Reddit, including creation date, submission (post or comment) text, community interactions (likes, upvotes, and downvotes), and more. Each API had different capabilities but generally provided access to organized and cataloged Reddit data that could be queried to obtain datasets for analysis.

A previous analysis of r/birthcontrol used data obtained from Pushshift's Reddit API [60], as have various other analyses focused on health-related use of Reddit [54-56,61-63]. Given changes to Pushshift's Reddit API in December 2022, specifically issues preventing direct access to any post

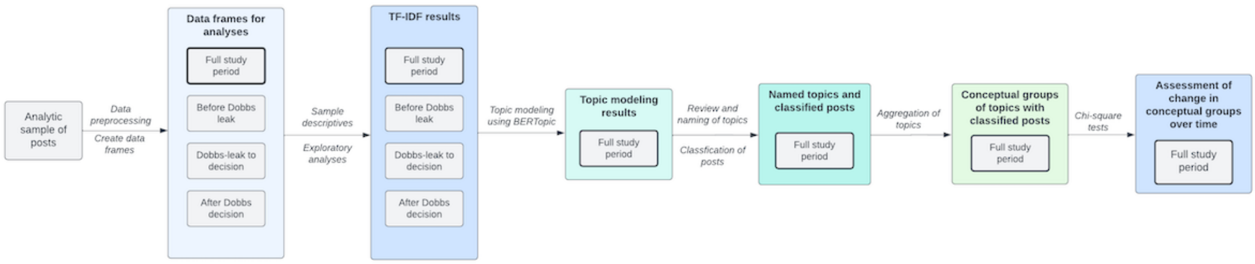
XSL•FO
RenderX

submission data from before November 3, 2022, we used 2 different data access approaches to obtain data for this analysis—one for data from December 24, 2021, to August 23, 2022 (gathered and stored for cleaning and analysis in 2022), and another for data from August 23, 2022, to December 23, 2022 (gathered and stored for cleaning and analysis in 2023). Notably, in April 2023, Reddit changed its data access policies, so public access to complete Reddit data through APIs, and the methodology outlined in this manuscript, is obsolete. We chose a start date of December 24, 2021, for this analysis to provide data for 6 months before and 6 months following the Dobbs decision. Details of the data collection approach and methods used for this analysis are presented in Multimedia Appendix 1 [64-68].

A visual overview of the analytic approach used in this research is presented in Figure 1. After procuring complete data from Reddit's API, we implemented additional restrictions to obtain the analytic sample used for this research. These restrictions aimed to provide a sample of posts theoretically within the public domain, containing sufficient text to support contextualized NLP analysis. Sequentially, we excluded posts if they were removed, were deleted, contained only an image, contained only a link, or contained <30 characters. We then cleaned content in this analytic sample to remove usernames, which were replaced with a unique submission ID created for this analysis. These deidentified data were used for the analyses described. The broad sample of posts obtained provides a holistic view of the topics discussed by r/abortion community members in relation to their abortion-related questions, experiences, challenges, and more following the Dobbs decision and related changes in abortion access in the United States in 2022. Comments responding to original posts were not included in this analysis.

To prepare our data for NLP analyses, we cleaned data from posts using part-of-speech tagging, removing links, removing punctuation, changing all text to lower case, removing words with ≤2 characters, and removing stop words.

Figure 1. Methods process flow for natural language processing of posts to r/abortion from December 24, 2021, to December 23, 2022. TF-IDF: term frequency–inverse document frequency.



Analysis

We began this analysis by counting all included posts submitted to r/abortion during the study period by date to summarize the community's overall use during the year. We then described the textual data from posts in our analytic sample for the year-long study period using the following: count of posts (by day and study period), average word count of text in posts, number of unique accounts that posts were made from, and average number of posts per author.

We determined distinctive words and phrases in posts using count vectorization [64,65], applying word frequency analysis to posts from the year and within 3 subperiods of interest in 2022 for this research: *before the Dobbs leak* (December 24, 2021-May 1, 2022), Dobbs leak to decision (period from the *Dobbs* leak to the decision; May 2, 2022-June 23, 2022), and *after the Dobbs decision* (June 24, 2022-December 23, 2022).

Topic Modeling

Overview

Topic modeling can discover latent topics, or themes, in documents. For this analysis, each post was defined as a document for topic modeling, and each data frame represented a set of documents. BERTopic was used given its capacity to account for the contexts of words in sentences in text, extending traditional topic modeling approaches that do not account for the semantic relationships between words [66]. It also provides various opportunities to explore topic hierarchies, topic visualization, and topic analyses. BERTopic was applied to preprocessed post text. Topic numbers, labels (top 10 words), and associated documents were reviewed by a research assistant and the lead researcher to assess model coherence. We manually reviewed a random sample of approximately 20% of all posts for a topic unless a topic had <50 posts, in which case review was completed for all posts in that topic. During this review process, we also developed and assigned a descriptive name for each topic based on the topic label and the raw text of representative submissions. Outlier documents were grouped into an outlier topic group (labeled as −1) in the modeling process and excluded from further review and analysis [67].

While BERTopic generates topics under the assumption that documents can include multiple topics, our categorization approach only assigned each submission to the topic that it had the highest probability of being in (rank 1 topic classification). This approach provided a single topic assignment for each document, reflecting the single most dominant topic describing that post. We reviewed posts classified under each topic to check the quality of assignments and then obtained counts to ascertain the number of posts classified under each topic. Using the trained models, we categorized all r/abortion posts using predictions in BERTopic [68]; classification was conducted using raw post text.

XSL•FO
RenderX

Topic modeling was carried out inductively, yielding results that we reviewed to determine a rigorous and consistent approach to topic aggregation. After reviewing topic labels and representative documents, it was clear that, while many topics used distinctive words, they described similar concepts. As such, we were interested in aggregating topics into *conceptual groups*, or clusters of related topics grouped into broader themes or domains that described key concepts represented by the topics in each model. Conceptual group assignments were manually determined for all topics using an approach that combined dendrogram results (quantitative) with a manual review of submission texts (qualitative), with a focus on defining conceptually meaningful and interpretable groups of topics. Post counts for each conceptual group were obtained based on the summed rank 1 classification assignments for topic modeling results, summarizing the commonality of submissions in each conceptual group.

Assessing Changes in Conceptual Group Frequency Over Time

We used counts of posts in each conceptual group for the year to assess the proportion in each group and differences between the subperiods of interest (*before the Dobbs leak*, *Dobbs leak to decision*, and *after the Dobbs decision*). We assessed statistically significant differences in the proportion of posts in a conceptual group using chi-square tests in R (*chisq.test* package; R Foundation for Statistical Computing). This compared the proportion of posts in a conceptual group versus those not in that conceptual group across the 3 study subperiods, assessing whether there was a change in the frequency of posts primarily focused on that concept throughout 2022. This component of the analysis sought to describe any changes in the primary focus of r/abortion submissions during 2022 considering the dramatic changes to the legal and social environment for abortion during the year

Ethical Considerations

The collection and analysis of these data was exempted from review by the Committee for Protection of Human Subjects at the University of California, Berkeley (2022-08-15585). As data were publicly available and Reddit content is accessible without an account, we did not obtain informed consent from r/abortion users. However, users may still engage with this platform with the expectation of privacy, and ethical principles related to informed consent, participant confidentiality, and privacy still arise when using data from Reddit [69-71]. These are of particular concern given the sensitivity of abortion narratives and the potential for digital data to be used to support abortion-related prosecution [72,73]. While past research using Reddit data has used techniques such as the exclusion of usernames and rephrasing or paraphrasing of post text to protect user identities [53,74,75], these approaches have generally been found to be insufficient to protect user identities, with the capabilities of various online search tools to discover users through the content they posted [76]. The current best practice when conducting research using Reddit data is to engage in a "heavy disguise" process with rigorous testing to produce disguised stories, as described by Reagle [76]. These concerns are especially pressing in the post-*Dobbs* era, where digital information may be weaponized in legal proceedings.

To mitigate these risks, we adhered to the current best practices for ethical internet research using sensitive data on health topics. We maintained the name of the subreddit that the data were obtained from given its visibility as an abortion community on Reddit and its relatively high number of members, offering some level of shielding for individual users [69]. While all analyses were carried out using complete text from submissions, we used an ethical fabrication process to develop representative narratives called "composite quotes" [76,77]. We gathered representative posts for a specific topical area of our findings and used them to generate a composite narrative related to that particular facet of an abortion experience. We replaced keywords, adjusted sentence ordering, and combined details of experiences across individuals. Once a composite quote was generated, its discoverability was checked by searching the overall quote and each sentence on Google (including the terms "Reddit AND abortion AND [the searched text]"), Reddit, and the Pushshift Reddit Search Tool [76,78]. Once composite quotes were checked via searches, a plagiarism checker was used as a final check to ensure that these stories were effectively nondiscoverable using current tools [79]. As such, all stories presented throughout this paper are composite quotes intended to reflect and represent the narratives shared in r/abortion but include a degree of ethical abstraction. This approach reflects a broader ethical stance that recognizes the public and private ambiguity of online data and the evolving legal risks of sharing abortion experiences online. Research in this area must not only comply with institutional review board standards but also grapple with new forms of digital vulnerability, especially in politically charged contexts.

An interdisciplinary team of researchers and advocates conducted this research, grounded in the principles of reproductive justice and the premise that abortion access is a health and human rights concern. To ensure transparency and provide context for our perspectives, we offer individual positionality statements. EP is a female-identifying individual who resided in an abortion-protective US state during the research. She is a Reddit user who primarily reads content. NP is a female-identifying individual who lived in an abortion-protective US state at the time of the study. She does not use Reddit but brings extensive experience working on sexual and reproductive health in abortion-restrictive countries. CM is a female-identifying person who lived in an abortion-protective US state at the time of this research; she is not a Reddit user. UU is a female-identifying person who lived in an abortion-protective US state at the time of this research; she is a Reddit user who primarily reads content. CC is a male-identifying individual who resided in an abortion-protective US state during the research period. He is an active Reddit user.

Our team brought varying levels of experience with Reddit, which enriched our engagement with data from r/abortion and informed our analysis. While all team members were living in abortion-protective states during the study, several of us have personal or professional ties to regions with more restrictive abortion policies. These experiences influenced how we

XSL•FO
RenderX

approached the data, interpreted the results, and framed our findings.

Results

Descriptive Statistics

Of the 12,509 posts shared in r/abortion during the study period (December 24, 2021-December 23, 2022), 2351 (18.79%) were deleted by the users, 1349 (10.78%) were deleted by the user and removed, and 1484 (11.86%) were removed (Table 1). Therefore, 41.44% (5184/12,509) of all posts were elided content and not eligible for analysis. Of the eligible posts, only 0.01% (1/12,509) contained only a link, and a small portion (51/12,509, 0.4%) contained <30 characters. This provided an analytic sample of 7273 eligible posts for the study year.

Table 1. Analytic samples of posts for the year and each subperiod, with counts of elided and excluded submissions (N=12,509).

	Before the *Dobbs* leak (December 24, 2021-May 1, 2022; n=3650)	*Dobbs* leak to decision (May 2, 2022-June 23, 2022; n=1828)	After the *Dobbs* decision (June 24, 2022-December 23, 2022; n=7031)	Year total
Elided content[a], n (%)	1586 (43.45)	829 (45.35)	2769 (39.38)	5184 (41.44)
Deleted by user[b]	572 (15.67)	267 (14.61)	1512 (21.5)	2351 (18.79)
Deleted by user+removed[b]	867 (23.75)	288 (15.75)	194 (2.76)	1349 (10.78)
Removed[b]	147 (4.03)	274 (14.99)	1063 (15.12)	1484 (11.86)
Posts excluded from the sample, n (%)	11 (0.3)	4 (0.22)	37 (0.53)	52 (0.42)
Contained only an image	0 (0)	0 (0)	0 (0)	0 (0)
Contained only a link	1 (0.03)	0 (0)	0 (0)	1 (0.01)
<30 characters	10 (0.27)	4 (0.22)	37 (0.53)	51 (0.41)
Elided+excluded posts, n (%)	1597 (43.75)	833 (45.57)	2806 (39.91)	5236 (41.9)
Posts in sample, n (%)	2053 (56.25)	995 (54.43)	4225 (60.09)	7273 (58.14)
Posts per day, median (IQR)	16.0 (5.5)	18.0 (5.0)	23.0 (8.0)	19.0 (8.0)

[a]The proportion of content removed by Reddit in this case aligned with previous accounts from Reddit on their removal of content (6% in 2020). While there is no systematic accounting of content deleted from Reddit, research on the 3 most popular sensitive-topic subreddits found that approximately half of the content was deleted by users [80], a higher proportion than that observed in these data from r/abortion.

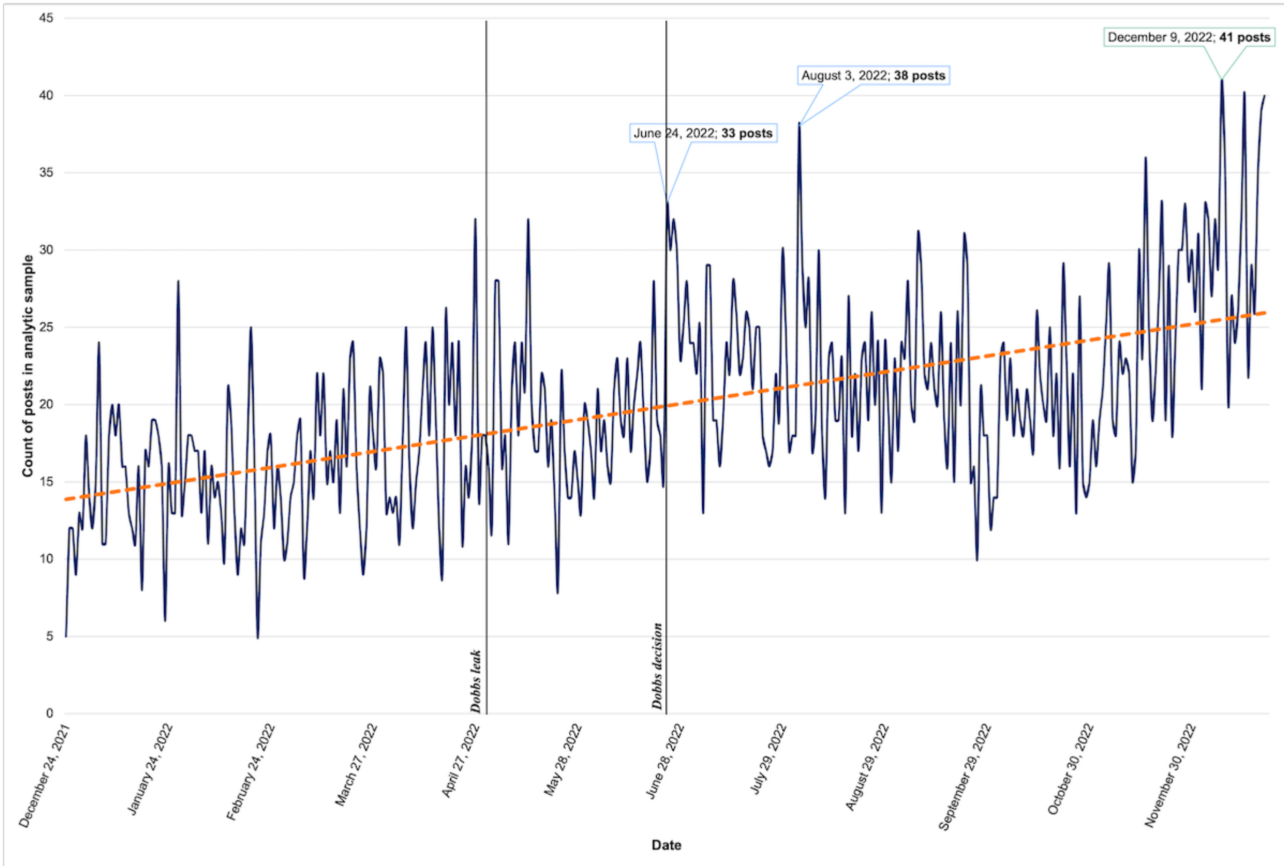
[b]Deleted content refers to posts removed from Reddit by the original poster. Removed refers to posts that were removed by Reddit or r/abortion moderators, and Deleted by user+removed refers to content that was both removed by Reddit or r/abortion moderators and then deleted by the user. Categories are exclusive.

The highest daily volume of posts was on December 9, 2022, with 41 eligible submissions (Figure 2). As there were no substantive policy, political, or social events related to abortion at this time (except for the legalization of abortion in Argentina), we believe that this peak is a random date in a larger upward trend. Other peak dates for post submissions were June 24, 2022 (33 posts; the date of the *Dobbs* decision), and August 3, 2022 (38 posts). The median number of eligible posts throughout the year was 19 (range: 5-41, IQR 8) per day, and the median number of posts per period increased throughout the year, with 16 posts per day before the *Dobbs* leak, 18 posts per day between the leak and the decision, and 23 posts per day following the *Dobbs* decision. Posts in the analytic sample had a median of 53 words per submission across the year (range 2-3204).

Posts were shared from 4468 unique accounts during the year, with posts made by 1379 users before the leak, 702 users between the leak and the *Dobbs* decision, and 2620 users following the *Dobbs* decision. On average, posters shared <2 posts in the r/abortion community during the year from a single Reddit user account (mean 1.6; IQR 0.5-1.6), with slight variation across subperiods.

XSL•FO
RenderX

Figure 2. Line graph showing daily volume of posts to r/abortion in 2022, with markers for the dates of the Dobbs leak and decision (N=7273 posts). The highest daily volume of posts was observed on December 9, 2022 (n=41 posts). Other peak dates for post submissions were June 24, 2022 (n=33 posts; the date of the Dobbs decision), and August 3, 2022 (n=38 posts).



Analyses of Cleaned Submission Text

Exploratory Data Analyses: Count Vectorizer Results

We obtained distinguishing sequences of consecutive words of different lengths, known as uni-, bi-, and trigrams ("n-grams"), from posts. These topics included abortion decision-making timing and feelings, which commonly included "feel," "want," "baby," "know," "pregnant," and other words (see Multimedia Appendix 2) [81] for the year and by study period in relation to the *Dobbs* leak and decision. For the year overall (N=7273 posts), n-grams reflect discussions of pregnancy and pregnancy confirmation, abortion timing and decision-making, and access through clinic-based and online care. Throughout 2022—in yearly results and across subperiods—n-grams reflect discussions of pregnancy and abortion decision-making and experiences during abortion processes. These n-grams also highlight consideration of changes related to *Dobbs* and the overturning of *Roe v Wade*, particularly access while living in a restrictive setting, with the emergence of related language starting from *Dobbs* leak.

Topic Modeling

Overview

Posts from the yearly sample (N=7273) were described through 55 inductively generated topics with 3 outlier topics (in the −1 outlier group, see Table 2). These topics contained between 10 and 472 words, reflecting a range of word density across topics. On the basis of a review of topic labels and representative documents, these topics reflect that the analyzed posts described a wide range of abortion-related experiences and concerns—summarized through topic names. These topics included *abortion decision-making timing and feelings*, which commonly included "feel," "want," "baby," "know," "pregnant," and other words. *Appointments and cost barriers* was another topic describing yearly posts, including "abortion," "help," "get," "get abortion," and "appointment" as common words. In addition, people posted about *medication abortion*, *physical process bleeding, and clots* using words such as "bleeding," "clots," "cramps," "took," "passed," and "blood." There were also topics such as *Aid Access shipping*, where posters used words such as "package," "customs," "tracking," "usps," and "aid access."

Document Classification Using Topics and Aggregation Into Conceptual Groups

Topics describing posts from the study year (n=55) were used to classify all posts (N=7273) into their highest-probability topic. The count and proportion of posts classified under each topic are presented in Table 2. Topics with the highest number of classified posts were *seeking access to medication abortion* (2408/7273, 33.11% of posts); *abortion decision-making, sharing stories, and seeking support* (657/7273, 9.03% of posts); and *medication abortion process, timing of pills* (328/7273, 4.51% of posts).

We aggregated these topics into 8 conceptual groups: *abortion decision-making*, *navigating access barriers*, *clinical abortion*

XSL•FO
RenderX

care, *medication abortion process*, *postabortion physical experiences*, *potential pregnancy*, and *SMA process*. On the basis of conceptual group classification counts (summed topic classification counts), *navigating access barriers* was the conceptual group with the largest number of posts (2446/7273, 33.63% of all posts), which included posts describing various experiences with challenges seeking and having an abortion. Posts focused on aspects of the *medication abortion process* were the second most common (1807/7273, 24.85%), followed by *abortion decision-making* (974/7273, 13.39%), *postabortion physical experiences* (672/7273, 9.24%), *clinical abortion care* (670/7273, 9.21%), *potential pregnancy* (550/7273, 7.56%), and *SMA process* (151/7273, 2.08%).

Composite versions of representative posts for each conceptual group are presented in Multimedia Appendix 2. Select quotes for each conceptual group are presented in this section, representing the content of posts in each group for the year overall and the comparable subperiods. Some posts in the *navigating access barriers* conceptual group described difficulty obtaining an appointment at an abortion clinic, sharing stories such as the following:

> *My boyfriend and I live in the south with our 9-month-old baby, and we just found out I'm pregnant again. We really can't afford to and aren't prepared to have another child. Abortion is illegal in our state and we live 4 hours from the nearest clinic in the nearest legal state. Do we have any other options?* [Composite quote]

Others focused on challenges in accessing medication abortion, often describing specific barriers and sometimes asking for information and resources to help them access it, as with posts such as the following:

> *My girlfriend and I are in urgent need of abortion pills. She's about 3 or 4 weeks pregnant and we want to get through this asap and for not a lot of money, preferably for a few hundred dollars or less. Can anyone provide some info and walk us through the process?* [Composite quote]

Others specifically referenced concerns or issues with abortion access related to *Roe v Wade* and the *Dobbs* decision, with narratives such as the following:

> *I think I might be pregnant and I'm really scared, especially after hearing Roe v Wade might be overturned. I'm thinking of having an abortion but I live in a state where abortion will likely become illegal right away. Does anyone know if I will still be able to get an abortion somehow? I've been trying to research online and have been seen a lot of different information. Are there any online clinics? Would it be illegal to travel to another state to go to a clinic in-person? I've heard that Colorado or Illinois might be safe options but I'm just really terrified. I just want to be sure I can access an abortion no matter what happens.* [Composite quote]

Among posts describing the *medication abortion process*, there were many asking questions about or descriptions of the use of pills to have a medication abortion. These posts sometimes included specific questions about correct timing in the use of medication abortion, sharing stories such as the following:

> *I ordered pills through AidAccess and they arrived today. I've seen online that you're supposed to wait 24 hours after you take the mifepristone to then take the misoprostol, but on the box it doesn't mention timing and the instructions are very vague. I don't want to screw this up at all so can someone clarify the timing and what exactly I should do?* [Composite quote]

People also commonly posted about *abortion decision-making.* In these posts, people often described their pregnancy experiences and abortion story while seeking community support, with narratives such as the following:

> *Can anyone share about their experiences during and after their abortions around 6 weeks? I'm trying to pick between medication or surgical. My last abortion was surgical and it was not a great experience. I just want this to be over with because ideally, I wouldn't want to have to go through this, but I am not in a place to support a child financially.* [Composite quote]

Posters also wrote about making their abortion decision in relation to their relationship dynamics, with narratives such as the following:

> *I can't take this anymore, I'm going to get an abortion. I'm just not ready to have a baby and I know that this will be for the best. I'm not looking forward to having to face my partner after this. I'm afraid he's going to hate me for having an abortion and break up with me.* [Composite quote]

Posts asking questions about and describing *postabortion physical experiences* were also common, with posters describing experiences with bleeding or clots after their abortion and often seeking input from the community on how normal their experiences were:

> *I took my miso last night and have been bleeding for almost 12 hours now. I had cramping for the first few hours then a big flow of blood where I passed two big clots the size of small lemons. But those are the only clots I've had and I haven't been bleeding any more. I know you're supposed to pass a lot of clots based on what I've been reading. Should I be worried?* [Composite quote]

Posts also described aspects of *clinical abortion care*, including specific facets of clinical care—particularly ultrasound—and fears and concerns about having an abortion. Posts also focused on *potential pregnancy*, sometimes describing sexual experience and possible risk of pregnancy. In contrast, others focused on potential pregnancy after an abortion in relation to still testing positive weeks later. In addition, some posts described *experiences with SMA*, often focused on the process of ordering medication and having it shipped while asking for support from the community in navigating those processes.

XSL•FO
RenderX

Table 2. BERTopic topic modeling classifications and conceptual group aggregations for yearly posts with topic names, rank 1 classification counts, conceptual groups, and summed rank 1 conceptual group classification counts (N=7270 posts; 3 posts excluded in the −1 outlier group).

Conceptual group and topic name	Topic posts, n (%)[a,b]	Conceptual group posts, n (%)[c]
Navigating access barriers		2446 (33.7)
“Appointments and cost barriers”	4 (0.2)[d]	
“Seeking access to medication abortion”	2408 (98.5)[d]	
“Dobbs decision barriers”	21 (0.9)[d]	
“Appointments and travel barriers”	13 (0.5)[d]	
Medication abortion process		1807 (24.9)
“Medication abortion process, timing of pills”	328 (18.2)[e]	
“Medication abortion normal physical process”	122 (6.8)[e]	
“Medication abortion physical process misoprostol”	299 (16.6)[e]	
“Medication abortion physical process miso”	51 (2.8)[e]	
“Medication abortion process pill administration”	10 (0.6)[e]	
“Medication abortion process timing”	7 (0.4)[e]	
“Medication abortion process taking pills”	276 (15.3)[e]	
“Medication abortion physical process bleeding and clots”	31 (1.7)[e]	
“Medication abortion physical process mifepristone and misoprostol”	90 (5)[e]	
“Medication abortion physical process and completion”	205 (11.3)[e]	
“Medication abortion process, seeking information about access and use experience”	304 (16.8)[e]	
“Medication abortion physical process pain and clots”	7 (0.4)[e]	
“Nausea during pregnancy and abortion”	77 (4.3)[e]	
“Abortion decision-making postabortion grief and regret”	31 (1.7)[e]	
“Abortion decision-making postabortion timing and regret”	3 (0.2)[e]	
“Abortion decision-making, sharing stories and seeking support”	657 (36.4)[e]	
“Abortion decision-making support (presence and lack)”	70 (3.9)[e]	
“Abortion decision-making, reflection and navigating challenges”	49 (2.7)[e]	
“Abortion decision-making and pregnancy confirmation”	4 (0.2)[e]	
“Abortion decision-making fear and support”	2 (0.1)[e]	
“Abortion decision-making relationship dynamics and possible pregnancy”	36 (2)[e]	
“Abortion decision-making relationship dynamics”	122 (6.8)[e]	
Postabortion physical experiences		672 (9.2)
“Post medication abortion bleeding and menstruation”	89 (13.2)[f]	
“Post abortion bleeding”	58 (8.6)[f]	
“Post abortion bleeding and menstruation”	44 (6.5)[f]	
“Post abortion bleeding and clots”	50 (7.4)[f]	
“Post procedural abortion physical experiences”	29 (4.3)[f]	
“Post abortion breast changes”	51 (7.6)[f]	

XSL•FO
RenderX

Conceptual group and topic name	Topic posts, n (%)[a,b]	Conceptual group posts, n (%)[c]
"Post abortion menstruation"	35 (5.2)[f]	
"Post procedural abortion bleeding"	123 (18.3)[f]	
"Post abortion clots"	193 (28.7)[f]	
Clinical abortion care		670 (9.3)
"Medication abortion clinical care"	1 (0.1)[g]	
"Procedural abortion fears"	59 (8.8)[g]	
"Accessing clinical abortion care"	72 (10.7)[g]	
"Clinic protestors"	4 (0.6)[g]	
"Clinical abortion care recovery"	15 (2.2)[g]	
"Abortion completion and ultrasound"	39 (5.8)[g]	
"Abortion experience fears"	244 (36.4)[g]	
"Abortion and ultrasound"	236 (35.2)[g]	
Potential pregnancy		550 (7.6)
"Potential pregnancy testing"	226 (41.1)[h]	
"Abortion completion pregnancy testing"	87 (15.8)[h]	
"Pregnancy risk and sex"	144 (26.2)[h]	
"Pregnancy risk and advice"	2 (0.4)[h]	
"Potential confirmation negative tests"	91 (16.5)[h]	
SMA [i] process		151 (2.1)
"SMA in illegal settings"	5 (3.3)[j]	
"Aid Access shipping"	50 (33.1)[j]	
"Aid Access ordering"	43 (28.5)[j]	
"Aid Access ordering and credibility"	11 (7.3)[j]	
"Aid Access ordering and shipping"	6 (4)[j]	
"SMA online navigation"	22 (14.6)[j]	
"SMA outside of the US"	14 (9.3)[j]	

[a]*Topic posts* refers to the count of posts classified under this topic based on rank 1 (highest-probability) classification.

[b]In total, 3 topics generated had 0 posts classified under them based on rank 1 (highest probability) and were dropped from the results. These topics were abortion decision-making timing and feelings, abortion decision-making desired children, and pregnancy risk and menstruation.

[c]"Conceptual group posts" refers to the count of all posts in this conceptual group. The percentages presented indicate the proportion of all posts in the sample in this conceptual group based on topic classification and conceptual group aggregation.

[d]n=2446.

[e]n=1807.

[f]n=672.

[g]n=670.

[h]n=550.

[i]SMA: self-managed abortion, defined for this research as taking action to end a pregnancy outside of the formal health care system with or without clinical support, which includes the use of safe medications such as misoprostol and mifepristone but also potentially harmful or ineffective methods [81].

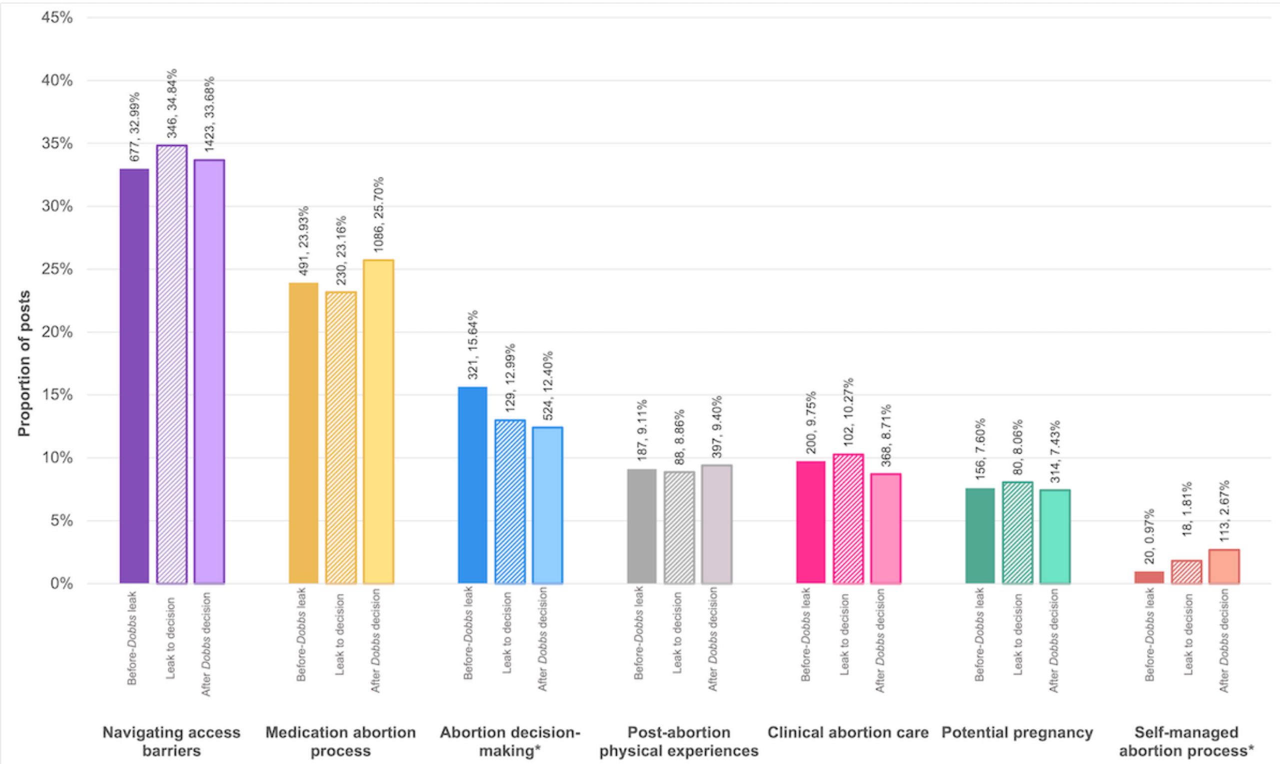
[j]n=151.

XSL•FO
RenderX

Assessing Changes in Conceptual Group Frequency Over Time

Counts and proportions of posts in each conceptual group across study subperiods based on the highest-probability (rank 1) classification, along with comparisons of proportions, are presented in Figure 3, with additional information in Multimedia Appendix 2. Overall, the number of posts to r/abortion increased after *Dobbs*. The proportion of posts related to *abortion decision-making* changed significantly across study periods (*P*=.002); with 974 posts in this group, 33% were from before the *Dobbs* leak (n=321); 13% from *Dobbs* leak to decision (n=129); and 54% after the *Dobbs* decision (n=524). In addition, while posts related to SMA were the least common, the proportion of posts in this group also changed significantly across study subperiods *(P*<.001), with 151 posts in this group, 13% were from before the *Dobbs* leak (n=20), 12% from *Dobbs* leak to decision (n=18), and 75% after the *Dobbs* decision (n=113). Changes across the study subperiods were not significant for other conceptual groups: *navigating access barriers* (*P*=.60), *medication abortion process* (*P*=.13), *post-abortion physical experiences* (*P*=.85), *clinical abortion care* (*P*=.19), *potential pregnancy* (*P*=.80).

Figure 3. Bar graph showing the proportion of r/abortion posts in each conceptual group (7 in total) in each study subperiod—before the Dobbs leak, Dobbs leak to decision, and after the Dobbs leak. The largest proportion of posts was in the navigating access barriers conceptual group, followed by medication abortion process and abortion decision-making. The proportion of posts about abortion decision-making changed significantly across subperiods, decreasing over time. Posts about self-managed abortion were the least common, but the proportion also changed across subperiods—increasing over time (N=7270 posts). *Indicates that the proportion of posts for this conceptual group varied significantly across study subperiods based on a Pearson chi-square test.



Discussion

Principal Findings

Given the rapidly changing landscape of abortion policy and access in 2022, with uncertainty and increasingly constrained pathways to care, online resources play a key role in providing access to abortion information, support, and services. Online communities in particular provide unique support for members of the public, as well as opportunities for researchers to understand emerging concerns and experiences of populations that may be hard to study through other methods [57-59]. We sought to contribute to the body of research that uses online information and leverages methods to work efficiently with large datasets to accelerate knowledge generation and provide novel insights into changing abortion-related experiences surrounding the *Dobbs* decision in 2022.

People posted on r/abortion about many different abortion-related concerns in 2022, coming to the community at various stages in their abortion trajectories. This ranged from posts seeking community input on pregnancy confirmation, abortion decision-making, pathways to access abortion, assessing abortion completion and postabortion experiences (physical and emotional), and support during and after their abortion in managing the process. Overall, the volume of posts and the breadth of content shared in them reflect that r/abortion is a community used abundantly and in alignment with its stated mission: “If you’re pregnant and don’t want to be, we can help you get an abortion. This is a proabortion, stigma-free space to ask questions, get information, and share your experiences” [82]. Given the increases in the use of r/abortion throughout 2022, this pseudonymous community may be an increasingly important online resource as abortion access continues to be restricted after *Dobbs*. Notably, other subreddits have

XSL•FO
RenderX

substantially larger volumes of community members and posts; for example, r/TwoXChromosomes had >13 million members and between 80 and 450 posts per day in 2022 compared to r/abortion with 34,000 members and between 5 and 41 posts per day [83,84]. However, r/abortion is an active and uniquely abortion-supportive community on Reddit facilitated by an active network of moderators (with the OARS), playing an important role in the landscape of abortion resources.

Overall, we observed that the conceptual group with the highest volume of posts focused on navigating barriers to abortion access, and the proportion of posts in that group did not change significantly across subperiods in 2022. This illustrates that r/abortion posters described barriers to abortion access even before *Dobbs* introduced national challenges to the legality of abortion. Previous research extensively documents pre-*Dobbs* barriers to abortion access and their adverse impacts on timely abortion access and the health of birthing people and their children; our findings illustrate the sustained influence of barriers on the experiences of r/abortion posters [10,28,29,85-101]. Posts describing navigating barriers to abortion access largely described challenges in accessing medication abortion based on topic modeling results, with some discussing the *Dobbs* decision, appointments and travel barriers, and costs. The overwhelming focus on challenges in accessing medication abortion in our sample may reflect broader trends, with use of medication abortion increasing since it became available in the United States in 2001 and now accounting for >60% of all abortions [102]. It may also reflect the unique priorities of the population using r/abortion and the potential for online resources to facilitate direct access to medication abortion. These findings are echoed and expanded by a qualitative analysis by our team using a subset of data from r/abortion following the *Dobbs* leak, which found that people described a variety of structural and social barriers to abortion access, including emerging challenges such as concerns about legal risks associated with accessing abortion. These findings also highlight the negative impacts of barriers on timely access to the desired modality of abortion care and mental health, as well as self-management of abortion out of necessity [36].

We also found that the proportion of posts related to SMA and abortion decision-making changed significantly throughout 2022. Abortion decision-making posts were more common before the *Dobbs* leak, whereas those related to self-management increased following the leak and decision. The relative decrease in posts related to decision-making may reflect shifting interests in the community when faced with increasing challenges. The increasing interest in SMA in our sample is notable, highlighting increasing interest in this abortion modality following the *Dobbs* leak and echoing other research indicating increased online interest in medication abortion and use of SMA during this time [103-105]. Taken together, these findings suggest that, as legal access to abortion is increasingly constrained, people may be focusing less on *whether* to have an abortion and more on *how* to access it under legal constraints. The rise in discussions about SMA implies that more individuals may be pursuing abortion outside formal health care systems, often due to legal, logistical, or financial barriers, raising concerns about legal risks and equitable access to preferred, supported abortion care.

In the SMA conceptual group, posts covered a range of topics related to procuring and receiving medications, including concerns about the credibility of online ordering platforms. The topics in the SMA conceptual group align with previous research describing the experiences of using online platforms (often Aid Access), highlighting concerns about scams, ordering, shipping, and potential surveillance [106]. Additional work by our team expands these findings related to SMA, highlighting the many questions posted about accessing and using medications to self-manage and critical gaps in posters' access to medical and legal information [107]. Notably, in the time since this study, the number of online platforms for ordering abortion medications has grown substantially, and abortion providers have started offering these medications legally from some states within the United States under shield laws [108]. As SMA becomes increasingly common and facilitated by more diverse pathways for obtaining medications, which is expected particularly in areas where abortion access is restricted [11,109], continued research is needed to understand emerging needs and challenges for people using online platforms to facilitate safe and satisfactory abortion experiences at home with limited clinical support. The experiences of people using the array of platforms facilitating access to medications for abortion are relatively unknown and warrant further research, particularly given likely differences in customer support and service quality across websites. In addition, efforts to protect and expand access to accurate online SMA information are critical as people rely on fully digital pathways to access services and face stigma, isolation, and legal risks that further limit access to information from nondigital sources.

This research explores abortion as a health and social concern on Reddit, with discussions influenced by the changing sociopolitical context in the United States. Previous qualitative research has explored how Reddit and r/abortion have been used for specific abortion-related concerns (eg, abortion costs and self-management barriers [28-33]), finding that analysis of Reddit data can provide meaningful insights into these concerns grounded in the unsolicited narratives shared in publicly available data. In addition, a recent study by Valdez et al [48] used NLP (specifically topic modeling with BERTopic) to describe discussions in a subset of posts to r/abortion and another Reddit community focused on abortion debate (r/AbortionDebate) and found that r/abortion was commonly used to seek and share social support, in contrast to r/AbortionDebate being used to discuss changing views on abortion. However, this analysis presents a novel approach to working with a theoretically complete sample of data from r/abortion during a period of sociopolitical changes that generated extreme uncertainty, fear, and constraint regarding abortion access in the United States. Our findings speak to the changing use of r/abortion during 2022 and the power of leveraging innovative research approaches grounded in content expertise to explore abortion as a critical public health concern.

Strengths and Limitations

There are several important limitations to keep in mind when interpreting these findings. First, while using data from an abortion subreddit leverages the power of Reddit as a pseudonymous platform known to invite discussion of sensitive

XSL•FO
RenderX

health concerns [24-26], we have no systematic demographic information about the r/abortion posters contributing to our sample. While most Reddit users across the entire platform are aged between 18 and 29 years (64%), male (64%), White individuals (70%), and based in the United States (52%) [110], research indicates that the demographics of members vary across subreddits, and in r/abortion, most users identify as female (86.1%) [111]. We also know that the population of Reddit and r/abortion users is unique [110,111] and likely does not represent people considering, seeking, and having abortions across the United States. As such, our findings cannot be interpreted as reflective of population-level experiences or concerns beyond r/abortion. Notably, in the post-*Dobbs* context, the access experiences of young people and those of a lower socioeconomic status are likely uniquely challenged as the impacts of *Dobbs* are experienced more intensely in these groups compared to others [1]. While the lack of demographic information limits our ability to determine who used r/abortion in 2022, there were stories shared in the community that represented a range of abortion experiences—including people who had not yet reached clinical abortion care, deciding between clinical abortion and SMA, and who never interacted with health care providers during their abortion processes. This representation of a diverse set of abortion experiences is a substantial contribution to our knowledge based on user-driven narratives and priorities.

Furthermore, although almost half of the users on Reddit live in the United States [20], it is a global platform. In qualitative analyses of a random subsample of these data, posts from people living outside of the United States were identified and excluded, accounting for 9% of the posts [37]. However, we took no similar steps for the current NLP analysis, introducing uncertainty about the connection of our results to the US abortion access context and policy change assuming that approximately 10% of all posts to r/abortion were from outside the United States. Despite this, we believe that US policy has a large enough impact on global perceptions of abortion access that this limitation does not substantially detract from our findings.

In addition, not all posts were correctly classified by BERTopic into topics, and perfect accuracy in classification is not expected. There are also methodological concerns related to our decision to use BERTopic model results to classify posts into a single topic as BERTopic operates under the assumption that documents can fall into multiple topics simultaneously. Choosing classification into a single topic allowed us to effectively make direct comparisons in the volume of posts per topic and conceptual group across subperiods but reduced consideration of the nuanced ways in which people talk about these topics within posts—often discussing multiple concerns that fall under different topics and perhaps different conceptual groups within 1 submission. It is plausible that this approach resulted in a substantial underestimation of the commonality of some topics within the corpus based on classification. However, in choosing to classify based on the topic with the highest probability for each post, we effectively captured the topics discussed most substantively in each submission—providing a simple and interpretable representation of the most dominant topic and related conceptual group described in each post.

Furthermore, while we attempted to be rigorous and precise in creating topic titles and aggregating topics into conceptual groups, the process was based on content expertise applied during subjective manual review guided by quantitative measures of topic similarity. Division of topics into conceptual groups was sometimes difficult, with posts in topics sometimes not exclusively describing clear concepts, underscoring concerns about using a single classification approach for each post. Delineating between topics related to medication abortion and SMA was conservative, erring on the side of only considering a topic as pertaining to SMA if it was very clearly focused on accessing, ordering, shipping, and receiving medications from an online platform. As such, this classification and aggregation approach likely underestimated the commonality of posts about SMA and perhaps the relative increase in those posts following the *Dobbs* leak and decision. This is further supported by the likelihood that people, despite the protection provided by pseudonymity on Reddit, may have limited their public sharing of information that overtly indicated that they were self-managing out of fear of prosecution [72,73].

Furthermore, the conceptualization of 3 subperiods in 2022 was based on key moments in national abortion policy but does not account for the rolling changes in abortion access within each of those periods. Even before the *Dobbs* leak, states were implementing abortion restrictions. Following the *Dobbs* decision, the enactment and enforcement of restrictions and bans across states has been progressive rather than the *Dobbs* decision functioning as a clear change point in all policies. As such, aggregating data into subperiods makes assumptions about the homogeneity of experiences of r/abortion users. Despite this concern, under the premise of exploring use of r/abortion and changes over time in the impacts of national abortion policy and related uncertainty about abortion access across states, using 3 subperiods as we did is a sound approach.

Conclusions

Our analysis provides a holistic view of the content of post submissions to r/abortion in 2022. In this research, we were able to merge content expertise and machine learning tools to describe people's posts to an online community for abortion during a time of extreme change and uncertainty in abortion access in the United States. Our findings highlight the critical role of r/abortion as an abortion-supportive resource, providing an online community for people to voice a vast array of concerns, questions, and experiences. They also illustrate how the use of r/abortion changed in 2022, speaking to the increased importance of SMA following the Dobbs leak. Overall, our findings highlight the need for further research exploring this trend across online platforms facilitating access to abortion information, support, and services—with particular focus on those providing access to abortion medications. As policies and pathways to abortion access continue to change across the United States, approaches leveraging NLP with sufficiently large samples of textual data present opportunities for timely monitoring, with the potential to reflect a broad range of abortion experiences, including those of people who have limited or no interaction with clinical abortion care.

XSL•FO
RenderX

Acknowledgments

The authors thank Aparna Roy for her contributions to data acquisition and analysis. This study was partially funded by the Society of Family Planning Emerging Scholars in Family Planning award (grant SFPRF16-ES9) to EP. This work was also supported by the Wallace Center for Maternal, Child, and Adolescent Health and the Bixby Center for Population, Health, and Sustainability in the School of Public Health at the University of California, Berkeley.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

EP contributed to conceptualization, methodology, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision, project administration, and funding acquisition. NP contributed to conceptualization, methodology, writing—review and editing, and funding acquisition. UU contributed to conceptualization, methodology, resources, and writing—review and editing. CM contributed to conceptualization, methodology, and writing—review and editing. CC contributed to conceptualization, methodology, writing—review and editing, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed processes for data acquisition and analyses.
[DOCX File , 25 KB - infodemiology_v5i1e72771_app1.docx]

Multimedia Appendix 2
Detailed results supplementing those in the manuscript text.
[DOCX File , 43 KB - infodemiology_v5i1e72771_app2.docx]

References

1. Fuentes L. Inequity in US abortion rights and access: the end of Roe is deepening existing divides. Guttmacher Institute. 2023 Jan 17. URL: https://www.guttmacher.org/2023/01/inequity-us-abortion-rights-and-access-end-roe-deepening-existing-divides [accessed 2025-01-30]
2. Felix M, Sobel L, Salganicoff A. Legal challenges to state abortion bans since the Dobbs decision. The Henry J. Kaiser Family Foundation. 2023 Jan 20. URL: https://www.kff.org/womens-health-policy/issue-brief/legal-challenges-to-state-abortion-bans-since-the-dobbs-decision/ [accessed 2025-04-29]
3. Abortion bans across the country: tracking restrictions by state. New York Times. 2024. URL: https://www.nytimes.com/interactive/2024/us/abortion-laws-roe-v-wade.html [accessed 2025-04-29]
4. #WeCount report April 2022 to June 2023. Society of Family Planning. 2023 Oct 24. URL: https://societyfp.org/wp-content/uploads/2023/10/WeCountReport_10.16.23.pdf [accessed 2025-01-30]
5. Grossman D, Joffe C, Kaller S, Kimport K, Kinsey E, Lerma K, et al. Care post-Roe: documenting cases of poor-quality care since the Dobbs decision. Advancing New Standards in Reproductive Health. 2023. URL: https://www.ansirh.org/sites/default/files/2023-05/Care%20Post-Roe%20Preliminary%20Findings.pdf [accessed 2025-01-30]
6. Sabbath EL, McKetchnie SM, Arora KS, Buchbinder M. US obstetrician-gynecologists' perceived impacts of post-Dobbs v Jackson state abortion bans. JAMA Netw Open 2024 Jan 02;7(1):e2352109 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.52109] [Medline: 38231510]
7. Wade overturned: our latest resources. Guttmacher Institute. URL: https://www.guttmacher.org/abortion-rights-supreme-court [accessed 2023-07-25]
8. Zhu DT, Zhao L, Alzoubi T, Shenin N, Baskaran T, Tikhonov J, et al. Implications of Dobbs v. Jackson for patients and providers: a scoping review. medRxiv Preprint posted online on July 12, 2023 [FREE Full text] [doi: 10.21203/rs.3.rs-3147593/v1]
9. Chor J, Tusken M, Young D, Lyman P, Gilliam M. Factors shaping women's pre-abortion communication with members of their social network. J Community Health 2019 Apr 10;44(2):265-271 [FREE Full text] [doi: 10.1007/s10900-018-0582-1] [Medline: 30306448]

XSL•FO
RenderX

10. Kavanaugh ML, Jerman J, Frohwirth L. "It's not something you talk about really": information barriers encountered by women who travel long distances for abortion care. Contraception 2019 Jul;100(1):79-84 [FREE Full text] [doi: 10.1016/j.contraception.2019.03.048] [Medline: 30980828]
11. Aiken AR, Starling JE, Scott JG, Gomperts R. Requests for self-managed medication abortion provided using online telemedicine in 30 US states before and after the Dobbs v Jackson women's health organization decision. JAMA 2022 Nov 01;328(17):1768-1770 [FREE Full text] [doi: 10.1001/jama.2022.18865] [Medline: 36318139]
12. Ralph L, Foster DG, Raifman S, Biggs MA, Samari G, Upadhyay U, et al. Prevalence of self-managed abortion among women of reproductive age in the United States. JAMA Netw Open 2020 Dec 01;3(12):e2029245 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.29245] [Medline: 33337493]
13. Upadhyay UD, Cartwright AF, Grossman D. Barriers to abortion care and incidence of attempted self-managed abortion among individuals searching Google for abortion care: a national prospective study. Contraception 2022 Mar;106:49-56 [FREE Full text] [doi: 10.1016/j.contraception.2021.09.009] [Medline: 34560051]
14. Altshuler AL, Gerns Storey HL, Prager SW. Exploring abortion attitudes of US adolescents and young adults using social media. Contraception 2015 Mar;91(3):226-233. [doi: 10.1016/j.contraception.2014.11.009] [Medline: 25537853]
15. Margo J, McCloskey L, Gupte G, Zurek M, Bhakta S, Feinberg E. Women's pathways to abortion care in South Carolina: a qualitative study of obstacles and supports. Perspect Sex Reprod Health 2016 Dec 28;48(4):199-207. [doi: 10.1363/psrh.12006] [Medline: 27893185]
16. Carroll E, White K. Abortion patients' preferences for care and experiences accessing services in Louisiana. Contracept X 2020;2:100016 [FREE Full text] [doi: 10.1016/j.conx.2019.100016] [Medline: 32550531]
17. Guendelman S, Yon E, Pleasants E, Hubbard A, Prata N. Shining the light on abortion: drivers of online abortion searches across the United States in 2018. PLoS One 2020 May 21;15(5):e0231672 [FREE Full text] [doi: 10.1371/journal.pone.0231672] [Medline: 32437369]
18. Reis BY, Brownstein JS. Measuring the impact of health policies using internet search patterns: the case of abortion. BMC Public Health 2010 Aug 25;10:514 [FREE Full text] [doi: 10.1186/1471-2458-10-514] [Medline: 20738850]
19. Johnston AC, Worrell JL, Di Gangi PM, Wasko M. Online health communities: an assessment of the influence of participation on patient empowerment outcomes. Inf Technol People 2013;26(2):213-235. [doi: 10.1108/ITP-02-2013-0040]
20. Reddit user and growth stats. Backlinko. URL: https://backlinko.com/reddit-users [accessed 2025-07-02]
21. Barney LJ, Griffiths KM, Banfield MA. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. BMC Psychiatry 2011 May 18;11:88. [doi: 10.1186/1471-244x-11-88]
22. De Choudhury M, De S. Mental health discourse on Reddit: self-disclosure, social support, and anonymity. Proc Int AAAI Conf Web Soc Media 2014;8(1):71-80. [doi: 10.1609/icwsm.v8i1.14526]
23. Newman MW, Lauterbach D, Munson SA, Resnick P, Morris ME. It's not that i don't have problems, i'm just not putting them on Facebook: challenges and opportunities in using online social networks for health. In: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work. 2011 Presented at: CSCW '11; March 19-23, 2011; Hangzhou, China. [doi: 10.1145/1958824.1958876]
24. Chen YY, Li CM, Liang JC, Tsai CC. Health information obtained from the internet and changes in medical decision making: questionnaire development and cross-sectional survey. J Med Internet Res 2018 Mar 12;20(2):e47 [FREE Full text] [doi: 10.2196/jmir.9370] [Medline: 29434017]
25. Sowles SJ, McLeary M, Optican A, Cahn E, Krauss MJ, Fitzsimmons-Craft EE, et al. A content analysis of an online pro-eating disorder community on Reddit. Body Image 2018 Mar;24:137-144 [FREE Full text] [doi: 10.1016/j.bodyim.2018.01.001] [Medline: 29414146]
26. Tamersoy A, De Choudhury M, Chau DH. Characterizing smoking and drinking abstinence from social media. HT ACM Conf Hypertext Soc Media 2015 Sep;2015:139-148 [FREE Full text] [doi: 10.1145/2700171.2791247] [Medline: 26640831]
27. Online Abortion Resource Squad. URL: https://www.abortionsquad.org [accessed 2025-03-01]
28. Higgins JA, Lands M, Valley TM, Carpenter E, Jacques L. Real-time effects of payer restrictions on reproductive healthcare: a qualitative analysis of cost-related barriers and their consequences among U.S. abortion seekers on Reddit. Int J Environ Res Public Health 2021 Aug 26;18(17):9013 [FREE Full text] [doi: 10.3390/ijerph18179013] [Medline: 34501602]
29. Richards NK, Masud A, Arocha JF. Online abortion empowerment in absence of parental and medical support: a thematic analysis of a Reddit community’s contributions to decision-making and access. Research Square Preprint posted online on May 1, 2021 [FREE Full text] [doi: 10.21203/rs.3.rs-523420/v1]
30. Jump L. “Abortion will never stop... It’s just safe abortion that will”: a uses and gratifications approach to examining abortion discourse on Reddit. University of Nebraska at Omaha. 2021 Jan. URL: https://www.proquest.com/docview/2556372220/abstract/2CA48EDEEEE7419EPQ/1 [accessed 2022-05-06]
31. Lewis RE, Messing AJ. Has anyone bought pills online? I'm afraid it's a scam”: US Reddit users’ experiences acquiring abortion pills online during COVID-19. Contraception 2021 Oct;104(4):457-458. [doi: 10.1016/j.contraception.2021.07.042]
32. Duarte NI, Moraes LL, Andrade CB. Abortion experience in the media: analysis of abortive paths shared in an online community. Cien Saude Colet 2018 Oct;23(10):3337-3346 [FREE Full text] [doi: 10.1590/1413-812320182310.14062018] [Medline: 30365853]

XSL•FO
RenderX

33. Jacques L, Carpenter E, Valley T, Alvarez B, Higgins J. Medication or surgical abortion? An exploratory study of patient decision making on a popular social media platform. Am J Obstet Gynecol 2021 Sep;225(3):344-347 [FREE Full text] [doi: 10.1016/j.ajog.2021.05.011] [Medline: 34022196]
34. Lands M, Carpenter E, Valley T, Jacques L, Higgins K. "Am I the only one who feels like this?": needs expressed online by abortion seekers. Soc Work 2023 Mar 16;68(2):103-111 [FREE Full text] [doi: 10.1093/sw/swad011] [Medline: 36795036]
35. Pleasants E, Parham L, Weidert K, Anderson E, Dolgins E, Prata N, et al. Waiting to start abortion: a qualitative exploration of narratives of waiting shared in a Reddit community for abortion post-Dobbs leak in 2022. Soc Sci Med 2024 May;349:116877 [FREE Full text] [doi: 10.1016/j.socscimed.2024.116877] [Medline: 38657319]
36. Pleasants E, Weidert K, Parham L, Anderson E, Dolgins E, Cheshire C, et al. Abortion access barriers shared in "r/abortion" after : a qualitative analysis of a Reddit community post-Dobbs decision leak in 2022. Sex Reprod Health Matters 2024 Dec 06;32(1):2426921 [FREE Full text] [doi: 10.1080/26410397.2024.2426921] [Medline: 39513330]
37. Pleasants E, Parham L, Weidert K, Anderson E, Dolgins E, Prata N, et al. “Is this normal?”: a qualitative exploration of narratives of social support in a Reddit-based community for abortion following the Dobbs leak in 2022. Wom Reprod Health 2024 Oct 18;12(1):157-178. [doi: 10.1080/23293691.2024.2413474]
38. Delnevo G, Mirri S, Monti L, Prandi C, Putra M, Roccetti M. Patients reactions to non-invasive and invasive prenatal tests: a machine-based analysis from Reddit posts. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2018 Presented at: ASONAM 2018; August 28-31, 2018; Barcelona, Spain. [doi: 10.1109/asonam.2018.8508614]
39. Ammari T, Schoenebeck S, Romero D. Self-declared throwaway accounts on Reddit: how platform affordances and shared norms enable parenting disclosure and support. Proc ACM Hum Comput Interact 2019 Nov 07;3(CSCW):1-30. [doi: 10.1145/3359237]
40. Amith M, Cohen T, Cunningham R, Savas LS, Smith N, Cuccaro P, et al. Mining HPV vaccine knowledge structures of young adults from Reddit using distributional semantics and pathfinder networks. Cancer Control 2020 Jan 08;27(1):1073274819891442 [FREE Full text] [doi: 10.1177/1073274819891442] [Medline: 31912742]
41. Nobles AL, Leas EC, Althouse BM, Dredze M, Longhurst CA, Smith DM, et al. Requests for diagnoses of sexually transmitted diseases on a social media platform. JAMA 2019 Nov 05;322(17):1712-1713 [FREE Full text] [doi: 10.1001/jama.2019.14390] [Medline: 31688875]
42. Gold KJ, Boggs ME, Mugisha E, Palladino CL. Internet message boards for pregnancy loss: who's on-line and why? Womens Health Issues 2012 Jan;22(1):e67-e72 [FREE Full text] [doi: 10.1016/j.whi.2011.07.006] [Medline: 21907592]
43. Pleasants E, Ryan JH, Ren C, Prata N, Gomez AM, Marshall C. Exploring language used in posts on r/birthcontrol: case study using data from Reddit posts and natural language processing to advance contraception research. J Med Internet Res 2023 Jun 30;25:e46342 [FREE Full text] [doi: 10.2196/46342] [Medline: 37389907]
44. Kalabikhina I, Zubova E, Loukachevitch N, Kolotusha A, Kazbekova Z, Banin E, et al. Identifying reproductive behavior arguments in social media content users’ opinions through natural language processing techniques. Popul Econ 2023 Jun 30;7(2):40-59. [doi: 10.3897/popecon.7.e97064]
45. Gao S, Pandya S, Agarwal S, Sedoc J. Topic modeling for maternal health using Reddit. In: Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis. 2021 Presented at: LOUHI@EACL 2021; April 19, 2021; Online URL: https://aclanthology.org/2021.louhi-1.8
46. Geiss HJ, Sakketou F, Flek L. OK boomer: probing the socio-demographic divide in echo chambers. In: Proceedings of the Tenth International Workshop on Natural Language Processing for Social Media. 2022 Presented at: SocialNLP'22; July 14-15, 2022; Seattle, Washington. [doi: 10.18653/v1/2022.socialnlp-1.8]
47. Beel J, Xiang T, Soni S, Yang D. Linguistic characterization of divisive topics online: case studies on contentiousness in abortion, climate change, and gun control. Proc Int AAAI Conf Web Soc Media 2022 May 31;16(1):32-42 [FREE Full text] [doi: 10.1609/icwsm.v16i1.19270]
48. Valdez D, Mena-Meléndez L, Crawford BL, Jozkowski KN. Analyzing Reddit forums specific to abortion that yield diverse dialogues pertaining to medical information seeking and personal worldviews: data mining and natural language processing comparative study. J Med Internet Res 2024 Mar 14;26:e47408 [FREE Full text] [doi: 10.2196/47408] [Medline: 38354044]
49. Machine learning, explained. MIT Sloan School of Management. 2021 Apr 21. URL: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained [accessed 2023-07-26]
50. Grimmer J, Roberts ME, Stewart BM. Text as data: a new framework for machine learning and the social sciences. Princeton University Press. 2022. URL: https://politicalscience.stanford.edu/publications/text-data-new-framework-machine-learning-and-social-sciences [accessed 2025-01-30]
51. Monarch R. Human-in-the-Loop Machine Learning: Active Learning and Annotation for Human-Centered AI. Shelter Island, NY: Manning Publications; 2021.
52. Chang T, DeJonckheere M, Vydiswaran VG, Li J, Buis LR, Guetterman TC. Accelerating mixed methods research with natural language processing of big text data. J Mixed Methods Res 2021 Jun 16;15(3):398-412. [doi: 10.1177/15586898211021196]

XSL•FO
RenderX

53. Latack KR, Patel J, Moog D, Spencer D, Nguyen BT. Withdrawal method inquiries and user experiences: an analysis of content posted on 4 gendered forums on Reddit. Contraception 2021 Aug;104(2):170-175. [doi: 10.1016/j.contraception.2021.04.005] [Medline: 33852898]
54. Meacham MC, Nobles AL, Tompkins DA, Thrul J. "I got a bunch of weed to help me through the withdrawals": naturalistic cannabis use reported in online opioid and opioid recovery community discussion forums. PLoS One 2022 Feb 8;17(2):e0263583 [FREE Full text] [doi: 10.1371/journal.pone.0263583] [Medline: 35134074]
55. Balsamo D, Bajardi P, Salomone A, Schifanella R. Patterns of routes of administration and drug tampering for nonmedical opioid consumption: data mining and content analysis of Reddit discussions. J Med Internet Res 2021 Jan 04;23(1):e21212 [FREE Full text] [doi: 10.2196/21212] [Medline: 33393910]
56. Smith I, Youssef GJ, Shatte A, Teague SJ, Knight T, Macdonald JA. “You are not alone”: a big data and qualitative analysis of men's unintended fatherhood. SSM Qual Res Health 2022 Dec;2:100085. [doi: 10.1016/j.ssmqr.2022.100085]
57. Upadhyay UD, McCook AA, Bennett AH, Cartwright AF, Roberts SC. State abortion policies and Medicaid coverage of abortion are associated with pregnancy outcomes among individuals seeking abortion recruited using Google Ads: a national cohort study. Soc Sci Med 2021 Apr;274:113747 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113747] [Medline: 33642070]
58. Moseson H, Wollum A, Seymour JW, Zuniga C, Thompson TA, Gerdts C. Comparison of Facebook, Google Ads, and Reddit for the recruitment of people who considered but did not obtain abortion care in the United States: cross-sectional survey. JMIR Form Res 2021 Mar 24;5(2):e22854. [doi: 10.2196/22854] [Medline: 33625368]
59. O'Donnell J, Goldberg A, Lieberman E, Betancourt T. "I wouldn't even know where to start": unwanted pregnancy and abortion decision-making in Central Appalachia. Reprod Health Matters 2018 Dec 29;26(52):1513270 [FREE Full text] [doi: 10.1080/09688080.2018.1513270] [Medline: 30360690]
60. Pushshift Reddit API documentation. GitHub. URL: https://github.com/pushshift/api [accessed 2022-05-06]
61. Biester L, Matton K, Rajendran J, Provost EM, Mihalcea R. Understanding the impact of COVID-19 on online mental health forums. ACM Trans Manage Inf Syst 2021 Sep 08;12(4):1-28. [doi: 10.1145/3458770]
62. Šćepanović S, Aiello LM, Zhou K, Joglekar S, Quercia D. The healthy states of America: creating a health taxonomy with social media. Proc Int AAAI Conf Web Soc Media 2021 May 22;15:621-632. [doi: 10.1609/icwsm.v15i1.18089]
63. Johnson AK, Bhaumik R, Nandi D, Roy A, Mehta SD. Sexually transmitted disease-related Reddit posts during the COVID-19 pandemic: latent Dirichlet allocation analysis. J Med Internet Res 2022 Oct 31;24(10):e37258 [FREE Full text] [doi: 10.2196/37258] [Medline: 36219757]
64. Sparck Jones K. A statistical interpretation of term specificity and its application in retrieval. J Document 1972 Jan;28(1):11-21. [doi: 10.1108/eb026526]
65. Ren C, Bloemraad I. New methods and the study of vulnerable groups: using machine learning to identify immigrant-oriented nonprofit organizations. Socius 2022 Feb 18;8. [doi: 10.1177/23780231221076992]
66. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online on March 11, 2022 [FREE Full text]
67. Grootendorst M. Quick Start - BERTopic. BERTopic. URL: https://maartengr.github.io/BERTopic/getting_started/quickstart/quickstart.html [accessed 2023-07-17]
68. Grootendorst M. BERTopic. URL: https://maartengr.github.io/BERTopic/api/bertopic.html [accessed 2024-04-05]
69. Proferes N, Jones N, Gilbert S, Fiesler C, Zimmer M. Studying Reddit: a systematic overview of disciplines, approaches, methods, and ethics. Soc Media Soc 2021 May 26;7(2). [doi: 10.1177/20563051211019004]
70. Betts D, Dahlen HG, Smith CA. A search for hope and understanding: an analysis of threatened miscarriage internet forums. Midwifery 2014 Jun;30(6):650-656. [doi: 10.1016/j.midw.2013.12.011] [Medline: 24439850]
71. Mannheimer S, Hull EA. Sharing selves: developing an ethical framework for curating social media data. Int J Digit Curation 2018 Apr 18;12(2):196-209. [doi: 10.2218/ijdc.v12i2.518]
72. Sandvik R. How US police use digital data to prosecute abortions. TechCrunch. 2023 Jan 27. URL: https://techcrunch.com/2023/01/27/digital-data-roe-wade-reproductive-privacy/ [accessed 2023-05-02]
73. Dym B, Fiesler C. Ethical and privacy considerations for research using online fandom data. Transform Works Cult 2020;33 [FREE Full text] [doi: 10.3983/twc.2020.1733]
74. Pleasure ZH, Frohwirth LF, Li N, Polis CB. A content analysis of Reddit users' posts about challenges to contraceptive care-seeking during COVID-19-related restrictions in the United States. J Health Commun 2022 Oct 03;27(10):746-754. [doi: 10.1080/10810730.2022.2157911] [Medline: 36519832]
75. Moreno MA, Goniu N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):708-713 [FREE Full text] [doi: 10.1089/cyber.2012.0334] [Medline: 23679571]
76. Reagle J. Disguising Reddit sources and the efficacy of ethical research. Ethics Inf Technol 2022 Sep 10;24(3):41 [FREE Full text] [doi: 10.1007/s10676-022-09663-w] [Medline: 36105629]
77. Bruckman A. Ethical guidelines for research online. Georgia Institute of Technology. URL: https://faculty.cc.gatech.edu/~asb/ethics/ [accessed 2025-01-30]
78. Pushshift Reddit search. Reddit. URL: https://redditsearch.io/ [accessed 2023-05-02]

XSL•FO
RenderX

79. Plagiarism checker. Grammarly. URL: https://www.grammarly.com/plagiarism-checker [accessed 2023-05-02]
80. Reagle J. Even pseudonyms and throwaways delete their Reddit posts. First Monday 2023 Jun 15;28(6) [FREE Full text] [doi: 10.5210/fm.v28i6.13193]
81. Moseson H, Herold S, Filippa S, Barr-Walker J, Baum SE, Gerdts C. Self-managed abortion: a systematic scoping review. Best Pract Res Clin Obstet Gynaecol 2020 Mar;63:87-110 [FREE Full text] [doi: 10.1016/j.bpobgyn.2019.08.002] [Medline: 31859163]
82. r/abortion. Reddit. URL: https://www.reddit.com/r/abortion/?rdt=55062 [accessed 2024-02-12]
83. r/TwoXChromosomes: You are the community. You have all the power of the internet to mold it. Subreddit. URL: https://subredditstats.com/r/twoxchromosomes [accessed 2024-03-08]
84. r/abortion stats. Subreddit. URL: https://subredditstats.com/r/abortion [accessed 2024-03-08]
85. Foster DG. The Turnaway Study: Ten Years, a Thousand Women, and the Consequences of Having—or Being Denied—an Abortion. New York, NY: Scribner; 2020.
86. Ostrach B, Cheyney M. Navigating social and institutional obstacles: low-income women seeking abortion. Qual Health Res 2014 Jul 16;24(7):1006-1017. [doi: 10.1177/1049732314540218] [Medline: 24970251]
87. Jerman J, Frohwirth L, Kavanaugh ML, Blades N. Barriers to abortion care and their consequences for patients traveling for services: qualitative findings from two states. Perspect Sex Reprod Health 2017 Jun 10;49(2):95-102 [FREE Full text] [doi: 10.1363/psrh.12024] [Medline: 28394463]
88. Biggs MA, Upadhyay UD, McCulloch CE, Foster DG. Women's mental health and well-being 5 years after receiving or being denied an abortion: a prospective, longitudinal cohort study. JAMA Psychiatry 2017 Mar 01;74(2):169-178. [doi: 10.1001/jamapsychiatry.2016.3478] [Medline: 27973641]
89. Biggs MA, Kaller S, Ralph L. Barriers accessing abortion care and their association with psychological well-being. Contraception 2020 May;101(5):355. [doi: 10.1016/j.contraception.2020.03.010]
90. Roberts SC, Gould H, Kimport K, Weitz TA, Foster DG. Out-of-pocket costs and insurance coverage for abortion in the United States. Womens Health Issues 2014 Mar;24(2):e211-e218. [doi: 10.1016/j.whi.2014.01.003] [Medline: 24630423]
91. Roberts SC, Fuentes L, Kriz R, Williams V, Upadhyay UD. Implications for women of Louisiana's law requiring abortion providers to have hospital admitting privileges. Contraception 2015 May;91(5):368-372. [doi: 10.1016/j.contraception.2015.02.001] [Medline: 25744615]
92. Doran F, Nancarrow S. Barriers and facilitators of access to first-trimester abortion services for women in the developed world: a systematic review. J Fam Plann Reprod Health Care 2015 Jul 22;41(3):170-180 [FREE Full text] [doi: 10.1136/jfprhc-2013-100862] [Medline: 26106103]
93. Fried MG. Abortion in the United States: barriers to access. Health Human Rights 2000;4(2):174-194. [doi: 10.2307/4065200]
94. Mello K, Smith MH, Hill BJ, Chakraborty P, Rivlin K, Bessett D, et al. Federal, state, and institutional barriers to the expansion of medication and telemedicine abortion services in Ohio, Kentucky, and West Virginia during the COVID-19 pandemic. Contraception 2021 Jul;104(1):111-116. [doi: 10.1016/j.contraception.2021.04.020]
95. Upadhyay UD. Innovative models are needed for equitable abortion access in the USA. Lancet Public Health 2017 Nov;2(11):e484-e485. [doi: 10.1016/s2468-2667(17)30181-0]
96. Upadhyay UD, Weitz TA, Jones RK, Barar RE, Foster DG. Denial of abortion because of provider gestational age limits in the United States. Am J Public Health 2014 Sep;104(9):1687-1694. [doi: 10.2105/AJPH.2013.301378] [Medline: 23948000]
97. Dennis A, Henshaw SK, Joyce TJ, Finer LB, Blanchard K. The impact of laws requiring parental involvement for abortion: a literature review. Guttmacher Institute. 2009 Mar. URL: https://www.guttmacher.org/report/impact-laws-requiring-parental-involvement-abortion-literature-review [accessed 2021-07-29]
98. Joyce TJ, Henshaw SK, Dennis A, Finer LB, Blanchard K. The impact of state mandatory counseling and waiting period laws on abortion: a literature review. Guttmacher Institute. 2009 Apr. URL: https://www.guttmacher.org/report/impact-state-mandatory-counseling-and-waiting-period-laws-abortion-literature-review [accessed 2025-01-30]
99. Myers CK. Cooling off or burdened? The effects of mandatory waiting periods on abortions and births. SSRN. 2021. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3865459 [accessed 2025-07-02]
100. de Londras F, Cleeve A, Rodriguez MI, Farrell A, Furgalska M, Lavelanet A. The impact of mandatory waiting periods on abortion-related outcomes: a synthesis of legal and health evidence. BMC Public Health 2022 Jun 21;22(1):1232 [FREE Full text] [doi: 10.1186/s12889-022-13620-z] [Medline: 35725439]
101. Bennett AH, Marshall C, Kimport K, Deardorff J, Gómez AM. "Have you ever wanted or needed an abortion you did not get?" Data from a 2022 nationally representative online survey in the United States. Contraception 2023 Jul;123:110007 [FREE Full text] [doi: 10.1016/j.contraception.2023.110007] [Medline: 36931550]
102. Jones RK, Friedrich-Karnik A. Medication abortion accounted for 63% of all US abortions in 2023—an increase from 53% in 2020. Guttmacher Institute. 2024. URL: https://www.guttmacher.org/2024/03/medication-abortion-accounted-63-all-us-abortions-2023-increase-53-2020 [accessed 2025-05-05]
103. Poliak A, Satybaldiyeva N, Strathdee SA, Leas EC, Rao R, Smith D, et al. Internet searches for abortion medications following the leaked supreme court of the United States draft ruling. JAMA Intern Med 2022 Sep 01;182(9):1002-1004 [FREE Full text] [doi: 10.1001/jamainternmed.2022.2998] [Medline: 35767270]

XSL•FO
RenderX

104. Gupta S, Perry B, Simon K. Trends in abortion- and contraception-related internet searches after the US Supreme Court overturned constitutional abortion rights: how much do state laws matter? JAMA Health Forum 2023 Apr 07;4(4):e230518 [FREE Full text] [doi: 10.1001/jamahealthforum.2023.0518] [Medline: 37115538]
105. #WeCount Report April 2022 to December 2022. Society of Family Planning. 2023 Apr 11. URL: https://societyfp.org/wp-content/uploads/2023/03/WeCountReport_April2023Release.pdf [accessed 2025-01-30]
106. Madera M, Johnson DM, Broussard K, Tello-Pérez LA, Ze-Noah CA, Baldwin A, et al. Experiences seeking, sourcing, and using abortion pills at home in the United States through an online telemedicine service. SSM Qual Res Health 2022 Dec;2:100075 [FREE Full text] [doi: 10.1016/j.ssmqr.2022.100075] [Medline: 37503356]
107. Weidert K, Pleasants E, Parham L, Dolgins E, Anderson E, Prata N. Have others had this experience? A qualitative analysis of posts on self-managed abortion to US-based Reddit community. Perspect Sex Reprod Health 2025 Jun 30;57(2):175-184. [doi: 10.1111/psrh.70011] [Medline: 40444799]
108. Websites that sell pills. Plan C. URL: https://www.plancpills.org/websites-that-sell-pills [accessed 2024-03-08]
109. Aiken AR, Tello-Pérez LA, Madera M, Starling JE, Johnson DM, Broussard K, et al. Factors associated with knowledge and experience of self-managed abortion among patients seeking care at 49 US abortion clinics. JAMA Netw Open 2023 Apr 18;6(4):e238701. [doi: 10.1001/jamanetworkopen.2023.8701]
110. Sattelberg W. The demographics of Reddit: who uses the site? Alphr. 2021 Apr 6. URL: https://www.alphr.com/demographics-reddit/ [accessed 2022-03-03]
111. Burkhart B. Subreddit gender ratios. Subreddit. URL: https://bburky.com/subredditgenderratios/ [accessed 2022-04-03]

Abbreviations

API: application programming interface
NLP: natural language processing
OARS: Online Abortion Resource Squad
SMA: self-managed abortion

Edited by T Mackey; submitted 17.02.25; peer-reviewed by C Ren, M Meacham, VSK Kancharla, LP Gorrepati; comments to author 17.04.25; revised version received 21.05.25; accepted 10.06.25; published 09.07.25.

Please cite as:
Pleasants E, Prata N, Upadhyay UD, Marshall C, Cheshire C
Using Natural Language Processing to Describe the Use of an Online Community for Abortion During 2022: Dynamic Topic Modeling Analysis of Reddit Posts
JMIR Infodemiology 2025;5:e72771
URL: https://infodemiology.jmir.org/2025/1/e72771
doi:10.2196/72771
PMID:

©Elizabeth Pleasants, Ndola Prata, Ushma D Upadhyay, Cassondra Marshall, Coye Cheshire. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 09.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Differential Analysis of Age, Gender, Race, Sentiment, and Emotion in Substance Use Discourse on Twitter During the COVID-19 Pandemic: A Natural Language Processing Approach

Julina Maharjan[1], BE; Ruoming Jin[1], Prof Dr; Jennifer King[2], Prof Dr; Jianfeng Zhu[1], PhD; Deric Kenne[2], Prof Dr

[1]Department of Computer Science, Kent State University, Kent, OH, United States

[2]Department of Public Health, Kent State University, Kent, OH, United States

Corresponding Author:
Julina Maharjan, BE
Department of Computer Science
Kent State University
800 E. Summit St.
Kent, OH, 44242
United States
Phone: 1 3305931365
Email: jmaharja@kent.edu

Abstract

Background: User demographics are often hidden in social media data due to privacy concerns. However, demographic information on substance use (SU) can provide valuable insights, allowing public health policy makers to focus on specific cohorts and develop efficient prevention strategies, especially during global crises such as the COVID-19 pandemic.

Objective: This study aimed to analyze SU trends at the user level across different demographic dimensions, such as age, gender, race, and ethnicity, with a focus on the COVID-19 pandemic. The study also establishes a baseline for SU trends using social media data.

Methods: The study was conducted using large-scale English-language data from Twitter (now known as X) over a 3-year period (2019, 2020, and 2021), comprising 1.13 billion posts. Following preprocessing, the SU posts were identified using our custom-trained deep learning model (Robustly Optimized Bidirectional Encoder Representations From Transformers Pretraining Approach [RoBERTa]), which resulted in the identification of 9 million SU posts. Then, demographic attributes, such as user type, age, gender, race, and ethnicity, as well as sentiments and emotions associated with each post, were extracted via a collection of natural language processing modules. Finally, various qualitative analyses were performed to obtain insight into user behaviors based on demographics.

Results: The highest level of user participation in SU discussions was observed in 2020, with a 22.18% increase compared to 2019 and a 25.24% increase compared to 2021. Throughout the study period, male users and teenagers increasingly dominated the SU discussions across all substance types. During the COVID-19 pandemic, user participation in prescription medication discussions was notably higher among female users compared to other substance types. In addition, alcohol use increased by 80% within 2 weeks after the global pandemic declaration in 2020.

Conclusions: This study presents a large-scale, fine-grained analysis of SU on social media data, examining trends by age, gender, race, and ethnicity before, during, and after the COVID-19 pandemic. Our findings, contextualized with sociocultural and pandemic-specific factors, provide actionable insights for targeted public health interventions. This study establishes social media data (powered with artificial intelligence and natural language processing tools) as a valuable platform for real-time SU surveillance and prevention during crises.

(JMIR Infodemiology 2025;5:e67333) doi:10.2196/67333

KEYWORDS

substance use; social media; deep learning; natural language processing; NLP; COVID-19; age; gender; race; sentiment; emotion; artificial intelligence; AI

XSL•FO
RenderX

Introduction

Overview

Substance use (SU) prevalence varies across demographics such as age, gender, race, and ethnicity. During the COVID-19 pandemic, these differences became more pronounced. The pandemic not only increased global SU, with overdose deaths rising by 29.4% [1], but also exacerbated societal and racial inequalities [2,3] and significantly impacted mental health [4-7]. As people often turn to substances as a coping mechanism during crises [8,9], the pandemic likely led to increased SU [10], particularly among populations considered vulnerable [11]. Investigating how these trends shifted across different demographic groups during the pandemic is crucial for understanding public health challenges and developing targeted interventions.

Background

Gender, Age, and Racial Disparities in SU

According to the National Center for Drug Abuse Statistics (NCDAS) [12], men are more likely than women to use illicit drugs. In 2020, 22% of male individuals and 17% of female individuals used illegal drugs or misused prescription drugs within the last year, and the highest prevalence was among individuals aged 18 to 25 (39%), followed by those aged 26 to 29 (34%) [13]. Racial and ethnic disparities have always been prevalent in the history of drug use. For instance, White individuals were more likely to misuse prescription drugs, while other races were more likely to use other illicit drugs [14]. Similarly, opioid overdose death rates were higher in Black individuals [15]. Furthermore, the disparities by race and ethnicity were also found to be varied with age. For most SU disorders, estimated prevalence was higher for White participants at younger ages and Black participants at older ages [16].

Importance of Studying SU During the COVID-19 Pandemic

Given the preexisting disparities in SU, the COVID-19 pandemic likely exacerbated these trends. According to the Centers for Disease Control and Prevention [1], COVID-19 mortality rates from January 1, 2020, to May 31, 2024, varied significantly by age, gender, race, and ethnicity. Non-Hispanic White individuals accounted for 67% of deaths, individuals aged ≥75 years represented approximately 54% of deaths, and male individuals comprised 54% of the mortality rate. Simultaneously, the COVID-19 pandemic brought significant social and economic changes, disproportionately affecting minoritized populations and those considered underprivileged [16,17]. The rapid spread of the virus overwhelmed health care services, leading to lower priority for treatment for racial and ethnic minority people and individuals considered economically disadvantaged [18,19]. This discrimination exacerbated mental health issues [4], also highlighted by the US Centers for Disease Control and Prevention [1], which noted disparities in mental health and substance misuse among racial and ethnic minority populations due to unequal access to care, psychosocial stress, and social determinants of health. Given the disparities in COVID-19 mortality rates by age, gender, race, and ethnicity and the social and economic challenges exacerbated by the pandemic, studying SU trends across different demographic groups requires high attention. The disproportionate impact on minoritized populations and those considered underprivileged highlights the need to understand how these factors influenced SU, which will aid in developing targeted public health strategies to address the specific needs of populations considered affected.

Natural Language Processing and Its Application in Health Care

The advent of advanced natural language processing (NLP) techniques, particularly deep learning models, has revolutionized the health care field, enabling researchers to extract meaningful insights from vast and complex datasets for advanced decision-making. Health care data, which include unstructured sources such as medical reports, electronic health records, clinical trials, and social media, has traditionally posed significant challenges for analysts due to its volume, variability, and complexity. Recent advancements in NLP have addressed these challenges by facilitating tasks such as health information retrieval and extraction, text summarization, sentiment and emotion analysis, and the construction of medical ontologies and knowledge graphs. For instance, studies have demonstrated the utility of NLP in analyzing social media data to monitor public health trends, such as SU and mental health discussions during crises [20]. Similarly, NLP has been applied to identify patterns in clinical texts and patient narratives, enabling personalized health care interventions and improved decision-making [21]. These applications highlight the transformative potential of NLP in health care, particularly in leveraging unstructured data to address pressing public health challenges.

Related Study

The study of SU prevalence across demographics has predominantly relied on survey-based research conducted by national agencies such as the Substance Abuse and Mental Health Services Administration (SAMHSA) [13] and the National Institute on Drug Abuse (NIDA) [22]. For example, the National Survey on Drug Use and Health (NSDUH) [12], administered by SAMHSA, provides comprehensive data on SU and mental health issues among the US population aged ≥12 years. Similarly, the Monitoring the Future [23] survey, funded by NIDA, focuses on SU patterns among youth by surveying middle and high school students (grades 8, 10, and 12). Both surveys provide detailed reports on the use of illicit and nonillicit drugs, disaggregated by age, gender, race, and ethnicity at a national level. In addition to these national surveys, various individual studies [16,24] have also explored SU disparities across demographics such as age, gender, race, and ethnicity.

While these surveys offer valuable insights, their scope is often limited by the diversity of true populations and the duration of the studied period. Traditional survey methods often rely on self-reported data, which can be affected by social desirability bias and recall errors. In addition, surveys are typically conducted annually or biennially, providing only periodic snapshots of SU trends. Moreover, the COVID-19 pandemic

XSL•FO
RenderX

posed additional challenges for data collection; for example, SAMHSA 2020 was only able to collect data for the first and fourth quarters due to restrictions on in-person activities [13].

In contrast, social media data addresses many of these limitations. Social media platforms capture real-time, user-generated content that often reflects more authentic behaviors and sentiments. In addition to this, researchers have also shown the prevalence of SU discussions on social media [25-32], possibly due to its anonymity feature. Likewise, the continuous stream of data allows researchers [26,28,33-35] to monitor trends as they evolve, providing insights that are not possible with traditional survey methods. In addition, the vast amount of data available on social media enables a more detailed analysis across a large population [34], including those that might be underrepresented in surveys [32,36].

Despite the extensive research on SU trends, there remains a gap in understanding how these trends vary across different demographic groups, especially in the context of the COVID-19 pandemic. Existing studies have primarily relied on less diverse survey data or short-term real-time data, often overlooking the dynamic and nuanced shifts in SU behavior during global crises. This study aimed to address this gap by leveraging a large-scale social media dataset to provide a more granular and continuous analysis of SU trends across diverse demographics before, during, and after the pandemic. The following research questions (RQs) are designed to explore these trends in detail, offering insights into how age, gender, race, ethnicity, and emotional factors have influenced SU patterns during the study period:

1. What are the statistical distributions of overall SU posts and their users, categorized by key demographic variables in prepandemic, pandemic, and postpandemic periods?
2. To what extent do the SU patterns and demographic distributions observed in Twitter (now known as X) discourse from 2019 to 2021 correspond with or differ from the baseline trends reported by the NCDAS and other research?
3. What are the temporal trends in SU posts across different substance types throughout the study period, and how does the frequency of user posting behavior vary over time for each substance type?
4. How did the number of individuals discussing alcohol change within the first 2 weeks following the pandemic declaration compared to other users, and what short-term trends in user behavior emerged across different age groups, genders, races, and sentiments during this period?
5. What are the trends in user participation across different demographics for each substance type?
6. What emotional expressions are prevalent across all substance types?

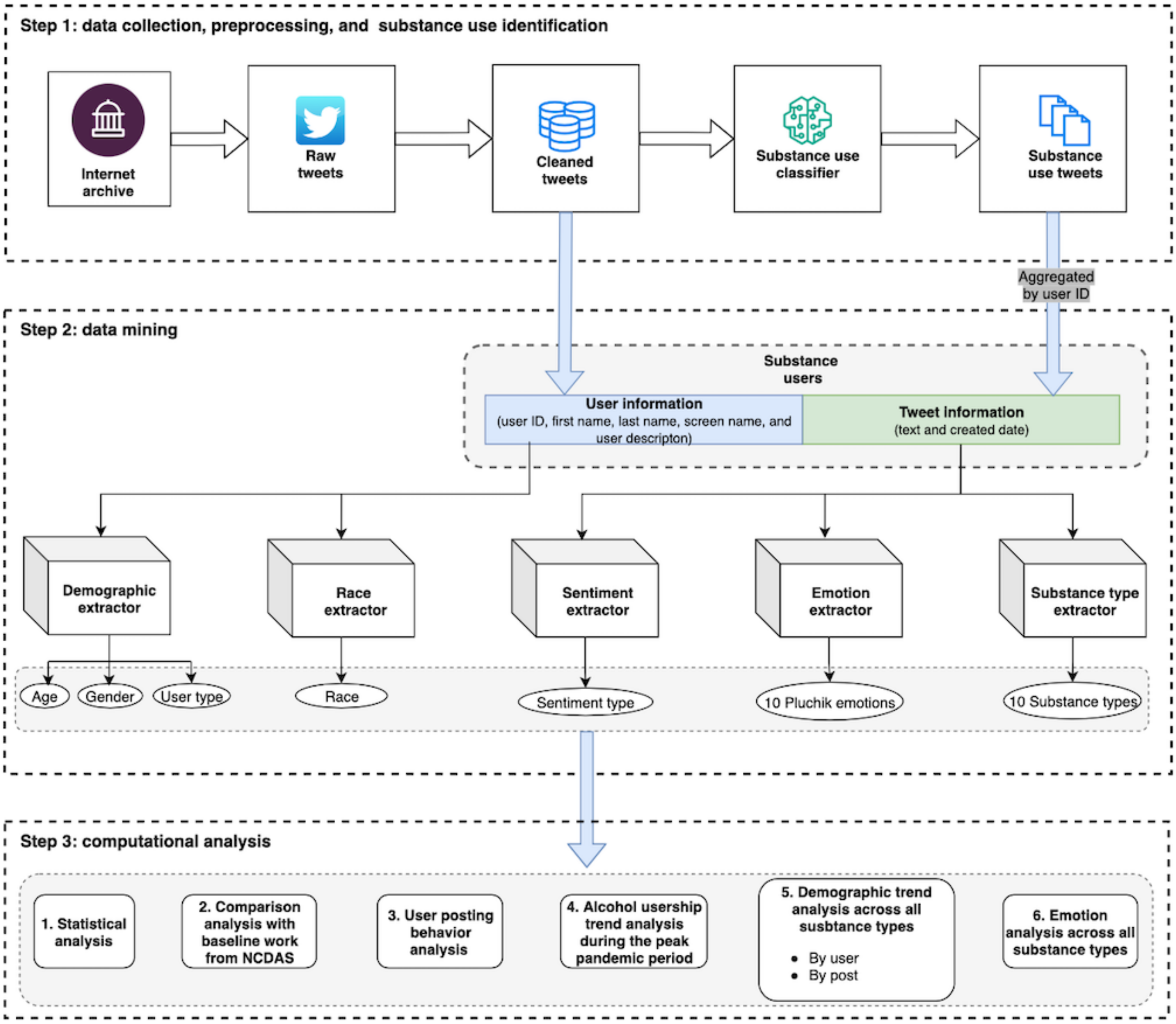
Methods

Overview

This study extends our previous research [37], which developed and validated a SU classifier using Robustly Optimized Bidirectional Encoder Representations From Transformers Pretraining Approach (RoBERTa) model [38] to identify SU posts from a cleaned dataset of 1.13 billion English-language posts (2019-2021). While our previous research [37] focused on analyzing SU at the post level, this study focuses on the user level. Thus, the key task of this study is to formulate the user base of SU posts from the previous study, followed by data mining (step 2) and computational analysis (step 3). As illustrated in Figure 1, step 1 encompasses the data collection, preprocessing, and SU identification modules. Step 2 involves user base formation and data mining methods for extracting demographic and emotional information associated with users and posts. Finally, step 3 involves computational analysis of the extracted data for users across various dimensions.

XSL•FO
RenderX

Figure 1. Methodology of the study design. NCDAS: National Center for Drug Abuse Statistics.



Data Collection, Preprocessing, and SU Identification

We downloaded the raw tweet data from the Internet Archive [39], covering the period from January 2019 to December 2021. Unlike our previous study, which extracted only tweet information, this study also retrieved user information necessary for user-level analysis. Thus, the downloaded tweets had 2 types of information: tweet information and user information. The text information comprises text and the created date of a tweet, while user information comprises user ID, screen name, first name, last name, and user description.

The data processing of tweets was carried out in multiple steps. Initially, we filtered out all non-US tweets and duplicate or retweeted tweets to focus our research on English-language tweets and reduce redundancy, respectively. Then, we cleaned the text data by removing punctuation and stop words using the Natural Language Toolkit package and converted all characters to lowercase to maintain uniformity and prevent discrepancies caused by case sensitivity. Subsequently, we also replaced all the usernames, URLs, and hashtags in the post with the keywords *USER*, *HTTPURL*, and *HASHTAG* to hide the users' identity and ease semantic understanding. Then, we performed lemmatization using the Natural Language Toolkit package to reduce words to their base form (eg, *drinking* to *drink*) to standardize text and improve consistency. Finally, we removed tweets containing <3 words, as these were deemed too brief to provide substantive insights. This comprehensive preprocessing approach resulted in a refined dataset of 1.13 billion cleaned tweets (3.1 million /1.13 billion, 26.84% in 2019; 4.5 million/1.13 billion, 40.05% in 2020; and 3.8 million/1.13 billion, 33.11% in 2021).

The cleaned tweets were then analyzed using a deep learning–based SU classifier, a pivotal element of our earlier research. This classifier was developed using a state-of-the-art model, RoBERTa [40], in conjunction with various techniques, such as transfer learning and a human-in-the-loop approach [41], to enhance its performance, achieving an accuracy rate of 80%. The successive validation on the sample data is also presented in the study by Maharjan et al [37]. The entire workflow for data collection, preprocessing, and SU identification is depicted in step 1 of Figure 1. This comprehensive phase concluded with 2.8 million, 3.5 million, and 2.5 million SU posts identified for the years 2019, 2020, and 2021, respectively, which we used in this study.

XSL•FO
RenderX

Data Mining

Overview

Data mining constitutes a critical initial phase for conducting this research. At first, we formulated the SU dataset by users (or user base dataset) required in this work. Subsequently, we used this user base dataset to extract additional variables, where we used 5 different analytical modules: 2 focused on demographic variables (M3-Inference [35] and Ethicolr [40]) and 3 targeted at other relevant variables (Valence Aware Dictionary and Sentiment Reasoner [VADER] [42], SpanEmo [43], and a substance type extractor developed in our previous research [37]). Specifically, M3-Inference [44] and Ethicolr [40] were used to extract demographic information, such as age, gender, user type, race, and ethnicity from the user SU dataset. Alongside, VADER [42], SpanEmo [43], and the substance type extractor [37] were used to extract sentiment, emotional content, and substance type, respectively, from the post SU dataset. In the following sections, we provide detailed descriptions of the user base dataset and each of the 5 extraction modules.

SU User Base Dataset

For this study, we considered unique users who had posted at least 1 SU-related post. If the user has never posted any SU-related posts, their posts would never be included in SU posts, and the user would not be considered a substance user. We used a metadata field called *user ID* which served as a unique identifier for all users on the Twitter platform. Technically, we used identified SU posts from step 1 and aggregated them by unique user IDs to obtain the unique user base for this study, as shown in Figure 1. For example, if we had 10 SU posts, where 3 posts belonged to user A, 2 posts to user B, and 5 posts to 5 distinct users (C, D, E, F, and G), then we aggregated these posts by unique user IDs, such that there would be 7 unique SU users—A, B, C, D, E, F, and G. This step resulted in user base datasets of 2,131,457 for 2019, 2,604,123 for 2020, and 2,553,235 for 2021. After the user dataset was formulated, we further retrieved the user metadata for each user by performing a joint operation on the aggregated user base dataset and the original cleaned dataset. Thus, the final dataset used in this study included both tweet and user information. Tweet data included the content and creation date, while user data comprised user ID, first name, last name, screen name, and other profile details. This information served as an input for successive modules to extract additional variables such as age, gender, race, sentiment, and emotion.

Demographic (Age, Gender, and User Type) Extraction Using M3-Inference

In this study, we implemented the M3-Inference model [35] to extract demographic information, specifically age group, gender, and user type, from Twitter accounts. M3-Inference is an open-source Python implementation of a multimodal deep learning system, trained on extensive datasets, including Twitter, IMDB, and Wikipedia. The model's architecture enables it to simultaneously predict 3 key demographic attributes: multimodal capabilities, which allow processing of both image and text features (we only used text features to perform our work); multilingual support, which includes 32 languages; and multiattribute prediction, which facilitates simultaneous forecasting of age, gender, and user type.

In terms of classification, the model treats gender (female or male) and user type (human or organization) as binary classification tasks, while age is categorized into 4 distinct groups: ≤18, 19 to 29, 30 to 39, and ≥40 years. For our analysis, we used a text-only pipeline to derive demographic predictions. This pipeline involved generating character-based embeddings for each textual input (username, screen name, and biography) and passing them through a 2-layer bidirectional character-level long short-term memory network.

To validate the M3-Inference model's efficacy in predicting demographic attributes, we collected profile information from 50 known Twitter users (as detailed in Table S1 in Multimedia Appendix 1). For the age classification, we combined the 19 to 29, 30 to 39, and ≥40 years age groups into a single nonteenager category, thereby reformulating the age prediction as a binary classification task. The model's performance metrics on the collected validation data indicated an accuracy of 99.05% for user type, 95% for gender, and 89% for age classification, with corresponding F_1-scores of 0.98, 0.94, and 0.73, respectively.

Race and Ethnicity Extraction Using Ethnicolr

To infer the racial and ethnic backgrounds of individuals from their names, we used the Ethnicolr Python library [40]. This tool leverages several models based on different datasets, including US census data, Wikipedia entries, and Florida voter registration records, to predict the likelihood of an individual's race and ethnicity. The model has 3 models depending on the type of dataset it is trained on. In our case, we used the Census Last Name Model, which was trained on US census data [45] from the years 2000 and 2010. This model estimates the percentage likelihood that an individual belongs to 4 main racial and ethnic categories, such as White, Black, Asian or Pacific Islander, or Hispanic. The predictions are appended as additional columns in the dataset, providing a probabilistic breakdown of racial composition. We verified the model on our sample data, where the model achieved an accuracy of 90%. The sample-predicted data are presented in Table S2 in Multimedia Appendix 1. However, in our study, we were able to extract the race information for only those posts that had the first name and last name present in the tweets; otherwise, the identification was not accomplished. Hence, approximately 65% (4.3 million/6.6 million) of the total users were identified, as detailed in Table S6 in Multimedia Appendix 1.

Sentiment Extraction Using VADER

VADER [42] is an open-source sentiment analysis tool designed specifically for analyzing social media text. It combines a lexicon-based approach with contextual rules to determine the sentiment of text as positive, negative, or neutral. VADER's lexicon assigns sentiment scores to words on a scale from −4 (very negative) to +4 (very positive). Contextual adjustments are made through several mechanisms; punctuation, such as exclamation points, can amplify sentiments; capitalization highlights intensity, with all-caps being more emphatic; degree modifiers, such as "very," strengthen sentiment; and

XSL•FO
RenderX

conjunctions, such as “but,” can alter sentiment direction. The tool calculates a compound score, ranging from −1 (very negative) to +1 (very positive), by summing these adjusted scores. This method enables VADER to effectively capture both explicit and nuanced emotional expressions, providing a quick and reliable measure of overall sentiment in large volumes of text. The predicted sentiments for sample tweets are presented in Table S3 in Multimedia Appendix 1.

Emotion Extraction Using SpanEmo

SpanEmo [43] is a deep learning–based multilabel emotion recognition model. It analyzes text segments (spans) and classifies each span according to the emotions it conveys. The keywords associated with each emotion class are presented in Table S4 in Multimedia Appendix 1. This is particularly useful in complex texts where different parts may express different emotions. The tool uses NLP techniques to understand the context and semantic meanings of words and phrases, which allows it to accurately detect emotions even in nuanced or mixed emotional content. The model is based on bidirectional encoder representations from transformers, which takes the number of emotion classes (|C|=10) and a sequence “s” as inputs formatted with standard tokens (start_of_token [CLS] and separator_token [SEP]) as [CLS] + [C] + [SEP] + s. The encoding of emotion classed in the input makes the model learn the association between the emotion classes and the words in the input sentence, which is why it outperforms existing emotion classifiers. The model outputs 10 multiemotion classes, namely, anger, anticipation, disgust, fear, joy, love, optimism, pessimism, sadness, surprise, and trust. Before using this module, we finetuned this model on the SemEval-2018 multilabel emotion classification dataset [45] and achieved a F_1-micro score of 0.70. The predicted emotions for sample tweets are presented in Table S5 in Multimedia Appendix 1.

Substance Type Identification

The substance type identification module is also based on our previous research [37]. In previous research [37], we considered the 10 primary substance types categorized together based on their pharmacological and behavioral effects and used the list of keywords from NIDA [22] to formulate keyword-based identification. The 10 types of substances were tobacco, alcohol, cannabinoids, opioids, stimulants, club drugs, hallucinogens, dissociative drugs, prescription medications, and other compounds.

Theme Identification

In this study, we used 6 key themes that we formulated in our previous study [37]. The 6 key themes were COVID-19, economic factors, social influences, mental health, supply chain disruptions, and health care disruptions. The themes were formulated using latent Dirichlet allocation token analysis and were based on significant COVID-19 factors, such as stress and concerns related to COVID-19, economic instability, social dynamics, mental health issues, and disruptions in drug supply and health care services.

Computational Analysis

In this study, we used 2 primary statistical techniques: trend and comparison analysis, along with sentiment and emotion analysis.

Trend and Comparison Analysis

To explore temporal patterns in SU discussions, we conducted a trend analysis, examining the frequency of posts over time. This allowed us to compare SU trends before, during, and after the pandemic. We further performed comparative analysis to assess differences in SU discussions across demographic categories, including age, gender, and race, identifying key disparities and dominant trends.

Sentiment and Emotion Analysis

We applied the VADER model to perform sentiment analysis, classifying the overall tone of posts (positive, negative, or neutral) related to SU. In addition, the SpanEmo model was used for emotion detection, allowing us to identify and categorize emotional expressions (eg, joy, anger, and sadness) linked to specific substances.

These methods provided insight into both the temporal dynamics of SU discussions and the emotional context in which they occurred.

Ethical Considerations

To ensure the privacy and confidentiality of individuals whose data were analyzed, all study data underwent a rigorous deidentification process before analysis. The data for this study were sourced from publicly available platforms [39], containing no identifiable personal information. In addition, the sample posts were preprocessed to transform them into tokens, effectively obscuring any details that could reveal users’ identities. Our research was supported by the SAMHSA Strategic Prevention Framework-19 (grant 6H79SP081502), which was approved by the institutional review board at Kent State University (IRB20-182).

Results

Overview

In this study, we present a fine-grained demographic analysis of SU discourse on Twitter from a dual perspective: by post and by user. After preprocessing and identifying SU-related content, our final dataset included 2,799,726; 3,502,171; and 2,553,235 posts and 2,131,457; 2,604,123; and 1,946,742 users in 2019, 2020, and 2021, respectively. In the following sections, we first present a substantial summary of SU trends across all demographic dimensions: user type, gender, age group, race and ethnicity, and sentiment. Then, we compare our results with survey-based baseline research from NCDAS. Furthermore, we analyze the user trends on different substances, where alcohol users were found to be the prime users during the peak pandemic (March 2020 to June 2020) period. Hence, we performed a detailed analysis of alcohol users and posts during this time. In addition to this, we present the trends of all SU by users across 5 dimensions for each substance type. Finally, we present radar

XSL•FO
RenderX

plots to understand the associated emotions with each substance type.

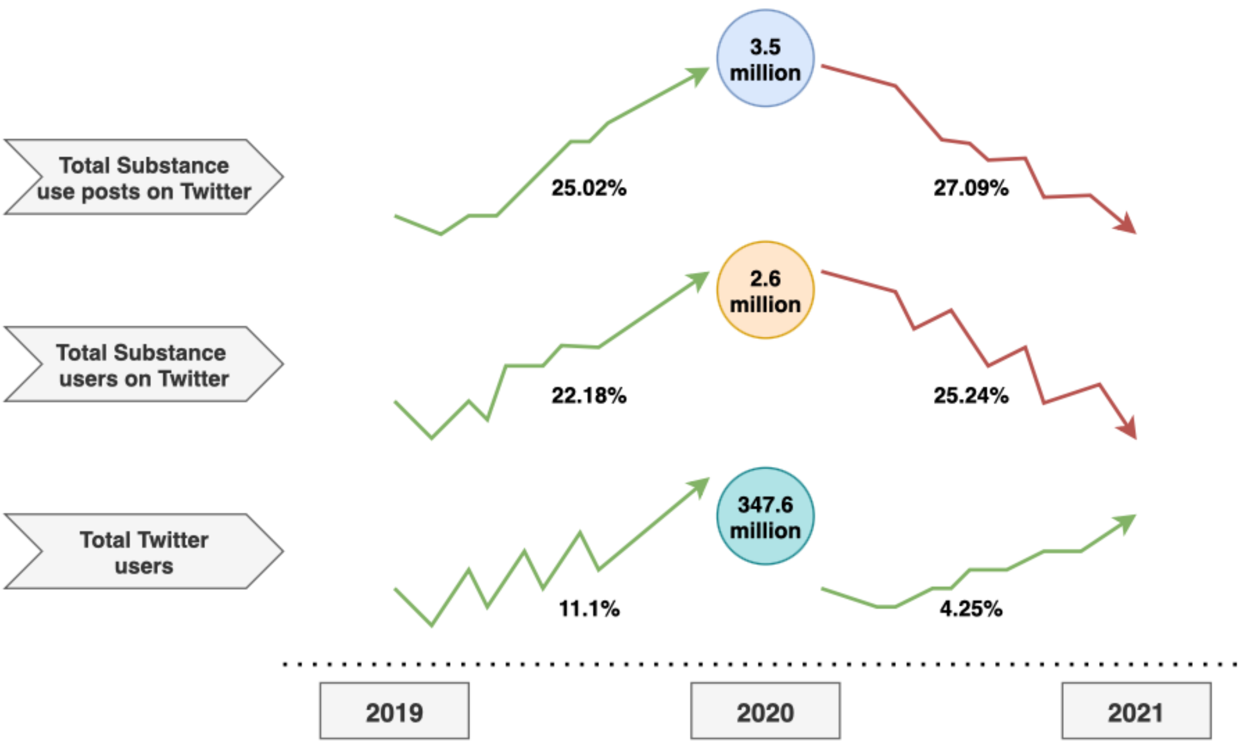
RQ 1: What Are the Statistical Distributions of Overall SU Posts and Their Users, Categorized by Key Demographic Variables in Prepandemic, Pandemic, and Postpandemic Periods?

Our key findings from the statistical analysis are presented in Table S6 in Multimedia Appendix 1, which summarizes the distribution of identified SU posts by both posts and users, further segmented by various categories, including user type, gender, age group, sentiment, race, and ethnicity. The Twitter user base has been expanding annually, with increases of 11.1% in 2020 and 4.25% in 2021 [46]. This growth is contextualized in Figure 2, where we illustrate the trends in the Twitter user base and SU in 2020, comparing both posts and users to prepandemic and postpandemic years. Notably, despite the increase in Twitter users, the marginal decline in SU posts and users in 2021 implies that SU was significantly higher in 2020.

Figure 3 presents the line plots depicting SU among the Twitter user base, categorized by gender, age group, race, and ethnicity. Statistics from Twitter users by gender [46] revealed that male users consistently outnumber female users on Twitter, with a distribution of 68% male users and 32% female users in 2020. In contrast, our analysis indicates that among substance users in 2020, 52% were male users and 48% were female users. This suggests that, despite a smaller female user base, female substance users represent a significant proportion of the overall female demographic on Twitter. As shown in Figure 3 (by gender), SU among female users increased from 2019 to 2021, whereas male users showed a declining trend over the same period.

Similarly, according to statistics from Twitter users [46], the highest levels of user engagement are found in the age group of 18 to 35 years. Moreover, our analysis also revealed a similar trend in substance users, as shown in Figure 3 (by age group), where a greater number of younger users were identified as substance users. Notably, our analysis indicated an increasing trend in SU among teenagers (≤18 years), alongside a decline among individuals aged 19 to 29 years. Our analysis identified race and ethnicity for 64.71% (1,811,516/2,131,457) of users in 2019, 64.99% (2,275,943/2,604,123) in 2020, and 66.51% (1,723,470/1,946,742) in 2021, representing approximately two-thirds of the total user base. Among these identified users, White individuals were the most prevalent across all years, as shown in Figure 3 (by race and ethnicity). The potential reasons and mitigation techniques are further discussed in the Limitations section. In addition to this, the sentiment distribution of SU posts revealed that the posts during 2019 and 2020 were highly associated with negative comments compared to 2021, as shown in Figure 4. Nevertheless, the increasing trend in positive and neutral comments after the pandemic suggests that 2020 was marked by relatively higher negative influences on SU.

Figure 2. Overview of trends among Twitter users, substance users, and related posts.



XSL•FO
RenderX

Figure 3. Trends in substance use by gender, age group, race, and ethnicity. API: Asian or Pacific islander.

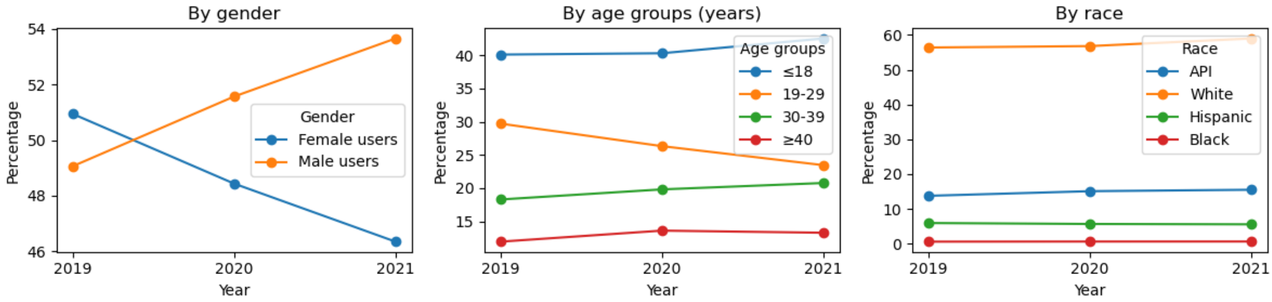
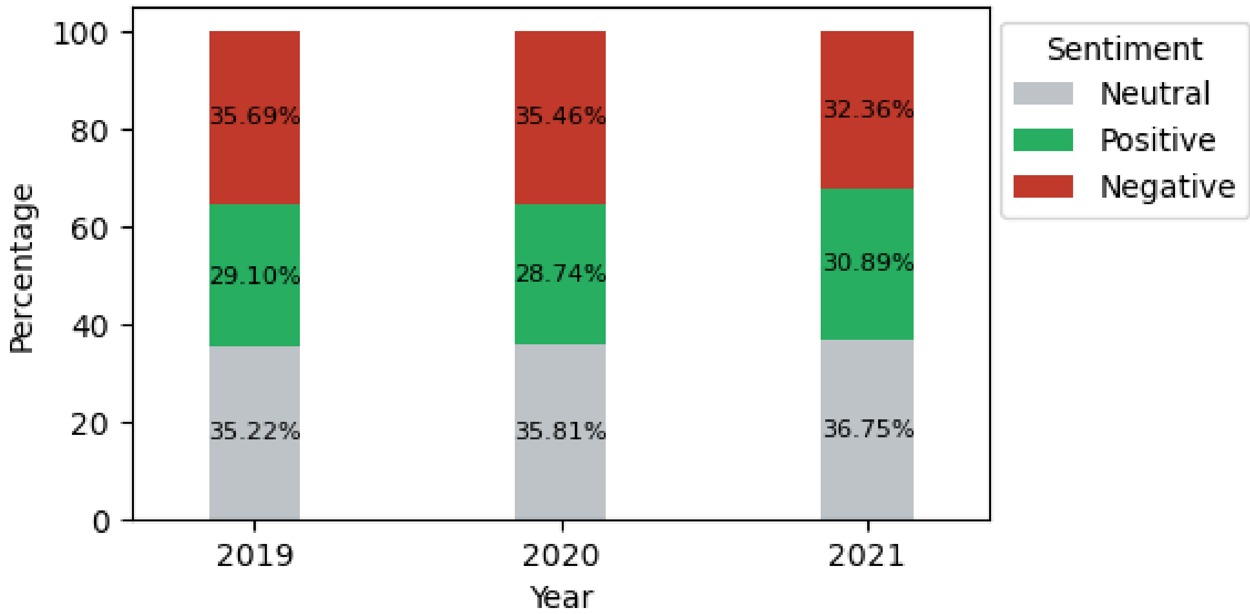


Figure 4. Sentiment distribution in substance use posts from 2019 to 2021.



RQ 2: To What Extent Do the SU Patterns and Demographic Distributions Observed in Twitter (Now Known as X) Discourse From 2019 to 2021 Correspond With or Differ From the Baseline Trends Reported by the NCDAS and Other Research?

Overview

To evaluate the extent to which the SU patterns and demographics observed in Twitter discourse align with or deviate from baseline trends reported by NCDAS, we conducted a comparative analysis of both datasets. The NCDAS provides comprehensive annual reports on SU across various demographics, which serve as a benchmark for understanding broader trends. Our analysis focuses on comparing these established trends with the data extracted from Twitter posts spanning 2019 to 2021. This comparison aims to identify consistencies or discrepancies in SU trends and demographic patterns between the 2 sources.

Key Findings From Twitter Discourse

The user distribution identified substance types and highlighted cannabinoids, stimulants, and opioids as the top 3 illicit substances discussed on Twitter (Figure 5). Demographically, male users dominated all substance types in all studied periods (Figure 6). Across the age group, SU was observed to be highest in teenagers aged ≤18 years (Figure 7). User participation in cannabinoid discussions remained the highest among all substance types, though it showed a declining trend among both adults and teenagers. Teenagers (aged ≤18 years) showed declines of 0.29% and 0.07% from 2019 to 2020 and 2020 to 2021, respectively, while adults (aged >18 years) showed declines of 0.52% and 0.07% over the same periods. For both opioids and stimulants, adults aged ≥40 years were observed to be highly involved among all age groups in all studied periods. Teenagers (aged ≤18 years) showed a decline in opioid use from 2019 to 2021. Similarly, the effect of the COVID-19 lockdown was evidenced in alcohol users profoundly (also supported by biweekly distribution charts in RQ 3; Figure 9), which increased by 80% in just 2 weeks after the global pandemic was declared on March 15, 2020.

XSL•FO
RenderX

Figure 5. Trends in substance use discourse on Twitter in 2019, 2020, and 2021.

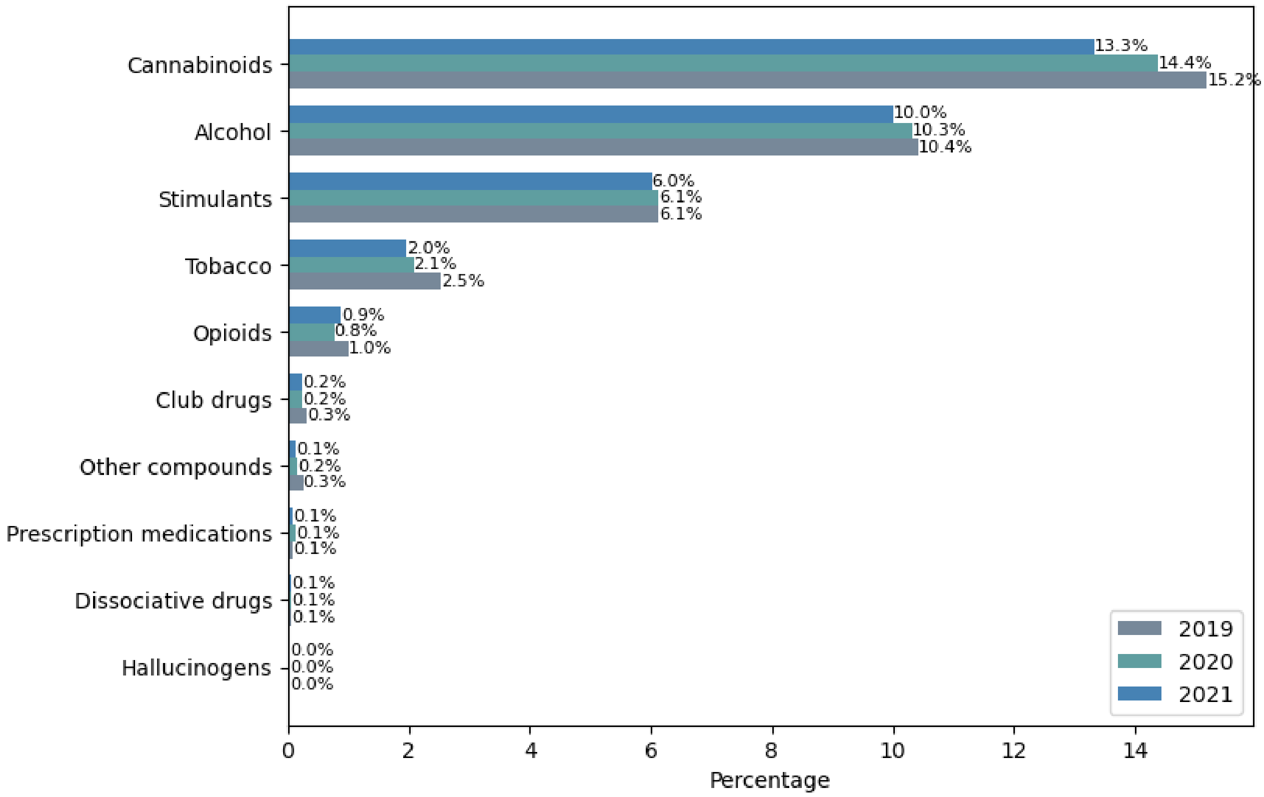


Figure 6. Gender-based trends in substance use discourse on Twitter in 2019, 2020, and 2021.

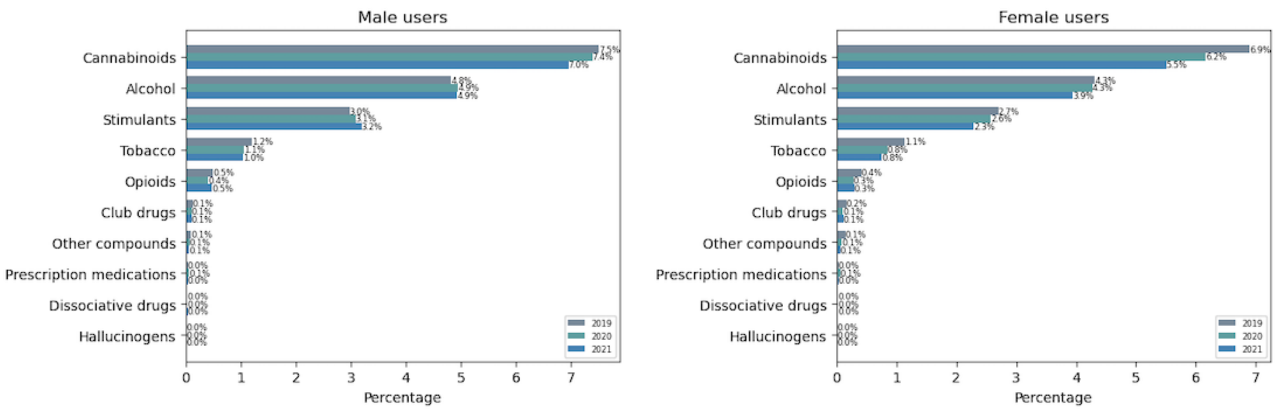
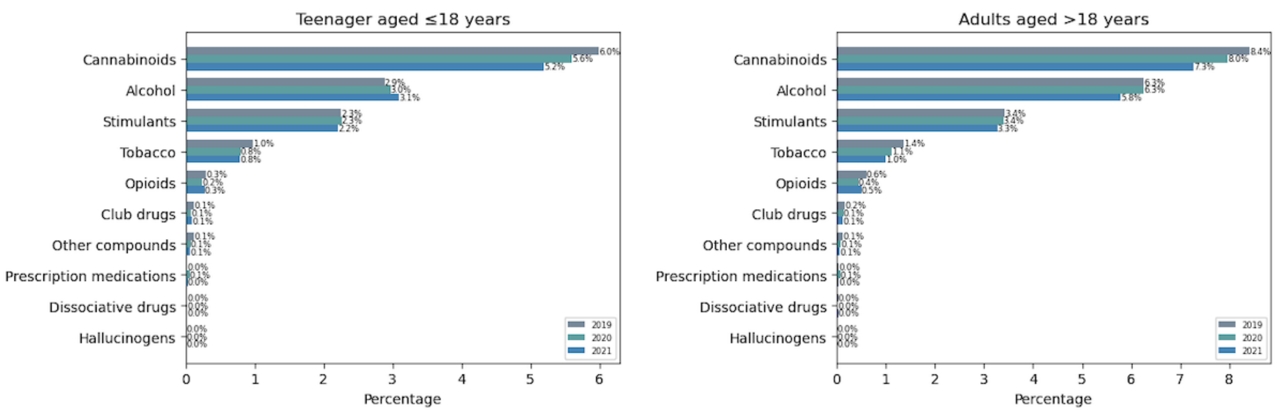


Figure 7. Age-based trends in substance use discourse on Twitter in 2019, 2020, and 2021.



XSL•FO
RenderX

Comparison With NCDAS Trends

Both the reports from NCDAS and our analysis have highlighted cannabinoids and stimulants as the top 2 illicit drugs in the study period. While these alignments suggest convergent validity, we note that Twitter discourse reflects both personal experiences and public commentary, whereas NCDAS measures self-reported use through standardized surveys. Although opioid-related usership on Twitter did not rank among the top substance discussions, as reported in NCDAS [12], both our study and the SAMHSA 2020 report [13] showed a declining trend. From 2019 to 2020, opioid mentions declined by 8.1% in the SAMHSA report and by 25% in our study. This difference in magnitude may reflect Twitter's real-time sensitivity to news events versus surveys' annualized behavioral data. Our result shows that opioid use was mostly prevalent in adults (aged >30 years) compared to teenagers (aged <18 years). This is likely supported by the overdose deaths report [22], where 75% of overdose deaths in adults were from opioids. Likewise, club drugs are widely known to be more commonly used by young people in higher-income settings. The same can be seen in this study, where teenagers were observed to be highly involved compared to other age groups. A similar trend of SU in terms of gender was observed in both studies. Male users were actively involved in all substance types, except for a few. The exception was prescription medication, which showed a higher prevalence among female users in both studies as shown in Table 1.

Table 1. User demographics in 2020.

Drug	By gender, (n=2,604,123), n (%)		By age group (y), (n=2,604,123), n (%)	
	Female users	Male users	≤18 years	>18 years
Cannabinoids	160,668 (6.17)	192,436 (7.39)	145,801 (5.6)	207,303 (7.96)
Alcohol	111,308 (4.27)	128,797 (4.95)	77,319 (2.97)	162,786 (6.25)
Stimulants	66,939 (2.57)	80,283 (3.08)	58,848 (2.26)	88,374 (3.39)
Tobacco	22,044 (0.85)	27,554 (1.06)	20,619 (0.79)	28,979 (1.11)
Opioids	7315 (0.28)	10,441 (0.4)	6235 (0.24)	11,521 (0.44)
Club drugs	2785 (0.11)	2805 (0.11)	1952 (0.07)	3638 (0.14)
Other compounds	2227 (0.09)	1900 (0.07)	1885 (0.07)	2242 (0.09)
Prescription medications	1751 (0.07)	1649 (0.06)	1378 (0.05)	2022 (0.08)
Dissociative drugs	618 (0.02)	657 (0.03)	480 (0.02)	795 (0.03)
Hallucinogens	260 (0.01)	348 (0.01)	252 (0.01)	356 (0.01)

Trend in Alcohol Users in 2020 From Other Survey-Based Research

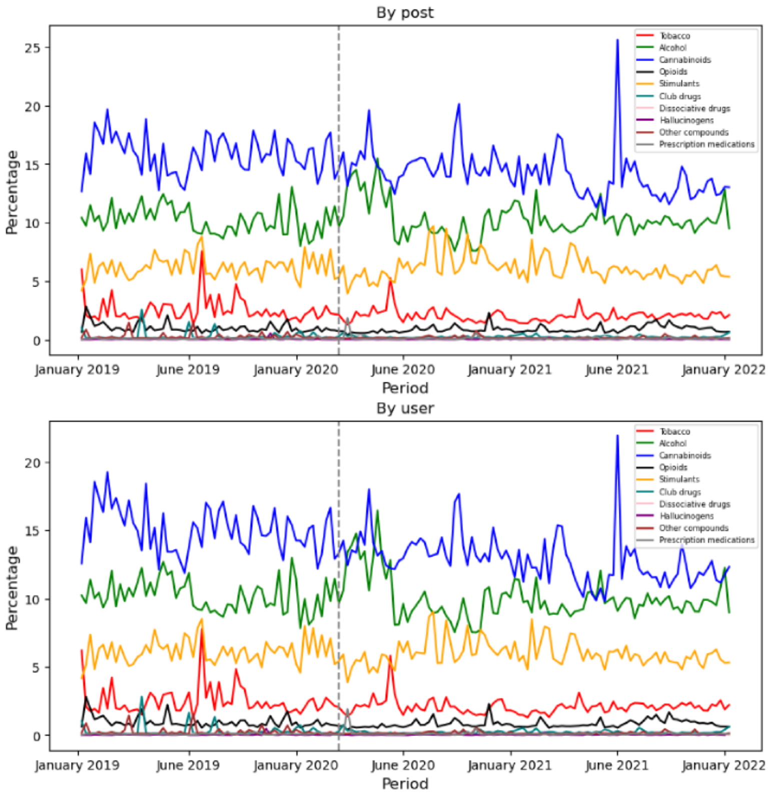
The rise in alcohol use observed in our analysis during the peak pandemic period is supported by multiple studies [45,47]. Notably, our social media data detected this surge within weeks, while surveys [45,47] reported it months later, highlighting Twitter's value for rapid monitoring, though with different population biases. A study by the United States Census Bureau [45] reported that alcohol consumption was observed to be highest as soon as college was closed during the pandemic lockdown. The study found that alcohol use was high for the users with mental health issues and low for those who received social support during the peak time; however, these patterns did not persist over time. A similar result was observed from our theme analysis detailed in RQ 4, where both social and mental health themes were observed as highly associated with alcohol posts during the peak pandemic period, March 15, 2020, to March 31, 2020. Likewise, a study by Lechner et al [48] also demonstrated that alcohol consumption was high during the peak pandemic period, which they associated with COVID-19–related stress, followed by availability of alcohol and boredom.

RQ 3: What Are the Temporal Trends in SU Posts Across Different Substance Types Throughout the Study Period, and How Does the Frequency of User Posting Behavior Vary Over Time for Each Substance Type?

The temporal trend analysis gives the nuance of change of proportions with respect to time. In our study, we plotted weekly trends for all substance types for both users and posts (Figure 8). First, we identified substance type for each post using our keyword-based methods, as detailed in our previous research [37]. Moreover, to further analyze user posting behavior, we aggregated posts by unique users. Our analysis covers the period from 2019 to 2021 and presents data on a weekly basis, capturing both short-term fluctuations and long-term trends. The plot highlights cannabinoids as the most constantly discussed substance among all, followed by alcohol, stimulants, tobacco, and others. While the cannabinoid posts were the most frequent across all study periods, the alcohol users' proportion increased sharply after the pandemic declaration day (March 15, 2020; as shown after the dotted gray line in Figure 8), demanding a detailed focus. Therefore, we present a detailed analysis of alcohol users during this period in our next question.

XSL•FO
RenderX

Figure 8. Weekly distribution across all substance types at the post and user level.



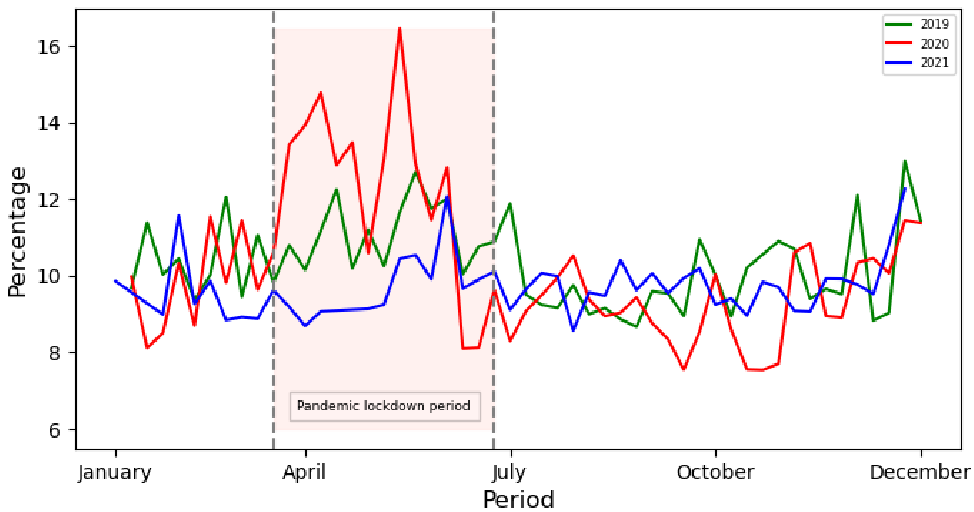
RQ 4: How Did the Number of Individuals Discussing Alcohol Change Within the First 2 Weeks Following the Pandemic Declaration Compared to Other Users, and What Short-Term Trends in User Behavior Emerged Across Different Age Groups, Genders, Races, and Sentiments During This Period?

Trend Analysis on Alcohol Users During the Peak Pandemic Period

Our weekly trend analysis from RQ 3 highlighted that after the pandemic declaration, alcohol users surpassed all other substance users, including cannabinoid users (which was the highest discussed substance throughout the study period). Hence, we drill down on the alcohol users to understand if the increase in trend is associated with COVID-19.

First, we compared weekly user trends during the pandemic year with the preceding year (2019) and following year (2021), as shown in Figure 9. The highlighted period from March 15 to June 15 marks the "pandemic lockdown period." The visualization shows a dramatic increase in alcohol-related users during the pandemic year, while proportions in other periods remained consistent with the preceding and following years.

Figure 9. Alcohol: weekly user distribution in 2019, 2020, and 2021.

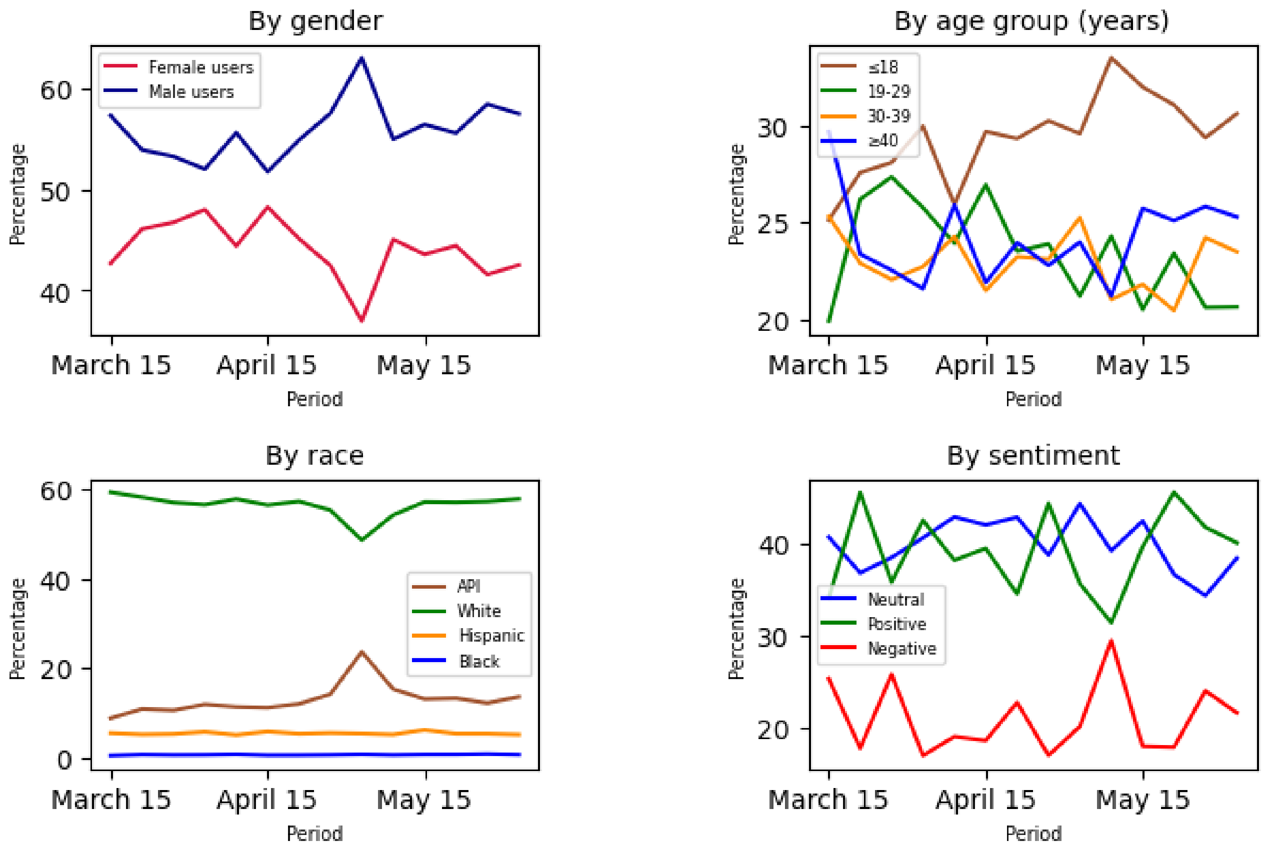


XSL•FO
RenderX

Demographic Trends on Alcohol Users During the Peak Pandemic Period

To further analyze patterns among alcohol users, we examined the pandemic lockdown period (from March 15, 2020, to June 15, 2020) in a weekly manner, segmented by age group, gender, race, and sentiment, as shown in Figure 10. Each subplot allows a comparative analysis within these specific demographic or sentiment groups. The gender and age analysis during this period show that male users and teenagers aged ≤18 years) were more involved in alcohol discussions compared to female users and other age groups, respectively. Likewise, increasing trends were observed among male users and teenagers, as well as among White users in the race analysis. However, the sentiments during this period were mostly neutral and positive.

Figure 10. User distribution during the peak pandemic lockdown period (March 15, 2020, to June 15, 2020). API: Asian or Pacific islander; neg: negative; neu: neutral; pos: positive.



Posts' Content (Theme and Topic) Analysis on Alcohol Use During the Peak Pandemic Period

We performed a detailed analysis on the content of posts, where we derived the underlying themes (COVID-19, economic, social, mental health, supply disruption, and medical disruption) associated with the posts using a keyword method from our previous research [37]. The weekly distribution of alcohol posts in each theme is presented in Table 2. The distribution showed that alcohol-related discussions on peaked during the second week of March (March 15, 2020) across all themes, except the economic theme. Further analysis showed a significant increase in alcohol-related discussions during the week of March 15, 2020, particularly within the themes of social impact, mental health, supply disruption, and medical disruption.

XSL•FO
RenderX

Table 2. Weekly distribution of alcohol-related posts across all themes from February 2020 to May 2020.

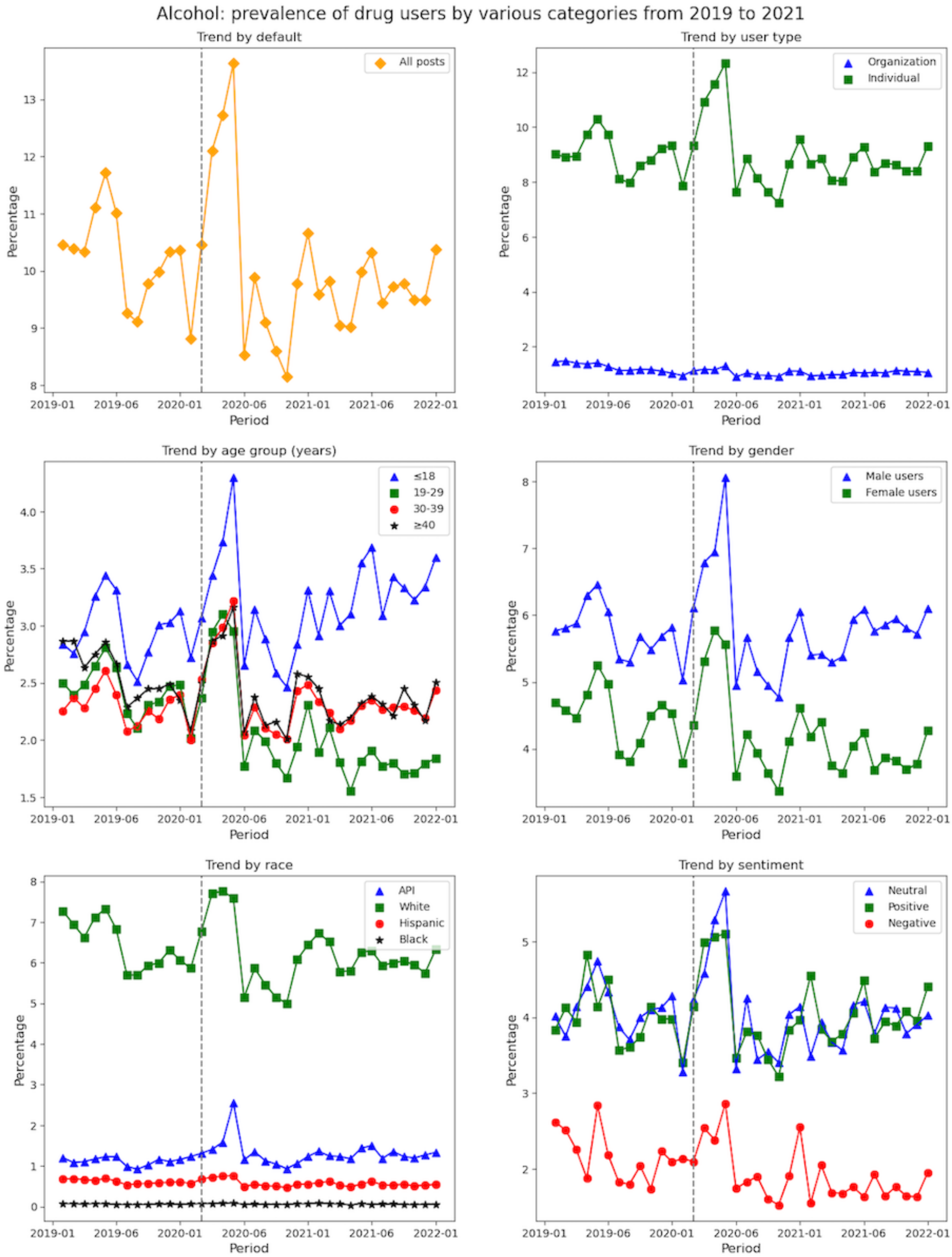
Week	COVID-19, n (%)	Economic, n (%)	Social, n (%)	Mental health, n (%)	Supply distribution, n (%)	Medical disruption, n (%)
February 1, 2020 (n=6635)	95 (1.45)	120 (1.81)	43 (0.65)	48 (0.72)	189 (2.85)	20 (0.3)
February 15, 2020 (n=5383)	81 (1.5)	73 (1.31)	38 (0.71)	35 (0.65)	153 (2.84)	17 (0.32)
March 1, 2020 (n=12,863)	1434 (11.15)	606 (4.71)	245 (1.9)	157 (1.22)	374 (2.91)	41 (0.32)
March 15, 2020 (n=26,550)	3213 (12.1)	427 (1.61)	2275 (8.57)	1226 (4.62)	1966 (7.4)	1091 (4.11)
April 1, 2020 (n=20,772)	1105 (5.32)	323 (1.55)	706 (3.4)	219 (1.05)	789 (3.8)	131 (0.63)
April 15, 2020 (n=21,027)	1022 (4.86)	389 (1.85)	871 (4.14)	198 (0.94)	787 (3.74)	97 (0.46)
May 1, 2020 (n=20,571)	2151 (10.46)	593 (2.88)	795 (3.86)	202 (0.98)	1145 (5.57)	106 (0.52)
May 15, 2020 (n=18,128)	505 (2.79)	293 (1.62)	286 (1.58)	132 (0.73)	654 (3.61)	82 (0.45)

RQ 5: What Are the Trends in User Participation Across Different Demographics for Each Substance Type?

The monthly trend diagram shown in Figure 11 provides the nuance of user trend in alcohol use by 6 dimensions, namely, overall (or by default), user type, age group, gender, race, and sentiment. The trend for other substance types can be found in Figures S2-S10 in Multimedia Appendix 1. Similarly, the trends of SU (by posts) across 6 dimensions can be found in Figures S11-S20 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 11. Alcohol users' distribution across 6 categories from 2019 to 2021. API: Asian or Pacific islander.

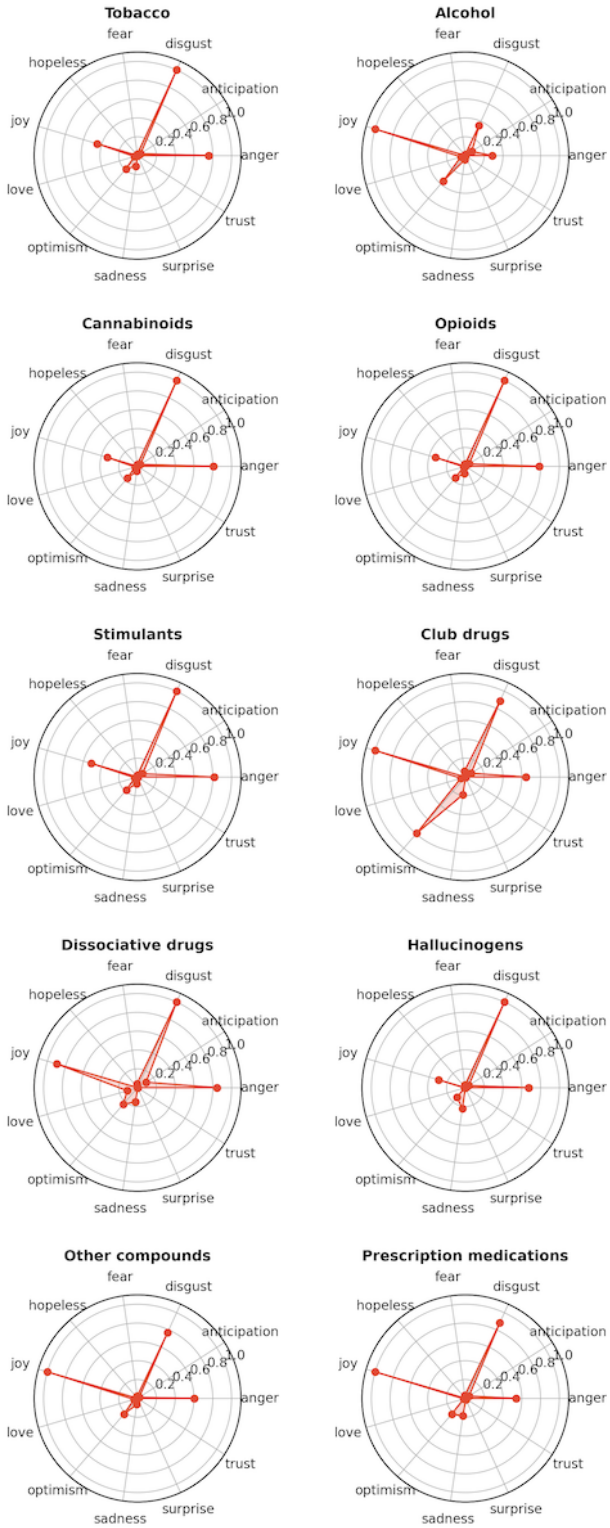


RQ 6: What Emotional Expressions Are Prevalent Across All Substance Types?

We applied the SpanEmo [35] model to perform emotion detection based on Plutchik Emotion Theory [36], which includes 10 main emotion categories (ie, "anger," "anticipation," "disgust," "fear," "hopeless," "joy," "love," "optimism," "sadness," "surprise," and "trust"). The detected emotions were processed further to calculate the mean intensity scores, which are presented in the radar plot in Figure 12. In our results, we present emotions for each substance type. Our results showed that SU-related discussions most frequently expressed emotions of joy, disgust, and anger.

XSL•FO
RenderX

Figure 12. Plutchik emotion analysis across all substance types.



Discussion

Principal Findings

This study established a foundation for analyzing SU across different demographics using web data, with a particular focus on the COVID-19 pandemic year. In addition to the substantial findings as a result, we made a comprehensive comparison with existing survey-based reports from NCDAS and other research works. Successively, our work has found a notable alignment with survey-based reports, as discussed in RQ 2. Notably, users' involvement in SU-related discussions surged in 2020, with users increasing by 22.18% compared to 2019 and 25.24% compared to 2021. Demographically, male users overtook female users in discussions, with their share of posts increasing from 48.87% before the pandemic to 53.4% after the pandemic. The youngest age group (aged ≤18 years) remained the most active, with their proportion growing over time, from 39.56%

XSL•FO
RenderX

in 2019 to 41.72% in 2021. Among posts with identified racial and ethnic data (64.7%-66.5% of total posts), White users predominated; however, this likely reflects platform demographics and limitations in inference methods rather than actual population distributions.

Each year, cannabinoids, alcohol, stimulants, and tobacco were the most frequently discussed substances (in ascending order), while dissociative drugs and hallucinogens were the least discussed. An overall annual decline was observed across all top substances, except for opioids, which showed a 20% drop only in 2020. A demographic breakdown for 2020 revealed that adults (aged >18 years) and male users dominated discussions on most substances. However, prescription medications and other compounds (edible substances) were more commonly discussed by female users, while tobacco use was more prevalent among teenagers (aged ≤18 years).

An increase in alcohol users was observed following the global pandemic declaration. In just a 2-week period, the alcohol users grew by 80%. Most male teenagers (aged ≤18 years) were involved in alcohol discourse, which is also supported by 2 studies [45,48]. Both studies highlighted that alcohol consumption increased during the peak pandemic lockdown period, driven by factors such as mental health challenges, social isolation, COVID-19–related stress, boredom, and easy availability. These findings are supported by our thematic analysis of alcohol-related discussions.

Another remarkable pattern was observed in the discussion of prescription medication, where female users were more involved in social media discourse. This finding is supported by Peteet et al [28], who reported that female users are more likely to use prescription medication compared to other recreational drugs.

Furthermore, our emotion analysis revealed that alcohol was the only substance strongly associated with the emotion of joy, while all other substances were mostly linked to emotions of disgust and anger. This suggests that alcohol use may be driven by positive or celebratory motives, while other substances are more often associated with negative emotional contexts.

Interpretation and Public Health Implications

Overview

This study provides critical insights into SU discourse on Twitter during the COVID-19 pandemic, revealing key demographic trends, emotional associations, and substance-specific patterns. By contextualizing these findings within sociocultural, behavioral, and pandemic-specific factors, we offer actionable strategies for public health stakeholders to design targeted interventions and policies.

Teenagers (Aged ≤18 Years): Early Exposure and Prevention

The youngest age group demonstrated the highest engagement in discussions about alcohol, tobacco, and cannabinoids. This trend may be attributed to increased screen time during lockdowns, amplified peer influence via social media, and the normalization of SU in popular culture. These findings underscore the need for age-specific prevention programs, such as school-based interventions, peer education initiatives, and social media campaigns tailored to teenagers. Addressing the accessibility of substances and providing structured recreational activities during crises are essential to reduce early exposure and experimentation.

Male Users: Gender Norms and Coping Mechanisms

Male users dominated discussions about alcohol and opioids, potentially reflecting sociocultural norms that associate SU with masculinity, as well as behavioral responses to pandemic-related stress. Men may have turned to substances as coping mechanisms due to societal expectations that discourage emotional expression and help-seeking behaviors. Male-centric messaging should emphasize healthy coping strategies and challenge harmful gender norms. Community-based programs, such as sports and recreational activities, can provide alternative outlets for stress relief and reduce reliance on substances. This aligns with NSDUH gender disparities but may be amplified by Twitter's male-skewed user base (estimated 62% male users in 2021).

Female Users: Prescription Medication Misuse

Female users were more likely to engage in discussions about prescription medications, aligning with existing research on higher rates of prescription drug misuse among women. This pattern may be influenced by higher rates of chronic pain conditions, greater likelihood of being prescribed medications, and pandemic-related stressors such as increased caregiving responsibilities. Gender-specific educational programs should address the risks of prescription drug misuse and promote alternative pain management strategies. Health care providers should be trained to recognize and address gender-specific risk factors during patient consultations.

Alcohol: Positive Perceptions and Social Coping

Alcohol was uniquely associated with the emotion of joy, suggesting it is often perceived positively and used as a social and coping mechanism. This perception may be reinforced by social media content that glamorizes alcohol consumption. Public health campaigns should reframe societal perceptions of alcohol, highlighting its negative health consequences and promoting nonalcoholic alternatives for stress relief and social interaction. Policies limiting alcohol accessibility, particularly for underage individuals, should also be considered.

Opioids and Other Substances: Distress and Self-Medication

Substances such as opioids, cannabinoids, and stimulants were linked with negative emotions, indicating that their use may be driven by distress or self-medication for underlying mental health issues. Harm reduction strategies, such as increasing access to addiction treatment services and mental health support, are critical. Integrating mental health screening into SU prevention programs can help identify individuals considered to be at risk and provide early intervention.

Leveraging Social Media for Public Health

This study highlights the potential of social media as a real-time surveillance tool for monitoring SU trends. Platforms such as Twitter can be leveraged to disseminate prevention messages, identify emerging trends, and engage populations considered

XSL•FO
RenderX

to be at risk. Policies such as stricter alcohol regulations for underage individuals and enhanced prescription drug monitoring programs can further mitigate substance misuse. Integrating mental health support into SU prevention and treatment programs is also essential, given the strong link between negative emotions and SU.

By connecting these patterns to sociocultural, behavioral, and contextual factors, this study not only advances the understanding of SU discourse during the COVID-19 pandemic but also provides a road map for public health stakeholders to design targeted interventions, policies, and campaigns. Future research should explore the integration of multiplatform data and multilingual analyses to further enhance the generalizability and applicability of these findings.

Limitations

This study builds upon our previous research [37]. As in our previous research, there is data skewness in certain months due to missing data in the original source [39]. This could potentially deviate from the actual results. Likewise, the SU identifier developed in our previous research [37], which used advanced deep learning tools such as RoBERTa and human-in-the-loop methods, achieved an accuracy of 80%. Thus, the user base analysis studied in this study does not account for all the substance users.

Second, we recognize that our study's focus on English-language posts may have limited the generalizability of our findings. By excluding non–English-speaking populations, we may have overlooked diverse racial groups and specific age groups in SU discourse. The focus on English-language content, combined with the lack of geocoding analysis, likely overrepresents English-speaking regions, particularly the United States, while underrepresenting global SU patterns. In addition, the user base identified in the study does not investigate the frequency of posts. Analyzing retweeted posts or frequent users could reveal deeper insights that may help public health policy makers develop more targeted and effective strategies.

Third, for the demographic identification, we relied on various machine learning models, such as m3inference and EthicolrM. Although we validated the tools with our ground data, the bias in these models still exists. For age and gender, we only achieve 80% validation accuracy using the m3inference model, which perhaps is due to the specific training dataset that did not fully capture the diversity of Twitter users. For race, our Ethicolr-based pipeline classified only on average of 65% (4.3 million/6.6 million) of users across study years, with unclassified users potentially skewing results if racial groups differed in their likelihood of providing identifiable metadata. This limitation is compounded by the model's training on US census data, which may not generalize to global or non-Anglicized naming conventions, and our reliance on only English-language posts. Thus, we acknowledge that these tools may have introduced biases in the extraction of age, gender, and race. This could have potentially affected the representativeness and accuracy of our findings.

Fourth, we acknowledge that the observed demographic trends (eg, male vs female participation) may be confounded by the inherent demographic distribution of Twitter users. Normalizing these trends by Twitter user base proportions for each demographic could provide a more accurate representation; however, such data are not publicly available, limiting our ability to perform this analysis.

Fifth, our analysis is limited to a single platform, Twitter, which might not fully represent the broader spectrum of SU discourse across other social media platforms such as Facebook, Instagram, TikTok, and Reddit.

Sixth, we acknowledge limitations regarding comparative analyses with survey data. While our findings show meaningful alignments with reports such as NSDUH and NCDAS, important methodological differences must be considered: our real-time social media data capture immediate discourse rather than the annualized behaviors reported in surveys; clinical surveys use validated screening tools such as Alcohol Use Disorders Identification Test, whereas our classifier detects broader public discourse, including news and advocacy content; and NSDUH's nationally representative sampling contrasts with Twitter's self-selected user base. These factors suggest that our temporal trends should be interpreted as complementary to, rather than confirmatory of, survey findings.

Finally, we acknowledge the limitation of the studied themes. While we referenced key COVID-19 factors derived from our primary study [37], the trends could have been influenced by other societal factors such as political tensions. Furthermore, the keywords used in identifying themes may have been too narrow, potentially leading to an overrepresentation of certain themes in our results.

Future Work

This study focused on analyzing SU differences from the perspectives of age, gender, and race. To enhance understanding, future work could incorporate geolocation data to analyze trends and patterns, enabling exploration of region-specific influencing factors. This would enable the development of targeted intervention strategies to prevent SU based on geographic location. Furthermore, extracting additional information, such as socioeconomic, mental, and physical health status, could significantly enhance the use of social media as a prominent platform for studying public health–related issues. In addition, analyzing user-based personality traits could provide valuable insights for the public health sector, allowing for the identification of specific characteristics that can inform prevention strategies, even in the absence of demographic information.

To address the limitations identified in this study, future research could also expand the analysis to include multiple languages, leveraging multilingual NLP models to capture a more comprehensive understanding of SU discourse across diverse linguistic and cultural contexts. Developing and using more robust, inclusive, and transparent demographic inference models trained on diverse datasets would further enhance the representativeness of findings. Future studies could also explore normalizing demographic trends by the underlying user base proportions of social media platforms, if such data become available. This would provide a more accurate representation

XSL•FO
RenderX

of demographic participation and strengthen the generalizability of findings. In addition, expanding the scope to include other social media platforms (eg, Facebook, Reddit, and Instagram) and offline data sources (eg, surveys and interviews) would provide a more holistic view of SU discourse during public health crises. Collaborations with researchers fluent in non-English languages, as well as social scientists and ethicists, could help refine these tools and methodologies to better account for intersectional identities, cultural contexts, and global perspectives.

Conclusions

Social media platforms, combined with advanced NLP technologies, offer a valuable alternative research space for uncovering insightful trends and patterns in SU discourse. This study has successfully demonstrated the potential of leveraging Twitter's data to analyze SU trends during the COVID-19 pandemic, aligning our findings with notable survey-based reports, such as NCDAS and Monitoring the Future. Our results highlight significant demographic shifts, such as the increased engagement of teenagers and male users in substance-related discussions, as well as substance-specific patterns, including the rise in alcohol discourse and the gender disparities in prescription medication discussions.

These insights offer actionable strategies for public health stakeholders, enabling targeted interventions for groups considered to be high risk and substance-specific harm reduction. By leveraging social media as a real-time surveillance tool, stakeholders can monitor trends, disseminate prevention messages, and engage populations considered to be at risk without relying on traditional surveying methods. This study underscores the potential of social media data to inform public health strategies, particularly during global crises, and emphasizes the need for future research to include multiple platforms and languages to enable more inclusive and impactful interventions.

Acknowledgments

This study is funded by the Substance Abuse and Mental Health Services Administration SPF-19 grant (6H79SP081502).

Data Availability

The data and code supporting this study are publicly available on GitHub [49].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary tables and figures.

[DOCX File , 5583 KB - infodemiology_v5i1e67333_app1.docx]

References

1. Drug overdose deaths in the United States, 2003–2023. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/nchs/products/databriefs/db522.htm [accessed 2025-07-04]
2. Lee J. Mental health effects of school closures during COVID-19. Lancet Child Adolesc Health 2020 Jun;4(6):421 [FREE Full text] [doi: 10.1016/S2352-4642(20)30109-7] [Medline: 32302537]
3. Addo IY. Double pandemic: racial discrimination amid coronavirus disease 2019. Soc Sci Humanit Open 2020;2(1):100074 [FREE Full text] [doi: 10.1016/j.ssaho.2020.100074] [Medline: 34173502]
4. Thomeer MB, Moody MD, Yahirun J. Racial and ethnic disparities in mental health and mental health care during the COVID-19 pandemic. J Racial Ethn Health Disparities 2023 Apr;10(2):961-976 [FREE Full text] [doi: 10.1007/s40615-022-01284-9] [Medline: 35318615]
5. Chae DH, Yip T, Martz CD, Chung K, Richeson JA, Hajat A, et al. Vicarious racism and vigilance during the COVID-19 pandemic: mental health implications among Asian and Black Americans. Public Health Rep 2021 May 25;136(4):508-517 [FREE Full text] [doi: 10.1177/00333549211018675] [Medline: 34034574]
6. Gerrard M, Stock ML, Roberts ME, Gibbons FX, O'Hara RE, Weng CY, et al. Coping with racial discrimination: the role of substance use. Psychol Addict Behav 2012 Sep;26(3):550-560 [FREE Full text] [doi: 10.1037/a0027711] [Medline: 22545585]
7. Zhu J, Yalamanchi N, Jin R, Kenne DR, Phan N. Investigating COVID-19's impact on mental health: trend and thematic analysis of Reddit users' discourse. J Med Internet Res 2023 Jul 12;25:e46867 [FREE Full text] [doi: 10.2196/46867] [Medline: 37436793]
8. Zhu J, Jin R, Kenne DR, Phan N, Ku WS. User dynamics and thematic exploration in r/Depression during the COVID-19 pandemic: insights from overlapping r/SuicideWatch users. J Med Internet Res 2024 May 20;26:e53968 [FREE Full text] [doi: 10.2196/53968] [Medline: 38767953]

XSL•FO
RenderX

9. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985;98(2):310-357. [doi: 10.1037//0033-2909.98.2.310]
10. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. MMWR Morb Mortal Wkly Rep 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: 10.15585/mmwr.mm6932a1] [Medline: 32790653]
11. Chacon NC, Walia N, Allen A, Sciancalepore A, Tiong J, Quick R, et al. Substance use during COVID-19 pandemic: impact on the underserved communities. Discoveries (Craiova) 2021 Dec 31;9(4):e141 [FREE Full text] [doi: 10.15190/d.2021.20] [Medline: 35261922]
12. Drug abuse statistics. National Center of Drug Abuse Statistics. URL: https://drugabusestatistics.org/ [accessed 2023-10-20]
13. Key substance and mental health indicators in the United States: results from the 2020 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration. 2020. URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt35325/NSDUHFFRPDFWHTMLFiles2020/2020NSDUHFFR1PDFW102121.pdf [accessed 2023-08-13]
14. Lee YH, Woods C, Shelley M, Arndt S, Liu CT, Chang YC. Racial and ethnic disparities and prevalence in prescription drug misuse, illicit drug use, and combination of both behaviors in the United States. Int J Ment Health Addict 2023 May 19:1-17 [FREE Full text] [doi: 10.1007/s11469-023-01084-0] [Medline: 37363760]
15. Larochelle MR, Slavova S, Root ED, Feaster DJ, Ward PJ, Selk SC, et al. Disparities in opioid overdose death trends by race/ethnicity, 2018–2019, from the HEALing communities study. Am J Public Health 2021 Oct;111(10):1851-1854. [doi: 10.2105/ajph.2021.306431]
16. Vasilenko SA, Evans-Polce RJ, Lanza ST. Age trends in rates of substance use disorders across ages 18-90: differences by gender and race/ethnicity. Drug Alcohol Depend 2017 Nov 01;180:260-264 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.08.027] [Medline: 28938183]
17. Alcendor DJ. Racial disparities-associated COVID-19 mortality among minority populations in US. J Clin Med 2020 Jul 30;9(8):2442 [FREE Full text] [doi: 10.3390/jcm9082442] [Medline: 32751633]
18. Couch KA, Fairlie RW, Xu H. Early evidence of the impacts of COVID-19 on minority unemployment. J Public Econ 2020 Dec;192:104287 [FREE Full text] [doi: 10.1016/j.jpubeco.2020.104287] [Medline: 32952224]
19. Acosta AM, Garg S, Pham H, Whitaker M, Anglin O, O'Halloran A, et al. Racial and ethnic disparities in rates of COVID-19-associated hospitalization, intensive care unit admission, and in-hospital death in the United States from March 2020 to February 2021. JAMA Netw Open 2021 Oct 01;4(10):e2130479 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.30479] [Medline: 34673962]
20. Hao T, Huang Z, Liang L, Weng H, Tang B. Health natural language processing: methodology development and applications. JMIR Med Inform 2021 Oct 21;9(10):e23898 [FREE Full text] [doi: 10.2196/23898] [Medline: 34673533]
21. Elbattah M, Arnaud É, Gignon M, Dequen G. The role of text analytics in healthcare: a review of recent developments and applications. In: Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies. 2021 Presented at: BIOSTEC 2021; February 11-13, 2021; Online URL: https://www.scitepress.org/PublishedPapers/2021/104145/104145.pdf [doi: 10.5220/0010414508250832]
22. COVID-19 and substance use. National Institutes of Health National Institute on Drug Abuse. URL: https://nida.nih.gov/research-topics/covid-19-substance-use [accessed 2023-08-13]
23. Johnston LD, Miech RA, O'Malley PM, Bachman JG, Schulenberg JE, Patrick ME. Monitoring the future national survey results on drug use, 1975-2020: overview, key findings on adolescent drug use. Institute for Social Research. 2021. URL: https://eric.ed.gov/?id=ED611736 [accessed 2025-06-28]
24. Schuler MS, Rice CE, Evans-Polce RJ, Collins RL. Disparities in substance use behaviors and disorders among adult sexual minorities by age, gender, and sexual identity. Drug Alcohol Depend 2018 Aug 01;189:139-146 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.05.008] [Medline: 29944989]
25. Daniulaityte R, Carlson R, Falck R, Cameron D, Perera S, Chen L, et al. "I just wanted to tell you that loperamide WILL WORK": a web-based study of extra-medical use of loperamide. Drug Alcohol Depend 2013 Jun 01;130(1-3):241-244 [FREE Full text] [doi: 10.1016/j.drugalcdep.2012.11.003] [Medline: 23201175]
26. Griffiths M. Internet addiction - time to be taken seriously? Addict Res 2009 Jul 11;8(5):413-418. [doi: 10.3109/16066350009005587]
27. Schifano F, Deluca P, Baldacchino A, Peltoniemi T, Scherbaum N, Torrens M, et al. Drugs on the web; the Psychonaut 2002 EU project. Prog Neuropsychopharmacol Biol Psychiatry 2006 Jun;30(4):640-646. [doi: 10.1016/j.pnpbp.2005.11.035] [Medline: 16458404]
28. Peteet B, Mosley C, Miller-Roenigk B, McCuistian C, Dixon S. Transnational trends in prescription drug misuse among women: a systematic review. Int J Drug Policy 2019 Jan;63:56-73. [doi: 10.1016/j.drugpo.2018.10.005] [Medline: 30500475]
29. Murguía E, Tackett-Gibson M, Lessem A. Real Drugs in a Virtual World: Drug Discourse and Community Online. London, UK: Bloomsbury Academic; 2007.
30. Griffiths P, Mounteney J. Drug trend monitoring. In: Miller PG, Strang J, Miller PM, editors. Addiction Research Methods. Hoboken, NJ: Blackwell Publishing Ltd; 2010.

XSL•FO
RenderX

31. Mounteney J, Fry C, McKeganey N, Haugland S. Challenges of reliability and validity in the identification and monitoring of emerging drug trends. Subst Use Misuse 2010 Dec 21;45(1-2):266-287. [doi: 10.3109/10826080903368598] [Medline: 20025453]
32. Miller PG, Sønderlund AL. Using the internet to research hidden populations of illicit drug users: a review. Addiction 2010 Sep 05;105(9):1557-1567. [doi: 10.1111/j.1360-0443.2010.02992.x] [Medline: 20626378]
33. Sarker A, O'Connor K, Ginn R, Scotch M, Smith K, Malone D, et al. Social media mining for toxicovigilance: automatic monitoring of prescription medication abuse from Twitter. Drug Saf 2016 Mar;39(3):231-240 [FREE Full text] [doi: 10.1007/s40264-015-0379-4] [Medline: 26748505]
34. Correia RB, Li L, Rocha LM. Monitoring potential drug interactions and reactions via network analysis of instagram user timelines. Pac Symp Biocomput 2016;21:492-503 [FREE Full text] [Medline: 26776212]
35. Wang Z, Hale SA, Adelani D, Grabowicz PA, Hartmann T, Flöck F, et al. Demographic inference and representative population estimates from multilingual social media data. ArXiv. Preprint posted online on May 15, 2019 2025 [FREE Full text] [doi: 10.1145/3308558.3313684]
36. Plutchik R. Emotions, evolution, and adaptive processes. In: Arnold M, editor. Feelings and Emotions: The Loyola Symposium. New York, NY: Academic Press; 1970.
37. Maharjan J, Zhu J, King J, Phan N, Kenne D, Jin R. Large-scale deep learning-enabled infodemiological analysis of substance use patterns on social media: insights from the COVID-19 pandemic. JMIR Infodemiology 2025 Apr 17;5:e59076 [FREE Full text] [doi: 10.2196/59076] [Medline: 40244656]
38. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. ArXiv. Preprint posted online on July 26, 2019 2025 [FREE Full text]
39. Archive team: the Twitter stream grab. Internet Archive. URL: https://archive.org/details/twitterstream [accessed 2022-09-30]
40. Parasurama P. raceBERT -- a transformer-based model for predicting race and ethnicity from names. ArXiv. Preprint posted online on December 7, 2021 2025 [FREE Full text]
41. Mosqueira-Rey E, Hernández-Pereira E, Alonso-Ríos D, Bobes-Bascarán J, Fernández-Leal Á. Human-in-the-loop machine learning: a state of the art. Artif Intell Rev 2022 Aug 17;56:3005-3054. [doi: 10.1007/s10462-022-10246-w]
42. Hutto C, Gilbert E. VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media. 2014 Presented at: ICWSM-14; June 1-4, 2014; Ann Arbor, MI. [doi: 10.1609/icwsm.v8i1.14550]
43. Alhuzali H, Ananiadou A. SpanEmo: casting multi-label emotion classification as span-prediction. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics. 2021 Presented at: EACL 2021; April 19-23, 2021; Online. [doi: 10.18653/v1/2021.eacl-main.135]
44. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. ArXiv. Preprint posted online on June 12, 2017 2025 [FREE Full text]
45. U.S. Census Bureau homepage. United States Census Bureau. URL: https://www.census.gov/ [accessed 2024-04-27]
46. X/Twitter: number of users worldwide 2024. Statista. URL: https://www.statista.com/statistics/303681/twitter-users-worldwide/ [accessed 2024-08-28]
47. Grossman ER, Benjamin-Neelon SE, Sonnenschein S. Alcohol consumption and alcohol home delivery laws during the COVID-19 pandemic. Subst Abus 2022 Jan 01;43(1):1139-1144. [doi: 10.1080/08897077.2022.2060432] [Medline: 35471927]
48. Lechner WV, Laurene KR, Patel S, Anderson M, Grega C, Kenne DR. Changes in alcohol use as a function of psychological distress and social support following COVID-19 related university closings. Addict Behav 2020 Nov;110:106527 [FREE Full text] [doi: 10.1016/j.addbeh.2020.106527] [Medline: 32679435]
49. JulinaM / DemographicAnalysis. GitHub. URL: https://github.com/JulinaM/DemographicAnalysis [accessed 2025-06-28]

Abbreviations

NCDAS: National Center for Drug Abuse Statistics
NIDA: National Institute on Drug Abuse
NLP: natural language processing
NSDUH: National Survey on Drug Use and Health
RoBERTa: Robustly Optimized Bidirectional Encoder Representations From Transformers Pretraining Approach
RQ: research question
SAMHSA: Substance Abuse and Mental Health Services Administration
SU: substance use
VADER: Valence Aware Dictionary and Sentiment Reasoner

XSL•FO
RenderX

Edited by T Mackey; submitted 08.10.24; peer-reviewed by M Agbede, M Elbattah, Q Ng, O Akhadelor; comments to author 05.12.24; revised version received 13.03.25; accepted 27.05.25; published 28.07.25.

Please cite as:
Maharjan J, Jin R, King J, Zhu J, Kenne D
Differential Analysis of Age, Gender, Race, Sentiment, and Emotion in Substance Use Discourse on Twitter During the COVID-19 Pandemic: A Natural Language Processing Approach
JMIR Infodemiology 2025;5:e67333
URL: https://infodemiology.jmir.org/2025/1/e67333
doi:10.2196/67333
PMID:

©Julina Maharjan, Ruoming Jin, Jennifer King, Jianfeng Zhu, Deric Kenne. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 28.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Interventions to Address Infodemics Through Epidemiological, Socioecological, and Environmental Health Models: Framework Analysis

Jennifer N John[1,2], BA; Sara Gorman[2], MPH, PhD; David Scales[2,3], MD, PhD, MPhil

[1]Perelman School of Medicine, Penn Medical Communication Research Institute, University of Pennsylvania, Smilow Center for Translational Research Room 12-136, 3400 Civic Center Boulevard, Philadelphia, PA, United States

[2]Critica, Bronx, NY, United States

[3]Department of Medicine, Division of General Internal Medicine, Weill Cornell Medicine, Cornell University, New York, NY, United States

Corresponding Author:
Jennifer N John, BA
Perelman School of Medicine, Penn Medical Communication Research Institute, University of Pennsylvania, Smilow Center for Translational Research Room 12-136, 3400 Civic Center Boulevard, Philadelphia, PA, United States

Abstract

Background: The COVID-19 pandemic was accompanied by a barrage of false, misleading, and manipulated information that inhibited effective pandemic response and led to thousands of preventable deaths. Recognition of the urgent public health threat posed by this infodemic led to the development of numerous infodemic management interventions by a wide range of actors. The need to respond rapidly and with limited information sometimes came at the expense of strategy and conceptual rigor. Given limited funding for public health communication and growing politicization of countermisinformation efforts, responses to future infodemics should be informed by a systematic and conceptually grounded evaluation of the successes and shortcomings of existing interventions to ensure credibility of the field and evidence-based action.

Objectives: This study sought to identify gaps and opportunities in existing infodemic management interventions and to assess the use of public health frameworks to structure responses to infodemics.

Methods: We expanded a previously developed dataset of infodemic management interventions, spanning guidelines, policies, and tools from governments, academic institutions, nonprofits, media companies, and other organizations, with 379 interventions included in total. We applied framework analysis to describe and interpret patterns within these interventions through their alignment with codes derived from 3 frameworks selected for their prominence in public health and infodemic-related scholarly discourse: the epidemiological model, the socioecological model, and the environmental health framework.

Results: The epidemiological model revealed the need for rigorous, transparent risk assessments to triage misinformation. The socioecological model demonstrated an opportunity for greater coordination across levels of influence, with only 11% of interventions receiving multiple socioecological codes, and more robust partnerships with existing organizations. The environmental health framework showed that sustained approaches that comprehensively address all influences on the information environment are needed, representing only 19% of the dataset.

Conclusions: Responses to future infodemics would benefit from cross-sector coordination, adoption of measurable and meaningful goals, and alignment with public health frameworks, which provide critical conceptual grounding for infodemic response approaches and ensure comprehensiveness of approach. Beyond individual interventions, a funded coordination mechanism can provide overarching strategic direction and promote collaboration.

(JMIR Infodemiology 2025;5:e67119) doi:10.2196/67119

KEYWORDS

infodemics; misinformation; disinformation; Covid-19; infodemic management; health communication; pandemic preparedness

Introduction

Background

The COVID-19 pandemic entailed an outbreak not only of viral illness but also of viral rumors. This so-called infodemic, defined by the World Health Organization as an overabundance of accurate and inaccurate information [1], had tangible public health consequences. As of April 2022, 24% of COVID-19 mortality, or 234,000 deaths, was vaccine-preventable [2], and misinformation and disinformation cost the United States between US $50,000,000 and US $300,000,000 each day during the pandemic in health care spending and economic losses [3].

XSL•FO
RenderX

These impacts demonstrated the necessity of addressing misinformation as part of public health responses [4].

A wide range of stakeholders globally including governments, nongovernmental organizations, academic institutions, professional societies, and technology companies rapidly developed and deployed a large number of interventions to mitigate the perceived harms of the infodemic. These interventions varied substantially in their foci and impacts and addressed both the infodemic itself and the social problems related to the infodemic, such as vaccine hesitancy and institutional distrust. For example, in the New York City Department of Health and Mental Hygiene, the misinformation response unit disseminated culturally specific communication materials in response to emerging web-based COVID-19 rumors through partnerships with community organizations [4]. YouTube and Google also prioritized credible health information sources in search results based on criteria developed by organizations including the World Health Organization, the National Academy of Medicine, and the Council of Medical Specialty Societies [5,6].

Given the inevitability and growing threat of future infodemics, it is critical to learn from the successes and shortcomings of the growing body of infodemic management interventions. Prior studies have evaluated the effectiveness of these interventions, their fundamental characteristics, and the psychological concepts underlying them [7-9]. However, these studies were limited in the scope of interventions examined, only considered 1 framework, or focused on individual-level factors. Little research has explored the areas of emphasis, both intended and unintended, and strategies revealed and gaps left by these interventions in aggregate. Such an analysis is needed to provide funders, government agencies, public health leaders, and other stakeholders that set priorities for infodemic responses with insights to inform proactive, sustainable, and coordinated efforts that effectively use limited resources. Given increasing politicized attacks on public health and misinformation research in recent years, it is particularly important to avoid infodemic management practices that lead to or exacerbate public mistrust. For example, in the United States, Republicans are disproportionately likely to consider the removal of false articles on social media, a key component of Facebook's COVID-19 misinformation policy [10], to be censorship [11].

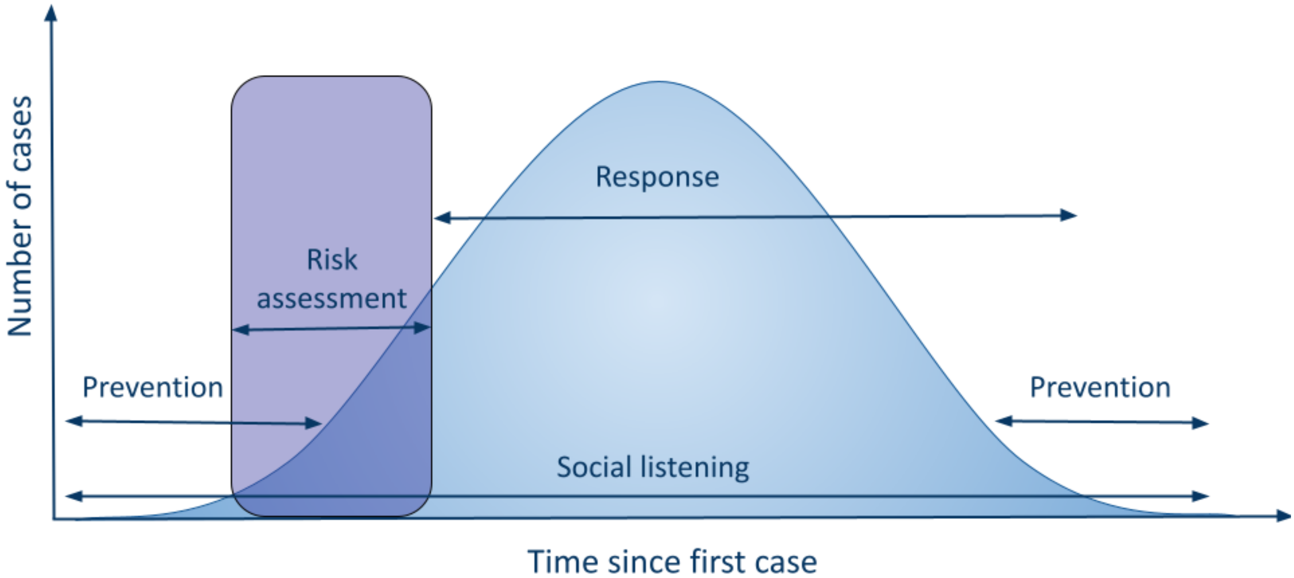
In public health, conceptual frameworks serve as lenses that systematically illuminate gaps, patterns, and opportunities in programs and policies [12-14]. Frameworks are not exhaustive or mutually exclusive, and multiple frameworks are necessary to comprehensively interrogate complex topics. Applying public health frameworks to infodemic interventions offers an opportunity to explore their theoretical foundations and inform the design of future interventions. Certain public health metaphors, particularly analogies to epidemics of disease, are frequently invoked in and often dominate discussions of misinformation in academia and public media. However, the use of these frameworks and the validity of their underlying assumptions in this setting have yet to be rigorously evaluated [15]. As a result, other promising mechanisms of impact supported by alternative paradigms may be overlooked [15]. In the following sections, we outline the 3 frameworks applied in this study and their applications to infodemics. These frameworks were selected because they are well established in public health or are often referenced, implicitly or explicitly, in infodemic-related discourse. Public health frameworks were prioritized to reflect the growing application of public health perspectives to address misinformation during the pandemic.

Epidemiological Model

Epidemiological models describe the spread of disease over time within a population. The epidemiological model frames misinformation as a contagion (Figure 1) [16]. As the epidemiological model is currently a dominant paradigm in discourse about misinformation [15], it is critical to assess how well suited previously developed interventions are to this model. Areas of engagement in the information ecosystem are drawn analogously from responses based on public health approaches to infectious diseases: social listening, risk assessment, response, and prevention [17]. Risk assessment can take place either as a one-time evaluation or a continuous assessment at various points along the epidemiological curve.

XSL•FO
RenderX

Figure 1. Epidemiological model.



Socioecological Model

The socioecological model illustrates the health impacts of various components of society and the environment (Figure 2) [18]. Given its widespread application in health promotion and public health [19-22], it is important to evaluate its use in health misinformation. Counterinfodemic activities fit within this paradigm as the information environment is an increasingly recognized determinant of health influenced at multiple levels, from clinical interactions to social media regulation [23]. This perspective indicates a need to comprehensively target misinformation throughout the socioecological spectrum [8], reflected in the US Surgeon General's "whole of society" response to misinformation [24,25] and reports from the World Health Organization and other public health experts [26,27].

Figure 2. Socioecological model.

Environmental Health Framework

Environmental health is an area of public health focused on the health impacts of the natural and built environment. Despite its decades of use, the term "information environment," previously defined as the space where people receive and process information to make sense of the world [28,29], has only recently been applied to misinformation. In national defense, it was conceptualized to facilitate (often clandestine) information operations [30]. Political science literature has examined to

XSL•FO
RenderX

what extent the information environment is conducive to political knowledge, civil discourse, and other democracy-relevant outcomes [31]. In both instances, the implied orientation of the information environment is toward information producers, rather than information consumers.

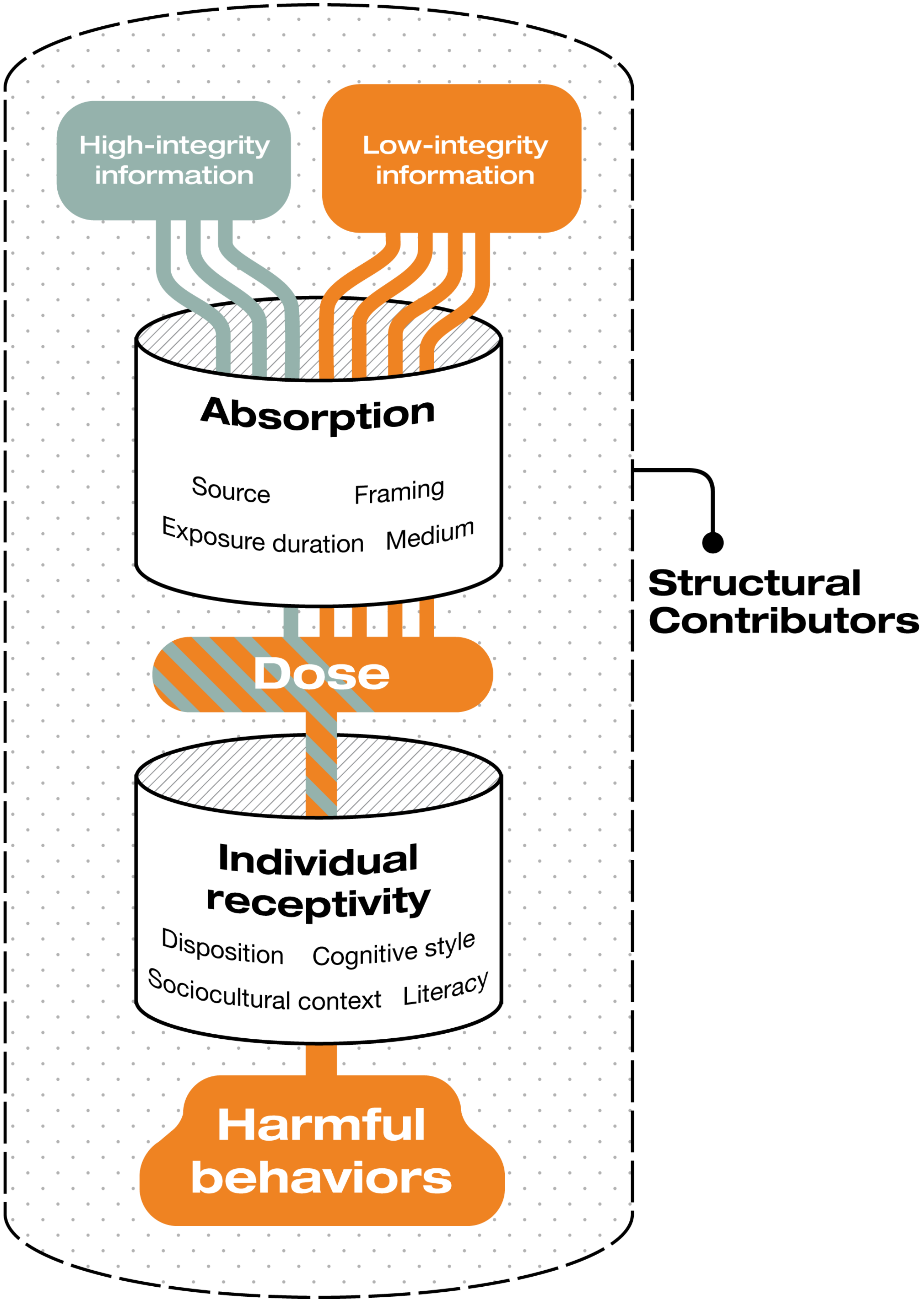
Environmental analogies about health-related information challenges have expanded, as scholars have alluded to the toxic effects of a polluted media environment [32,33]. In 2021, the US Surgeon General included the subtitle “Building a Healthy Information Environment” in his special advisory on misinformation [24]. The New York City Commissioner of Public Health, Ashwin Vasan, and the New York City Mayor, Eric Adams, recently urged public health authorities to “treat social media as a toxin, ever present in our daily environments” [34]. Here, the implied orientation is toward information consumers.

Despite the use of environmental metaphors, environmental health frameworks have been underused to understand public health–related information challenges. From a public health perspective, the information environment has been defined as an adaptive space that includes content from traditional and web-based media and in-person sources and technology to access and process this content [35]. This paradigm highlights several mechanisms of misinformation spread and corresponding opportunities for intervention: altering the dose of information of variable integrity to which an individual is exposed, influencing an individual’s receptivity to toxic misinformation, assessing the threat posed by a claim or narrative (referred to as hazard identification), and mitigating the harms of information hazards through multipronged approaches (hazard management) (Figure 3). Detailed definitions and examples of each of these intervention types are provided in the “Methods” section.

By applying these 3 models, this study sought to identify gaps and opportunities in an aggregate view of pandemic-related infodemic management interventions and to assess the use of public health frameworks to broadly structure and strategize responses to infodemics.

XSL•FO
RenderX

Figure 3. Environmental health framework.



XSL•FO
RenderX

Methods

Data Collection

This analysis drew on a dataset of infodemic management interventions aiming to address the effects and spread of misinformation that was previously developed as part of a report commissioned by the National Academies of Sciences, Engineering, and Medicine, which ultimately led to a peer-reviewed publication [8,36,37]. These interventions, which were identified between October 2022 and January 2023, include guidelines, policies, and tools from local and federal governments, public health departments, nonprofits, universities, technology and media companies, and other organizations [8]. The original authors identified these interventions through searches of the following sources: academic literature about infodemics and infodemic management; gray literature from organizations including federal agencies, nongovernmental organizations, and technology companies; and websites from state and local health departments [8,36]. Interviews with key informants were used by the original authors to identify additional interventions [36]. We expanded this dataset to include additional interventions lacking from the original dataset, focusing on interventions performed by professional societies that were identified through a similar search strategy, and reviewing the websites and resources of medical and scientific societies. As many societies' interventions were undertaken without either a publication or a description of such interventions on the societies' websites, the goal with this expansion was to be illustrative of these interventions and not exhaustive. The final dataset consisted of 379 interventions and can be made available upon request.

Data Analysis

We used framework analysis, a form of qualitative content analysis useful for applied health policy research [12,13]. Framework analysis provides a comprehensive and systematic approach to describe, interpret, and identify patterns in policies and procedures [12,14]. Codes based on thematic frameworks are applied to cases, allowing data to be compared across and within cases [13]. Through applying frameworks to a given topic, framework analysis can assess the relevance of public health analogies that are frequently applied to health infodemics but have yet to be rigorously defined in this context. Studying multiple frameworks allows for a more comprehensive lens to examine the many dimensions to an issue such as misinformation [13].

Five steps are involved in framework analysis: (1) familiarization, in which the researchers become immersed in the data and reflect on patterns; (2) identifying the thematic framework, based on emerging themes; (3) indexing, or coding components of the data that correspond to themes; (4) charting, which involves rearranging data based on themes; and (5) mapping and interpretation, when themes are analyzed through the charts [12]. We first familiarized ourselves with the data by reviewing the intervention descriptions and websites in the dataset. The thematic frameworks were identified based on prior literature cited in the introduction that provide a range of perspectives to conceptualize misinformation. We developed a coding scheme of deductive codes drawn a priori from the components of the frameworks. This coding scheme accommodated additional inductive codes that emerged through the coding process.

The epidemiological model included the following codes: prevention, social listening, risk assessment, and response. Prevention activities proactively protect populations and information networks from the adverse effects of an infodemic. Social listening activities identify and track harmful (web-based) narratives [38]. Risk assessments determine which narratives require intervention based on factors such as its spread over time, the channels in which it is disseminated, and the communities it affects, with the goal of avoiding expending limited resources on or giving oxygen to low-impact narratives [38]. For example, narratives about vaccines causing infertility that are disseminated widely in the press and on social media during a pandemic would be considered high risk [38]. Finally, rapid responses curtail the spread of harmful information.

The codes derived from the socioecological model included individual, interpersonal, community, organization, and public policy, referring to the societal level at which influence was exerted on the information environment [19]. A public policy intervention was considered to be "a choice made by government to undertake some course of action" involving goals and means of reaching them [39].

The following codes were applied for the environmental health model: dose (which could be further specified as increasing high-integrity information exposure, decreasing low-integrity information exposure, or influencing absorption), receptivity, hazard identification, and hazard management. Drawing from toxicology, "dose" refers to the concentration of low-integrity information compared with high-integrity information, defined as information that is "trustworthy; distinguishes fact from fiction, opinion, and inference; acknowledges uncertainties; and is transparent about its level of vetting," [40] and the degree of absorption of this content [41]. Hazard identification and management are conducted by organizational and governmental entities engaged in infodemic management and information integrity protection. Analogously to toxicology approaches, hazard identification refers to assessing the health effects of an information toxin [42]; hazard management describes multipronged approaches to evaluating and mitigating the threats posed by such a toxin. While structural determinants (eg, health care access or socioeconomic marginalization) influence the information environment, we did not code for this domain in order to focus on the individual components of the information environment that are specific to this model. Examples of each of these codes are given in Table 1.

XSL•FO
RenderX

Table . Example interventions corresponding to each environmental health code.

Code	Examples	Examples
Dose	InVID assists journalists in assessing the reliability of videos on social media, thus facilitating the sharing of high-integrity videos while inhibiting the further spread of low-integrity videos.	[43]
Receptivity	Interland is a game developed by Google that teaches young children to distinguish truths from misinformation on the web.	[44]
Hazard identification	Logically tracks misinformation campaigns to understand threats to national security, corporations, nonprofits, and elections.	[45]
Hazard management	CrossCheck, a program run by First Draft, promotes collaboration and resource sharing for journalists responding to misinformation. The Vaccination Community Navigator Program similarly takes a multipronged approach in educating community health workers to boost vaccine confidence.	[46]

Coding was conducted in an iterative, discursive process. One author (JNJ) coded the entire dataset in batches, documenting evolving code definitions and interpretations of the data, where relevant, multiple codes were applied to the same intervention. After each batch, 2 of the authors met to discuss uncertainties and insights that arose, such as ambiguities in the code definitions and emerging patterns in the data. Coding was conducted iteratively, until thematic saturation was reached [47]. Then, DS independently coded a random sample of approximately 20% of the dataset. Codes were reviewed to ensure alignment and discrepancies were resolved through discussions between both authors.

Results

Overview

Overall, 379 interventions were included in the final analysis, including 14 interventions from professional societies that were identified through the expanded search. The 3 frameworks lended distinct insights into the functions and capacities of the interventions (Table 2). The applications of each of the frameworks are described in detail in the subsequent sections. For further details on the coding results and representative interventions, see Multimedia Appendix 1.

XSL•FO
RenderX

Table . Insights from the 3 frameworks.

Key finding	Framework	Supporting evidence	Infodemic management recommendations
Risk assessments are often value-based or poorly defined.	Epidemiological framework	Vague or absent language about how risk assessments are conducted.	Risk assessments should be rigorous, objective, and transparent about how community values are incorporated into decision-making.
Interventions are skewed toward acting at the individual level and often focus on only 1 level of influence.	Socioecological model	Most interventions were focused on either individuals alone or individual members of organizations, rather than implementing structural change with community, interpersonal, organizational, or policy interventions. Only 11% of interventions received more than 1 socioecological code.	Interventions acting at the interpersonal, community, organizational, and policy levels should be explored, and structural barriers to implementing interventions at these levels should be identified and overcome. Collaborations should involve interventions targeting multiple levels of the socioecological spectrum.
Interventions often lack mechanisms to reach their intended audiences (ie, the Field of Dreams Fallacy) [48].	Socioecological model	Abundance of resources and tools that lacked connections to established workflows and organizations within the socioecological spectrum.	Interventions should be developed in partnerships with the organizations that are intended to use them.
Interventions place a greater emphasis on increasing high-integrity information rather than decreasing low-integrity information.	Environmental health framework	More than 3 times as many interventions address high-integrity as low-integrity content.	Interventions that decrease the spread of low-integrity information should be developed.
Demographic factors are emphasized when addressing receptivity to misinformation, while psychological factors are overlooked.	Environmental health framework	Focus on targeting racial, cultural, or age-related communities.	Interventions should consider approaches to segmenting audiences based on personas and psychobehavioral factors.
Interventions that address receptivity tend to involve a one-time action rather than longitudinal education.	Environmental health framework	Prevalence of self-contained courses, games, handouts, etc, that lack mechanisms to reinforce instruction over time.	Media literacy initiatives should incorporate mechanisms for longitudinal instruction on detecting and responding to misinformation.
Few organizations are equipped to implement hazard management approaches, despite increasing awareness that such approaches are critical.	Environmental health framework	Overrepresentation of tool kits, handbooks, and other resources lacking direct action in the hazard management category.	Media, public health, and government agencies should adopt hazard management approaches.

Epidemiological Framework

By distinguishing between the stages of an infodemic, the epidemiological framework highlighted critical distinctions in the foci of interventions that emerged in response to a specific ongoing or predicted infodemic. This framework was less relevant to interventions that addressed general components of misinformation that were agnostic of a particular crisis, such as tools providing assessments of the credibility of information sources. The framework also did not apply to interventions that lacked a clear audience or mechanism of impact.

In total, 50% (189/379) of interventions were engaged in activities intended to prevent an infodemic itself, in contrast to preventing an individual from falling for misinformation amid an ongoing infodemic. Prevention activities were most prominent when the amount of misinformation was low. Moreover, 19% (73/379) of interventions conducted social listening, monitoring conversations, concerns, claims, and news, online or offline [49]. Social listening tools most often analyzed social media feeds and datasets. The degree of analysis varied widely, from tracking misinformation with artificial intelligence to descriptive statistics on rumor spread. Seven percent (28/379) were risk assessment interventions that assessed the severity or status of an infodemic to inform whether and to what extent a response was needed. These interventions not only provided data that could be relevant to a risk assessment, such as the amount of spread of a rumor, but conducted the risk assessment itself. Most interventions (286/379, 76%) responded to an ongoing infodemic, primarily through fact-checking, debunking, and amplifying reliable information and sources. They also conducted prebunking to address topics for which misinformation is already widespread.

Socioecological Model

The socioecological model allowed for a better understanding of the key groups and audiences that are affected by or are in a position to address misinformation. Applying this framework revealed a skew toward interventions that acted at the individual level, rather than the interpersonal or community levels. While most interventions were directed toward organizations, they required exposure or uptake by individual members, rather than spurring structural change within the organization overall. In addition, interventions often lacked a clearly defined target

XSL•FO
RenderX

group and means of reaching this audience. By revealing these shortcomings, the socioecological model shed light on opportunities to align valuable resources with the groups with the greatest capacity to leverage them.

We identified 150 (40%) interventions that acted at the individual level. These interventions included media literacy and prebunking initiatives, repositories of reliable information, fact-checks and debunks, and tools evaluating the credibility of claims and sources, when these tools were intended for use by the general public. Interventions acting at the interpersonal level, such as an app that provides guidance about discussing vaccines with friends, were the least common, representing only 2% (9/379) of this dataset. Eleven percent (42/379) of interventions were community-level, targeting groups based on educational systems, geographic regions, and racial or ethnic identities, as well as social networks. The interventions often included content or dissemination strategies tailored to a community's needs. The 178 (47%) organization-level interventions primarily provided resources and tools that were intended for members of a profession, such as journalists, researchers, physicians, teachers, librarians, policy makers, or organizational bodies. These resources included infodemic management tool kits, communication materials, social listening platforms, media literacy curricula, reporting guidelines, and social media policies. There were 39 (10%) public policy interventions. Most of these policies were developed by federal governments. Two came from the United States; other regions included Singapore, Australia, the United Kingdom, France, Egypt, Germany, and the European Union.

Environmental Health Framework

The environmental health framework allowed for a more nuanced perspective on the mechanisms through which interventions interacted with the information environment. By outlining a variety of components that contribute to the information environment, this framework underscored the importance of contextualizing misinformation within information networks and audiences.

Most interventions (244/379, 64%) targeted the dose of high- and low-integrity information. More interventions increased the amount of high-integrity information (155/379, 41%) rather than decreasing the volume of low-integrity information (44/379, 12%). We identified 61 (16%) interventions that addressed receptivity to misinformation. Most of these interventions involved media literacy education, including curricula, games, infographics, and web-based courses. Seventeen percent (65/379) of interventions conducted hazard identification by assessing the dose or toxicity of misinformation. These interventions were primarily resources and tools for professionals, particularly infodemic managers, public health communicators, and journalists. The interventions involved content verification, social listening, credibility assessments, and fact-checking. Seventy (19%) hazard management interventions took a comprehensive and higher-level approach to addressing misinformation that went beyond any 1 particular intervention. They often took the form of tool kits, handbooks, field guides, and frameworks intended to inform professional hazard management activities, rather than conducting hazard management themselves.

Crosscutting Insights

We identified several findings that suggest opportunities for future interventions relating to the use of technology, coordination, and sustainability that surfaced from a combination of all 3 frameworks (Table 3). For example, some interventions such as artificial intelligence–powered chatbots suggested an overzealous application of new technologies that lacked grounding in user needs. Perhaps owing to the urgent and unprecedented nature of the COVID-19 pandemic, interventions were often duplicative and short-lived.

Table . Crosscutting insights.

Key finding	Supporting evidence	Infodemic management recommendations
Greater strategic direction to align theories of change with desired impact is needed.	Unclear distinctions between efforts to address acute compared with endemic misinformation as well as efforts engaged in prevention versus response. The intended audiences of interventions also tended to be poorly defined.	Interventions should specify the nature of the infodemics they are intended to address, intentionally select a guiding framework, and address the unmet needs of a specific audience.
Technological tools are often built and used without adequate need finding.	Prominence of tools such as chatbots enabled by technology that do not clearly fill a well-defined need.	The design process for interventions should center around identified needs rather than the tool.
Lack of coordination or pervasive duplication of efforts.	Very few initiatives included cross-sector collaboration; those that did were not sustainably funded to persist beyond the pandemic. A number of initiatives duplicate work and effort (eg, see "tool kits").	Sustainable cross-disciplinary or sector coordination mechanisms may be required to support effective and ethical infodemic management initiatives [50].
Short-term funding opportunities early on in the COVID-19 pandemic were not conducive to long-term sustainability.	Many interventions had concluded or had websites that had not been recently updated.	Sustainability given funding trends should be a key consideration when developing interventions. Funding programs should include support to sustain efforts beyond immediate crises and collect longitudinal data.
The role of incidental information exposure compared with intentional information consumption was rarely accounted for.	Interventions frequently made unsupported assumptions about the degree of agency individuals have in the information they encounter.	Future frameworks should incorporate the distinction between incidental information exposure and intentional consumption.

XSL•FO
RenderX

Discussion

In our analysis, the epidemiological, socioecological, and environmental health frameworks shed light on trends, gaps, and opportunities among counterinfodemic interventions. The epidemiological framework revealed an opportunity to implement more robust and transparent risk assessment measures in partnership with communities to triage rumors and allocate resources, particularly as more evidence emerges on the threats posed by various claims and narratives. By relying on value judgments, the risk assessments in the interventions in this dataset risk undermining trust and expending limited resources on low-impact efforts. Instead, the World Health Organization recommends developing risk assessment matrices that synthesize considerations such as the timing of a narrative, its spread on various platforms, and the impacted communities to categorize narratives as high, moderate, or low risk, and positive sentiment [38].

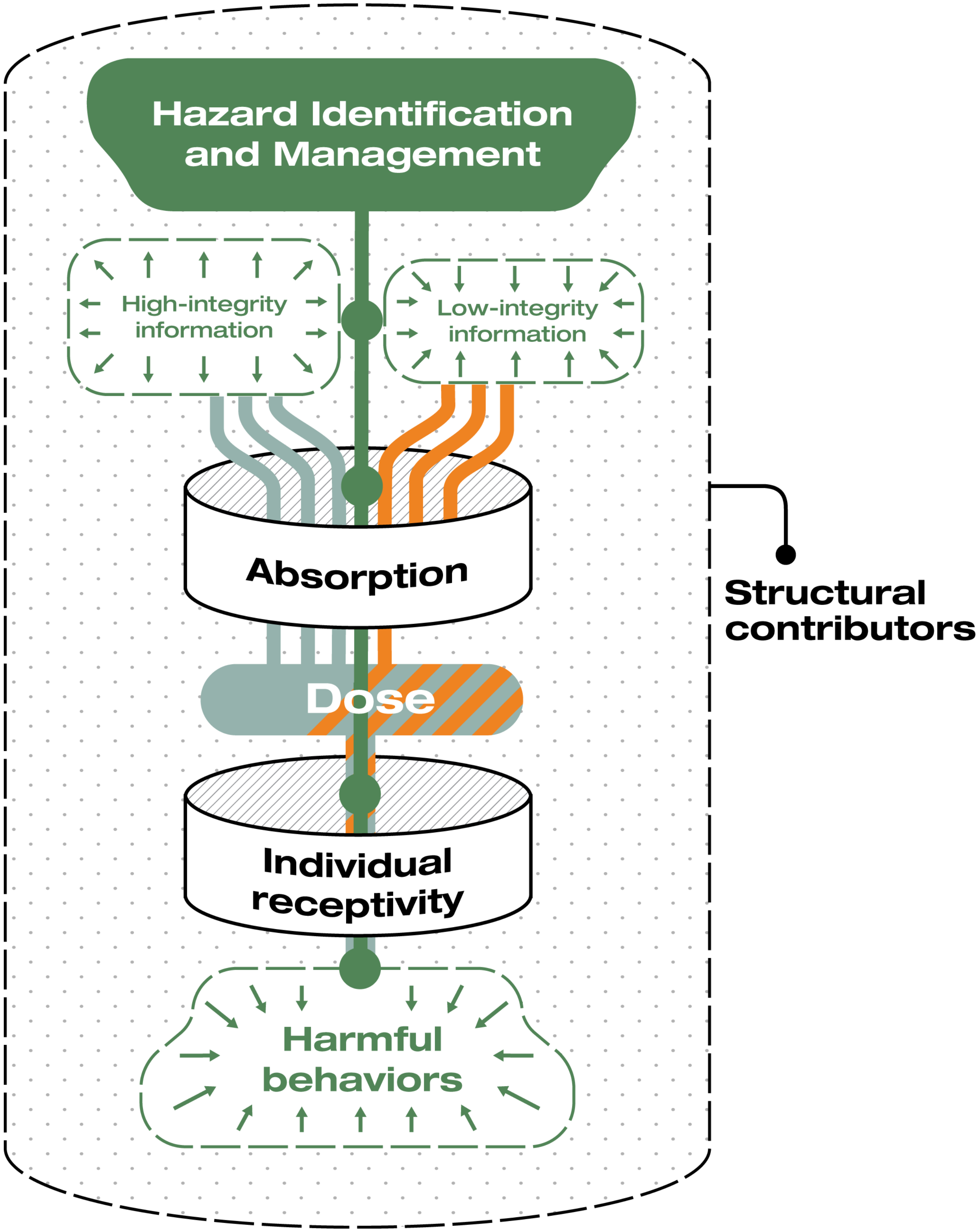
The socioecological framework demonstrated the need to target higher levels of influence through collaborations spanning multiple levels, reinforcing a finding from the original analysis of this dataset [8]. Scholars have recently argued that the outsized attention given to individually framed behavioral interventions "pollutes" the discourse and diverts attention from structural interventions [51,52]. This trend was replicated in our dataset, where structural change through public policy or enduring platform adjustments was rarely the priority. As with other complex public health challenges such as diabetes or drug overdoses, structural-level interventions coordinated with efforts acting at other levels of the socioecological spectrum are likely to be more effective and sustainable than individual-level efforts in the case of infodemic management. Policy efforts to protect children from social media–related harms have garnered significant attention, most notably in the US Surgeon General's recommendation to display warning labels on social media [53]. Despite their limitations, related legislation, such as the Stop Addictive Feeds Exploitation [54], offers potential models for analogous efforts to mitigate the harms of digital infodemicshttps://www.zotero.org/google-docs/?FlRQ40

The socioecological framework additionally revealed the importance of avoiding the Field of Dreams Fallacy [48], as many interventions neglected to specify mechanisms to reach their intended audiences. While the speed of a response is often prioritized in an emergency, the resulting lack of alignment with existing efforts may prove harmful in infodemic management due to the resource and trust barriers to maintaining strong relationships with community partners. Sustaining proactively developed partnerships is needed to increase the uptake and sustainability of infodemic interventions, particularly the tool kits and other resources that were often deployed independently of established partnerships in this dataset.

The environmental health framework provided a structure for systems-level, multipronged approaches that influence the information environment as a whole (Figure 4). A key finding was that reducing exposure to low-integrity information, which digital platforms can implement through content moderation, deplatforming, and algorithmic adjustments, was a notable gap. Amid the growing politicization of content moderation, many social media platforms have recently rolled back these efforts [11,55]. Differing perceptions of trustworthiness and integrity may also reduce the efficacy of content moderation or even lead to further polarization [6,56]. Regulating algorithmic recommendation and amplification may encourage platforms to prioritize high-integrity content while protecting First Amendment rights [57]. While the answer to bad speech was once considered to be "more speech" [58], in the social media era, it is now recognized that freedom of speech does not equate to freedom of reach [59]. Current revenue models incentivize platform architectures and algorithms that promote content that provokes negative emotional reactions, particularly anger [60].

XSL•FO
RenderX

Figure 4. Opportunities for intervention based on the environmental health framework. Points of intervention within this framework are represented by green nodes; for example, interventions can modify individual receptivity to misinformation. The shift in the composition of the information environment toward high-integrity information and subsequent reduction in harmful behaviors as a result of these interventions is indicated with green arrows. Structural contributors influence these dynamics but were not a focus of the present analysis.



While many interventions used demographic characteristics to target the information environments of particular communities, psychobehavioral segmenting may allow for more precise tailoring of messages to individuals uniquely receptive to misinformation (eg, those who engage in absolutist thinking) [61,62]. An information environment perspective additionally suggests that initiatives based on inoculation theory could expand their impact through longitudinal rather than one-time

XSL•FO
RenderX

modes of engagement and by reaching a saturation point that displaces low-integrity information. Hazard management approaches are critical to address an issue as pervasive as an infodemic. Such approaches were uncommon in our dataset, however, likely due to the funding, coordination, and sustainability challenges. Strong governance and financial support are needed to enable key stakeholders, including media, public health, environmental scientists, and government, to create and sustain hazard management approaches, potentially following models such as the Elections Infrastructure Information Sharing and Analysis Center [50].

Several key crosscutting considerations emerged (Table 3). Infodemic management interventions could benefit from greater strategic direction regarding the theories of change applied in various settings. The intended mechanism and audience of an intervention should be informed by a framework that aligns with the relevant type of information distortion. For example, while misinformation is often considered as part of acute infodemics, endemic misinformation unrelated to particular health events may require different theories of change, use of alternate frameworks (eg, socioecologic or environmental), and corresponding interventions. Too often, the development of tools using novel technologies such as generative artificial intelligence centered the technology itself, rather than a need they are intended to address. Need-finding processes must be incorporated into the design of technologically enabled interventions to maximize their potential impact. Design-thinking principles, for example, provide an approach to explore stakeholders' needs and develop tailored solutions [63].

Funders and stakeholders involved in the interventions were often fragmented and uncoordinated, leading to duplication and unstrategic allocation of resources. Well-governed and funded coordination mechanisms, perhaps modeled on Elections Infrastructure Information Sharing and Analysis Center, offer an opportunity to streamline resources while diversifying efforts. Since many efforts to counter the COVID-19 infodemic were not sustained after the immediate threat of the pandemic subsided, funding structures that support longitudinal and crisis-agnostic efforts are needed. Interventions rarely accounted for the distinction between incidental exposure and intentional information consumption. While a consumptive lens suggests that individuals make conscious decisions about the information they encounter, from an exposure-based perspective, individuals are subject to influence by information within their environments. Incorporating this distinction into future frameworks may illuminate new approaches for interventions. Overall, by testing these frameworks in our dataset, we identified their strengths and weaknesses, allowing for iterative adaptation to the infodemic management context.

Our analysis was limited in that not all components of the interventions that we considered, such as reach and distribution, were typically reported. As a result, it was sometimes necessary to make inferences about goals and impacts. Many interventions lacked information about time and scale, which resulted in organizing the data in a way that gave the same prominence to small- and large-scale initiatives. This lack of information biased the data toward smaller-scale initiatives, although large-scale initiatives likely had a broader impact. Many of the codes we applied were subjective, not mutually exclusive, and reliant on interpretation, a limitation that was exacerbated when details of an intervention were not available. For example, for the epidemiological framework, prevention and response entail critically distinct activities, but we were unable to distinguish between these 2 foci when information about the stage of the infodemic at which an intervention was deployed was not provided. There was also at times overlap in the insights derived from each framework; our analysis attempted to focus on the dominant framework that surfaced a given insight. The dataset used in this study is not exhaustive; notably, given the focus on terms such as "infodemic management," a term that emerged during the COVID-19 pandemic, interventions that predate the pandemic may have been underrepresented. Our study was designed to be illustrative, not exhaustive, so it did not use systematic search criteria. This study considered only 3 frameworks, which were chosen based on their prominence in public health and misinformation discourse; future work should consider additional frameworks to illuminate further findings. For example, recent work has adapted a public health prevention framework to infodemic management [64]. Finally, we acknowledge that the feasibility of our recommendations may be limited given resource constraints and an evolving evidence base.

In this study, we used a framework analysis using 3 public health frameworks to illuminate emphases and gaps in interventions to address the COVID-19 infodemic. While many opportunities to expand the reach and impact of interventions were identified, it was also clear that the landscape of infodemic management approaches lacks an overarching strategy and entity responsible for coordinating and evaluating activities. In preparation for future infodemics, emphasis should be placed on multisector collaboration, alignment with measurable and meaningful goals, and top-down approaches to determining and implementing strategies.

Acknowledgments

The authors thank Jack Gorman, Claire Wardle, and Dimitri Prybylski for helpful comments on and discussions about the manuscript. Jenna Sherman, Céline Gounder, and other members of the Reproductive Health Misinformation Working Group contributed to thoughtful discussions of the environmental health framework. The authors also thank Nina Martinez for her graphic design contributions. JNJ conducted the data analysis and composed the manuscript. SG contributed to the development and revision of the manuscript. DS conceptualized, designed, and oversaw the study; contributed to the data analysis; and supported the composition of the manuscript. This work was supported by a grant from the Robert Wood Johnson Foundation (grants 76935 and 78084). An earlier, less rigorous version of this analysis was supported by and published as a report with the American Board

XSL•FO
RenderX

of Internal Medicine Foundation. JNJ was supported by a research fellowship at the Penn Medical Communication Research Institute. DS received funding from the Weill Cornell Medicine JumpStart award. The opinions expressed in this study do not necessarily reflect those of the Robert Wood Johnson Foundation, Weill Cornell Medicine, or the Penn Medical Communication Research Institute.

Conflicts of Interest

JNJ is an employee of Roon. SG and DS have no conflicts of interest to disclose.

Multimedia Appendix 1
Example interventions.
[DOCX File, 10 KB - infodemiology_v5i1e67119_app1.docx]

References

1. World Health Organization. Managing the COVID-19 infodemic: promoting healthy behaviours and mitigating the harm from misinformation and disinformation. : World Health Organization; 2020. URL: https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation [accessed 2024-04-17]
2. Amin K, Ortaliza J, Cox C, Michaud J, Kates J. COVID-19 mortality preventable by vaccines. Peterson-KFF Health System Tracker. 2022. URL: https://www.healthsystemtracker.org/brief/covid19-and-other-leading-causes-of-death-in-the-us/ [accessed 2025-01-25]
3. Bruns R, Hosangadi D, Trotochaud M, Sell K. COVID-19 vaccine misinformation and disinformation costs an estimated $50 to $300 million each day.: Johns Hopkins Center for Health Security; 2021. URL: https://centerforhealthsecurity.org/sites/default/files/2023-02/20211020-misinformation-disinformation-cost.pdf
4. Knudsen J, Perlman-Gabel M, Uccelli IG, Jeavons J, Chokshi DA. Combating misinformation as a core function of public health. NEJM Catalyst 2023 Jan 18;4(2):CAT.22.0198. [doi: 10.1056/CAT.22.0198]
5. Burstin H, Curry S, Ranney ML, et al. Identifying credible sources of health information in social media: phase 2-considerations for non-accredited nonprofit organizations, for-profit entities, and individual sources. NAM Perspect 2023;2023:10. [doi: 10.31478/202305b] [Medline: 37916063]
6. Lalani HS, Laine C. The credibility conundrum: can social media companies define credibility for users? Ann Intern Med 2023 May;176(5):721-722. [doi: 10.7326/M23-0490] [Medline: 36877967]
7. Smith R, Chen K, Winner D, Friedhoff S, Wardle C. A systematic review of COVID-19 misinformation interventions: lessons learned. Health Aff (Millwood) 2023 Dec;42(12):1738-1746. [doi: 10.1377/hlthaff.2023.00717] [Medline: 37967291]
8. Sundelson AE, Jamison AM, Huhn N, Pasquino SL, Sell TK. Fighting the infodemic: the 4 i Framework for Advancing Communication and Trust. BMC Public Health 2023 Aug 30;23(1):1662. [doi: 10.1186/s12889-023-16612-9] [Medline: 37644563]
9. Ziemer CT, Rothmund T. Psychological underpinnings of misinformation countermeasures. J Media Psychol 2024 Nov;36(6):397-409. [doi: 10.1027/1864-1105/a000407]
10. Broniatowski DA, Simons JR, Gu J, Jamison AM, Abroms LC. The efficacy of Facebook's vaccine misinformation policies and architecture during the COVID-19 pandemic. Sci Adv 2023 Sep 15;9(37):eadh2132. [doi: 10.1126/sciadv.adh2132]
11. Appel RE, Pan J, Roberts ME. Partisan conflict over content moderation is more than disagreement about facts. Sci Adv 2023 Nov 3;9(44):eadg6799. [doi: 10.1126/sciadv.adg6799] [Medline: 37922349]
12. Thomson SB, Srivastava A. Framework analysis: a qualitative methodology for applied policy research. J Adm Gov 2009;4(2):72-79 [FREE Full text]
13. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
14. Goldsmith L. Using framework analysis in applied qualitative research. Qual Rep 2021;26(6):2061-2076. [doi: 10.46743/2160-3715/2021.5011]
15. Eadon YM, Wood SE. Combating contamination and contagion: embodied and environmental metaphors of misinformation. Convergence 2024;0. [doi: 10.1177/13548565241255347]
16. Scales D, Gorman J, Jamieson KH. The Covid-19 infodemic—applying the epidemiologic model to counter misinformation. N Engl J Med 2021 Aug 19;385(8):678-681. [doi: 10.1056/NEJMp2103798]
17. SEARO Regional Office for the South East Asia. Strategic framework for action for strengthening surveillance, risk assessment and field epidemiology for health security threats in the WHO South-East Asia region. : World Health Organization; 2023. URL: https://www.who.int/publications/i/item/9789290210030 [accessed 2025-03-04]
18. Bronfenbrenner U. Toward an experimental ecology of human development. Am Psychol 1977;32(7):513-531. [doi: 10.1037//0003-066X.32.7.513]

XSL•FO
RenderX

19. Dahlberg LL, Krug EG. Violence a global public health problem. Ciênc Saúde Coletiva 2006 Jun;11(2):277-292. [doi: 10.1590/S1413-81232006000200007]
20. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]
21. Townsend N, Foster C. Developing and applying a socio-ecological model to the promotion of healthy eating in the school. Public Health Nutr 2013 Jun;16(6):1101-1108. [doi: 10.1017/S1368980011002655] [Medline: 22014870]
22. Mehtälä MAK, Sääkslahti AK, Inkinen ME, Poskiparta MEH. A socio-ecological approach to physical activity interventions in childcare: a systematic review. Int J Behav Nutr Phys Act 2014 Feb 22;11(1):22. [doi: 10.1186/1479-5868-11-22] [Medline: 24559188]
23. Graham G, Goren N, Sounderajah V, DeSalvo K. Information is a determinant of health. Nat Med 2024 Apr;30(4):927-928. [doi: 10.1038/s41591-023-02792-9] [Medline: 38351186]
24. Murthy V. Confronting health misinformation: the U.S. Surgeon General's advisory on building a healthy information environment. : Office of the US Surgeon General; 2021. URL: https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf [accessed 2025-03-04]
25. Murthy V. A community toolkit for addressing health misinformation. : Office of the US Surgeon General; 2021. URL: https://www.hhs.gov/sites/default/files/health-misinformation-toolkit-english.pdf [accessed 2025-03-04]
26. World Health Organization Regional Office for Europe. Digital solutions to health risks raised by the COVID-19 infodemic: policy brief. : Regional Office for Europe; 2022. URL: https://iris.who.int/handle/10665/356315 [accessed 2024-04-15]
27. Donovan J, Friedberg B, Lim G, Leaver N, Nilsen J, Dreyfuss E. Mitigating medical misinformation: a whole-of-society approach to countering spam, scams, and hoaxes. : Technology and Social Change Project; 2021. URL: https://mediamanipulation.org/research/mitigating-medical-misinformation-whole-society-approach-countering-spam-scams-and-hoaxes [accessed 2024-04-15]
28. Wanless A. Seeing the disinformation forest through the trees: how to begin cleaning up the polluted information environment. : Carnegie Endowment for International Peace; 2023. URL: https://www.oecd-forum.org/posts/seeing-the-disinformation-forest-through-the-trees-how-to-begin-cleaning-up-the-polluted-information-environment?utm_source=user_mailer&utm_medium=email&utm_campaign=send_publish_notification_to_follower [accessed 2024-04-15]
29. Wanless A. The more things change. : King's College London; 2023. URL: https://kclpure.kcl.ac.uk/portal/en/studentTheses/the-more-things-change [accessed 2024-04-15]
30. Joint Chiefs of Staff. Joint publication 3-13 information operations. 2012. URL: https://defenseinnovationmarketplace.dtic.mil/wp-content/uploads/2018/02/12102012_io1.pdf [accessed 2025-03-04]
31. Jerit J, Barabas J, Bolsen T. Citizens, knowledge, and the information environment. Am J Pol Sci 2006 Apr;50(2):266-282. [doi: 10.1111/j.1540-5907.2006.00183.x]
32. Phillips W. The toxins we carry. : Columbia Journalism Review; 2019. URL: https://www.cjr.org/special_report/truth-pollution-disinformation.php/ [accessed 2024-04-15]
33. Phillips W, Milner RM. You Are Here: A Field Guide for Navigating Polarized Speech, Conspiracy Theories, and Our Polluted Media Landscape: MIT Press; 2021. [doi: 10.7551/mitpress/12436.001.0001]
34. Vasan A. We must treat social media like the toxin that it is. : The Hill; 2023. URL: https://thehill.com/opinion/technology/4058281-we-need-to-treat-social-media-like-the-toxin-that-it-is [accessed 2024-04-15]
35. Shapiro AW, Jacob N. A CERN model for studying the information environment. : Carnegie Endowment for International Peace; 2022. URL: https://carnegieendowment.org/2022/11/17/cern-model-for-studying-information-environment-pub-88408 [accessed 2024-04-17]
36. Sundelson A, Huhn N, Jamison A, Pasquino SL, Kirk Sell T. Infodemic management approaches leading up to, during, and following the COVID-19 pandemic. : Johns Hopkins Center for Health Security; 2023. URL: https://centerforhealthsecurity.org/sites/default/files/2023-04/230407-nasempaper.pdf [accessed 2025-03-04]
37. Google Sheets. Combatting mis/disinfo- best/existing practices. URL: https://docs.google.com/spreadsheets/d/1VjEwkR3tjgcPmd3ve6N03pqiHMF8hMEojUkRWx8QJFc/edit?gid=872458924&usp=embed_facebook [accessed 2024-10-02]
38. World Health Organization, United Nations Children's Fund. How to build an infodemic insights report in 6 steps. : World Health Organization and United Nations Children's Fund; 2023. URL: https://www.who.int/publications-detail-redirect/9789240075658 [accessed 2024-05-23]
39. Howlett M, Cashore B. Conceptualizing public policy. In: Engeli I, Allison CR, editors. Comparative Policy Studies: Conceptual and Methodological Challenges: Palgrave Macmillan; 2014:17-33. [doi: 10.1057/9781137314154_2]
40. National Science & Technology Council. Roadmap for researchers on priorities related to information integrity research and development. : Office of Science and Technology Policy; 2022. URL: https://web.archive.org/web/20250116082552/https://www.whitehouse.gov/wp-content/uploads/2022/12/Roadmap-Information-Integrity-RD-2022.pdf
41. Kirk MA, Deaton ML. Bringing order out of chaos: effective strategies for medical response to mass chemical exposure. Emerg Med Clin North Am 2007 May;25(2):527-548. [doi: 10.1016/j.emc.2007.02.005] [Medline: 17482031]

XSL•FO
RenderX

42. Fowle III JR, Dearfield KL. Risk characterization handbook. U.S. : Environmental Protection Agency Science Policy Council; 2000. URL: https://www.epa.gov/sites/default/files/2015-10/documents/osp_risk_characterization_handbook_2000.pdf [accessed 2025-03-04]
43. The invid technologies. InVID. URL: https://www.invid-project.eu/ [accessed 2025-03-14]
44. INTERLAND. be internet awesome. Google. URL: https://beinternetawesome.withgoogle.com/en_us/interland [accessed 2025-03-14]
45. AI powered information advantage. Logically. URL: https://logically.ai/ [accessed 2025-03-14]
46. CrossCheck: together, now. First Draft. URL: https://firstdraftnews.org/tackling/crosscheck-together-now/ [accessed 2025-03-14]
47. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
48. Karsh BT, Weinger MB, Abbott PA, Wears RL. Health information technology: fallacies and sober realities. J Am Med Inform Assoc 2010;17(6):617-623. [doi: 10.1136/jamia.2010.005637] [Medline: 20962121]
49. Sommariva S, Mote J, Ballester Bon H, et al. Social listening in Eastern and Southern Africa, a UNICEF risk communication and community engagement strategy to address the COVID-19 infodemic. Health Secur 2021;19(1):57-64. [doi: 10.1089/hs.2020.0226] [Medline: 33606573]
50. Lalani HS, DiResta R, Baron RJ, Scales D. Addressing viral medical rumors and false or misleading information. Ann Intern Med 2023 Aug;176(8):1113-1120. [doi: 10.7326/M23-1218] [Medline: 37459614]
51. Chater N, Loewenstein G. The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. Behav Brain Sci 2022 Sep 5;46:e147. [doi: 10.1017/S0140525X22002023] [Medline: 36059098]
52. Maani N, van Schalkwyk MC, Petticrew M, Buse K. The pollution of health discourse and the need for effective counter-framing. BMJ 2022 May 4;377:o1128. [doi: 10.1136/bmj.o1128] [Medline: 35508319]
53. Office of the Surgeon General (OSG). Social media and youth mental health: the U.S. Surgeon General's advisory. : US Department of Health and Human Services; 2023. URL: http://www.ncbi.nlm.nih.gov/books/NBK594761 [accessed 2025-02-02]
54. Gounardes A. Establishes the Stop Addictive Feeds Exploitation (SAFE) for Kids act prohibiting the provision of addictive feeds to minors. URL: https://www.nysenate.gov/legislation/bills/2023/S7694/amendment/A [accessed 2025-02-02]
55. Myers SL, Grant N. Combating disinformation wanes at social media giants. : The New York Times; 2023. URL: https://www.nytimes.com/2023/02/14/technology/disinformation-moderation-social-media.html [accessed 2024-08-02]
56. Otala JM, Kurtic G, Grasso I, Liu Y, Matthews J, Madraki G. Political polarization and platform migration: a study of Parler and Twitter usage by United States of America Congress members. 2021 Presented at: Companion Proceedings of the Web Conference 2021; Ljubljana, Slovenia p. 224-231. [doi: 10.1145/3442442.3452305]
57. Rubinstein I, Kenneth T. Taming online public health misinformation. Harv J Legis 2022;60:201-218. [doi: 10.2139/ssrn.4192903]
58. Brandeis L. Whitney v. People of state of California.(supreme court 1927). URL: https://www.law.cornell.edu/supremecourt/text/274/357 [accessed 2024-08-02]
59. DiResta R. Free speech is not the same as free reach. : Wired; 2018. URL: https://www.wired.com/story/free-speech-is-not-the-same-as-free-reach/ [accessed 2024-08-02]
60. Munn L. Angry by design: toxic communication and technical architectures. Humanit Soc Sci Commun 2020;7(1):1-11. [doi: 10.1057/s41599-020-00550-7]
61. Salvi C, Iannello P, Cancer A, et al. Going viral: how fear, socio-cognitive polarization and problem-solving influence fake news detection and proliferation during COVID-19 pandemic. Front Commun 2021;5. [doi: 10.3389/fcomm.2020.562588]
62. Surgo Ventures. How do we get America vaccinated? URL: https://surgoventures.org/vaccine-persona-explainer [accessed 2024-04-29]
63. Bender-Salazar R. Design thinking as an effective method for problem-setting and needfinding for entrepreneurial teams addressing wicked problems. J Innov Entrep 2023;12(1):24. [doi: 10.1186/s13731-023-00291-2]
64. Ishizumi A, Kolis J, Abad N, et al. Beyond misinformation: developing a public health prevention framework for managing information ecosystems. Lancet Public Health 2024 Jun;9(6):e397-e406. [doi: 10.1016/S2468-2667(24)00031-8] [Medline: 38648815]

XSL•FO
RenderX

Edited by T Purnat; submitted 05.11.24; peer-reviewed by J Kolis, S Mandić-Rajčević; revised version received 05.02.25; accepted 06.02.25; published 24.03.25.

Please cite as:
John JN, Gorman S, Scales D
Understanding Interventions to Address Infodemics Through Epidemiological, Socioecological, and Environmental Health Models: Framework Analysis
JMIR Infodemiology 2025;5:e67119
URL: https://infodemiology.jmir.org/2025/1/e67119
doi:10.2196/67119

© Jennifer N John, Sara Gorman, David Scales. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 24.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Geosocial Media's Early Warning Capabilities Across US County-Level Political Clusters: Observational Study

Dorian Arifi[1,2], MSc; Bernd Resch[1,2,3], Prof Dr; Mauricio Santillana[4,5], Prof Dr; Weihe Wendy Guan[3], PhD; Steffen Knoblauch[6,7,8], MSc; Sven Lautenbach[6,8], Prof Dr; Thomas Jaenisch[9,10], Prof Dr; Ivonne Morales[10,11], PhD; Clemens Havas[12], Prof Dr

[1]Department of Geoinformatics, University of Salzburg, Salzburg, Austria
[2]Interdisciplinary Transformation University Austria, Linz, Austria
[3]Center for Geographic Analysis (CGA), Harvard University, Cambridge, MA, United States
[4]Machine Intelligence Group for the Betterment of Health and the Environment, Northeastern University, Boston, MA, United States
[5]Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, United States
[6]GIScience Research Group, Heidelberg University, Heidelberg, Germany
[7]Interdisciplinary Centre of Scientific Computing (IWR), Heidelberg University, Heidelberg, Germany
[8]HeiGIT (Heidelberg Institute for Geoinformation Technology) GbmH, Heidelberg, Germany
[9]Center for Global Health, Colorado School of Public Health, Aurora, CO, United States
[10]Heidelberg Institute of Global Health, Heidelberg University Hospital, Heidelberg, Germany
[11]Department of Infectious Disease and Tropical Medicine, Heidelberg University Hospital, Heidelberg, Germany
[12]Salzburg University of Applied Sciences, Puch/Salzburg, Austria

Corresponding Author:
Dorian Arifi, MSc
Department of Geoinformatics
University of Salzburg
Kapitelgasse 4/6
Salzburg, 5020
Austria
Phone: 43 662 80440
Email: dorian.arifi@plus.ac.at

Abstract

Background: The novel coronavirus disease (COVID-19) sparked significant health concerns worldwide, prompting policy makers and health care experts to implement nonpharmaceutical public health interventions, such as stay-at-home orders and mask mandates, to slow the spread of the virus. While these interventions proved essential in controlling transmission, they also caused substantial economic and societal costs and should therefore be used strategically, particularly when disease activity is on the rise. In this context, geosocial media posts (posts with an explicit georeference) have been shown to provide a promising tool for anticipating moments of potential health care crises. However, previous studies on the early warning capabilities of geosocial media data have largely been constrained by coarse spatial resolutions or short temporal scopes, with limited understanding of how local political beliefs may influence these capabilities.

Objective: This study aimed to assess how the epidemiological early warning capabilities of geosocial media posts for COVID-19 vary over time and across US counties with differing political beliefs.

Methods: We classified US counties into 3 political clusters, democrat, republican, and swing counties, based on voting data from the last 6 federal election cycles. In these clusters, we analyzed the early warning capabilities of geosocial media posts across 6 consecutive COVID-19 waves (February 2020-April 2022). We specifically examined the temporal lag between geosocial media signals and surges in COVID-19 cases, measuring both the number of days by which the geosocial media signals preceded the surges in COVID-19 cases (temporal lag) and the correlation between their respective time series.

Results: The early warning capabilities of geosocial media data differed across political clusters and COVID-19 waves. On average, geosocial media posts preceded COVID-19 cases by 21 days in republican counties compared with 14.6 days in democrat counties and 24.2 days in swing counties. In general, geosocial media posts were preceding COVID-19 cases in 5 out of 6 waves across all political clusters. However, we observed a decrease over time in the number of days that posts preceded COVID-19

XSL•FO
RenderX

cases, particularly in democrat and republican counties. Furthermore, a decline in signal strength and the impact of trending topics presented challenges for the reliability of the early warning signals.

Conclusions: This study provides valuable insights into the strengths and limitations of geosocial media data as an epidemiological early warning tool, particularly highlighting how they can change across county-level political clusters. Thus, these findings indicate that future geosocial media based epidemiological early warning systems might benefit from accounting for political beliefs. In addition, the impact of declining geosocial media signal strength over time and the role of trending topics for signal reliability in early warning systems need to be assessed in future research.

(JMIR Infodemiology 2025;5:e58539) doi:10.2196/58539

KEYWORDS

spatiotemporal epidemiology; geo-social media data; digital disease surveillance; political polarization; epidemiological early warning; digital early warning

Introduction

On March 12, 2020, the World Health Organization (WHO) declared the novel coronavirus disease COVID-19 a pandemic [1]. Its high infectiousness and severity posed a great threat to large populations worldwide, ultimately causing an estimated 15.9 million pandemic-related deaths [2], challenging health care professionals, hospitals, and authorities alike. Thus, decision makers around the world sought to unravel and predict the spreading dynamics of this novel coronavirus. Consequently, researchers explored various ways of adjusting and improving existing epidemiological early warning systems, with complementary internet-based data sources being one such method to better monitor and anticipate how this new disease would affect different geographies around the world [3-5].

Multiple studies have already emphasized the role of geosocial media data in improving early warning of epidemiological phenomena. For instance, geosocial media data were used to improve real-time reporting on diseases like Zika and Ebola [6] or to enhance the prediction of dengue fever [7]. Accordingly, various recent examples further emphasize the ability of geosocial media data for real-time surveillance and early warning in the context of COVID-19 [8,9]. In this regard, Kogan et al [10] observed that in the beginning of the pandemic, increases in geosocial media activity, among other digital data sources, preceded surges in COVID-19 cases by 2 to 3 weeks on state level. Similarly, Zhang et al [11] used geosocial media posts in a linear regression model to predict COVID-19 signals on state-level. Yet, an increasing trend in epidemiological analysis focuses on ever finer spatial scales in the hopes of gaining a more distinct understanding of infection patterns. In this regard, Stolerman et al [12] investigated the value of X posts (formerly known as Twitter) for COVID-19 early warning on a representative subset of US counties. However, the authors only investigated a comparably small sample of counties (n=97), raising questions with respect to the generalizability of the presented results. Thus, in this study, we extended their investigation on the early warning capabilities of geosocial media data to all US counties.

Furthermore, geosocial media data garnered notable attention across various fields to answer research questions related to mental health or public attitudes, during the COVID-19 pandemic [13]. For instance, researchers investigated how language in Reddit posts reflected real-world pandemic-driven events like lockdowns, revealing significant psychological shifts among users which coincided with tendencies toward decreased analytical thinking [14]. Similarly, Swain et al [15] developed a machine learning model leveraging geosocial media data to predict disruptions in mental well-being caused by the COVID-19 pandemic. Beyond that, researchers explored geosocial media users' attitudes and concerns toward COVID-19 vaccines for the United States and the United Kingdom [16]. They observed that geosocial media derived results correlated broadly with nationwide surveys. In essence, the previous results suggest that geosocial media exchange during the COVID-19 pandemic was likely influenced by real-world public attitudes and even users' mental health. Similarly, a variety of studies indicate that the language used and the topics of interest of geosocial media users vary based on political beliefs [17-19]. This further supports our underlying assumption that differences in political beliefs are likely to be reflected in geosocial media behavior, which could, in turn, correspond to differences in geosocial media's early warning capabilities for COVID-19 cases.

However, even before the surge of the COVID-19 pandemic, researchers observed the emergence of echo chambers when analyzing pro and antivaccination attitudes on Facebook (Meta), which in their opinion might have caused further polarization [20]. In this regard, Howard et al [21] found that X was particularly prone to misinformation and polarizing content compared with professionally produced news during the 2016 presidential election. They even found more misinformation being prevalent in swing states. Such spread of misinformation and emerging political polarization on geosocial media should be of further concern for health experts and policy makers. In particular, since many researchers illustrated that diverging political beliefs can not only influence exchange on geosocial media [17-19], but also real-world individual behavior such as vaccine up-take [22] or the usage of nonpharmaceutical interventions such as mask wearing [23]. This is in line with previous findings [24], which highlight significant variation between individuals with different political beliefs with respect to self-estimated COVID-19 risks, self-reported adherence to COVID-19 health care measures, and expectations on the future course of the pandemic. In addition, researchers observed that US counties that voted in favor of the republican presidential candidate in the 2016 election, experienced up to 3 times higher mortality due to COVID-19 during the winter of 2020 [25].

XSL•FO
RenderX

Hence, in essence it can be assumed that individuals may respond differently on geosocial media to a swiftly politicized epidemic event like the COVID-19 pandemic [26], corresponding to their political beliefs. Evidence further suggests that differences in political beliefs do not only influence online and offline behavior, but they might indeed coincide with higher COVID-19 cases and death rates [25,27,28]. In summary, these results highlight the need to understand and adjust geosocial media based early warning systems with respect to political beliefs. Thus, within the scope of this paper, we seek to answer the following 2 research questions with a particular focus on geosocial media posts:

1. How do the early warning capabilities of geosocial media data change across consecutive epidemiological waves of COVID-19 cases?
2. What differences across US county-level political clusters can be observed with respect to geosocial media's early warning capabilities for COVID-19 cases?

To explore the early warning capabilities of geosocial media data, we determined the correlation between geosocial media posts and COVID-19 cases and the number of days by which signals in geosocial media data preceded actual COVID-19 cases (temporal lag). Furthermore, we specifically examined the temporal lag and the correlation in the context of political clusters based on US county voting data and over the course of 6 consecutive waves of COVID-19 cases.

Methods

Data Collection

We used 2 main data sources in this study. First, we gathered official data on confirmed COVID-19 cases in the United States and we obtained geolocated posts (Tweets) from the geosocial media network X. The time frame for which we collected our data ranges from February 28, 2020, the beginning of the pandemic in the United States, to April 27, 2022, which denotes the end of the first major Omicron wave that began in November 2021 [29]. This time frame covers the main COVID-19 waves, time periods before and after the availability of vaccines, and was selected based on retrospective knowledge on the course of the pandemic. The contiguous United States was chosen as our study area. Furthermore, to gain a more refined understanding of the underlying spatial patterns, we decided to use US counties as our finest spatial analysis resolution, on which we identified politically similar clusters, advancing previous research that was mostly performed on national or state levels.

COVID-19 Case Data

We downloaded officially confirmed COVID-19 cases for the United States in csv format from the not-for-profit public data aggregator USAFacts [30]. The COVID-19 cases csv file contained daily cumulated COVID-19 cases, which we transformed into daily incidence data. In addition, we applied a 14-day moving average to account for possible reporting delays and differing update cycles across states.

Geosocial Media Data

Furthermore, we collected geolocated posts from the geosocial media network X through their official application programming interfaces (APIs) during our investigation time frame [10,12], when academic access for researchers was still available. In particular, we used the Twitter REST and Streaming API access points to gather about 727 million geosocial media posts. The REST API allowed us to retrieve posts from the previous 7 days, with a limit of 450 requests per 15-minute window. In contrast, the Streaming API provided a continuous, real-time stream of posts. For both API endpoints we applied filters to capture only posts containing a geolocation. Thus, each collected geosocial media post includes a geolocation, which can either be the Global Navigation Satellite System position of the device through which the post was shared, or a user-defined location. Furthermore, locations can consist of polygons (eg, city, state level polygons) or point locations. We excluded geosocial media posts with polygon or point geometries that were not located within the county-level geometries, which left us with 242 million posts.

Next, to obtain geosocial media posts that are relevant to the analysis of COVID-19, we performed keyword filtering on the remaining 242 million posts located within county geometries. Therefore, we defined keywords based on the knowledge of geosocial media and health experts, with the goal to properly capture geosocial media trends relevant to the COVID-19 pandemic (Textbox 1). For some keywords only their word stem was used to allow for different variations of the word to be detected.

Textbox 1. Keywords used for relevant post extraction.

COVID-19 keywords:

covid, corona, sarscov, sars-cov, sars, epidemic, pandemic, influenza, virus, viral, infect, spread, 2019-ncov, Delta variant, Omicron, H1N1, H3N2, Wuhan, sickness, transmission, contagio, Illness, outbreak, super spread, incubation, quarantine, lockdown, vaccin, fever, cough, headache, fatigue, body aches, loss of taste, loss of smell, no smell, no taste, respirator, face mask, masks.

After the keyword extraction, the posts were aggregated on US county-level and a 14-day moving average was applied. Finally, to cope with differing amounts of geosocial media posts over time and space, we normalized the amount of relevant filtered geosocial media posts over the amount of all geosocial media posts on county level. In the remainder of this study, we solely used this ratio, that is, the proportion of relevant posts over all posts per county. This allows us to account for spatially clustered population and post density. In total, the semantic filtering procedure left us with 3.3 million relevant posts.

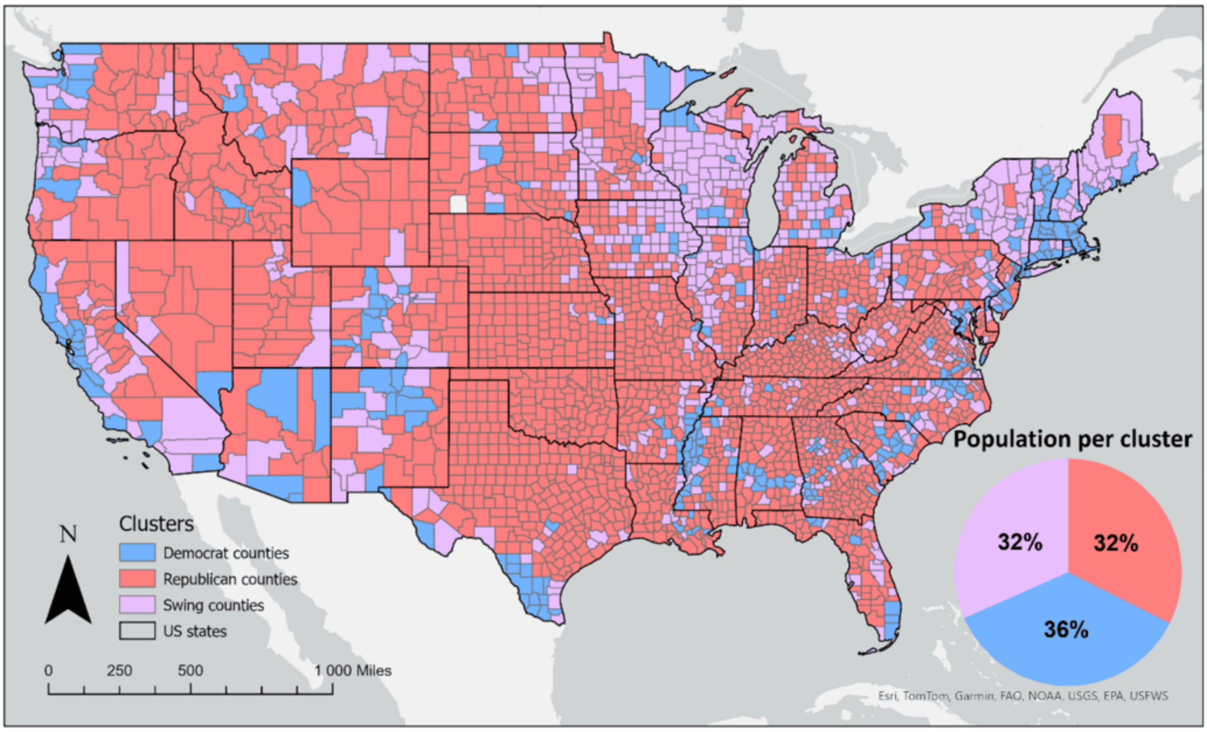
Political Clusters

To examine the differences between the various political beliefs, we based our analysis on voting data from the last 6 US presidential elections. The voting data were obtained from the Harvard Dataverse [31]. We classified US counties into 3

XSL•FO
RenderX

different clusters depending on their historical vote share for either the republican or the democrat party. In the political sciences literature, swing states are traditionally defined through a variety of quantitative and qualitative indicators. However, most of these definitions such as the bellwether status of a state [32], or it being perceived as a battleground [32], are not directly transferable to county-level analysis. Thus, we decided to base the classification into republican, democrat, or swing county clusters, on the so-called flippability of a county [32]. We chose to assess the flippability of a county on its last 6 federal election cycles. Concretely, we classified a county as belonging to a specific party, if said party had won at least 5 consecutive elections in the last 6 elections cycles. All other counties were considered as flipping between political parties and thus classified as swing counties. This division yielded political clusters, each of which representing approximately one third of the US population (Figure 1).

Figure 1. Geospatial distribution of political belief clusters on county level based on the last 6 election cycles.



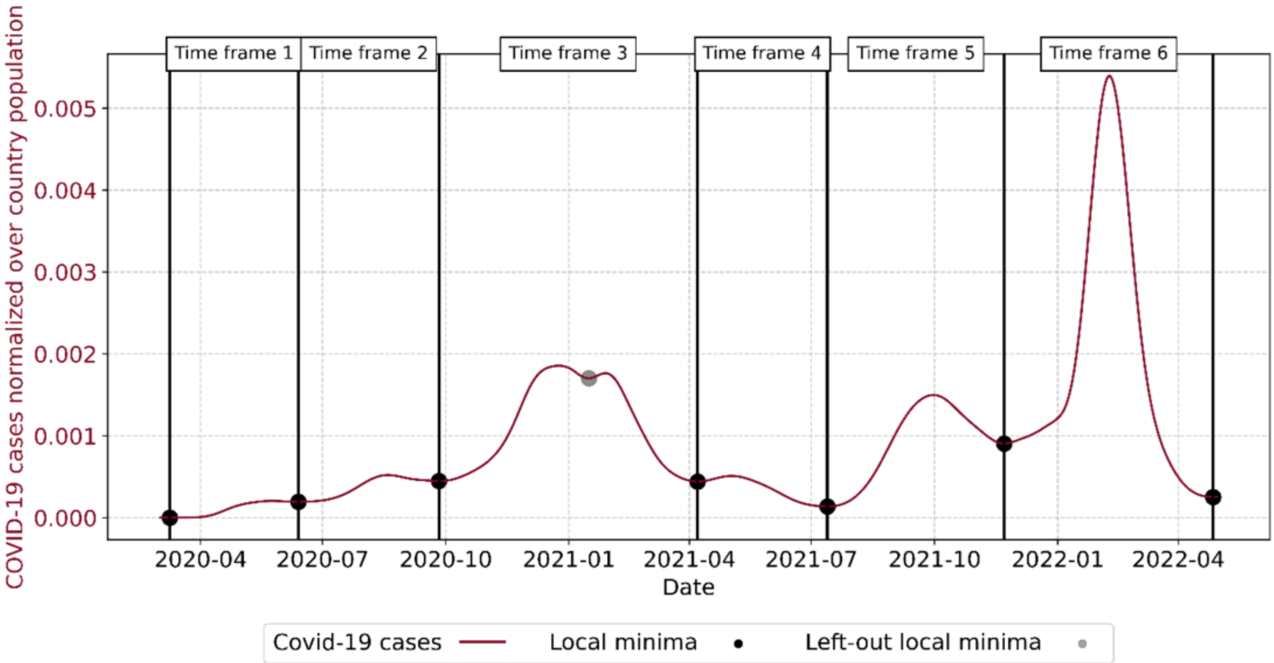
Defining COVID-19 Waves

We split the COVID-19 cases time series into smaller time frames, to capture individual epidemiological waves. However, there exist multiple approaches to define epidemic waves ranging from statistical methods using, for instance, exponential growth [10,33] or the effective reproduction number R [12,34]. In contrast, other authors tried to identify statistics and guiding principles on the duration of COVID-19 waves based on empirical data [35]. Nevertheless, all these approaches are based on strong assumptions and subjective definitions on what thresholds characterize an epidemic wave. Thus, similarly to [35], we based our definition of COVID-19 waves on a rule-based approach using the local minima on a 21-day moving average of the COVID-19 cases, which was informed through retrospective knowledge on the course of the pandemic.

We defined these time frames based on COVID-19 cases for the entire United States, rather than defining them individually for each political cluster. Furthermore, our procedure yielded 7 different time frames (Figure 2). Nonetheless, these 7 time frames did not accurately reflect all epidemic waves. In particular, the wave ranging roughly from October 2020 to April 2021, was split into 2. As a result, we decided to combine the original time frames 3 and 4 into 1 epidemic wave, which left us with 6 epidemic waves in total. This decision enabled us to capture the epidemic waves more accurately (Figure 2).

XSL•FO
RenderX

Figure 2. COVID-19 case waves for the entire US primarily defined through local minima.



Early Warning Capabilities

Finally, we quantified the early warning capabilities separately for each of the epidemic waves. We defined early warning capabilities twofold: (1) as the Pearson correlation between the time series of COVID-19 related geosocial media posts and COVID-19 cases, and (2) the number of days by which geosocial media posts preceded COVID-19 cases. However, the more important measure for early warning is the correlation between the 2 time series. Put differently, this means that if the temporal lag is high, however a correlation close to zero is present, it is obviously not reasonable to attribute any early warning capabilities to geosocial media data.

Furthermore, to identify the maximal correlation and the corresponding temporal lag, we shifted the geosocial media posts time series between 7 and 42 days into the future to determine the highest possible early warning capabilities. This procedure is repeated for each individual political cluster and epidemic wave, respectively. The decision to investigate a temporal lag between 7 and 42 days into the future was based on previous results [12], in which an early warning model, using, among others, geosocial media data, was able to predict COVID-19 cases between 1 and 6 weeks in advance.

Ethical Considerations

The study was carried out in accordance with the Declaration of Helsinki and with the ethical regulations in place at the Paris Lodron University of Salzburg, and complies with the General Data Protection Regulation legislation of the European Union. We only used publicly available data, which were collected in accordance with the terms of service of the respective geosocial media platform X at the time of data collection. Furthermore, no identifiable information was revealed in this study. Specifically, the user-provided geographic locations and semantic content were spatially aggregated to ensure user privacy and anonymity. Thus, we did not need to seek ethical approval from our institution for this study.

Results

Democrat Counties

Figure 3 depicts the Pearson correlation for different temporal lags between the time series of COVID-19 cases and geosocial media posts in democrat counties. In particular, the y-axis represents the individual waves of COVID-19 cases as introduced in Figure 2, while the x-axis denotes the number of days the posts time series was shifted into the future. The coloring of individual windows reflects the Pearson correlation between COVID-19 cases and the shifted posts time series. Furthermore, Figure 4 illustrates the corresponding COVID-19 cases, the post time series and the post time series shifted by the correlation maximizing temporal lag for each individual epidemic wave.

XSL•FO
RenderX

Figure 3. Depicting the Pearson correlation between COVID-19 cases and geosocial media post time series for each epidemiological wave when stepwise shifting the geosocial media post time series into the future for democrat counties.

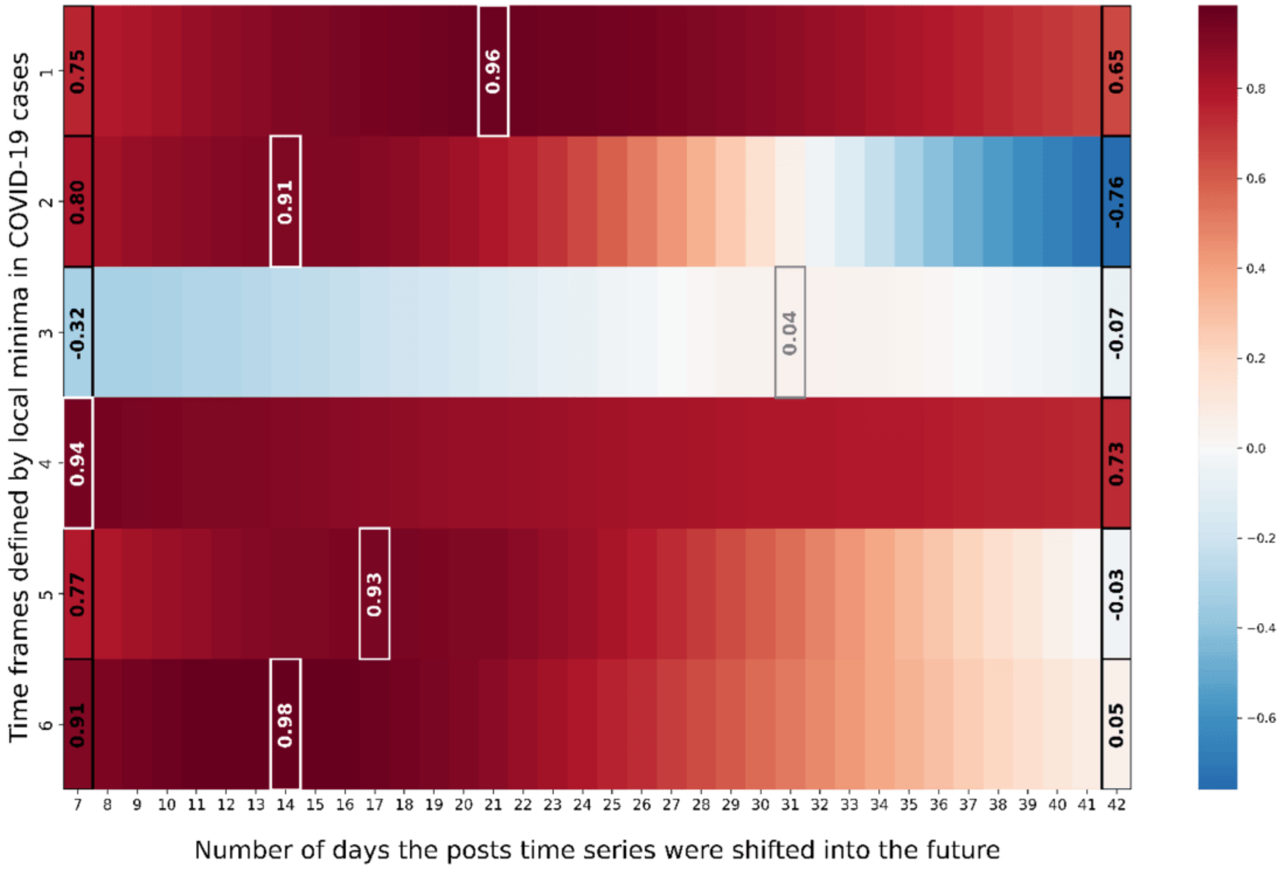
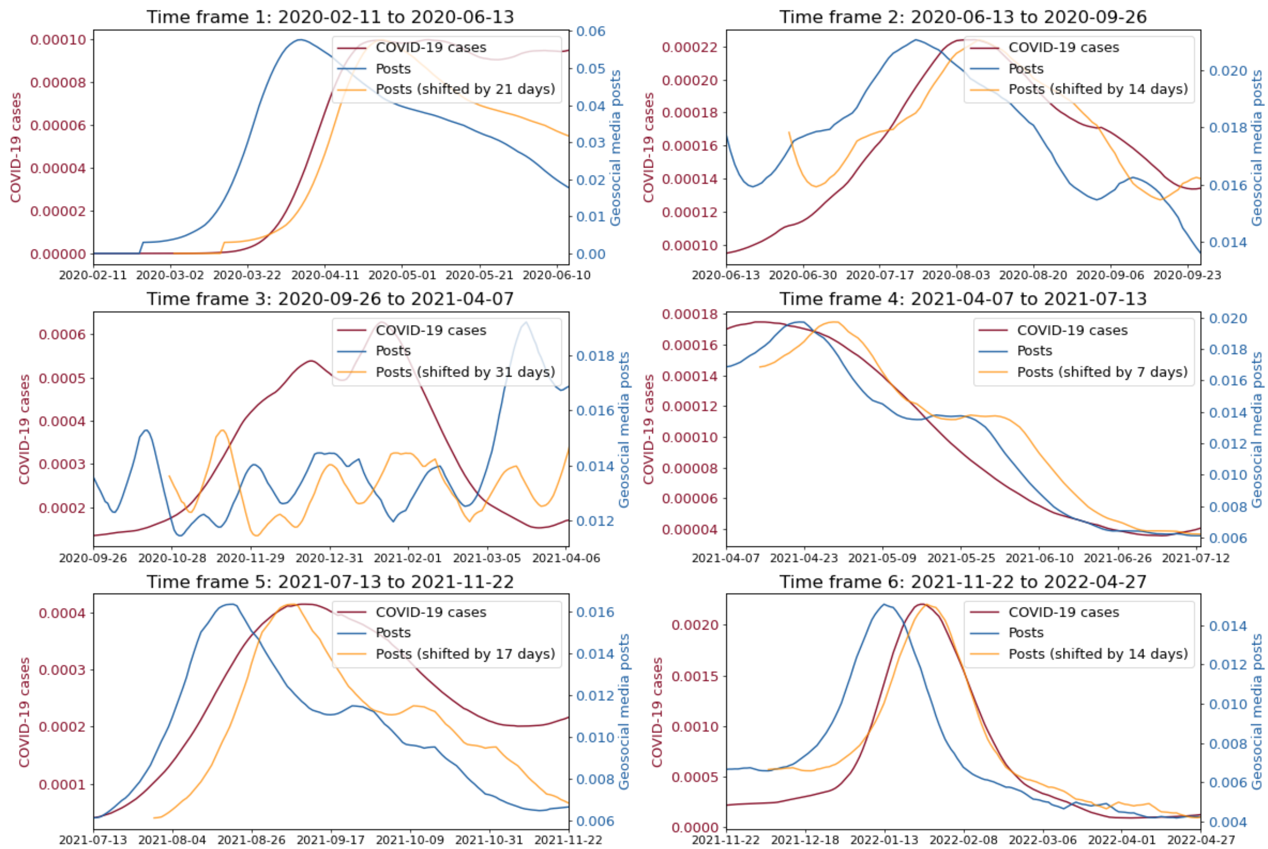


Figure 4. Depicting COVID-19 cases and the geosocial media posts time series and the shifted geosocial media posts time series across individual epidemic waves for democrat counties.



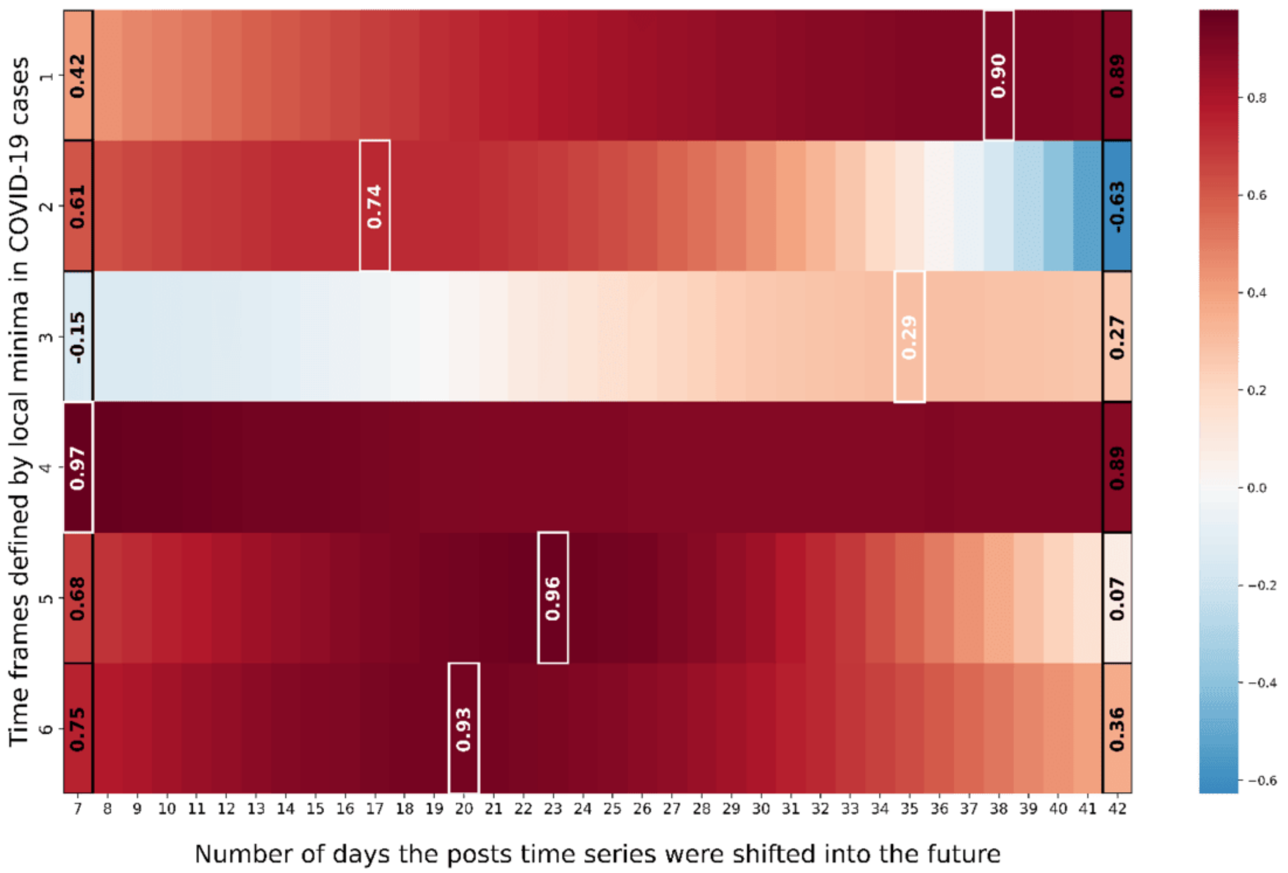
XSL•FO
RenderX

The results for democrat counties in Figure 3 indicate the highest Pearson correlations between posts and COVID-19 cases time series in 5 out of 6 epidemic waves, for a shift of 7 to 21 days (time frames 1, 2 and 4-6). For the same 5-time frames, the Pearson correlations ranged between 0.91 to 0.98. Furthermore, Figures 3 and 4 suggest that only for time frames 1, 2 and 4-6, geosocial media data exhibited actual early warning capabilities. For these time frames, signals in COVID-19 cases were clearly preceded by signals in X data, while for time frame 3 no clear early warning signal in geosocial media data was apparent. Nevertheless, in the beginning of the pandemic (time frames 1 and 2) geosocial media posts showcased a clear increase up to 21 (time frame 1) and 14 days (time frame 2) ahead increases in COVID-19 infections, with Pearson correlations of 0.96 and 0.91. In addition, the COVID-19 wave from mid of July 2021 to the end of November 2021 (time frame 5) was reflected in geosocial media posts up to 17 days earlier than an increase in COVID-19 cases, with a Pearson correlation of 0.93. Also, the Omicron wave (time frame 6) starting in mid of November 2021 [29] was accurately reflected 14 days in advance in the geosocial media time series (Pearson correlation of 0.98). Beyond that, Figure 4 clearly illustrates that the ratio of geosocial media posts related to COVID-19 decreased significantly over the course of the pandemic. Specifically, the percentage of relevant geosocial media posts gradually decreased from 5.7% at its peak in the first time frame, to 1.5% in the last time frame.

Republican Counties

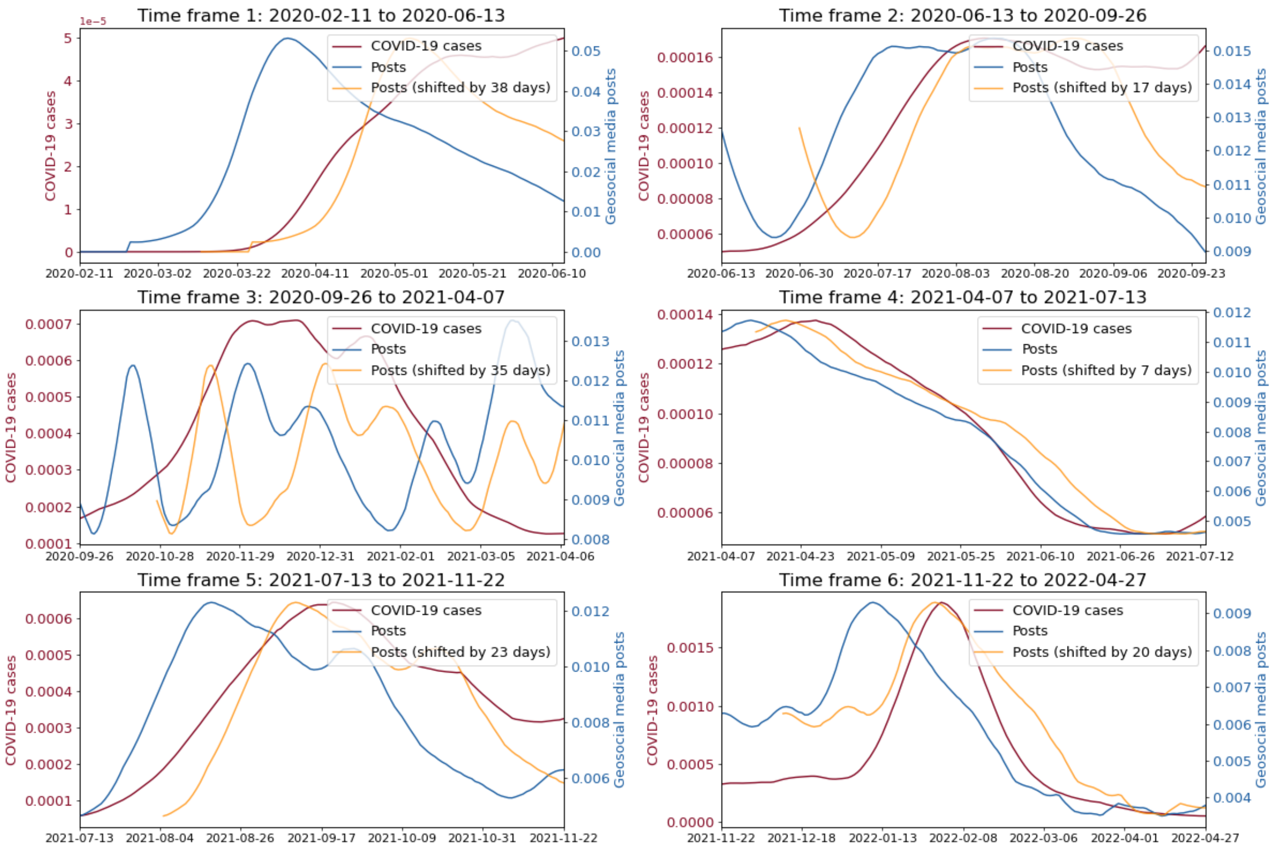
Figure 5 illustrates for the republican counties that in 5 out of 6 time frames the post time series exhibited the highest Pearson correlation with the COVID-19 cases 7 to 38 days ahead of time (time frames 1, 2, and 4-6). Furthermore, for these time frames the Pearson correlations between posts shifted 7 to 38 days into the future and COVID-19 cases were between 0.74 and 0.97. Furthermore, Figure 6 showcases that for republican counties, early warning signals in geosocial media posts could be observed for time frames 1, 2 and 4-6. Similarly to the democrat county cluster, the COVID-19 cases wave in time frame 3 was not captured in advance by the geosocial media time series. The fact that all time frames besides time frame 3, lend themselves for early warning is also consistent with the results for the democrat counties. Furthermore, it appears that in the republican counties, geosocial media data preceded COVID-19 cases time series a few days more in advance. On average over all 5 time frames for which we attest early warning capabilities (time frames 1, 2, and 4-6), the mean temporal lag in democrat counties is 14.6 days (average correlation 0.94) and for 21 days republican counties (average correlation 0.9). Furthermore, it appears that the ratio of relevant posts decreased over time for republican counties from roughly about 5.3% to 0.9%.

Figure 5. Depicting the Pearson correlation between COVID-19 cases and geosocial media post time series for each epidemiological wave when stepwise shifting the geosocial media post time series into the future for republican counties.



XSL•FO
RenderX

Figure 6. Depicting COVID-19 cases and the geosocial media posts time series and the shifted geosocial media posts time series across individual epidemic waves for republican counties.



Swing Counties

Figure 7 illustrates for swing counties that shifting the posts time series between 7 and 37 days into the future achieved the highest correlation for all time frames. Furthermore, for all time frames the maximal Pearson correlations between geosocial media posts and COVID-19 cases ranged between 0.52 and 0.96. Beyond that, Figure 8 shows that the time frames 1, 2 and 4-6 exhibited clear early warning signals in geosocial media data ahead increases in COVID-19 cases. Similarly to the republican and democrat counties, the COVID-19 wave in time frame 3 was not clearly captured in advance by geosocial media data. However, similar, to republican counties, Figure 8 showcases for swing counties that there actually existed a signal in geosocial media data which is in line with the COVID-19 data in time frame 3. Nevertheless, the actual early warning capabilities are still limited due to noise in the signal which coincides with the COVID-19 infection of former President Donald Trump. Overall, the posts time series preceded COVID-19 cases in swing counties across all time frames, excluding the third, on average by 24.2 days. Also, the intensity with which geosocial media data appears to precede COVID-19 waves clearly decreased for swing counties over the course of the pandemic (from 5.6% to 1.1%).

XSL•FO
RenderX

Figure 7. Depicting the Pearson correlation between COVID-19 cases and geosocial media post time series for each epidemiological wave when stepwise shifting the geosocial media post time series into the future for swing counties.

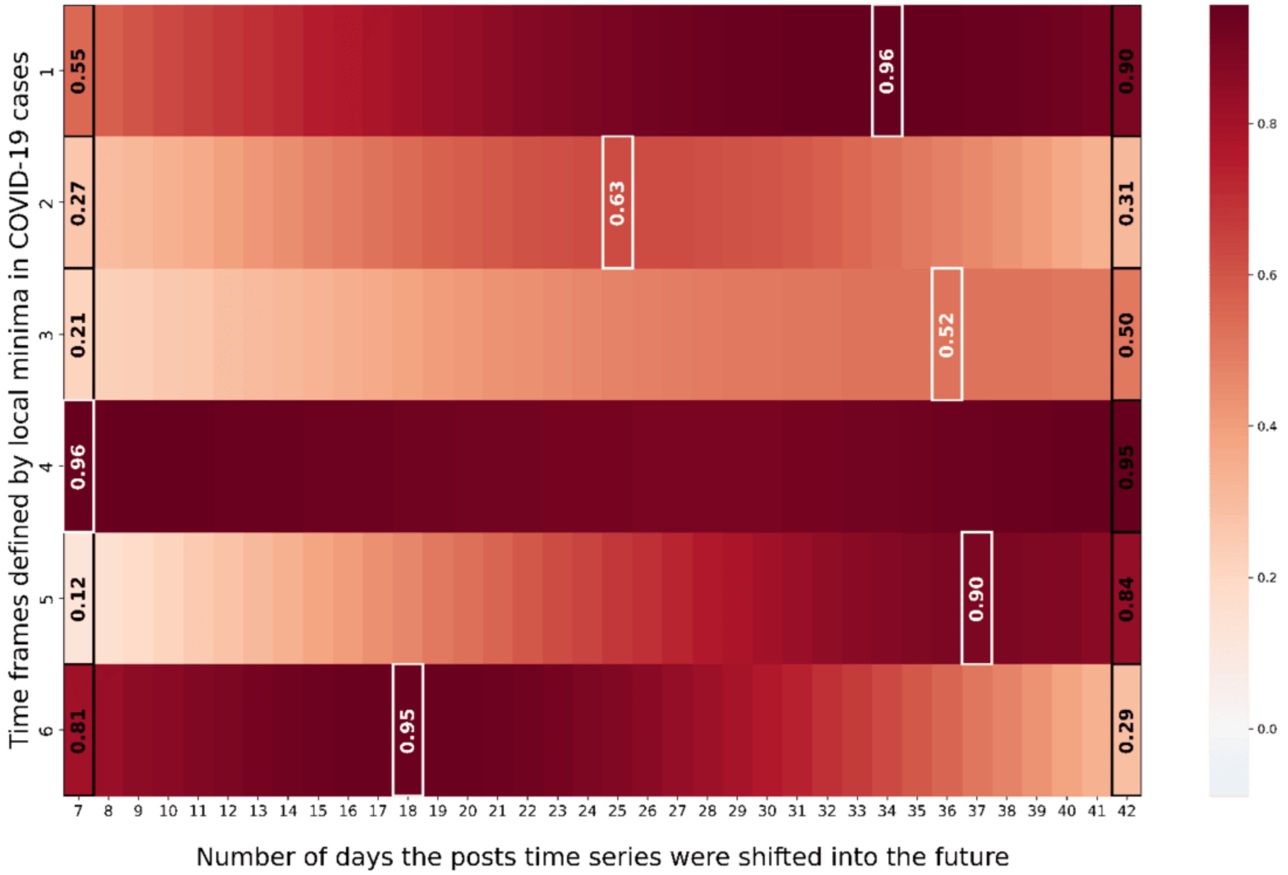
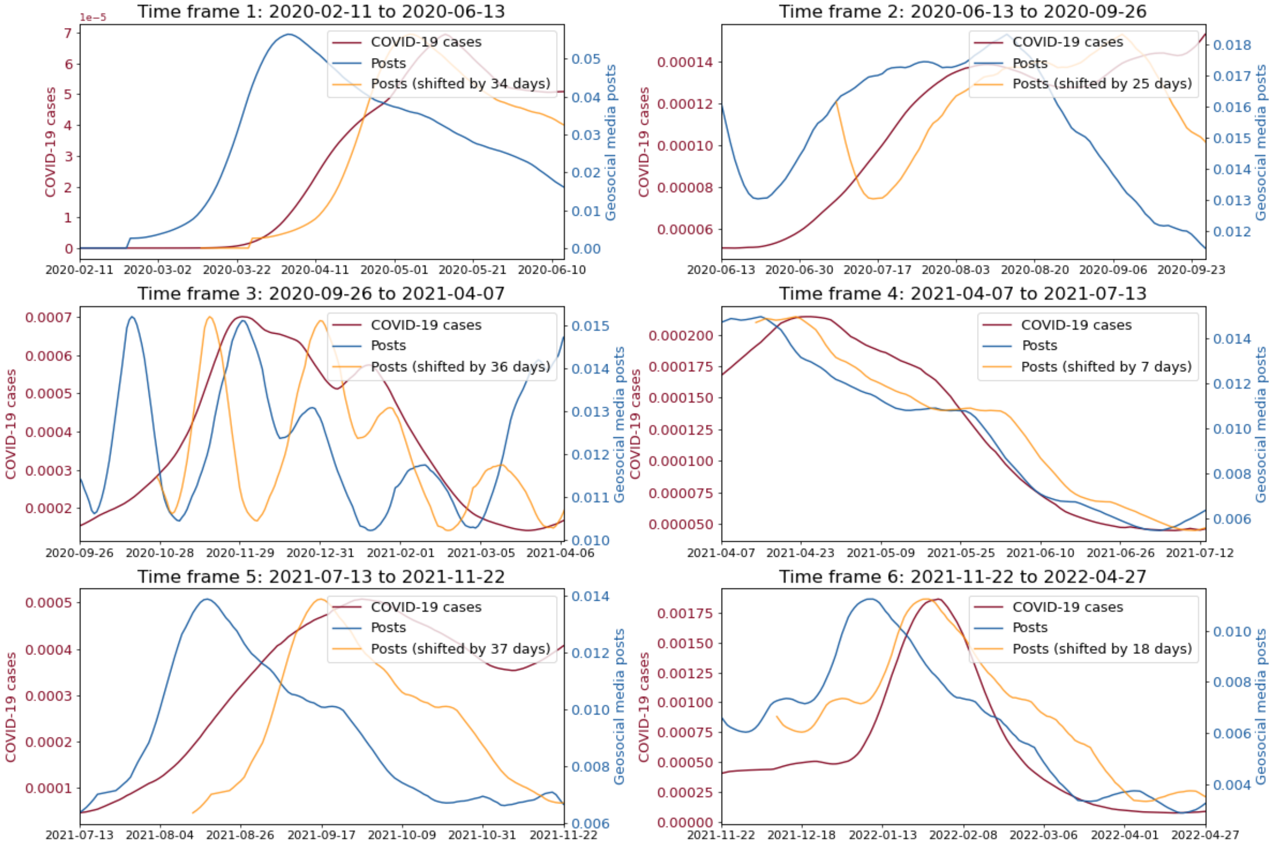


Figure 8. Depicting COVID-19 cases and the geosocial media posts time series and the shifted geosocial media posts time series across individual epidemic waves for swing counties.



XSL•FO
RenderX

Discussion

Principal Findings

The results of this study highlight how a deeper understanding of the relationship between COVID-19–related geosocial media data and confirmed COVID-19 cases, across politically distinct geographies, may help improve epidemiological early warning systems. Specifically, our analysis confirmed and expanded previous findings on the use of geosocial media posts as early indicators of disease activity [8-10,12]. However, we observed strong differences in the early warning capability of geosocial media data across different epidemiological waves. For example, geosocial media data were unable to reliably anticipate the third major COVID-19 wave, September 2020 to April 10, 2021 (time frame 3), across all political clusters. After significantly high COVID-19–related engagement on geosocial media in the first wave, it appears that the geosocial media signal lost some of its sensitivity in the third wave. The only event clearly detectable in COVID-19–related geosocial media posts in the third time frame is the COVID-19 infection of the former President Donald Trump in October 2020. The significance of this event might have reduced the sensitivity of the geosocial media users toward an increase in COVID-19 symptoms and infections. The reaction signal to this event was particularly visible in the republican and swing county clusters, while the democrat counties only registered a minor increase in geosocial media posts coinciding with the COVID-19 infection of President Trump. This further highlights how susceptible geosocial media data can be to politically charged trending topics and how these topics of interest might differ across political clusters. This is also in line with previous findings that the topics geosocial media users engage with and the language they use can differ depending on political beliefs [17-19]. Thus, we hypothesize that it might be key to identify different sets of keywords related to political beliefs and resulting trending topics, to capture geosocial media signals more accurately across political clusters. Therefore, future research should explore the influence of different geosocial media topics on early warning capabilities across political clusters and how such topics might change over time.

Furthermore, the findings of this study illustrate differences in the early warning capabilities of geosocial media posts for COVID-19 cases across counties with diverging political beliefs. This is particularly true for the number of days that geosocial media posts precede COVID-19 cases (temporal lag) and the Pearson correlation between these 2 time series for republican and democrat counties. For instance, geosocial media posts appear to anticipate COVID-19 cases in republican counties (21 days) on average 6.4 days earlier than in democrat counties (14.6 days). This difference in temporal lag might partly be caused by varying population densities between democrat and republican counties. In less densely populated republican counties [36], infection transmission might be slower [37], which could lead to a higher temporal lag between the onset of COVID-19 symptoms being observed and shared on geosocial media, to the eventual peak of infections in that region. However, it remains beyond the scope of this study to substantiate the actual underlying mechanisms which might cause these observed differences in early warning capability between political clusters. Despite that, the results of this study clearly emphasize the need to account for political beliefs in epidemiological early warning systems using geosocial media data. Nevertheless, the precise methodology to integrate political beliefs into real-time geosocial media-based early warning models remains the subject of future research.

The psychological effects of public health measures, such as lockdowns, might offer another explanation for the observed differences in early warning capabilities of geosocial media data across political clusters. These effects may be connected to the fact that public health measures were implemented and suspended at different points in time across political administrative areas. In this regard, Pettersen et al [38] associated more stringent public health and quarantine measures with increased mental distress in adults in Norway. Similarly, Ferwana and Varshney [39] observed a significant increase in visits to mental health facilities during the 2020 lockdown periods in the United States. While Ashokkumar and Pennebaker [14] even reported drops in analytical thinking and shifts in the emotional states of Reddit users coinciding with the start of lockdowns. Hence, it might be the case that the varying timing of public health measures across political regions caused various psychological effects, manifesting in changes of geosocial media behavior. However, our findings do not sufficiently verify this hypothesis. Although numerous studies have explored the psychological effects of public health measures, future research should focus on how these effects might influence the early warning capabilities of geosocial media data across the political spectrum.

In addition, we also found a clear decrease in the number of days with which geosocial media posts preceded COVID-19 cases and in the strength of the geosocial media post signal over time. Interestingly, yet to be explained, the decrease in temporal lag appears to be less pronounced in republican and swing counties. Nonetheless, this overall phenomenon might be caused by some sort of geosocial media and emotional COVID-19 fatigue. The association between self-reported depression symptoms and geosocial media usage [40], alongside potential factors contributing to social media fatigue [41-43] have already been explored in the context of the COVID-19 pandemic. For instance, recent findings by Li et al [43] indicate a direct relationship between social media overload during the COVID-19 pandemic and increased anxiety. Similarly, Sun and Lee [44] observe that COVID-19 information overload on social media directly contributes to fatigue toward pandemic related messages. Nevertheless, it remains beyond the scope of this study to substantiate whether the observed decreasing strength of the geosocial media post signal and temporal lag are robust and attributable to some form of geosocial media or COVID-19 fatigue. However, based on our observations, we advise caution, as the epidemiological early warning capabilities of geosocial media appear to change over time and depending on prevailing political beliefs. In this regard, it remains the task of future research to develop geosocial media-based early warning approaches, which can account for decreasing signal strength over time.

XSL•FO
RenderX

Furthermore, Howard et al [21] observed varying levels of misinformation and thus topics of interest, across states with different political beliefs. Interestingly, they found the highest rates of misinformation occurring in swing states. This is particularly noteworthy, as we found geosocial media data to be highly capable for early epidemiological warning in swing counties. Specifically, the average temporal lag of 24.2 days over all time frames in which we observed the highest early warning capabilities for swing counties, while mostly achieving high correlations (average correlation over all time frames with early warning capabilities 0.88). Thus, concluding from Howard et al [21] and our findings, it appears that it might not be the quality or factual correctness of the shared information on geosocial media that warrants its value for early warning purposes. Nevertheless, future research needs to further validate these findings in the context of different countries and their political ramifications as they might influence the relevance of shared information quality and factual correctness for epidemiological early warning capability.

Data and Methods

We acknowledge that using a simple linear correlation measure might not always reflect the actual similarity between time series accurately. However, in preliminary analysis we also used different nonlinear correlation measures, which yielded only neglectable differences in the actual results. In addition, we also tested more advanced time series matching algorithms such as dynamic time warping [45], the Fréchet distance [46], or mutual information [47]. Nevertheless, neither nonlinear correlation measures nor more advanced comparison algorithms outperformed conventional linear correlation measures for most of our analyses. We evaluated the performance of these different methods in their ability to match the peaks and onsets of geosocial media signals and COVID-19 cases. Nonetheless, we acknowledge that the alignment of peaks and onsets is not always feasible, as the time it takes from the onset to the peak may vary between geosocial media signals and COVID-19 cases. As a result, for some epidemic waves the determined temporal lag might not reflect the actual real-world early warning capabilities of geosocial media data. Despite that, our main objective in this study was not to assess the exact temporal lag and correlations, but rather to provide an algorithmic way to compare the early warning capabilities of geosocial media data across political clusters.

In addition, there is a need to discuss the definition of epidemiological waves based on COVID-19 cases of the entire United States as one might argue that this decision might potentially have caused the observed variations in the number of days and the correlation between the geosocial media and the COVID-19 cases time series. The reason for this is that the COVID-19 waves can have different starting points and intensities across states [48] and as our results show also across political clusters (Figures 4, 6, and 8). Therefore, it might appear reasonable to assume that variation in the starting points and intensities caused the underlying observed differences in temporal lag and correlation between geosocial media posts and COVID-19 cases across political clusters. However, upon testing this hypothesis by defining COVID-19 waves individually for each political cluster, the fundamental results of our study remained unchanged. Although minor discrepancies were present in the temporal lag (primarily ranging from 1-2 days) and the correlations between COVID-19 cases and geosocial media posts, their differences persisted across political clusters and time frames in a similar manner. For example, republican counties still exhibited on average a higher temporal lag than democratic counties and the decrease in geosocial media signals was also still prevalent across political clusters.

In addition, it is important to consider the choice of keywords used for our analysis, as they strongly influence the observed results. One might argue that some keywords relevant to the discourse related to the COVID-19 pandemic were left out. However, in this analysis we mainly focused on gathering less polarized keywords, topics, and hashtags. The reason for this is that certain words, topics and hashtags were predominantly used by 1 political faction [17,18], which might indeed introduce bias into the final comparison between early warning capabilities across political clusters from the start. Concretely, keywords used predominantly in republican counties and less in democrat counties might directly influence differences in early warning capability across political clusters. Therefore, we decided to use a condensed set of keywords, which was to our knowledge mostly not inherently politically charged or biased.

Furthermore, we acknowledge that some keywords which we used in the semantic filtering process of the geosocial media posts, might not be only COVID-19 specific. However, we argue that for most words there exists a baseline signal of how often these words are being used. Therefore, our underlying assumption is that a real-world epidemiological event causes a significant spike in the usage of relevant keywords. Indeed, our results confirmed this assumption. We observed a baseline fluctuation in geosocial media posts and significant spikes in filtered posts, which in most cases preceded COVID-19 cases.

We also tried to improve the semantic filtering by leveraging machine learning approaches such as BERTopic or Latent Dirichlet Allocation [49,50]. However, due to performance issues with our large dataset (600+ GB) and based on the insufficient results for subsample experiments, we decided to stick to traditional keyword filtering. Nevertheless, in future work large language models [51] might be a possibility to improve the process of identifying relevant geosocial media posts.

Limitations

The main limitation of this study stems from its retrospective nature. Our findings, while insightful for the past pandemic, may not be directly transferable to future epidemiological events. This limitation is partly due to the unpredictable nature of political polarization. Specifically, it is inherently difficult to predict whether a topic will become politically charged and, as a result, be discussed differently on social media across geographies with diverging political beliefs. In addition, social media behavior itself is influenced by various dynamic factors, for instance platform algorithms [52] or changing governance structures, which affect public engagement [53], all of which may differ significantly across social media platforms, future epidemiological events, and national borders. Although our study revealed differences in the epidemiological early warning

XSL•FO
RenderX

capabilities of geosocial media data across US county-level political clusters, these results should be treated with caution when considering future-use cases.

Conclusion

Our results confirmed the findings of previous research [9,10,12], demonstrating that geosocial media data can improve epidemiological early warning for consecutive waves of COVID-19 cases. In addition, we expand the existing literature by showing that the early warning capabilities of geosocial media data vary across US county clusters with differing political beliefs. For instance, geosocial media posts in republican counties (21 days) tend to precede increases in COVID-19 cases on average about 6.4 days earlier than in democrat counties (14.6 days). We hypothesize that this discrepancy in temporal lag between the geosocial media signal and the COVID-19 cases may stem from differences in the adoption of public health measures or population density variations across regions. In addition, we observed that the early warning capabilities of geosocial media data can be mitigated due to its susceptibility to a shift in trending topics and a decrease in signal strength over time.

Based on our findings, we would recommend that policy makers and researchers enhance and further investigate real-time geosocial media monitoring capabilities to improve epidemiological early warning systems. In addition, our findings suggest that it could be particularly beneficial for such systems to account for political beliefs prevalent across finer spatial scales such as county-level, given their potential to impact the early warning capabilities of geosocial media signals. Furthermore, since our results clearly highlight the value of geosocial media data for epidemiological early warning, we strongly encourage social media companies to grant researchers access to their data. Furthermore, future research should examine the early warning capabilities of different geosocial media topics specific to regional political beliefs and assess the transferability of our findings to other countries with different political environments. Furthermore, investigating the role of political communication strategies and potential improvements to social media algorithms to mitigate political polarization could enhance our understanding of how geosocial media data can be leveraged for future epidemiological events.

Acknowledgments

This research was funded in part by the Austrian Science Fund (FWF) Grant-DOI: 10.55776/I5117. For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

MS has been funded (in part) by contract 200-2016-91779 and cooperative agreement CDC-RFA-FT-23-0069 with the Centers for Disease Control and Prevention (CDC). The findings, conclusions, and views expressed are those of the author(s) and do not necessarily represent the official position of the CDC. MS was also partially supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM130668. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

WG has been sponsored by NSF award #1841403.

Conflicts of Interest

MS has received institutional research funds from the Johnson and Johnson foundation, from Janssen global public health, and Pfizer. All other authors declare no conflicts of interest.

References

1. Ciotti M, Ciccozzi M, Terrinoni A, Jiang WC, Wang CB, Bernardini S. The COVID-19 pandemic. Crit Rev Clin Lab Sci 2020;57(6):365-388. [doi: 10.1080/10408363.2020.1783198] [Medline: 32645276]
2. GBD 2021 Demographics Collaborators. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950-2021, and the impact of the COVID-19 pandemic: a comprehensive demographic analysis for the global burden of disease study 2021. Lancet 2024;403(10440):1989-2056 [FREE Full text] [doi: 10.1016/S0140-6736(24)00476-8] [Medline: 38484753]
3. Gahlot P, Alley KD, Arora S, Das S, Nag A, Tyagi VK. Wastewater surveillance could serve as a pandemic early warning system for COVID-19 and beyond. WIREs Water 2023;10(4):e1650. [doi: 10.1002/wat2.1650]
4. Kamalrathne T, Amaratunga D, Haigh R, Kodituwakku L. Need for effective detection and early warnings for epidemic and pandemic preparedness planning in the context of multi-hazards: Lessons from the COVID-19 pandemic. Int J Disaster Risk Reduct 2023;92:103724 [FREE Full text] [doi: 10.1016/j.ijdrr.2023.103724] [Medline: 37197332]
5. MacIntyre CR, Chen X, Kunasekaran M, Quigley A, Lim S, Stone H, et al. Artificial intelligence in public health: the potential of epidemic early warning systems. J Int Med Res 2023;51(3):3000605231159335 [FREE Full text] [doi: 10.1177/03000605231159335] [Medline: 36967669]
6. Al-Surimi K, Khalifa M, Bahkali S, EL-Metwally A, Househ M. The potential of social media and internet-based data in preventing and fighting infectious diseases: from internet to Twitter. In: Rezza G, Ippolito G, editors. Emerging and Re-Emerging Viral Infections: Advances in Microbiology, Infectious Diseases and Public Health Volume 6. Advances in Experimental Medicine and Biology. Midtown Manhattan, New York City: Springer International Publishing; 2017:131-139.

XSL•FO
RenderX

7. Albinati J, Meira W, Pappa G, Teixeira M, Marques-Toledo C. Enhancement of epidemiological models for dengue fever based on twitter data. : Association for Computing Machinery; 2017 Presented at: DH '17: International Conference on Digital Health; 2017 July 2-5; London United Kingdom p. 109-118. [doi: 10.1145/3079452.3079464]
8. Butt ZA. Early warning for emerging infectious disease outbreaks: digital disease surveillance for public health preparedness and response. In: Accelerating Strategic Changes for Digital Transformation in the Healthcare Industry. Amsterdam: Elsevier; 2023:309-320.
9. Comito C. How COVID-19 Information Spread in U.S.? The Role of Twitter as Early Indicator of Epidemics. IEEE Trans. Serv. Comput 2022;15(3):1193-1205. [doi: 10.1109/tsc.2021.3091281]
10. Kogan NE, Clemente L, Liautaud P, Kaashoek J, Link NB, Nguyen AT, et al. An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time. Sci Adv 2021;7(10):eabd6989 [FREE Full text] [doi: 10.1126/sciadv.abd6989] [Medline: 33674304]
11. Zhang Y, Chen K, Weng Y, Chen Z, Zhang J, Hubbard R. An intelligent early warning system of analyzing twitter data using machine learning on COVID-19 surveillance in the US. Expert Syst Appl 2022;198:116882 [FREE Full text] [doi: 10.1016/j.eswa.2022.116882] [Medline: 35308584]
12. Stolerman LM, Clemente L, Poirier C, Parag KV, Majumder A, Masyn S, et al. Using digital traces to build prospective and real-time county-level early warning systems to anticipate COVID-19 outbreaks in the United States. Sci Adv 2023;9(3):eabq0199 [FREE Full text] [doi: 10.1126/sciadv.abq0199] [Medline: 36652520]
13. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021;3(3):e175-e194 [FREE Full text] [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
14. Ashokkumar A, Pennebaker JW. Social media conversations reveal large psychological shifts caused by COVID-19's onset across U.S. cities. Sci Adv 2021;7(39):eabg7843 [FREE Full text] [doi: 10.1126/sciadv.abg7843] [Medline: 34550738]
15. Das Swain V, Ye J, Ramesh SK, Mondal A, Abowd GD, De Choudhury M. Leveraging social media to predict COVID-19-induced disruptions to mental well-being among university students: modeling study. JMIR Form Res 2024;8:e52316 [FREE Full text] [doi: 10.2196/52316] [Medline: 38916951]
16. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on facebook and twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. J Med Internet Res 2021;23(4):e26627 [FREE Full text] [doi: 10.2196/26627] [Medline: 33724919]
17. Guntuku SC, Purtle J, Meisel ZF, Merchant RM, Agarwal A. Partisan differences in twitter language among US legislators during the COVID-19 pandemic: cross-sectional study. J Med Internet Res 2021;23(6):e27300 [FREE Full text] [doi: 10.2196/27300] [Medline: 33939620]
18. Jing E, Ahn YY. Characterizing partisan political narrative frameworks about COVID-19 on twitter. EPJ Data Sci 2021;10(1):53 [FREE Full text] [doi: 10.1140/epjds/s13688-021-00308-4] [Medline: 34745825]
19. Sylwester K, Purver M. Twitter language use reflects psychological differences between democrats and republicans. PLoS One 2015;10(9):e0137422 [FREE Full text] [doi: 10.1371/journal.pone.0137422] [Medline: 26375581]
20. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. Vaccine 2018;36(25):3606-3612. [doi: 10.1016/j.vaccine.2018.05.040] [Medline: 29773322]
21. Howard P, Kollanyi B, Bradshaw S, Neudert LM. Social media, news and political information during the US election: was polarizing content concentrated in swing States? arXiv:1802.03573 2018. [doi: 10.48550/arXiv.1802.03573]
22. Lyu H, Wang J, Wu W, Duong V, Zhang X, Dye TD, et al. Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination. Intell Med 2022;2(1):1-12 [FREE Full text] [doi: 10.1016/j.imed.2021.08.001] [Medline: 34457371]
23. Kahane LH. Politicizing the mask: political, economic and demographic factors affecting mask wearing behavior in the USA. East Econ J 2021;47(2):163-183 [FREE Full text] [doi: 10.1057/s41302-020-00186-0] [Medline: 33424048]
24. Allcott H, Boxell L, Conway J, Gentzkow M, Thaler M, Yang D. Polarization and public health: partisan differences in social distancing during the coronavirus pandemic. Journal of Public Economics 2020;191:104254. [doi: 10.1016/j.jpubeco.2020.104254] [Medline: 32836504]
25. Kaashoek J, Testa C, Chen JT, Stolerman LM, Krieger N, Hanage WP, et al. The evolving roles of US political partisanship and social vulnerability in the COVID-19 pandemic from February 2020-February 2021. PLOS Glob Public Health 2022;2(12):e0000557 [FREE Full text] [doi: 10.1371/journal.pgph.0000557] [Medline: 36962752]
26. Hart PS, Chinn S, Soroka S. Politicization and polarization in COVID-19 news coverage. Sci Commun 2020;42(5):679-697 [FREE Full text] [doi: 10.1177/1075547020950735] [Medline: 38602988]
27. Neelon B, Mutiso F, Mueller NT, Pearce JL, Benjamin-Neelon SE. Associations between governor political affiliation and COVID-19 cases, deaths, and testing in the U.S. Am J Prev Med 2021;61(1):115-119 [FREE Full text] [doi: 10.1016/j.amepre.2021.01.034] [Medline: 33775513]
28. Gadarian SK, Goodman SW, Pepinsky TB. Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic. PLoS One 2021;16(4):e0249596 [FREE Full text] [doi: 10.1371/journal.pone.0249596] [Medline: 33826646]

XSL•FO
RenderX

29. Das S, Samanta S, Banerjee J, Pal A, Giri B, Kar SS, et al. Is Omicron the end of pandemic or start of a new innings? Travel Med Infect Dis 2022;48:102332 [FREE Full text] [doi: 10.1016/j.tmaid.2022.102332] [Medline: 35472451]
30. US COVID-19 cases and deaths by state.: USAFacts; 2020. URL: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ [accessed 2025-01-02]
31. County presidential election returns 2000-2020. MIT Election Data And Science Lab. 2018. URL: https://search.dataone.org/view/sha256:6c387220de1dcdaaf94c2e7ca85474badf51d3b7047902823946c8769acde2d5 [accessed 2024-12-31]
32. Schultz E, Jacob R, Melcher JP, Fried A. Presidential Swing States. Lanham: Lexington Books; 2018.
33. Kriston L. A statistical definition of epidemic waves. Epidemiologia (Basel) 2023;4(3):267-275 [FREE Full text] [doi: 10.3390/epidemiologia4030027] [Medline: 37489498]
34. Zhang SX, Arroyo Marioli F, Gao R, Wang S. A second wave? What do people mean by COVID waves? - A working definition of epidemic waves. Risk Manag Healthc Policy 2021;14:3775-3782 [FREE Full text] [doi: 10.2147/RMHP.S326051] [Medline: 34548826]
35. Ayala A, Villalobos Dintrans P, Elorrieta F, Castillo C, Vargas C, Maddaleno M. Identification of COVID-19 waves: considerations for research and policy. Int J Environ Res Public Health 2021;18(21):11058 [FREE Full text] [doi: 10.3390/ijerph182111058] [Medline: 34769577]
36. Gimpel JG, Lovin N, Moy B, Reeves A. The urban–rural gulf in American political behavior. Polit Behav 2020;42(4):1343-1368. [doi: 10.1007/s11109-020-09601-w]
37. Sy KTL, White LF, Nichols BE. Population density and basic reproductive number of COVID-19 across United States counties. PLoS One 2021;16(4):e0249271 [FREE Full text] [doi: 10.1371/journal.pone.0249271] [Medline: 33882054]
38. Pettersen JH, Hannigan LJ, Gustavson K, Lund IO, Pearson RM, Jensen P, et al. COVID-19 pandemic quarantines and mental health among adolescents in Norway. JAMA Netw Open 2024;7(7):e2422189. [doi: 10.1001/jamanetworkopen.2024.22189] [Medline: 38995642]
39. Ferwana I, Varshney LR. The impact of COVID-19 lockdowns on mental health patient populations in the United States. Sci Rep 2024;14(1):5689 [FREE Full text] [doi: 10.1038/s41598-024-55879-9] [Medline: 38454064]
40. Perlis RH, Green J, Simonson M, Ognyanova K, Santillana M, Lin J, et al. Association between social media use and self-reported symptoms of depression in US adults. JAMA Netw Open 2021;4(11):e2136113 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.36113] [Medline: 34812844]
41. Islam AKMN, Laato S, Talukder S, Sutinen E. Misinformation sharing and social media fatigue during COVID-19: an affordance and cognitive load perspective. Technol Forecast Soc Change 2020;159:120201 [FREE Full text] [doi: 10.1016/j.techfore.2020.120201] [Medline: 32834137]
42. Xiao H, Zhang Z, Zhang L. An investigation on information quality, media richness, and social media fatigue during the disruptions of COVID-19 pandemic. Curr Psychol 2023;42(3):2488-2499 [FREE Full text] [doi: 10.1007/s12144-021-02253-x] [Medline: 34511861]
43. Li K, Jiang S, Yan X, Li J. Mechanism study of social media overload on health self-efficacy and anxiety. Heliyon 2024;10(1):e23326 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e23326] [Medline: 38163164]
44. Sun J, Lee SK. "No more COVID-19 messages via social media, please": the mediating role of COVID-19 message fatigue between information overload, message avoidance, and behavioral intention. Curr Psychol 2023;2:1-15 [FREE Full text] [doi: 10.1007/s12144-023-04726-7] [Medline: 37359620]
45. Wang K, Gasser T. Alignment of curves by dynamic time warping. Ann. Statist 1997;25(3):1251-1276. [doi: 10.1214/aos/1069362747]
46. Cleasby IR, Wakefield ED, Morrissey BJ, Bodey TW, Votier SC, Bearhop S, et al. Using time-series similarity measures to compare animal movement trajectories in ecology. Behav Ecol Sociobiol 2019;73(11). [doi: 10.1007/s00265-019-2761-1]
47. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E Stat Nonlin Soft Matter Phys 2004;69(6 Pt 2):066138 [FREE Full text] [doi: 10.1103/PhysRevE.69.066138] [Medline: 15244698]
48. Kumar A, Kalra Y, Guan WW, Tibrewal V, Batta R, Chen A. COVID-19 impact on excess deaths of various causes in the United States. Annals of GIS 2021;28(4):457-478. [doi: 10.1080/19475683.2021.1982001]
49. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3(null):993-1022. [doi: 10.7551/mitpress/1120.003.0082]
50. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv:2203.05794 2022 [FREE Full text] [doi: 10.1002/asi.21166]
51. Pham CM, Hoyle A, Sun S, Iyyer M. TopicGPT: a prompt-based topic modeling framework. arXiv:2311.01449 2023. [doi: 10.18653/v1/2024.naacl-long.164]
52. Bellina A, Castellano C, Pineau P, Iannelli G, De Marzo G. Effect of collaborative-filtering-based recommendation algorithms on opinion polarization. Phys Rev E 2023;108(5-1):054304. [doi: 10.1103/PhysRevE.108.054304] [Medline: 38115540]
53. Schmidt S, Zorenböhmer C, Arifi D, Resch B. Polarity-based sentiment analysis of georeferenced tweets related to the 2022 Twitter acquisition. Information 2023;14(2):71. [doi: 10.3390/info14020071]

XSL•FO
RenderX

Abbreviations

API: application programming interface
WHO: World Health Organization

Edited by T Mackey; submitted 19.03.24; peer-reviewed by J Abbas, L Stolerman; comments to author 26.09.24; revised version received 17.11.24; accepted 23.11.24; published 30.01.25.

Please cite as:
Arifi D, Resch B, Santillana M, Guan WW, Knoblauch S, Lautenbach S, Jaenisch T, Morales I, Havas C
Geosocial Media's Early Warning Capabilities Across US County-Level Political Clusters: Observational Study
JMIR Infodemiology 2025;5:e58539
URL: https://infodemiology.jmir.org/2025/1/e58539
doi:10.2196/58539
PMID:

©Dorian Arifi, Bernd Resch, Mauricio Santillana, Weihe Wendy Guan, Steffen Knoblauch, Sven Lautenbach, Thomas Jaenisch, Ivonne Morales, Clemens Havas. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 30.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mental Health and Coping Strategies of Health Communicators Who Faced Online Abuse During the COVID-19 Pandemic: Mixed Methods Study

Lisa Wight[1], MPH; Chris Tenove[1], PhD; Saima Hirani[2], MScN, PhD; Heidi Tworek[1], PhD

[1]Centre for the Study of Democratic Institutions, School of Public Policy and Global Affairs, University of British Columbia, Vancouver, BC, Canada
[2]School of Nursing, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Chris Tenove, PhD
Centre for the Study of Democratic Institutions
School of Public Policy and Global Affairs
University of British Columbia
C.K. Choi Building, 251 – 1855 West Mall
Vancouver, BC, V6T 1Z2,
Canada
Phone: 1 6048223944
Email: christopher.tenove@ubc.ca

Abstract

Background: During the COVID-19 pandemic, health experts used social media platforms to share information and advocate for policies. Many of them faced online abuse, which some reported took a toll on their mental health and well-being. Variation in the impacts of online abuse on mental health, well-being, and professional efficacy suggest that health communicators may differ in their coping strategies and ultimately their resilience to such abuse.

Objective: We aimed to explore the impacts of online abuse on health communicators' mental health and well-being as well as their emotion- and problem-focused coping strategies.

Methods: We recruited health communicators (public health officials, medical practitioners, and university-based researchers) in Canada who engaged in professional online communication during the COVID-19 pandemic. In phase 1, semistructured interviews were conducted with 35 health communicators. In phase 2, online questionnaires were completed by 34 individuals before participating in workshops. Purposive recruitment resulted in significant inclusion of those who self-identified as racialized or women. Interview and workshop data were subjected to inductive and deductive coding techniques to generate themes. Descriptive statistics were calculated for selected questionnaire questions.

Results: In total, 94% (33/35) of interviewees and 82% (28/34) of questionnaire respondents reported experiencing online abuse during the study period (2020-2022). Most health communicators mentioned facing an emotional and psychological toll, including symptoms of depression and anxiety. Racialized and women health communicators faced abuse that emphasized their ethnicity, gender identity, and physical appearance. Health communicators' most common emotion-focused coping strategies were withdrawing from social media platforms, avoiding social media platforms altogether, and accepting online abuse as unavoidable. Common problem-focused coping strategies included blocking or unfriending hostile accounts, changing online behavior, formal help-seeking, and seeking peer support. Due to the impacts of online abuse on participants' mental health and well-being, 41% (14/34) of the questionnaire respondents seriously contemplated quitting health communication, while 53% (18/34) reduced or suspended their online presence. Our findings suggest that health communicators who used problem-focused coping strategies were more likely to remain active online, demonstrating significant professional resilience.

Conclusions: Although health communicators in our study implemented various emotion- and problem-focused coping strategies, they still faced challenges in dealing with the impacts of online abuse. Our findings reveal the limitations of individual coping strategies, suggesting the need for effective formal organizational policies to support those who receive online abuse and to sanction those who perpetrate it. Organizational policies could improve long-term outcomes for health communicators' mental health and well-being by mitigating online abuse and supporting its targets. Such policies would bolster professional resilience, ensuring that important health information can still reach the public and is not silenced by online abuse. More research is needed to determine whether gender, race, or other factors shape coping strategies and their effectiveness.

XSL•FO
RenderX

(JMIR Infodemiology 2025;5:e68483) doi:10.2196/68483

KEYWORDS

mental health; online harassment; online abuse; coping strategies; resilience; social media; online advocacy; public health communication; health communication

Introduction

Background

From the start of the COVID-19 pandemic, public health officials, medical practitioners, university-based researchers, health journalists, and other health experts played a crucial role in shaping public opinions and behaviors. To inform publics and counter poor-quality information, many health experts increased their use of social media platforms: frontline health care workers creating TikTok videos [1], medical professionals countering misinformation on Twitter [2], and physicians and researchers posting on Facebook, Instagram, and Twitter [3,4]. These health communicators could be considered an emergent community of practice, meaning they encountered many similar opportunities and challenges of engaging audiences through social media platforms.

Many health communicators were exposed to online harassment and abuse, ranging from trivial criticisms to sexual harassment and violent threats [5,6]. This abuse was faced by those communicating as individuals or on behalf of institutions. Health communicators were typically unprepared for the abuse that often follows online advocacy [7], which was exacerbated by a lack of existing institutional protections and supports [7,8].

Exposure to Online Abuse

In a survey conducted from February 2019 to March 2019, 23% of 464 physicians in the United States reported being personally attacked on social media, primarily for advocacy on topics such as vaccines, gun control, and abortion [9]. Research suggests that the online harassment experienced by health communicators worsened during the COVID-19 pandemic [10,11]. In a 2022 survey of 359 physicians, biomedical scientists, and trainees in the United States, 228 (64%) reported harassment on social media related to comments they had made about the COVID-19 pandemic [12]. Similarly, more than two-thirds of 321 scientists responding to a *Nature* survey in 2021, predominantly located in the United States, the United Kingdom, and Germany, reported negative experiences because of their media interviews or social media comments regarding COVID-19 [10]. In total, 22% of these respondents had received threats of physical or sexual violence, and 15% had received death threats [10]. Threats of violence illustrate that online abuse is not merely confined to the internet [13]. Escalating violence against health care workers in Canada during the COVID-19 pandemic prompted the Canadian Medical Association to call for legislation in 2021 that would protect health care workers from aggressive patients and protesters, both online *and* in-person [14].

Forms of online abuse may differ with individuals' gender identity, race or ethnicity, professional role, and other factors. For instance, women health communicators and journalists have faced more gendered or sexualized abuse than men [6,9,15].

Impacts of Online Abuse

Research suggests that online abuse can have serious negative consequences for health communicators' mental health and well-being: individuals who have experienced online harassment consistently report emotional distress and fear [2,9]. In a prepandemic survey, 62.4% (63/101) of prominent medical science communicators reported some negative mental health impacts, including depression, anxiety, and stress because of public engagement [6]. While most mental health impacts they reported were minor, 15% of them reported considerable or significant mental health ramifications [6]. Mental health consequences of online abuse can be disproportionately experienced due to gender, race or ethnicity, and other sociodemographic factors [15]. Throughout the COVID-19 pandemic, many health communicators have spoken out about the toll that negative comments and personal attacks have taken on their mental health [5,11]. In the *Nature* survey, more than 40% of 321 scientists reported experiencing emotional or psychological distress after commenting about the COVID-19 pandemic in traditional media interviews or on social media [10].

Some health communicators have expressed a desire to reduce or stop their online advocacy [6,10,12]. Thus, online abuse and its mental health consequences may undermine individuals' professional capacity, reducing the amount of engagement within the communication sector. If certain voices (eg, women and racialized individuals) are pushed out [16], then the diversity of perspectives within this sector will be in jeopardy.

Coping and Other Responses to Online Abuse

Variation in the impacts of online abuse on mental health, well-being, and professional efficacy suggests that health communicators differed in their coping strategies and ultimately their resilience to such abuse. A person who has experienced harassment will tend to adopt one or more strategies to "cope with it" [16]. Lazarus and Folkman [17] developed a seminal model to understand coping as either emotion- or problem-focused. Emotion-focused coping strategies help regulate emotional responses to a stressful situation, whereas problem-focused coping strategies aim to manage or alter the situation itself [17]. The effectiveness of emotion- and problem-focused coping strategies has been debated, but those who rely predominantly on emotion-focused coping strategies report significant negative emotions and poor mental health outcomes, such as depression [18].

Research on the coping strategies used by individuals who receive online abuse has primarily been conducted with journalists, scholars, and students [16,18-20], rather than health communicators. Many studies have further stratified emotion- and problem-focused coping strategies. For example, Scarduzio et al [21] put forth 11 "types" of emotion-focused strategies (eg, ignoring negative comments) and 5 "types" of

XSL•FO
RenderX

problem-focused strategies (eg, blocking hostile users) used by university students in response to online sexual harassment. In this study, we leveraged the model developed by Lazarus and Folkman as well as the examples of digital coping strategies given by Scarduzio et al to examine how health communicators responded to online abuse.

Objectives

Despite the broad adoption of social media by health care providers, scientists, and public health officials and the increasing recognition of online abuse they have received, no study has focused on the mental health consequences of *online* harassment for health communicators during the COVID-19 pandemic. Thus, our study aimed to explore (1) the impacts of online abuse on health communicators' mental health and well-being during the COVID-19 pandemic and (2) the emotion- and problem-focused coping strategies health communicators used to manage online abuse.

Methods

Study Design

This study is part of a larger participatory action research project on an emergent community of practice of health communicators. Relationships with health communicators were developed through direct outreach and through a partnership with ScienceUpFirst, a Canadian initiative that works with science and health experts to address misinformation. We used mixed methods, including semistructured interviews, an online questionnaire, and 2 workshops to examine the impacts of online abuse on Canadian health communicators during the COVID-19 pandemic.

Participant Recruitment

Health communicators for the larger participatory action research project were purposively recruited in 2 phases, emphasizing significant inclusion of those who self-identified as women or racialized. We use the term "racialized" rather than "visible minority groups" or the terms "Black," "Indigenous," and "people of color," following recommendations to use the former term by the Canadian government [22] and academics [23], arguing that race "does not exist as a biological concept to distinguish between human beings, but that social processes of racialization are inherently linked to major forms of historical, social, economic, and cultural oppression, including slavery and colonialism" [23].

During the first phase, research team members monitored traditional and social media platforms to identify people in Canada who were actively discussing public health measures online during the COVID-19 pandemic. Team members contacted a subset of these individuals by email, seeking participation from health communicators of diverse gender identities, ethnicities, and professional roles (eg, public health officials, health care professionals, university-based or civil society health experts, and health journalists). Health communicators who replied (N=35) were invited to participate in a virtual one-on-one semistructured interview, after informed consent was obtained.

During the second phase, team members recruited health communicators to participate in 2 workshops. Participants were recruited through #ScienceUpFirst affiliate groups and the authors' professional networks, seeking similar forms of diversity to the first phase. Journalists were intentionally excluded from the second phase because exploratory conversations with journalists and other health communicators suggested that combining these groups might put participants in awkward professional predicaments and because there has already been extensive research on online abuse of journalists [15,19,24]. In this phase, health communicators (N=34) first completed an online questionnaire to collect data on sociodemographic characteristics, communication activities, and experiences of harassment. Then, they participated in 1 of 2 virtual small-group 2-hour workshops. Observations from these workshops were taken from the research team members' notes because no audio-recordings or transcripts of the discussions were made.

Data Collection and Analysis

Quantitative Data

Given that there are no validated scales for online abuse of individuals in their professional capacities, we developed a questionnaire by drawing on existing questionnaires [13], including one previously created by the research team members [25] and another created by Ipsos to survey journalists' experiences of online harm [24]. Alongside the frequency, causes, and sources of online abuse, our questionnaire assessed how participants responded to online harm and how they changed their personal and professional work as health communicators due to online abuse. Simple descriptive statistics were calculated for specific questions using Microsoft Excel. All data were anonymized to protect participant confidentiality and were stored in a secure electronic database.

Qualitative Data

One-on-one interviews (N=35) were conducted by a research team member and recorded over Zoom (Zoom Communications, Inc) between December 2021 and June 2023. Each interview lasted between 40 and 90 minutes. A semistructured interview guide was used to conduct the interviews, which allowed researchers to compare participants' responses to set questions within the guide and explore other insights based on responses. Questions addressed issues, including the form and frequency of online abuse, professional and mental health impacts, and the online and offline actions that individuals took to respond to abuse. Zoom audio-recordings were transcribed verbatim.

For the first round of coding, our principal researcher created an initial list of deductive codes from literature on online harassment of politicians and journalists [16,25]. Three team members then coded approximately 15% (5/35) of the interview transcripts using this list, meeting regularly to discuss whether deductive codes and their definitions should be modified and whether new inductive codes should be added to the list. A revised codebook was created, and the remaining interview transcripts were independently coded using ATLAS.ti (Lumivero) software. Team members continued to discuss

XSL•FO
RenderX

coding during this process to ensure their shared fidelity to the revised codebook.

A second round of coding was undertaken to examine interview data about health communicators' mental health and well-being in more detail. One team member coded all interview excerpts that had been tagged with the "Impact: Mental health" code during the first round of coding. Inductive coding techniques were used to generate themes around mental health and resilience. Deductive coding techniques were used to identify emotion- and problem-focused coping strategies from the model developed by Lazarus and Folkman [17] and the study by Scarduzio et al [21]. Patterns across the excerpts were noted, and our team met 4 times to collectively categorize and interpret themes.

Ethical Considerations

This study was reviewed and approved by the University of British Columbia Behavioural Research Ethics Board (H21-01503 and #H22-01816). For one-on-one interviews, all participants received a consent statement outlining the study, potential risks and benefits, and measures to protect their personal information. The interviewer reviewed these details at the start of the interview and obtained verbal consent before proceeding. For individuals who completed surveys and participated in workshops, all participants signed a consent form that described the study, potential risks and benefits, and measures to protect their personal information.

Discussing online abuse can be difficult. Both the consent statement and consent form emphasized that participation was entirely voluntary, that individuals could stop participating in the interview or focus group at any time, and that they could withdraw from the study at any point. During interviews, if a participant expressed or displayed discomfort, the interviewer offered to pause, end the interview, or move to another question. For workshops, all participants received a community guidelines document in advance to ensure that group discussions were conducted safely and inclusively.

Data from this project are stored on password-protected hard drives and University of British Columbia–managed cloud storage, accessible only to core research team members registered as part of our research ethics board certification.

Results

Participant Characteristics

Table 1 outlines the sociodemographic characteristics of the interviewees (N=35) and questionnaire respondents (N=34). Most participants were living and working in either British Columbia or Ontario, Canada's 2 most populous English-speaking provinces. More women participated in both phases of this study than men, but both phases had an equal number of self-identified racialized and White participants.

XSL•FO
RenderX

Table 1. Sociodemographic characteristics of the participants from individual interviews and questionnaires.

Characteristics	Phase 1: Interviews (N=35), n (%)	Phase 2: Questionnaires (N=34), n (%)
Self-identified gender		
Woman	20 (57)	21 (62)
Man	15 (43)	13 (38)
Nonbinary	0 (0)	0 (0)
Prefer not to answer	0 (0)	0 (0)
Self-identified ethnicity		
Racialized	17 (49)	17 (50)
White	18 (51)	16 (47)
Prefer not to answer	0 (0)	1 (3)
Province of residence		
Alberta	5 (14)	1 (3)
British Columbia	15 (43)	11 (32)
Ontario	11 (31)	20 (59)
Quebec	3 (9)	1 (3)
Yukon	1 (3)	0 (0)
Other	0 (0)	1 (3)
Primary professional role[a]		
Public health official[b]	8 (23)	9 (26)
Medical professional	10 (29)	9 (26)
Health journalist	9 (26)	0 (0)
University-based or civil society expert	8 (23)	6 (18)
Other[c]	0 (0)	10 (29)

[a]While several participants belong to more than one category (eg, medical professional *and* university-based expert), we categorized participants based on their *primary* professional role.

[b]Employed by a public health agency, provincial government, or federal government.

[c]Employed by a nonprofit organization, research institute, community health center, science center, or self-employed.

Impacts of Online Harassment on Health Communicators' Mental Health and Well-Being

In the first phase, 94% (33/35) of the interviewees reported facing some form of online abuse since the pandemic began. Of the 2 interviewees who had not faced online abuse, one managed a health agency's general accounts and the other primarily relied on staff to manage her accounts. In the second phase, we asked the participants more detailed questions about the frequency of online harassment they experienced in the last 6 months. In total, 82% (28/34) of the questionnaire respondents received online threats, harassment, or false claims on multiple occasions in the last 6 months.

About 38% (13/34) of the questionnaire respondents claimed that their nationality or ethnic background was the reason they were targeted with online harassment. Racialized interviewees described how the online abuse they encountered differed from their White colleagues. For instance, a racialized woman stated:

> *There are random people out there that don't like you just because of who you are. Not necessarily because of what you say... I have female colleagues who are White... who say the same thing and don't get the reaction that I get.* [Civil society expert]

These sentiments were repeated by multiple racialized workshop participants.

In total, 18% (6/34) of the questionnaire respondents asserted that their gender identity or appearance was the reason they were targeted. Several women described how the abuse they received tended to emphasize their identity, including references to sexualized acts and derogatory remarks about their professional capabilities based on their gender. One woman explained how she inadvertently put her "whole self" online for scrutiny and, consequently, received many messages laced with "fatphobia and body shaming" (White woman, health journalist). She further explained, "It feels pretty vulnerable to be attacked that way as a young woman... I think it's really been probably the most intense misogyny I've ever faced in my life."

When interviewees recounted the online harassment they experienced, they often described negative emotions (Figure 1). Similarly, 41% (14/34) of the questionnaire respondents

XSL•FO
RenderX

experienced strong negative feelings in response to online harassment (Table 2).

Most health communicators mentioned multiple negative emotions around online abuse. Interviewees commonly mentioned feeling simultaneously frustrated and exhausted when they received a “constant barrage” of online harassment (White woman, public health official). Health communicators were “being fed this stream of negativity and abuse through...your phone all day” (White man, health journalist), and some felt that “even the best possible...person can’t deal with an assault of hostility 24-7” (White woman, public health official).

An interviewee reflected on how online harassers were misrepresenting his opinion on certain health topics:

> *That really does sting, and you find it frustrating and you waste...cognitive energy worrying about it.* [White man, university-based expert]

Several health communicators explicitly mentioned the “psychological toll” and “psychological exhaustion” of receiving and reading countless negative, hostile, and threatening e-mails (White man, university-based expert).

Interviewees also frequently expressed feeling sad and scared in the same instance:

> *When I get a hateful message, a...message that threatens violence against me, it makes me feel sad. ...it does scare me. It instills fear in me, and...it makes me feel really sad to know that this person...has taken time out of their day to...send me that message with the hope of hurting me somehow.* [Racialized woman, civil society expert]

While some of the interviewees described symptoms of depression, such as “not wanting to get out of bed...[in the] morning” (Racialized woman, civil society expert), others described symptoms of anxiety, such as difficulty “trying to unplug” (White woman, health journalist). Another interviewee, who received an email with “a message from someone...saying they hope that I get blood clots and like, basically die,” explained how these types of messages are anxiety-inducing: “sometimes it keeps you up at night, makes you very worried and concerned” (White woman, health journalist).

Racialized and women health communicators across interviews and workshops discussed how negative comments about their ethnicity, ancestry, and physical appearance impacted their mental health and well-being:

> *When it actually ends up being personal attacks on you as a person, on how you look, on the colour of your skin, where you’re originally from...It takes a toll on you.* [Racialized woman, civil society expert]

Interestingly, only 18% (6/34) of the questionnaire respondents claimed to be “struggling with mental health issues” (Table 2), whereas the interviewees repeatedly mentioned how their mental health and well-being had been impacted by health communication and subsequent harassment during the pandemic. This variation in the severity of mental health impacts reported suggests that certain participants implemented strategies to mitigate some of the mental health consequences of online abuse.

Figure 1. Examples of negative emotions expressed by interviewees because of online abuse.

XSL•FO
RenderX

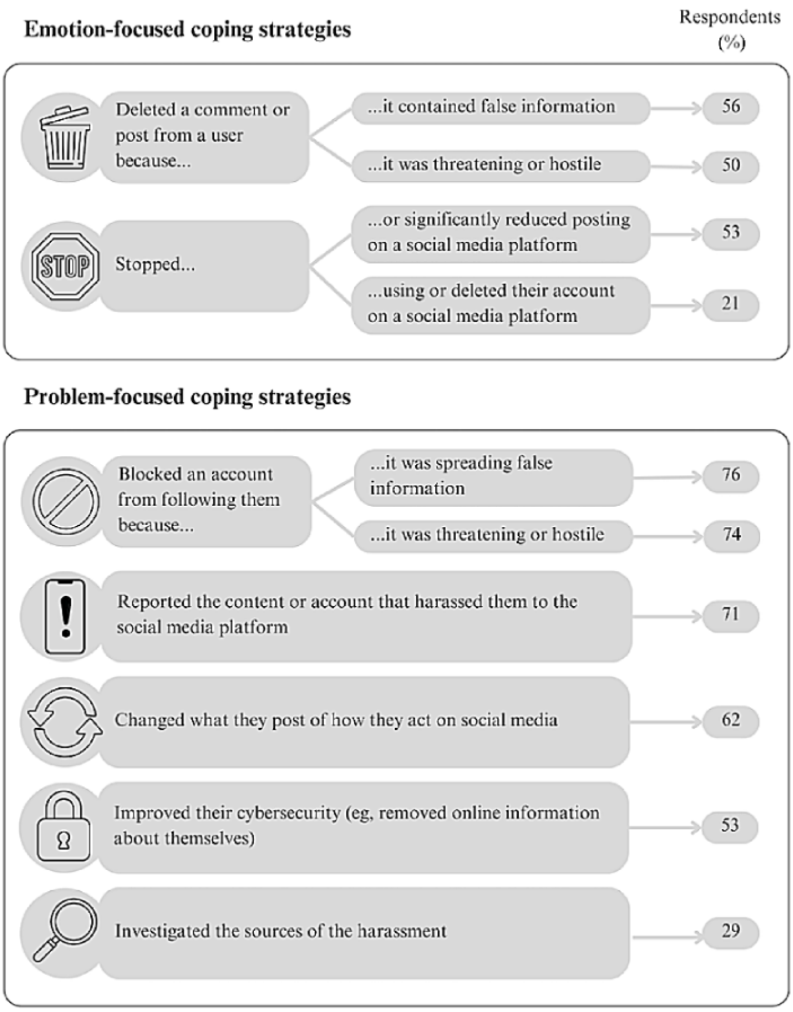
Table 2. Impacts of online harassment on questionnaire respondents.

	Questionnaire respondents (N=34), n (%)
Personal impacts	
Experienced strong negative feelings (eg, fear, horror, anger, guilt, or shame)	14 (41)
Felt scared for the safety of their family and friends	8 (24)
Felt scared for their physical safety	7 (21)
Struggled with mental health issues	6 (18)
Felt jumpy or easily startled	6 (18)
Experienced strong negative beliefs about themselves or other people	6 (18)
Repeated, disturbing dreams of the stressful experience	3 (9)
Professional impacts	
Avoided publicly addressing certain topics	21 (62)
Seriously considered quitting health communication	14 (41)
Requested a change in their professional role	5 (15)
Took a greater number of sick days than usual	2 (6)

Coping Strategies Implemented by Health Communicators

When health communicators experienced online harassment, they used a variety of coping strategies (Figure 2).

Figure 2. Responses to online harassment implemented by questionnaire respondents (N=34).



XSL•FO
RenderX

Emotion-Focused Coping Strategies

Health communicators' most common emotion-focused coping strategies were (1) withdrawing from social media planforms, (2) avoiding social media platforms altogether, and (3) accepting online harassment and abuse as unavoidable.

First, many participants withdrew from online public health communication because of harassment. In total, 53% (18/34) of questionnaire respondents noted that they significantly reduced or stopped posting on a social media platform (Figure 2). Interviewees mentioned that they "used to...be very active on Twitter," (White woman, health journalist) but they "made a conscious choice to distance myself from Twitter as a professional" due to persistent online harassment (White man, health journalist).

Second, 21% (7/34) of questionnaire respondents claimed that they have stopped using or deleted their account on a social media platform (Figure 2). Several health communicators explicitly stated that they avoided social media platforms to safeguard their mental health and well-being. "I felt for my mental well-being, I am avoiding this because it is like a triggering response to see all those notifications coming in again and again and again" (White woman, health journalist). One interviewee said he had "to turn if off" for his "own peace of mind...because there's been an overwhelming amount of negativity," (White man, public health official), and another firmly stated she was "not on Twitter for...[her] mental health" (White woman, public health official).

Several interviewees justified these emotion-focused coping strategies by explaining how sharing public health information was not part of their job descriptions, so they were "not paid to be on Twitter" (Racialized man, medical professional). A racialized civil society expert noted, "It's not as though this is my job...I'm not getting paid for any of this work...[and] on top of that, I get vitriol."

Third, several interviewees seemed to accept the possibility of receiving online harassment. One health communicator conceded that "because of the way that things are now, whatever you share, anything you post...you do open yourself up to some degree of abuse" (White man, health journalist). Another felt disheartened that she was "going to have to be dealing with [online harassment]...for a long time," particularly if she wanted to continue publishing in major outlets and working on "more polarizing" topics (White woman, health journalist).

Health communicators may have relied on emotion-focused coping strategies because they were unaware of problem-focused coping strategies or unable to implement them. One woman asserted that she did not "know how to get back on Twitter without...facing all this same garbage again" (Health journalist). Another admitted that she felt "ill-equipped to...be on the site anymore" (White woman, health journalist).

Problem-Focused Coping Strategies

Scarduzio et al [21] proposed 5 specific "types" of problem-focused coping strategies, including blocking or unfriending, changing online behavior, formal help-seeking, peer intervention, and confronting. Health communicators in our study implemented all 5 of these digital coping strategies to mitigate further online harassment and abuse.

In total, 74% (25/34) of the questionnaire respondents blocked a threatening or hostile account from following them, and 53% (18/34) of respondents improved their cybersecurity (Figure 2). Several interviewees explained the exact steps they took to improve their cybersecurity, such as "taking my face off of my Twitter profile" (White man, health journalist). One interviewee explained how making her "account private" helped her create "a bit of a safer space" online (Racialized woman, civil society expert).

In total, 62% (21/34) of the questionnaire respondents (Figure 2) and many interviewees changed what they posted and how they behaved online. One participant started making an extremely conscious effort to refrain from posting personal information online:

> *It has just made me more cognizant of what I share in general about my life...I never tweet about where I am. I never tweet about the neighbourhood I live in...I didn't want malicious people [to] have that information about me.* [White woman, health journalist]

Other health communicators refrained from sharing public health information because they were "worried about it becoming somehow a lightning rod for hate or harassment or just unwanted negative attention" (White woman, health journalist). For example, some interviewees explained how they no longer shared opinions on "controversial and important topics" (White woman, health journalist), and they "don't advocate outwardly for...[COVID-19] restrictions" (Racialized man, medical professional). Similarly, 62% (21/34) of the questionnaire respondents avoided publicly addressing certain topics (Table 2). As more health communicators resorted to self-censorship, crucial health information became less available to publics.

Other participants described setting boundaries on how they used social media platforms, not as a form of self-censorship but to balance personal well-being with professional efficacy. For instance, when online abuse made his social media engagements particularly stressful, a health official took time away from the platform to develop new strategies for how he would use it:

> *I just reactivated [my account] after a few weeks...Coming back with a few rules in mind, I felt much better.* [Racialized man, public health official]

Some health communicators sought assistance to manage online abuse by reporting the harassment directly via social networking sites' reporting mechanisms. In our study, 71% (24/34) of questionnaire respondents engaged in "formal help-seeking" by reporting the content or account that harassed them to the social media platform (Figure 2). Several interviewees stated that this action did not have a reliable effect, since in some situations the posts or accounts remained on platforms long after the health communicator had reported them. Our interviewees did not discuss how inconsistent action by social media platforms shaped their assessment of the efficacy of this problem-focused coping strategy.

XSL•FO
RenderX

In total, 41% (14/34) of the questionnaire respondents reported acts of online harassment to a supervisor or employer, and 18% (6/34) of respondents sought help from a supervisor or employer (Table 3). Two health journalists found support and directives from their employers to be effective. One journalist explained how the independent news website she writes for "has been really supportive and proactive" by clarifying "the conditions that my work is expected to continue under" (White woman, health journalist). The other journalist described how the mental health supports that her employers have provided were beneficial:

> *Both organizations that I've worked for have really been putting an emphasis on...getting mental health support. I think some big changes that were made structurally to benefits that were being offered in terms of how much mental health support was being covered made a really big difference.* [Racialized woman, health journalist]

These findings suggest that health journalists received more support from their employers than health communicators in other professions.

Other health communicators relied on colleagues and family members to help manage online abuse and stop harassers (ie, "peer intervention"). In response to online harassment, more than half of the questionnaire respondents (18/34, 53%) asked a colleague for help, and 47% (16/34) of respondents asked a friend or a family member for help (Table 3). For instance, members of a public health agency team regularly checked in with each other to provide emotional support, often in the form of humor, and to review potentially abusive messages.

If someone else, such as an employee, colleague, or significant other, could oversee their social media accounts, then health communicators would not have to read negative comments and messages themselves:

> *I was...telling with my campaign team that I actually want to hand...over the keys [to my social media accounts]...because I actually don't want to see it anymore.* [Racialized man, university-based expert]

One public health official described how her "partner joined Twitter partly because he took on the job [of] monitoring my account," but she acknowledged "there's a toll... when you read angry tweets about your partner every day" (White woman, public health official).

Confronting online harassers was not a popular problem-focused coping strategy used by health communicators. Only one interviewee explained that she wanted to engage with perpetrators of online harassment "because criticism and conflict eats at" her, and it was "empowering" to try to connect with those who were unnecessarily hostile (White woman, university-based expert). She noted that, "I've engaged twice, by phone and it actually worked."

Table 3. Sources of reporting and support for questionnaire respondents who experienced online harassment.

	Questionnaire respondents (N=34), n (%)
To whom did you report the acts of online harassment?	
Social media platforms	22 (65)
Supervisor or employer	14 (41)
Professional association or governing body	6 (18)
Police	4 (12)
Government or political representative	1 (3)
Unions	1 (3)
I did not report any acts of harassment	5 (15)
From whom did you seek support?	
Asked a colleague for help	18 (53)
Asked a friend or family member for help	16 (47)
Spoken publicly about the experience of being harassed, having your reputation attacked, or the sources of harassment	13 (38)
Looked for online resources to protect yourself or cope with harassment	9 (26)
Asked my supervisor, employer, or organization for help	6 (18)
Sought help from a professional organization or other civil society group	5 (15)
Sought legal advice	5 (15)
Sought medical or psychological help	4 (12)
I did not look for support	3 (9)

XSL•FO
RenderX

Continuation of Online Health Communication

Health communicators' experiences of online abuse prompted them to reflect on their long-term use of social media to engage publics. We identified 2 broad groups: those who felt the negative impacts of online harassment on their mental health and well-being outweighed the potential benefits of public health advocacy and those who expressed a desire to continue sharing relevant public health information despite online harassment. Health communicators discussed the internal conflict between reducing their engagement to protect their mental health and continuing their advocacy. One interviewee, after being targeted on Twitter for discussing abuse she and other journalists faced, stopped engaging with those issues online:

> *That was honestly really frustrating because I felt that at a time that I really needed to be vocal about these things, I couldn't without compromising my safety and my mental well-being. It kind of felt like there was no good option: either stay silent about what had been done or speak out and perhaps welcome more harm.* [White woman, health journalist]

Those in the first group generally reduced their online engagement to prioritize their mental health and well-being:

> *My mental health and well-being are more important than the hope that maybe...these people will learn...because they're not going to learn.* [Racialized man, medical professional]

Similarly, a participant who used online platforms to remain engaged with health issues reported that he began to believe those benefits are being outweighed by "the risks to my mental health and well-being...and the threats to my productivity" (Racialized man, university-based expert).

In total, 41% (14/34) of questionnaire respondents seriously considered quitting health communication (Table 2). An interviewee explained how burnout prompted him to take a step back from his profession during the pandemic:

> *I was just kind of tired, I guess...I certainly felt burnt out from reporting on the pandemic...There are still COVID stories that are important to tell and there's important journalism to be done, but I felt as if I didn't want it to be done by me anymore.* [White man, health journalist]

Conversely, interviewees in the second group explained why and how they would continue to engage in online health communication despite the challenges:

> *I do feel upset for a little bit (after experiencing online abuse), but it has never gotten to a point where I would think to myself, "I'm not going to do this ever again." I know of people who have given up, who've taken time off or just completely stopped engaging, but...at least for now, I haven't reached that point.* [Racialized woman, civil society expert]

Several participants emphasized that they continued to advocate online because they believed in the public benefit of sharing health information:

> *I'm just trying to tell (people) the facts. And to be targeted for being the messenger of those facts is not very fun. There have been times because of the backlash that I've thought, well, maybe I won't tweet as much. And I definitely had that thought a few times during the course of the pandemic. I really had to weigh...(is) me getting a few messages that are annoying more important than me trying to get information out to people? And for me, getting information out is always more important.* [Racialized woman, health journalist]

Although most health communicators implemented a variety of problem-focused and emotion-focused coping strategies (Figure 2), only some demonstrated a strong willingness to continue their online engagement, whereas others contemplated quitting public health communication. In the face of persistent online abuse, *continuing* to post online could be understood as an act of "professional resilience." Health communicators who faced challenges overcoming the mental health and well-being impacts of online abuse were more likely to reduce or abandon their online advocacy efforts.

Discussion

Impacts of Online Abuse

Throughout the COVID-19 pandemic, many health care providers, researchers, public health officials, and health journalists put extraordinary effort into engaging publics online, which often exposed them to unwanted harassment and abuse [10]. Although online harassment is often dismissed because it occurs in virtual environments, the consequences of such harassment can be very real, including psychological stress and burnout [13]. Among our questionnaire respondents, 82% (28/34) faced online harassment or abuse. Therefore, most participants reported negative emotions, including feeling fatigued, sad, distressed, and angry. Some participants shared symptoms of anxiety and depression, and some explicitly reported that they had been struggling with mental health issues. This emotional distress caused by online harassment has exacerbated the widespread burnout experienced by medical professionals during the COVID-19 pandemic [26].

Coping Strategies for Online Abuse

Participants in this study mentioned a variety of strategies to cope with the mental health impacts of online harassment. Drawing on the framework proposed by Lazarus and Folkman [17], we categorized these as emotion-focused and problem-focused coping strategies [21,27,28].

Participants' most common emotion-focused coping strategies were enduring or ignoring online harassment and disengaging or withdrawing from social media platforms. Several interviewees seemed to accept online harassment as something that came with the territory, rather than something that could be mitigated with problem-focused coping strategies. This sentiment aligns with other research findings that many scientists who publicly commented on the COVID-19 pandemic said they learned to cope with online harassment by "accepting it as an unpleasant but expected side effect of getting information to

XSL•FO
RenderX

the public" [10]. Furthermore, research suggests that many health communicators have purposefully ignored and not responded to social media trolls [10,11,29], since "engagement is their oxygen" [5]. Other emotion-focused coping strategies for online abuse reported by health communicators in our study and the literature include deleting negative comments, reducing engagement on social media platforms, avoiding certain social media platforms, and deleting social media accounts altogether [5,10,29]. Although self-blame has been identified as an emotion-focused coping strategy in other studies [16], no health communicator in our study described blaming themself for the online harassment they received.

Health communicators also used many problem-focused coping strategies to respond to online abuse. As opposed to reactive emotion-focused strategies, such as deleting negative comments, problem-focused strategies tend to be more proactive, such as blocking or reporting hostile users. For example, by preventing such users from sending messages directly to the communicator or seeking to have social media platforms enforce their terms of service against online harassment, health communicators have tried to limit the number of negative or threatening messages they will receive *in the future*. One academic told other researchers that she even blocked her abuser's followers to make it harder for them to target her [10]. Health communicators in our study and in the literature have taken several proactive steps to avoid receiving online abuse, including refining cybersecurity settings by making accounts private [11,29] and removing contact information from public websites [5].

Furthermore, rather than avoiding posting on social media platforms *entirely*, many health advocates became "more careful about how...[they] use" social media [10], making conscious efforts to strategically avoid posting about *specific* topics online. In fact, 63.5% (228/359) of the physicians, biomedical scientists, and trainees in the United States who reported experiencing any online harassment during the pandemic claimed that they have *changed* how they use social media [12]. Many health communicators, in the literature and our study, have begun compartmentalizing professional and personal identities online, avoiding "making comments that might be perceived as political" or controversial [10], or refusing to correct misinformation online [29].

When we examined help-seeking behaviors among our participants, we found a distinction between *reporting* online harassment and *seeking support* for such harassment. While almost three-quarters of questionnaire respondents (24/34, 71%) formally reported online harassment directly to social media platforms, when the same respondents recorded who they asked for help, colleagues (18/34, 53%) and friends and family members (16/34, 47%) were the most common sources of support. Several interviewees explained how their employees or loved ones helped them manage online abuse, limiting the number of negative comments and direct messages they read about themselves. Similarly, Hodson et al [30] reported that women scholars who experienced online harassment were most likely to try to deal with the problem by enlisting the help of spouses, close family members, or friends to help manage their online presence. Social media platforms allow users to report and block hostile users [31], but taking these actions may not be as effective in improving health communicators' mental health and well-being as receiving assistance and emotional support from friends and family members [30]. However, our study focused on health communicators' perspectives on the efficiency of various coping strategies, rather than examining social media platform activity. Thus, we cannot directly assess whether health communicators' reports of hostile users were acted on by platforms or whether these platform responses shaped health communicators' assessments on the efficacy of this strategy.

Some scholars have discussed how emotion- and problem-focused coping strategies can be difficult to discern, and we also found some overlap between the 2 categories. For example, Scarduzio et al [21] described asking friends and family members "for support and advice" as an active emotion-focused coping strategy yet asking friends and family members "to help stop the harasser" as a problem-focused coping strategy. Thus, we acknowledge that some coping strategies may fall into a "gray zone" between emotion-focused and problem-focused.

Although confrontation was an unpopular problem-focused coping strategy among health communicators in our study, some participants expressed a desire to have productive dialogues with their harassers. One physician said she occasionally responded to comments or messages but not when she was upset or angry [10]. However, confronting perpetrators may pose a risk of further abuse and negative mental health consequences [32], which could be one reason most health communicators in our study and the literature relied on other coping strategies.

We found some problem-focused coping strategies were individual in nature, while other strategies involved support through personal and professional relationships or official organizational policies. Individual strategies, like blocking hostile accounts, can lessen exposure to online abuse and, consequently, lessen the impacts of such abuse on mental health and well-being. Another strategy to lessen exposure is sharing the burden of monitoring and responding to hostile content with friends, loved ones, or colleagues. Beyond reducing exposure to online abuse, social and organizational support can strengthen a health communicator's ability to emotionally process abuse and rebuild mental health. For example, 2 journalists in our study, who demonstrated a willingness to continue their professional advocacy, highlighted the importance of access to expanded employee benefits, which enabled them to take time off work and receive counseling after experiencing online abuse. In the workshops, several public health officials described how their teams routinely discussed the hostility they received to address any sense of isolation or personal responsibility for these reactions and to share coping strategies. Conversely, health communicators in our study who worked as freelancers or in individual medical practices noted that a sense of isolation and lack of workplace support had exacerbated the mental health consequences of online abuse.

Professional Resilience Among Health Communicators

There are several opinions about the effectiveness of emotion- and problem-focused coping strategies, but many scholars argue

XSL•FO
RenderX

that problem-focused strategies are more beneficial in the long-term [17,33]. Our findings suggest that health communicators who used problem-focused coping strategies were more likely to continue their advocacy than health communicators who used emotion-focused coping strategies. In the face of persistent online harassment and abuse, Veletsianos et al [16] reported that women scholars who continued working and teaching "required determination and resilience." Thus, simply continuing their professional obligations became an act of resistance [16]. Similarly, we put forth that health communicators who remained "active" online demonstrated significant professional resilience, compared to those who censored or otherwise minimized their online presence. Several communicators in our study noted their commitment to sharing important health information broadly with publics was one reason for this resilience. Some of the clearest expressions of professional resilience were shared by racialized health communicators in our study, which warrants further investigation.

Importantly, we do not define professional resilience as an individual quality or character trait. An individual's capacity to continue using online spaces to inform and advocate publics is significantly shaped by the forms and intensity of online abuse they face as well as the interpersonal and institutional support they receive. Moreover, the extraordinarily high levels of online engagement by health communicators during the COVID-19 pandemic required many health experts to take on burdens that went beyond their job descriptions or were otherwise unsustainable.

Online and in-person abuse has contributed to burnout and high turnover among health communicators, particularly medical professionals and health journalists [11]. Many participants in our study contemplated reducing or ceasing their online health communication activities. This decision could have negative professional consequences, such as a reduction in opportunities to network and collaborate with other scholars [9]. These consequences were especially pronounced for women and racialized individuals, who have historically been excluded from academia. Women who have reported considerable online harassment, especially sexual harassment, have frequently responded by reducing and censoring their online participation as well as deleting their accounts on social media platforms [9,16,34], further limiting their opportunities for professional development.

There are also broader social consequences if health communicators reduce their engagement online. Notably, misinformation and disinformation may be left unchecked by those most qualified to counter it [12]. Experts who were attacked online said they were less likely to participate in future media interviews, highlighting the effectiveness of these attacks [7]. Similarly, scientists who reported the highest frequency of trolling in the *Nature* survey were most likely to report that their experiences have greatly affected their willingness to speak to the media in the future [10]. At a time when "we've never needed them so badly" [10], many health communicators are avoiding certain topics on social media or withdrawing from these platforms entirely. Furthermore, given the alarming amount of abuse reported by senior public health officials, it seems likely that the hostile online environment could dissuade up-and-coming health communicators from fully engaging in important discussions [6]. Consequently, we may see a reduction in the diversity of thoughts and opinions shared within academia and public discourse, especially if women and racialized academics are disproportionately pushed out of online spaces [16].

Policy Recommendations

Health communicators in our study implemented various emotion- and problem-focused coping strategies, many of which they implemented as individuals. Future studies should investigate the effectiveness of these digital coping strategies for health communicators' mental health, well-being, professional efficacy, and professional resilience, especially those who belong to gender and racial minorities.

Our findings also highlight the limitations of individual coping strategies, necessitating the development of organizational policies to support those who receive online abuse and sanction those who perpetrate it. While health communicators have taken many steps to mitigate the frequency and severity of harassment they experience on social media platforms, advocates argue that individuals should not have to "cope on their own" [10].

Advocates have asserted that there is much that institutions can do to assist scientists who are receiving online abuse [10]. Studies have suggested several actions for institutions that employ health communicators: creating formal policies to guide health communicators' digital interactions [19], hosting trainings for health communicators on how to engage with the media and what to expect from online trolls [5,10], and enlisting organizations' information technology departments to block consistent abusive emailers and report incidents to social media platforms and police [10]. These organizational efforts should address the potential for different forms of abuse based on individuals' gender identity, race or ethnicity, ancestry, and other sociodemographic characteristics. In our study, the strongest examples of organizational policies were provided by health journalists and communicators at public health agencies. These examples include (1) clear recognition from superiors that online abuse is a serious problem that requires action and (2) institutional programs providing psychological therapy, cybersecurity assistance, and peer support. This finding warrants an exploration of organizational polices across industries to ascertain and promote best practices. Programs are also needed to support health communicators who are not full-time employees of large organizations, such as family doctors, free-lance journalists, and others.

While individual actions may have immediate short-term outcomes, institutional policies and practices could have sustained long-term outcomes for health communicators' mental health and well-being by preventing online harassment or, at least, mitigating it. Organizational policies would support professional resilience, ensuring that important health information is "not silenced" by online abuse and can still reach publics [10].

XSL•FO
RenderX

Limitations

There are several limitations to this study. Our purposive sampling of health communicators who were highly engaged in online health communication provided findings from an important population but not necessarily a representative one. Our study was conducted in English, which might have precluded insights from Canada's significant French-speaking population. Moreover, because of the small sample size of our study, we could not quantitatively compare exposure to online harassment by gender identity, ethnicity, or professional role. A more comprehensive sample of health communicators across institution types, professional roles, and sociodemographic characteristics could identify broader patterns and gaps in our findings as well as greater insights into the experiences and coping strategies of members from marginalized populations. Finally, this study focuses on health communicators in Canada and their experiences during the first few years of the COVID-19 pandemic, when uncertainty and fear were heightened. Further comparative studies across countries are needed to measure the long-term impacts of online abuse and coping strategies on health communicators in different political and health care contexts.

Conclusions

This study elucidates the significant impacts of online abuse on health communicators during the COVID-19 pandemic, highlighting both mental health and professional consequences. Despite the variety of individual coping strategies used by health communicators, there remains a pressing need for organizational efforts that offer comprehensive protection against online abuse and support for those who receive it. Institutions must acknowledge that the burden of coping with online abuse should not fall solely on individuals and that they should be supported by formal organizational policies and practices that safeguard the mental health, well-being, and professional efficacy of health communicators. These efforts will support individuals at the forefront of public health communication to share critical information.

Acknowledgments

The authors are grateful for the interviewees and workshop participants who shared their insights and experiences, some of which were quite difficult. They deeply appreciate research assistance from Hannah Hett, Sabah Haque, and Oliver Zhang, as well as feedback from Jaigris Hodson, Elizabeth Dubois, and George Veletsianos. The authors would like to thank ScienceUpFirst, which helped them engage a broader community of health communicators. This study has been made possible by funding from the New Frontiers in Research Fund (NFRFR-2021-00289) by the government of Canada and from the Bridge Research Consortium, a part of Canada's Immuno-Engineering and Biomanufacturing Hub.

Authors' Contributions

CT, SH, and HT conceptualized the study. CT collected interview and survey data. CT and a research assistant coded the qualitative data. LW analyzed the quantitative and qualitative data with support from SH and CT. LW drafted the manuscript. CT, SH, and HT reviewed and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Al-Rawi A, Zemenchik K. Tiktoking COVID-19 with frontline workers. Digit Health 2023 Feb 08;9:20552076231152766 [FREE Full text] [doi: 10.1177/20552076231152766] [Medline: 36776404]
2. Royan R, Pendergrast TR, Del Rios M, Rotolo SM, Trueger NS, Bloomgarden E, et al. Use of Twitter amplifiers by medical professionals to combat misinformation during the COVID-19 pandemic. J Med Internet Res 2022 Jul 22;24(7):e38324 [FREE Full text] [doi: 10.2196/38324] [Medline: 35839387]
3. Azpiri L, Aylesworth L. UBC prof plays social media myth-buster on COVID-19 questions. Global News. URL: https://globalnews.ca/news/7876858/ubc-prof-covid-19-tiktok-questions/ [accessed 2024-04-29]
4. Mello S, Glowacki E, Fuentes I, Seabolt J. Communicating COVID-19 risk on Instagram: a content analysis of official public health messaging during the first year of the pandemic. J Health Commun 2023 Jan 02;28(1):38-52 [FREE Full text] [doi: 10.1080/10810730.2023.2175278] [Medline: 36760119]
5. Basky G. Health advocates want help handling online harassment. CMAJ 2021 Feb 22;193(8):E292-E293 [FREE Full text] [doi: 10.1503/cmaj.1095921] [Medline: 33619072]
6. Grimes DR, Brennan LJ, O'Connor R. Establishing a taxonomy of potential hazards associated with communicating medical science in the age of disinformation. BMJ Open 2020 Jul 05;10(7):e035626 [FREE Full text] [doi: 10.1136/bmjopen-2019-035626] [Medline: 32624466]
7. Arora VM, Bloomgarden E, Jain S. Supporting health care workers to address misinformation on social media. N Engl J Med 2022 May 05;386(18):1683-1685. [doi: 10.1056/nejmp2117180] [Medline: 35353977]
8. Nogrady B. Harassment of scientists is surging - institutions aren't sure how to help. Nature 2024 May 21;629(8013):748-750. [doi: 10.1038/d41586-024-01468-9] [Medline: 38773292]

XSL•FO
RenderX

9. Pendergrast TR, Jain S, Trueger NS, Gottlieb M, Woitowich NC, Arora VM. Prevalence of personal attacks and sexual harassment of physicians on social media. JAMA Intern Med 2021 Apr 01;181(4):550-552 [FREE Full text] [doi: 10.1001/jamainternmed.2020.7235] [Medline: 33393965]
10. Nogrady B. 'I hope you die': how the COVID pandemic unleashed attacks on scientists. Nature 2021 Oct 13;598(7880):250-253. [doi: 10.1038/d41586-021-02741-x] [Medline: 34645996]
11. Larkin H. Navigating attacks against health care workers in the COVID-19 era. JAMA 2021 May 11;325(18):1822-1824. [doi: 10.1001/jama.2021.2701] [Medline: 33881489]
12. Royan R, Pendergrast TR, Woitowich NC, Trueger NS, Wooten L, Jain S, et al. Physician and biomedical scientist harassment on social media during the COVID-19 pandemic. JAMA Netw Open 2023 Jun 01;6(6):e2318315 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.18315] [Medline: 37314808]
13. Gosse C, Veletsianos G, Hodson J, Houlden S, Dousay TA, Lowenthal PR, et al. The hidden costs of connectivity: nature and effects of scholars' online harassment. Learn Media Technol 2021 Jan 24;46(3):264-280. [doi: 10.1080/17439884.2021.1878218]
14. Cukier A, Vogel L. Escalating violence against health workers prompts calls for action. CMAJ 2021 Dec 13;193(49):E1896 [FREE Full text] [doi: 10.1503/cmaj.1095980] [Medline: 34903595]
15. Pearson M, Seglins D. A report on mental health, well-being and trauma among Canadian media workers. Canadian Journalism Forum on Violence and Trauma. 2022. URL: https://static1.squarespace.com/static/60a28b563f87204622eb0cd6/t/6285561b128d0447d7c373b2/1652905501967/TakingCare_EN.pdf [accessed 2024-04-29]
16. Veletsianos G, Houlden S, Hodson J, Gosse C. Women scholars' experiences with online harassment and abuse: self-protection, resistance, acceptance, and self-blame. New Media Soc 2018 Jun 22;20(12):4689-4708. [doi: 10.1177/1461444818781324]
17. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. New York, NY: Springer; 1984.
18. Houlden S, Hodson J, Veletsianos G, Gosse C, Lowenthal P, Dousay T, et al. Support for scholars coping with online harassment: an ecological framework. Fem Media Stud 2021 Feb 10;22(5):1120-1138. [doi: 10.1080/14680777.2021.1883086]
19. Harlow S, Wallace R, Chacón LC. Digital (in)security in Latin America: the dimensions of social media violence against the press and journalists' coping strategies. Digit Journal 2022 Oct 07;11(10):1829-1847. [doi: 10.1080/21670811.2022.2128390]
20. Weber DL, Bygholm AK, Kanstrup AM. Digital coping strategies of young people living with cognitive disabilities: using emotion-focused and problem-focused coping in digital relationships. Disabil Soc 2023 Feb 21;39(8):2007-2031. [doi: 10.1080/09687599.2023.2181767]
21. Scarduzio JA, Sheff SE, Smith M. Coping and sexual harassment: how victims cope across multiple settings. Arch Sex Behav 2018 Feb 21;47(2):327-340. [doi: 10.1007/s10508-017-1065-7] [Medline: 29164362]
22. Tolley E. Framed: Media and the Coverage of Race in Canadian Politics. Vancouver, BC: UBC Press; 2016.
23. Blackett A. A transformative framework to achieve and sustain employment equity. Employment and Social Development, Canada. 2024 Jul 05. URL: https://www.canada.ca/en/employment-social-development/corporate/portfolio/labour/programs/employment-equity/reports/act-review-task-force/chapter-3.html [accessed 2025-01-22]
24. Online harm in journalism. Ipsos. 2021. URL: https://www.ipsos.com/sites/default/files/ct/news/documents/2021-11/Online_Harm_in_Journalism-Report-2021-11-09-v1.pdf [accessed 2024-04-29]
25. Tenove C, Tworek H, Lore G, Buffie J, Deley T. Damage control: how campaign teams interpret and respond to online incivility. Pol Commun 2022 Nov 04;40(3):283-303 [FREE Full text] [doi: 10.1080/10584609.2022.2137743]
26. Prasad K, McLoughlin C, Stillman M, Poplau S, Goelz E, Taylor S, et al. Prevalence and correlates of stress and burnout among U.S. healthcare workers during the COVID-19 pandemic: A national cross-sectional survey study. EClinicalMedicine 2021 May;35:100879 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100879] [Medline: 34041456]
27. Prowten SD, Breitenstein RS. Coping in the digital age: do coping strategies mediate the link between online and in-person victimization and depressive symptoms? J Interpers Violence 2023 Oct 04;38(19-20):10723-10748. [doi: 10.1177/08862605231176790] [Medline: 37272011]
28. Völlink T, Bolman CA, Eppingbroek A, Dehue F. Emotion-focused coping worsens depressive feelings and health complaints in cyberbullied children. J Criminol 2013 Jun 20;2013:1-10. [doi: 10.1155/2013/416976]
29. Bautista JR, Zhang Y, Gwizdka J. US physicians' and nurses' motivations, barriers, and recommendations for correcting health misinformation on social media: qualitative interview study. JMIR Public Health Surveill 2021 Sep 01;7(9):e27715 [FREE Full text] [doi: 10.2196/27715] [Medline: 34468331]
30. Hodson J, Gosse C, Veletsianos G, Houlden S. I get by with a little help from my friends: the ecological model and support for women scholars experiencing online harassment. First Monday 2018;23(8):5 [FREE Full text] [doi: 10.5210/fm.v23i8.9136]
31. Wong RY, Cheung CM, Xiao B, Thatcher JB. Standing up or standing by: understanding bystanders' proactive reporting responses to social media harassment. Inf Syst Res 2021 Jun;32(2):561-581. [doi: 10.1287/isre.2020.0983]
32. Smith PH, Murray CE, Coker AL. The Coping Window: a contextual understanding of the methods women use to cope with battering. Violence Vict 2010 Feb 01;25(1):18-28. [doi: 10.1891/0886-6708.25.1.18] [Medline: 20229691]

XSL•FO
RenderX

33. D'Zurilla TJ, Nezu AM. Problem-solving therapies. In: Dobson KS, editor. Handbook of Cognitive-Behavioral Therapies. 3rd edition. New York, NY: Guilford Press; 2010:197-225.
34. Vitak J, Chadha K, Steiner L, Ashktorab Z. Identifying women's experiences with and strategies for mitigating negative effects of online harassment. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR p. 1231-1245 URL: https://dl.acm.org/doi/10.1145/2998181.2998337 [doi: 10.1145/2998181.2998337]

Edited by T Purnat; submitted 06.11.24; peer-reviewed by S Gordon, A Tarasov; comments to author 20.01.25; revised version received 03.02.25; accepted 03.02.25; published 02.04.25.

Please cite as:
Wight L, Tenove C, Hirani S, Tworek H
Mental Health and Coping Strategies of Health Communicators Who Faced Online Abuse During the COVID-19 Pandemic: Mixed Methods Study
JMIR Infodemiology 2025;5:e68483
URL: https://infodemiology.jmir.org/2025/1/e68483
doi:10.2196/68483
PMID:

©Lisa Wight, Chris Tenove, Saima Hirani, Heidi Tworek. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 02.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

How the General Public Navigates Health Misinformation on Social Media: Qualitative Study of Identification and Response Approaches

Sharmila Sathianathan[1,2*], BPharm (Hons), MClinPharm; Adliah Mhd Ali[1*], BPharm (Hons), MPharm(Clin), PhD; Wei Wen Chong[1*], BPharm (Hons), PhD

[1]Centre of Quality Management of Medicines, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

[2]Pharmacy Department, Hospital Tengku Ampuan Rahimah, Selangor, Malaysia

[*]all authors contributed equally

Corresponding Author:
Wei Wen Chong, BPharm (Hons), PhD
Centre of Quality Management of Medicines, Faculty of Pharmacy, Universiti Kebangsaan Malaysia
Jalan Raja Muda Abdul Aziz
Kuala Lumpur, 50300
Malaysia
Phone: 60 392897118
Email: weiwen@ukm.edu.my

Abstract

Background: Social media is widely used by the general public as a source of health information because of its convenience. However, the increasing prevalence of health misinformation on social media is becoming a serious concern, and it remains unclear how the general public identifies and responds to it.

Objective: This study aims to explore the approaches used by the general public for identifying and responding to health misinformation on social media.

Methods: Semistructured interviews were conducted with 22 respondents from the Malaysian general public. The theory of motivated information management was used as a guiding framework for conducting the interviews. Audio-taped interviews were transcribed verbatim and imported into ATLAS.ti software for analysis. Themes were identified from the qualitative data using a thematic analysis method.

Results: The 3 main themes identified were emotional responses and impacts of health misinformation, approaches used to identify health misinformation, and responses to health misinformation. The spread of health misinformation through social media platforms has caused uncertainty and triggered a range of emotional responses, including anxiety and feelings of vulnerability, among respondents who encountered it. The approaches to identifying health misinformation on social media included examining message characteristics and sources. Messages were deemed to be misinformation if they contradicted credible sources or exhibited illogical and exaggerated content. Respondents described multiple response approaches to health misinformation based on the situation. Verification was chosen if the information was deemed important, while misinformation was often ignored to avoid conflict. Respondents were compelled to take action if misinformation affected their family members, had been corrected by others, or if they were knowledgeable about the topic. Taking action involved correcting the misinformation and reporting the misinformation to relevant social media, enforcement authorities, and government bodies.

Conclusions: This study highlights the factors and motivations influencing the general public's identification and response to health misinformation on social media. Addressing the challenges of health misinformation identified in this study requires collaborative efforts from all stakeholders to reduce the spread of health misinformation and reduce the general public's belief in it.

(JMIR Infodemiology 2025;5:e67464) doi:10.2196/67464

KEYWORDS

approaches; response; general public; health misinformation; social media

XSL•FO
RenderX

Introduction

Background

Social media has increasingly been used by the general public for health-related purposes, primarily for receiving social support and searching for and sharing health information [1]. Social media platforms offer unlimited access to prompt and easily accessible health information, making them a preferred channel for seeking health information [2]. The most commonly used social media platforms for this purpose include WhatsApp (WhatsApp LLC), YouTube (Google LLC), Facebook (Meta Platforms, Inc), and X (formerly known as Twitter; X Corp) [2,3]. Studies have shown that between 76% and 85% of respondents, including the general public and patients aged ≥18 years from Saudi Arabia and the United States, search for health information on social media [2,4].

Although social media may offer some advantages in the dissemination of health information, there is growing concern about the prevalence of health misinformation on these platforms. Health misinformation is defined as a health-related claim that is false or misleading because of a lack of supporting scientific evidence at a given time [5]. Studies have shown that the quality of health information on social media is generally poor, based on criteria such as content accuracy, information design, credibility, disclosure of user information, and interactivity [6]. A systematic review found that the 3 most common topics with health misinformation on social media were vaccines (32%), drugs and smoking issues (22%), and noncommunicable diseases (19%) [5].

Health misinformation can be harmful as it can cause unnecessary fear and lead the general public to make inappropriate decisions about their health, which can worsen physical health and even lead to increased morbidity and mortality [7-10]. One example is the widespread antivaccine content on social media that has contributed to decreases in vaccine acceptance and vaccination rates as well as an increase in preventable disease outbreaks [7]. Physicians have also expressed concerns about the dangers of health misinformation, describing regular encounters with patients who were hesitant to take potentially lifesaving medications or adhere to prescribed treatments owing to misinformation on the internet and social media [8]. Moreover, the narratives found in health misinformation often instill fear, anxiety, and mistrust in health institutions [9]. For example, health misinformation circulating on social media during the Ebola virus outbreak created hostility toward health care workers and contributed to challenges in efforts to control the epidemic [11]. In addition, misinformation may cause the general public to have reduced adherence to government public health policies, rendering these efforts ineffective [12-16].

Previous research has demonstrated that various factors can influence the perceived credibility of health misinformation on social media. These factors may collectively contribute to the rapid spread of misinformation, particularly during health crises such as the COVID-19 pandemic. One identified factor is the inclusion of claims of background evidence, such as an attached link or source, in misinformation messages, which can make messages appear more credible to the general public [17]. Social endorsements, such as a higher number of likes and shares, also enhance believability [18-20]. The perceived credibility of the message's source is also important, with individuals tending to trust messages from respected authorities, close contacts, influencers, or celebrities, leading to further misinformation spread [20,21]. Health care professionals and organizations are often perceived as reliable sources of information; however, they are also exposed to misinformation on social media [17,22]. A qualitative study among physicians and nurses in the United States identified several cues for spotting health misinformation, including messages with exaggerated claims, grammar errors, unreliable links, and conspiracy content [22].

The general public's response to health misinformation on social media, if they are able to identify it, has been shown to be influenced by their beliefs, cultural factors, and social media use. Beliefs in conspiracy theories, such as government secrets, and cultural beliefs, such as trust in traditional treatments over scientific evidence, contribute to the spread and acceptance of misinformation, as observed during the COVID-19 pandemic [23,24]. General public reactions to misinformation on social media also vary; some people choose to disregard the information [25], while others challenge and report it [26]. Fact-checking has been shown to reduce the spread of misinformation, but heavy reliance on social media as information sources reduces the critical verification process, leading to the further spread of health misinformation [27].

Due to the urgency of this issue, various measures have been proposed and adopted by organizations, governments, health care workers, and researchers to combat health misinformation on social media. These measures include regulating social media content, skillfully communicating public health messages, promoting verification and correcting health misinformation, warning about the sources of information, promoting evidence-based medicine, educating the general public about health information, conducting fact-checking efforts, improving critical thinking skills, and enhancing media and health literacy [10,28-31]. For example, the World Health Organization (WHO), recognizing the mounting concerns about health misinformation on social media, has adopted numerous measures to combat it, such as providing an avenue for reporting health misinformation, creating a MythBusters web page to debunk misinformation with facts and figures, and working toward amendments in social media policies [32,33]. While this is recognized as a pressing issue requiring a multidisciplinary effort, effectively addressing it requires a deeper understanding of how the general public identifies and responds to health misinformation on social media platforms. This understanding will enable the identification of areas requiring concerted efforts from all stakeholders to mitigate the spread of health misinformation and reduce the general public's belief in it.

While previous studies on health misinformation have primarily focused on regions such as Europe, the Middle East, Africa, and parts of Asia, Southeast Asia remains underrepresented in this field of research [34,35]. Nevertheless, there has been notable rise in health misinformation on social media within the South East Asian region [35]. Given that culture plays a significant role in how health misinformation is perceived and

XSL•FO
RenderX

managed [36-38], it is important to explore this issue within a Southeast Asian context. As a multicultural and multilingual country, Malaysia offers a unique setting for such studies [39]. The diverse cultural backgrounds and traditions within the Malaysian population can influence health information–seeking behaviors [40] and provide additional insights into health misinformation on social media. However, no studies have specifically examined how the Malaysian general public identifies and responds to health misinformation on social media. Therefore, this study aims to explore the approaches used by the Malaysian general public to recognize health misinformation on social media and their responses to it.

Theory of Motivated Information Management

The theory of motivated information management (TMIM) serves as a guiding framework in this study. The TMIM explains how uncertainty influences information management behaviors [41] and has been widely applied to explore the barriers and motivations behind health information seeking [42-45]. While its application to health misinformation is relatively limited, it may offer valuable insights into how individuals engage with, process, and respond to health misinformation encountered on social media. It can also help explain how feelings of uncertainty drive health information–seeking behaviors in the context of health misinformation [46].

This theory is structured into 3 phases: interpretation, evaluation, and decision. In the context of health misinformation on social media, the interpretation phase occurs when individuals recognize discrepancies or uncertainties in health information or misinformation they encounter. This recognition may trigger emotional responses, such as anxiety, which can influence their approach to managing the information. Next is the evaluation phase, where individuals assess their ability to reduce uncertainty through health-related information seeking and consider the potential outcomes—both positive and negative—of obtaining new information. The likelihood of seeking information is influenced by individuals' perceived ability to cope with the expected outcomes (coping efficacy), their ability to manage the communication process involved in seeking information (communication efficacy), and the perceived reliability of the information source (target efficacy). Finally, the decision phase determines the individuals' course of action based on the evaluations from the previous phase, such as whether to search for additional health information or disregard the health misinformation encountered [41].

Methods

Study Design and Participants

This qualitative study involved individual semistructured interviews with members of the general public in Malaysia. This study design was selected to allow a deeper exploration of the factors influencing the management of health misinformation in this population, providing richer insights into the issue.

The respondents were initially recruited through advertisements posted on social media platforms, including Facebook, Instagram (Meta Platforms, Inc), WhatsApp, and X. Those who were interested in participating completed a Google Form providing their sociodemographic information and details about their social media activity. Potential respondents who met the inclusion criteria, which included being Malaysian, aged >18 years, and active social media users on the specified platforms, were then contacted. Individuals who were unable to participate in web-based interviews or could not speak English or Malay were excluded. For this study, active social media users were defined as individuals who engaged with content by liking, commenting, and sharing information on Facebook, Instagram, X, and WhatsApp. These social media platforms were selected because they are the most commonly used for seeking and sharing health information in Malaysia [47].

The respondents were then purposively selected by the researchers to ensure a diversity of sociodemographic characteristics, such as age, education, occupation, income, gender, state of residence, and religion. This was done to gather a broad range of views and perspectives on the issue of health misinformation in Malaysian social media and the responses toward it.

Ethical Considerations

This study was approved by the Human Research Ethics Committee of Universiti Kebangsaan Malaysia (JEP-2022-037). Before conducting the interviews, the purpose of the interviews was explained to the participants, and their written informed consent was obtained. To ensure privacy and confidentiality, all transcripts were anonymized through pseudonymization. Each participant was assigned a unique identifier, and all potentially identifying information was removed during transcription. Only the research team had access to the anonymized data, which was stored on a secure, password-protected server. Following the interviews, RM 50 (US $11) was credited into the e-wallets of the respondents as compensation for their participation in this study.

Data Collection

An interview guide was developed based on domains identified through literature review and guided by the TMIM framework, as outlined earlier [41]. Table 1 presents examples of interview questions aligned with the 3 key phases of the TMIM. In the interpretation phase, questions were designed to explore the types of health misinformation encountered on social media; the uncertainty recognized in the information; and the emotional responses triggered, such as anxiety. In the evaluation phase, interview questions focused on understanding how individuals identify health misinformation, their perceptions of social media as a source of health information, and their confidence in seeking accurate information and assessing its reliability. Finally, questions in the decision phase explored how individuals respond to health misinformation based on their evaluations, including disregarding the misinformation or using strategies for verifying health information. The interview guide was piloted with 2 respondents and amended to improve clarity and length. The pilot interviews were not included in the final analysis.

XSL•FO
RenderX

Table 1. Examples of questions used during the interview.

TMIM[a] phases [41]	Description	Examples of questions
Interpretation	Individuals recognize discrepancies or uncertainties in health misinformation, which may trigger emotional responses.	• Have you ever encountered information that conflicts with your beliefs? • If so, what was the information? • How did you feel upon encountering this (mis)information? • What was your initial reaction? • What do you think are the risks of health misinformation?
Evaluation	Individuals assess their ability to reduce uncertainty through health information seeking and evaluate potential outcomes.	• What would you do when you are unsure about health information encountered on social media, and why? • How did you determine whether it was health misinformation? • How do you assess whether an information source on social media is credible or not? How confident are you in your evaluation?
Decision	Individuals decide on the course of action based on their evaluations.	• What would you do if you discovered that health information encountered on social media is misinformation? • How would you respond, and why? • How would responding this way make you feel? • Would you seek further information, and why?

[a]TMIM: theory of motivated information management.

The individual interviews were conducted online via Zoom (Zoom Video Communications) between June 2022 and February 2023. The interviews were conducted in English and Malay. Each interview session lasted between 35 and 62 minutes, with an average duration of 40 minutes. All interviews conducted were audio recorded with the respondents' consent, and field notes were taken during the interview sessions. The audio recordings of the interviews were transcribed verbatim. Interviews conducted in Malay were first transcribed in Malay and then translated into English, with the translations reviewed by bilingual research team members. The transcripts were assigned a code, and potentially identifying information was removed. The transcripts were anonymized through a pseudonymization procedure. Data collection was conducted until data saturation was reached, which was determined by the repetition of themes and the absence of new insights. This was assessed collectively by the research team, who continuously compared data from interviews to identify recurring themes among respondents. Data collection continued until all emerging themes were fully explored and stopped when no new codes or themes emerged from the last 3 interviewees [48].

Data Analysis

The transcripts were imported into ATLAS.ti (ATLAS.ti Scientific Software Development GmbH) to facilitate the coding process and identify themes from the qualitative data. The reflexive thematic analysis method by Braun and Clarke [49] was used, following 6 phases, which included familiarization with the data, identifying initial codes, searching for themes, reviewing themes, classifying themes, and generating reports. Initially, inductive coding was conducted independently by SS and WWC to derive preliminary codes from the data. These initial codes were subsequently reviewed, compared, and discussed in depth during multiple iterative meetings between the researchers to ensure rigor and consistency. Following this, abductive coding, guided by the TMIM framework, was incorporated to interpret and refine the codes and themes. Specifically, TMIM was applied to interpret and refine the categorization of themes by incorporating insights into how individuals manage uncertainty and seek or avoid information. This theoretical lens helped clarify the motivations behind participants' behaviors, ensuring that the themes captured both empirical patterns and underlying motivational processes [50]. Final codes were determined through discussion-based consensus. Themes and subthemes were identified based on significant patterns observed in the data and were continuously revised and refined using the constant comparison approach [51]. Field notes were used to aid early analysis, and reflexivity was maintained by the researchers throughout the interviews and analysis phases via memos collected during and after the interviews [52]. All themes and subthemes were finalized through discussions among research team members until a consensus was achieved.

Results

Overview

A total of 22 respondents participated in this study. Respondents represented a diverse range of demographics, including age groups (ranging from 19 to 54 years), levels of education (ranging from secondary school to postgraduate degrees), incomes (ranging from <RM 1000 to >RM 6000), and professional backgrounds (including unemployed, self-employed, private sector, and government) and came from different states across Malaysia. The detailed characteristics of the respondents are included in Table 2.

XSL•FO
RenderX

Table 2. Respondents' demographics.

Respondent number	Age	Sex	Education	Employment sector	Family income[a]
1	27	Female	First degree	Private sector	RM 5000-RM 6000
2	32	Male	First degree	Private sector	>RM 6000
3	54	Female	First degree	Government sector	>RM 6000
4	36	Female	Postgraduate	Government sector	>RM 6000
5	19	Female	Diploma (postsecondary level)	None	>RM 6000
6	25	Male	First degree	Private sector	RM 3000-RM 4000
7	31	Female	First degree	Private sector	>RM 6000
8	23	Female	First degree	Private sector	>RM 6000
9	30	Male	Postgraduate	None	<RM 1000
10	49	Female	Diploma	Government sector	RM 4000-RM 5000
11	19	Male	Diploma	None	RM 3000-RM 4000
12	40	Male	First degree	Self-employed	RM 5000-RM 6000
13	29	Female	First degree	Government sector	>RM 6000
14	22	Female	Secondary School	Private sector	RM 1000-RM 2000
15	33	Female	Postgraduate	Private sector	>RM 6000
16	32	Male	Postgraduate	Government sector	RM 5000-RM 6000
17	37	Female	Diploma	Private sector	RM 2000-RM 3000
18	31	Female	First degree	Private sector	<RM 1000
19	23	Male	Secondary School	Private sector	RM 4000-RM 5000
20	34	Male	First degree	Private sector	>RM 6000
21	22	Male	First degree	None	RM 2000-RM 3000
22	29	Male	Postgraduate	None	<RM 1000

[a]RM 50=US $11.

In total, 3 main themes were identified and will be discussed in detail in subsequent paragraphs with representative quotes: emotional responses and impacts of health misinformation, approaches used to identify health misinformation, and responses to health misinformation.

Emotional Responses and Impacts of Encountering Health Misinformation on Social Media

Respondents generally regarded health misinformation as a serious issue with important and potentially harmful consequences. Many reported encountering such misinformation across various social media platforms, including WhatsApp, X, Facebook, and TikTok (ByteDance). The topics described included misleading claims about COVID-19 (eg, consumption of clove water, coconut water, or black pepper water as cures for COVID-19), conspiracy theories regarding health treatments (eg, pharmaceutical companies concealing cancer cures), vaccinations (eg, vaccines causing autism or COVID-19 vaccines containing tracking devices), and pseudoscientific claims related to traditional supplements (eg, turmeric curing all ailments).

Respondents described a variety of emotions when they encountered health misinformation on social media. They felt uncertainty because of the conflicting information being spread, leading to worry as they were unsure which was correct. In addition, when they came across health misinformation on social media, they experienced self-doubt, questioning their own knowledge as the health misinformation conflicted with their own understanding of the health issue at hand:

> *This is worrying as one person says something and another person says something else. So, we are not sure if it is right or wrong.* [Respondent 11]

> *It would get me thinking that whatever I studied, is it wrong?* [Respondent 7]

They explained that they felt scared as they saw even health care professionals spread health misinformation on social media, especially during the peak of the COVID-19 pandemic when everyone was uncertain about what was happening. This further led to feelings of vulnerability as they were not from a health care background, and with health care professionals themselves believing health misinformation, this put them in a more vulnerable position:

> *I saw videos with healthcare workers saying that vaccination could make our immune system worse. This is scary for me as if they can be uncertain about it, what about those who are not from a healthcare background?* [Respondent 7]

XSL•FO
RenderX

Besides that, some respondents described feeling angry as they found the health misinformation spreading on social media to be ridiculous. They further explained that this health misinformation was selling hope to those who were sick, which they felt was very irresponsible. Furthermore, another respondent noted that during the peak of the COVID-19 pandemic, when vaccination was promoted by authorities as a solution, the spread of antivaxxers on social media led to anxiety, as they felt they were exposed to the virus when they came across them in a public area:

> *Who says coconut water can kill Covid-19? They make me so angry.* [Respondent 1]
>
> *How can this kind of food recover the function of the kidneys? It's irresponsible to be spreading this information. They are selling hope. Those who are sick are willing to do anything or spend money to get their health back.* [Respondent 12]
>
> *It gives us anxiety when faced with anti-vaxxers, especially at the peak of the pandemic.* [Respondent 2]

Some respondents described feeling disappointed as they were lied to with promised recovery through miraculous treatment options that failed:

> *You just feel you're being lied to, and then you get your hopes down.* [Respondent 22]

However, a few respondents indicated that misinformation was rare and not serious or significant. This was because they rarely encountered misinformation and had only mild or no effects:

> *It is rare that there is misinformation at all.* [Respondent 18]

Approaches Used to Identify Health Misinformation

Respondents indicated the approaches they used to identify health misinformation, which can be divided into 2 subthemes: message characteristics and the source of the message on social media.

Message Characteristics

Respondents examined the characteristics of social media messages to decide whether they were health misinformation. This comprised the message content and layout.

Respondents decided that it could be misinformation when claims were illogical and exaggerated or when the message claimed that the product is a magical cure for every ailment and was too good to be true. In addition, it was potentially misinformation if the tips or treatment suggested were too simple compared with the severity of the illness at hand, whereby the information provided was deemed to be clearly false:

> *If it is either too good to be true or sounds ridiculous.* [Respondent 2]
>
> *It is not logical, how can this type of herb cure everything?* [Respondent 12]

Some respondents also expressed doubt about the products being promoted, as they lacked proper dosage guidelines in comparison to medications provided by health care professionals. One respondent highlighted that messages that appear visually unattractive should be investigated further, as official information would be created with much thought and graphics to ensure appropriate dissemination of information to the public. In addition, messages circulated on WhatsApp with the caption *forwarded many times* should be approached with caution:

> *As most of the time, it is simply forwarded many times. When we see this tag on WhatsApp, it has a high chance of being something that is circulated for the sake of circulating, while nobody actually knows whether it is correct.* [Respondent 22]

Source of Message

Overview

Respondents highlighted that health misinformation was commonly encountered on blogs, Facebook, X, and WhatsApp groups. They identified several key factors related to the source of a message when determining whether health information was misinformation. These factors can be grouped into 3 main areas: verifying misinformation through source credibility, which focuses on the objective validation of the source based on institutional backing or official endorsements; social trust and misinformation on social media, which emphasizes the role of perceived trustworthiness and social influence of the message source; and sociodemographic influences on source credibility, which considers how factors such as education level, geographic location, and age shape the perceived credibility of the source.

Verifying Misinformation Through Source Credibility

The perceived credibility of the information source was a major factor in identifying health misinformation. Respondents consistently mentioned that health misinformation was often associated with sources that lacked verification or evidence-based backing. Government bodies, such as the Malaysian Ministry of Health, and verified organizations (eg, nongovernmental organizations with blue-ticked social media accounts) were seen as more trustworthy. Messages that contradicted or lacked support from these credible sources were more likely to be considered misinformation. For instance, a lack of official approval or research backing from the Ministry of Health, or promotion of unregistered products, were key indicators that the information could be false. In addition, information not reported by traditional media sources was also deemed as possible misinformation:

> *Suppose they have at least a blue tick source, I will be more confident as some government pages and NGOs have a blue tick. You know they are verified and that gives a bit more trust factor. But if it is only some small shop that sells health ointments or whatever, like nothing is credible, no information; then for those, I will definitely have to search for extra information.* [Respondent 1]
>
> *Okay, I know it is [misinformation] because I never found the information in the newspaper or news, and the Honorable Minister of Health has never allowed its usage in health.* [Respondent 9]

XSL•FO
RenderX

Social Trust and Misinformation on Social Media

In general, respondents highlighted that health misinformation mainly comes from individuals rather than organizations, and any individual could potentially spread misinformation, including vaccine opponents, family members, friends, celebrities, health care professionals, and business-oriented individuals. Furthermore, those who lack knowledge of how to obtain accurate information were considered potential sources of misinformation. Those spreading misinformation about health were described as having a strong belief in their stance:

> *They have a lot of followers, it's like cults. For example, if they post something, and you comment with scientific evidence, they will [continue to] reject because they have been brainwashed.* [Respondent 17]

> *They do not know where to find the correct information, so whatever information they get, they feel it is true.* [Respondent 20]

Respondents highlighted important characteristics related to social trust in the source of a message, which significantly influenced their perception of health misinformation. They noted that individuals who have a strong social media presence and a trustworthy image are more likely to be believed. This trust is particularly important when assessing the credibility of health information. Respondents indicated that individuals with strong educational backgrounds, professional qualifications, or verified social media profiles were seen as more reliable, thus emphasizing the image portrayed on social media being of utmost importance. Conversely, individuals lacking these credentials were viewed as less trustworthy.

Sociodemographic Influences on Source Credibility

Respondents also discussed how sociodemographic factors influenced the perceived credibility of sources of health misinformation. Education level was identified as a key factor, with respondents perceiving that people with lower educational backgrounds were more likely to spread misinformation. Geographic location, particularly rural areas, was another significant influence, with respondents believing that people in rural areas had limited access to reliable health information. Age was also mentioned, with older individuals often seen as sources of misinformation, although respondents acknowledged that misinformation could originate from any age group:

> *Those less privileged, less educated, or [from] rural areas, they tend to share a lot of fake news.* [Respondent 13]

> *Especially parents, those in their 50s, share a lot of [such information], but they don't really know whether it is true or not.* [Respondent 22]

Challenges Associated With Identification of Health Misinformation

Respondents indicated that identifying misinformation becomes difficult when a message possesses certain characteristics. These characteristics included lengthy posts and messages that spread virally. Furthermore, messages that contain anecdotes and testimonials appeared convincing, as they had been tried by many people with positive results. Similarly, messages and videos that feature endorsements from academicians or health care professionals for treatments were also deemed challenging to differentiate. Some respondents perceived that they could trust the message if it went viral. Hence, viral messages can create confusing outcomes for the general public, as some believe that the message is reliable, while others believe that it might be spreading misinformation:

> *It looks like something genuine because they include anecdotes and testimonials. They include videos like I've tried this.... So, when you have these kinds of anecdotes, people are influenced by the power of emotions.* [Respondent 2]

> *They quote doctors claiming the latest evidence shows that high cholesterol is normal, safe, and can prolong lifespan. So, the public will believe they can lead a sedentary life and consume a high-fat meal.* [Respondent 4]

Responses To Health Misinformation

Responses to health misinformation can be divided into 3 themes: ignoring, verifying health information, and taking action.

Ignoring

Many respondents disclosed that they often chose to simply ignore or disregard health misinformation and cited several reasons for this.

Some respondents mentioned that they would ignore the misinformation when they are uncertain about the health topic, if it was controversial, or if it came from unknown strangers. In addition, many respondents believed that they lacked sufficient expertise in health matters to offer informed opinions and, thus, preferred not to risk appearing ignorant on social media. Some were also concerned that challenging the misinformation might lead the perpetrator to question their qualifications, and that could backfire:

> *I want to fix [correct] it. But the netizens will comment "Who are you to tell this?" So, I leave it alone.* [Respondent 10]

> *Because we do not have the knowledge, I am a public person, so we follow those who have studied. When they argue these types of topics, we simply leave it.* [Respondent 19]

Another reason cited by respondents for ignoring health misinformation was to avoid potential trouble and conflict. They expressed concerns that challenging the perpetrator might lead to cyberbullying or negative reactions. Others expressed feelings of fear or concern about offending or embarrassing the person sharing the misinformation, which led them to refrain from challenging the content, highlighting how emotions, such as empathy, fear, and worry, influenced their decision to remain passive. Furthermore, 1 respondent highlighted that they would ignore the misinformation if there were many supporters for those sharing misinformation:

> *I will not interfere with them. If the poster [person posting the misinformation] wants to look for trouble*

XSL•FO
RenderX

> *with you or the people who follow that person are really supportive of that person, then I am going to be in trouble.* [Respondent 22]

> *If there are many people in the group that you do not know and you do not know how they will react, I will just ignore them. It is because I try to avoid conflict and trouble.* [Respondent 22]

A few respondents mentioned that they would ignore the misinformation when it does not affect them personally. They believed that it would require too much effort to try to correct the misinformation and that they would not be able to make a difference in any case. Some simply believed that someone else would correct it, while others felt that it was the responsibility of the government to take corrective actions:

> *A lot of people seem to believe that [misinformation]. So, you cannot really change their minds.* [Respondent 1]

> *It is a mentality of how much difference am I going to make?* [Respondent 7]

In addition, 1 respondent believed that the perpetrator's opinion should be respected as a form of free expression allowed on social media. They added that condemning people for their opinions is not an acceptable practice.

Verifying Health Information

Most respondents described verifying health-related information that they come across on social media and shared various reasons and methods they used to verify the information. Verification usually involves checking both the content of the information and the background of the individual or organization sharing it. Many indicated the importance of verification, even if the source of the information is a health care professional or family member:

> *When they post something, I will search about it first, even if my family members share something.* [Respondent 14]

Respondents reported that they were more likely to check information if it pertained to a serious issue that could potentially affect them. They also mentioned that they would verify information if they were uncertain about it, for instance, if the information was difficult to understand or if they had never heard about it before. They were also more likely to verify information that came from fewer sources or sources with questionable reliability and trustworthiness. In addition, respondents said that they would verify information if conflicting information was presented or if they were considering spending money, for example, on a product advertised:

> *I will check if I never heard about it before.* [Respondent 21]

> *I would look for more information if it affects me or I have to spend money for it.* [Respondent 3]

> *When they post, usually I check their backup sources, from where they share their information...At least if there is research or link to reliable sources. I do not like it when they share something not to be done, without any evidence.* [Respondent 13]

When evaluating the credibility of health-related information, respondents typically used multiple sources, including social media, health care professionals, health organizations, traditional media, family members, acquaintances, websites, and studies. Some respondents mentioned that they could identify false information because the correct information had already gone viral. If the respondents still had doubts, they simply visited a doctor to clarify the information. The respondents emphasized the importance of being vigilant and having critical thinking skills to distinguish between true and false information:

> *I will read a lot more than just Facebook. Maybe I will Google first and then read. I will read the type of feedback provided by other people.* [Respondent 17]

> *I will check with doctor friends, pharmacists, and nurses.* [Respondent 9]

Taking Action

This theme relates to respondents describing the actions that they would take to address health misinformation on social media. If respondents were to take action, it would be either to report, block, or correct misinformation. Respondents also discussed who should be responsible for correcting the health misinformation.

Many respondents mentioned that they would report the misinformation despite perceiving limited regulations on social media. The reporting actions included notifying relevant social media administrators; law enforcement agencies, such as the National Pharmaceutical Regulatory Agency; and the Ministry of Health. Interestingly, 1 respondent said that they would tag the police department. These efforts were aimed at ensuring the misinformation was removed from the respective social media platforms:

> *If I am sure it is wrong, there is a button where you can press report. Yeah, I will just do that.* [Respondent 14]

Some respondents indicated that they did not want to ruin their social media page with content that they disagreed with. Thus, they would either mute or block the offender's account, while others deleted those who posted misinformation from groups.

Many respondents believed that correcting misinformation is important and expressed their willingness to do so if they encounter false information. However, respondents mentioned that they were more likely to provide corrections in certain situations. Generally, they were inclined to correct misinformation if they felt there could be no conflict or direct confrontation as a result. For example, if they knew the person posting misinformation, they would feel comfortable providing corrections, driven by a sense of familiarity and trust. They felt safe correcting misinformation through personal WhatsApp messages or group chats if they were familiar with everyone in the group. In addition, respondents described feeling more confident in correcting misinformation if they were knowledgeable about the topic and if others had already

XSL•FO
RenderX

corrected it before them. They were also more motivated to provide corrections and encourage others to verify the information source if they believed it could affect their family members, indicating how emotional factors such as care and responsibility can motivate actions against misinformation:

> *If there are people close to me, for example a group of five people that I am close to, close friends, then it will be easy for me to correct them.* [Respondent 22]

> *If there are others who have commented it is wrong, then I will comment. Otherwise, I will not.* [Respondent 14]

> *If it affects my family, I would shut it down immediately. I will tell them that they would want to check the sources. But when it comes to other people, I mostly would not budge [react].* [Respondent 7]

When correcting misinformation, respondents provided justifications for why the information was false and sought to persuade using data and facts. A few respondents also mentioned using general terms and simple language, similar to that used by those disseminating false information. In addition, respondents shared information from sources deemed reliable, including social media sources with blue verification ticks, to correct misinformation:

> *Try to speak things in terms of their perspective, like put yourself in their shoes and try to speak in their language. However, if things fail, then try to convince with facts and figures.* [Respondent 2]

> *Take the evidence from the blue tick source.* [Respondent 21]

Respondents described experiencing mixed outcomes when correcting misinformation. Some family members were thankful and believed the corrected information, while others deleted the false information post. Unfortunately, some respondents faced reprimand and subsequently chose to disregard it. Despite the possibility of disbelief from others, they believed it was their duty to educate them regarding the truth. However, some respondents reported that their comments went unanswered, and in some cases, the offenders shared even more questionable links. Nevertheless, it was emphasized that respondents had done their part in correcting the inaccurate information.

Who Should Correct Health Misinformation on Social Media

The respondents believed that the government should take action against the spread of misinformation. In addition, it was suggested that the government should adopt measures to educate the general public about the dangers of misinformation and to be cautious before sharing health information on social media. Furthermore, respondents proposed enforcing legal provisions and regulations against those spreading health misinformation. An interesting suggestion was made that the Ministry of Health could use nudges on social media, such as "Are you sure about this?" to encourage the general public to consider the validity of health information before sharing it:

> *I hope the government of Malaysia will tighten the law for those who like to spread wrong information, because if they do not, they will keep doing it.* [Respondent 19]

> *For example, Ministry of Health can comment "Are you sure this fact is true?" and people will be thinking, okay this may not be true because Ministry of Health commented.* [Respondent 22]

Respondents also recommended that the Ministry of Health establish a multidisciplinary team dedicated to verifying and correcting misinformation on social media. They also highlighted an opportunity to improve accessibility by providing government health information in languages beyond English and Malay to better accommodate a multicultural population. Furthermore, participants emphasized the importance of providing timely and transparent explanations for changes in health-related information, as observed during the COVID-19 pandemic, when frequent updates were necessary because of emerging evidence:

> *They should have a team or department that targets or even tracks all these rumors being spread around.* [Respondent 7]

> *I wish they would have more languages. At this point, it is mostly in English and Malay, and there are many elderly people who do not understand.* [Respondent 7]

> *During the Covid-19 pandemic, initially they [government] provided a recommendation [on] vaccines...then they provided different information...but they never explained whether the vaccines were really effective.* [Respondent 6]

> *The public may not understand that knowledge is progressive. Initially, the government recommended two doses to achieve herd immunity, but when cases were rising, they introduced booster doses. The public will hold on to the recommendations made two years ago.* [Respondent 16]

In addition to the government, respondents also believed that health care professionals have a responsibility to combat misinformation. It was suggested that health care professionals should be more proactive in correcting misinformation on social media and consider reaching out to populations considered vulnerable:

> *Healthcare workers such as doctors and pharmacists should correct people's understanding on medication. If more people share on social media, it is better so that the wrong and true information are in the same quantity.* [Respondent 13]

> *When doctors post on health information, they have to think how to reach those susceptible to this [misinformation].* [Respondent 13]

XSL•FO
RenderX

Discussion

Principal Findings

The spread of health misinformation through social media platforms has led to uncertainty and evokes a range of emotional responses in those who encounter it. This study provides insights into the various approaches used by the general public to identify health misinformation on social media as well as the reasons and motivations behind their responses to it. In addition, several challenges faced by the general public in identifying and addressing misinformation were identified.

The first phase of TMIM describes the uncertainty caused by health misinformation. Our study reflects this, showing that misinformation spread through social media platforms led to uncertainties and triggered emotional responses, such as worry, fear, self-doubt, anger, feelings of vulnerability, disappointment, and anxiety. Studies in other countries also showed this, where misinformation in South Africa caused panic, confusion, and anxiety [53]. Similarly, studies in Jordan and Spain reported elevated levels of anxiety in response to health misinformation [13,54]. These emotions significantly influence how individuals process and respond to misleading health information. Higher levels of anxiety and fear, for example, are associated with increased belief and willingness to share misinformation [55,56]. Anger, which often arises when users feel deceived by false claims, also contributes to intuitive actions and the further spread of misinformation [57,58]. Fear and anxiety tend to intensify when misinformation relates to health threats, increasing concerns and feelings of vulnerability over personal and general public health [59]. Uncertainty and self-doubt emerge when users encounter conflicting information, leading to cognitive dissonance that hinders informed decision-making [60]. Such emotional distress can lead to a general feeling of disappointment, especially when people realize that they have been misled by trusted sources or when accurate information is overshadowed by misinformation [57]. The cumulative effect of these emotions can hinder effective health decision-making and reduce trust in health information sources [12-15]. This may influence the evaluation phase, where individuals analyze and attempt to identify health misinformation on social media platforms.

Previous research has identified both internal and external factors that individuals use to determine the credibility of the information [61-63]. Internal factors include elements such as the source of the message, message characteristics, and individual personality traits. By contrast, external factors include institutional sources or interpersonal networks, such as verification from family, friends, and trusted institutions. This study also described the internal and external factors used by respondents to assess credibility and the challenges faced in identifying misinformation. One significant internal factor is the source of the message. Trust in government and health care professionals emerged as an important component, highlighting how credibility is tied to recognized, authoritative sources. This aligns with the MAIN (modality, agency, interactivity, and navigability) model, which suggests that technological aspects of digital media can influence credibility judgments, particularly when information comes from official authorities [64]. Supporting this, a study conducted in Malaysia found that respondents mainly selected the Ministry of Health as their preferred source of health information [65]. This trust may stem from cultural norms in Malaysia, where the conservative Asian context fosters greater respect for figures of authority [66]. Furthermore, respondents believed that the government has a critical role to play in combating this problem by implementing regulations and awareness campaigns. Beyond institutional sources, respondents also expressed trust in messages from those they believed to have trustworthy reputations and relevant educational backgrounds. This was also similar in India, with respondents placing trust in messages from local government representatives and community health workers deemed to have a wide knowledge of the area [62].

In contrast to previous studies conducted in Malaysia and the United States, where most respondents were unable to evaluate the accuracy of health-related information, our study identified specific characteristics of messages that respondents associated with misinformation [67,68]. Respondents pointed out that exaggerated or illogical claims, such as those promising miracle cures, were indicative of misinformation. This was similarly seen in a survey study in Austria, where exaggerated claims were met with skepticism [69]. Conversely, social media messages that were visually appealing were often perceived as more credible by respondents. Research has indicated that visual design plays a role in influencing people’s judgments about the credibility of health information found on the web [70]. This may be because of the attractive design of these messages, such as well-crafted infographics, which suggest that effort and thought have been invested in creating them, which in turn increases their trustworthiness [71].

Correction directed to those sharing the misinformation and individuals looking at misinformation was considered a strategy for combating health misinformation [30]. A content analysis of monkeypox on Instagram found that one-third of the content was debunked, with social media users actively correcting the misinformation [72]. It was interesting that in our study, many respondents chose to ignore misinformation, citing avoidance of conflict and perceived futility in correcting it. This behavior aligns with findings from another Malaysian study, which showed that most respondents tended to ignore fake news during the COVID-19 pandemic [68]. Such tendencies may be influenced by cultural norms, as previous research has found that Malaysians often prefer to avoid conflicts [73]. Nevertheless, this finding is also consistent with other studies [62,74-76], including a UK study that found avoidance of conflict to be a major barrier in preventing the correction of COVID-19 misinformation [76]. However, despite these challenges, some respondents in our study were willing to take action when their loved ones were affected or when they had a close relationship with those posting misinformation. This could be beneficial, as the literature suggests that correction by a close tie could be successful [77]. In addition, corrections that focus on collective interests have been shown to work better [78].

Respondents also were more inclined to correct misinformation when they felt that they knew the topic well. This is supported by previous studies where health care workers themselves face

XSL•FO
RenderX

challenges in correcting misinformation because of various factors, such as the belief that there would be no improvement after correction; lack of time to address the issue; fear of retaliation from those posting misinformation; and a lack of support, such as social media training, to handle these situations [79]. They opted to correct misinformation only if it was important to them or if the offender was someone close to them [74]. Furthermore, an experimental study among health experts in China showed that they were willing to correct health misinformation when the perceived threat to readers was severe, when they believed that they had the necessary skills to correct it, and when they wanted to maintain a reputation of kindness toward others [80]. This aligns with a study in India where respondents chose not to correct misinformation if they believed it would not cause harm [62].

Suggestions for Interventions to Tackle Health Misinformation

Multifaceted interventions involving multiple stakeholders, including government, health care providers, researchers, academicians, and social media administrators, are needed to effectively address health misinformation on social media. Previous studies have recommended measures to combat health misinformation in Malaysia, such as legal action against offenders, the establishment of fact-checking portals, and the continuous health information dissemination to the general public [81]. The findings from this study suggest additional measures that can be implemented to further address the spread of misinformation. Respondents identified several challenges in identifying misinformation, including those presented with anecdotes or testimonials. It has been acknowledged that misinformation can spread when compelling anecdotes are presented and data are misrepresented with fake experts [82]. Studies also revealed that correct information and misinformation both contained anecdotes, which explains why the general public may be confused in this regard [83]. This highlights the need for digital literacy training on how to search for, identify, verify, and share credible health information on social media, as indicated by other studies [84-87]. Improvement in digital literacy skills was seen to improve recognition of the quality of health information on the web [88-91]. Educational institutions should also focus on increasing awareness regarding the identification of health misinformation on social media [92-94]. This was shown to lead to fact-checking information before sharing [95].

In addition, messages that are viral need to be monitored, as they tend to cause confusion regarding credibility, as indicated by respondents in this study. Another study supports this, showing that among African American older adults, respondents were more likely to believe a message if they saw the information multiple times [96]. Furthermore, other studies have shown that health misinformation with a greater number of likes is often perceived as more credible, as it was viewed as social endorsements [18-20]. One potential strategy is to assemble a team of health experts to correct misinformation on social media and highlight the truth, with the support of multiple experts to prevent cyberbullying toward them [79,86,97]. In addition, health care workers can address health misinformation by correcting it when they encounter it with patients at clinics or in hospital settings. This can be done by educating the patients and encouraging 2-way communication between patients and health care workers [98,99]. Moreover, the government could focus on "prebunking" by addressing potential areas that could pose challenges in differentiating misinformation before it spreads [100,101]. As "prebunking" effect is limited, this strategy should be combined with other efforts to enhance its effectiveness in combating misinformation [102].

Access to multilingual health information has also been identified as a need, with potential implications for health equity. The WHO has acknowledged multilingualism as an area to improve equality in health information dissemination [103]. However, further investigation is required to determine the extent of implementation across countries and to assess the need for additional languages in health information dissemination to effectively combat misinformation.

Limitations

This study offers insights into public perceptions of health misinformation on social media in a Southeast Asian country, a region with limited research on the topic. The challenges described can be used to develop interventions to address the issue of health misinformation on social media. Nevertheless, this study has some limitations that need to be considered, including the possibility of sampling biases and social desirability biases. Although efforts were made to include a wide range of respondents from various sociodemographic backgrounds, it is possible that certain groups may have been missed, such as those with lower educational backgrounds or older respondents. Further quantitative research is recommended to better understand how different population groups identify and respond to health misinformation. The development and validation of tools to measure this, such as survey-based approaches to validate identified themes, would be valuable for enabling cross-country comparisons regarding responses to health misinformation.

Conclusions

The characteristics of a message and its source on social media are important factors that the public considers when identifying health misinformation. Various reasons and circumstances may affect individual responses toward health misinformation, which range from ignoring it to verifying health information and adopting measures to correct it. Digital literacy training may be useful in addressing the challenges faced by the public in identifying and responding to health misinformation. This study also highlights the need to further investigate populations considered vulnerable or at risk of health misinformation on social media and the factors influencing responses toward health misinformation, which will allow the development of targeted intervention strategies.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the Director General of Health Malaysia for his permission to publish this paper. This study was supported by the Ministry of Higher Education of Malaysia’s Fundamental Research Grant Scheme (grant FRGS/1/2020/SS0/UKM/02/11).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical concerns but are available from the corresponding author on reasonable request.

Authors' Contributions

SS was involved in the conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing (original draft), and writing (review and editing). AMA was involved in conceptualization, methodology, supervision, validation, and writing (review and editing). WWC was involved in conceptualization, data curation, data analysis, funding acquisition, methodology, project administration, resources, software, supervision, validation, visualization, writing (original draft), and writing (review and editing).

Conflicts of Interest

None declared.

References

1. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917 [FREE Full text] [doi: 10.2196/17917] [Medline: 33978589]
2. Sumayyia MD, Al-Madaney MM, Almousawi FH. Health information on social media. Perceptions, attitudes, and practices of patients and their companions. Saudi Med J 2019 Dec;40(12):1294-1298 [FREE Full text] [doi: 10.15537/smj.2019.12.24682] [Medline: 31828284]
3. AlMuammar SA, Noorsaeed AS, Alafif RA, Kamal YF, Daghistani GM. The use of internet and social media for health information and its consequences among the population in Saudi Arabia. Cureus 2021 Sep;13(9):e18338 [FREE Full text] [doi: 10.7759/cureus.18338] [Medline: 34722089]
4. Neely S, Eldredge C, Sanders R. Health information seeking behaviors on social media during the COVID-19 pandemic among American social networking site users: survey study. J Med Internet Res 2021 Jun 11;23(6):e29802 [FREE Full text] [doi: 10.2196/29802] [Medline: 34043526]
5. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
6. Afful-Dadzie E, Afful-Dadzie A, Egala SB. Social media in health communication: a literature review of information quality. Health Inf Manag 2023 Jan 04;52(1):3-17. [doi: 10.1177/1833358321992683] [Medline: 33818176]
7. Steffens MS, Dunn AG, Wiley KE, Leask J. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. BMC Public Health 2019 Oct 23;19(1):1348 [FREE Full text] [doi: 10.1186/s12889-019-7659-3] [Medline: 31640660]
8. Hill JA. Medical misinformation: vet the message!. Pacing Clin Electrophysiol 2019 Mar 08;42(3):299-300. [doi: 10.1111/pace.13616] [Medline: 30735253]
9. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
10. Kisa S, Kisa A. A comprehensive analysis of COVID-19 misinformation, public health impacts, and communication strategies: scoping review. J Med Internet Res 2024 Aug 21;26:e56931 [FREE Full text] [doi: 10.2196/56931] [Medline: 39167790]
11. Allgaier J, Svalastog AL. The communication aspects of the Ebola virus disease outbreak in Western Africa--do we need to counter one, two, or many epidemics? Croat Med J 2015 Oct;56(5):496-499 [FREE Full text] [doi: 10.3325/cmj.2015.56.496] [Medline: 26526888]
12. Nelson T, Kagan N, Critchlow C, Hillard A, Hsu A. The danger of misinformation in the COVID-19 crisis. Mo Med 2020;117(6):510-512 [FREE Full text] [Medline: 33311767]
13. Sallam M, Dababseh D, Yaseen A, Al-Haidar A, Taim D, Eid H, et al. COVID-19 misinformation: mere harmless delusions or much more? A knowledge and attitude cross-sectional study among the general public residing in Jordan. PLoS One 2020;15(12):e0243264 [FREE Full text] [doi: 10.1371/journal.pone.0243264] [Medline: 33270783]
14. Shah SM, Alpern JD, Stauffer WM. Medical misinformation targets vulnerable populations and threatens the nation's health. Minn Med 2017;100(5):22-24.

XSL•FO
RenderX

15. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman AL, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020 Oct;7(10):201199 [FREE Full text] [doi: 10.1098/rsos.201199] [Medline: 33204475]
16. Al-Aghbari AA, Hassan OE, Dar Iang M, Jahn A, Horstick O, Dureab F. Exploring the role of infodemics in people's incompliance with preventive measures during the COVID-19 in conflict settings (mixed method study). Healthcare (Basel) 2023 Mar 26;11(7):952 [FREE Full text] [doi: 10.3390/healthcare11070952] [Medline: 37046879]
17. Bapaye JA, Bapaye HA. Demographic factors influencing the impact of coronavirus-related misinformation on WhatsApp: cross-sectional questionnaire study. JMIR Public Health Surveill 2021 Jan 30;7(1):e19858 [FREE Full text] [doi: 10.2196/19858] [Medline: 33444152]
18. Ali K, Li C, Zain-ul-abdin K, Zaffar MA. Fake news on Facebook: examining the impact of heuristic cues on perceived credibility and sharing intention. Internet Res 2021 Jul 28;32(1):379-397. [doi: 10.1108/intr-10-2019-0442]
19. Lee H, Kim J. Norm perceptions about rumor sharing on genetically modified foods: the interaction between Facebook likes and a refuting comment. Health Commun 2023 Jun;38(7):1467-1476. [doi: 10.1080/10410236.2021.2012956] [Medline: 34894923]
20. Apuke OD, Omar B. Modelling the antecedent factors that affect online fake news sharing on COVID-19: the moderating role of fake news knowledge. Health Educ Res 2020 Oct 01;35(5):490-503 [FREE Full text] [doi: 10.1093/her/cyaa030] [Medline: 33090215]
21. Shin I, Wang L, Lu Y. Twitter and endorsed (fake) news: the influence of endorsement by strong ties, celebrities, and a user majority on credibility of fake news during the COVID-19 pandemic. Int J Commun 2022;16:4-23 [FREE Full text]
22. Bautista JR, Zhang Y, Gwizdka J. Healthcare professionals' acts of correcting health misinformation on social media. Int J Med Inform 2021 Apr;148:104375. [doi: 10.1016/j.ijmedinf.2021.104375] [Medline: 33461008]
23. Barua Z, Barua S, Aktar S, Kabir N, Li M. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. Prog Disaster Sci 2020 Dec;8:100119 [FREE Full text] [doi: 10.1016/j.pdisas.2020.100119] [Medline: 34173443]
24. Chen K, Luo Y, Hu A, Zhao J, Zhang L. Characteristics of misinformation spreading on social media during the COVID-19 outbreak in China: a descriptive analysis. Risk Manag Healthc Policy 2021;14:1869-1879 [FREE Full text] [doi: 10.2147/RMHP.S312327] [Medline: 34007225]
25. Bahl R, Chang S, McKay D, Buchanan G. Health misinformation across multiple digital ecologies: qualitative study of data from interviews with international students. J Med Internet Res 2022 Jul 06;24(7):e38523 [FREE Full text] [doi: 10.2196/38523] [Medline: 35727960]
26. Stewart R, Madonsela A, Tshabalala N, Etale L, Theunissen N. The importance of social media users' responses in tackling digital COVID-19 misinformation in Africa. Digit Health 2022 Mar 18;8:20552076221085070 [FREE Full text] [doi: 10.1177/20552076221085070] [Medline: 35321021]
27. Melki J, Tamim H, Hadid D, Makki M, El Amine J, Hitti E. Mitigating infodemics: the relationship between news exposure and trust and belief in COVID-19 fake news and social media spreading. PLoS One 2021 Jun 4;16(6):e0252830 [FREE Full text] [doi: 10.1371/journal.pone.0252830] [Medline: 34086813]
28. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. Health Commun 2018 Sep;33(9):1131-1140. [doi: 10.1080/10410236.2017.1331312] [Medline: 28622038]
29. Xiao X. Let's verify and rectify! Examining the nuanced influence of risk appraisal and norms in combatting misinformation. New Media Soc 2022 Jun 28;26(7):3786-3809. [doi: 10.1177/14614448221104948]
30. Vraga EK, Bode L. Correction as a solution for health misinformation on social media. Am J Public Health 2020 Oct;110(S3):S278-S280. [doi: 10.2105/AJPH.2020.305916] [Medline: 33001724]
31. Thompson EL, Preston SM, Francis JK, Rodriguez SA, Pruitt SL, Blackwell J, et al. Social media perceptions and internet verification skills associated with human papillomavirus vaccine decision-making among parents of children and adolescents: cross-sectional survey. JMIR Pediatr Parent 2022 Sep 14;5(3):e38297 [FREE Full text] [doi: 10.2196/38297] [Medline: 36103216]
32. Combatting misinformation. World Health Organization. URL: https://www.who.int/teams/digital-health-and-innovation/digital-channels/combatting-misinformation-online [accessed 2025-05-29]
33. How to report misinformation. World Health Organization. URL: https://www.who.int/campaigns/connecting-the-world-to-combat-coronavirus/how-to-report-misinformation-online [accessed 2025-05-29]
34. Hu B, Liu X, Lu C, Ju X. Prevalence and intervention strategies of health misinformation among older adults: a meta-analysis. J Health Psychol 2024 Nov 28:13591053241298362. [doi: 10.1177/13591053241298362] [Medline: 39607815]
35. Dang HL. Social media, fake news, and the COVID-19 pandemic: sketching the case of southeast Asia. Austrian J South East Asian Stud 2021;14(1):37-57. [doi: 10.14764/10.ASEAS-0054]
36. Kainja J, Makhumula C, Twabi HS, Gunde AM, Ndasauka Y. Cultural determinants of COVID-19 vaccines misinformation in Malawi. Humanit Soc Sci Commun 2024 Sep 12;11(1):1193. [doi: 10.1057/S41599-024-03743-6]
37. Chandrasekaran R, Sadiq TM, Moustakas E. Racial and demographic disparities in susceptibility to health misinformation on social media: national survey-based analysis. J Med Internet Res 2024 Nov 06;26:e55086 [FREE Full text] [doi: 10.2196/55086] [Medline: 39504121]

XSL•FO
RenderX

38. Kbaier D, Kane A, McJury M, Kenny I. Prevalence of health misinformation on social media-challenges and mitigation before, during, and beyond the COVID-19 pandemic: scoping literature review. J Med Internet Res 2024 Aug 19;26:e38786 [FREE Full text] [doi: 10.2196/38786] [Medline: 39159456]
39. Mohamadali NA, Oun CT, Ahmad M, Mohammed SA. Information needs and seeking behavior of patients and carers in palliative care: a case study from public hospital Malaysia. Indian J Sci Technol 2018;11(20):1-7 [FREE Full text] [doi: 10.17485/ijst/2018/v11i20/123588]
40. Dawood OT, Hassali MA, Saleem F, Ibrahim IR, Abdulameer AH, Jasim HH. Assessment of health seeking behaviour and self-medication among general public in the state of Penang, Malaysia. Pharm Pract (Granada) 2017;15(3):991 [FREE Full text] [doi: 10.18549/PharmPract.2017.03.991] [Medline: 28943981]
41. Afifi WA, Weiner JL. Toward a theory of motivated information management. Commun Theory 2004 May;14(2):167-190. [doi: 10.1111/j.1468-2885.2004.tb00310.x]
42. Hovick SR. Understanding family health information seeking: a test of the theory of motivated information management. J Health Commun 2014;19(1):6-23. [doi: 10.1080/10810730.2013.778369] [Medline: 24117214]
43. Droser VA. Information seeking and grief in bereaved emerging adults: a test of the theory of motivated information management. South Commun J 2019 Oct 12;85(1):28-42. [doi: 10.1080/1041794X.2019.1678194]
44. Kanter E, Bevan JL, Dorros SM. The use of online support groups to seek information about chronic illness: applying the theory of motivated information management. Commun Q 2018 Nov 07;67(1):100-121. [doi: 10.1080/01463373.2018.1539406]
45. Afifi WA, Weiner JL. Seeking information about sexual health: applying the theory of motivated information management. Human Comm Res 2006 Jan;32(1):35-57. [doi: 10.1111/j.1468-2958.2006.00002.x]
46. Sentell T, Saiki K, Abrams MA, Jones A, Melendez K, Chevrolet J, et al. Building community health literacy with a social network perspective. Eur J Public Health 2023;33(Suppl2):ckad160. [doi: 10.1093/eurpub/ckad160.048]
47. Siti Zuhaida H, Chung HC, Mohd Said F, Tumingan K, Sahar Shah N, Hanim S, et al. The best online tools based on media preference reflected by health information received on social media amongst diabetic patients in hospital Canselor Tuanku Muhriz, Kuala Lumpur, Malaysia. Malays J Med Sci 2021 Jun;28(3):118-128 [FREE Full text] [doi: 10.21315/mjms2021.28.3.11] [Medline: 34285650]
48. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
49. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
50. Thompson J. A guide to abductive thematic analysis. Qual Rep 2022 May 20;27(5):1410-1421. [doi: 10.46743/2160-3715/2022.5340]
51. Glaser BG. The constant comparative method of qualitative analysis. Soc Probl 1965 Apr;12(4):436-445. [doi: 10.2307/798843]
52. Phillippi J, Lauderdale J. A guide to field notes for qualitative research: context and conversation. Qual Health Res 2018 Feb 05;28(3):381-388. [doi: 10.1177/1049732317697102] [Medline: 29298584]
53. Schmidt T, Cloete A, Davids A, Makola L, Zondi N, Jantjies M. Myths, misconceptions, othering and stigmatizing responses to COVID-19 in South Africa: a rapid qualitative assessment. PLoS One 2020;15(12):e0244420 [FREE Full text] [doi: 10.1371/journal.pone.0244420] [Medline: 33351852]
54. Ruiz-Frutos C, Ortega-Moreno M, Dias A, Bernardes JM, García-Iglesias JJ, Gómez-Salgado J. Information on COVID-19 and psychological distress in a sample of non-health workers during the pandemic period. Int J Environ Res Public Health 2020 Sep 24;17(19):6982 [FREE Full text] [doi: 10.3390/ijerph17196982] [Medline: 32987712]
55. Freiling I, Krause NM, Scheufele DA, Brossard D. Believing and sharing misinformation, fact-checks, and accurate information on social media: the role of anxiety during COVID-19. New Media Soc 2023 Jan 22;25(1):141-162 [FREE Full text] [doi: 10.1177/14614448211011451] [Medline: 36620434]
56. Rosenzweig LR, Bago B, Berinsky AJ, Rand DG. Happiness and surprise are associated with worse truth discernment of COVID-19 headlines among social media users in Nigeria. Harv Kennedy Sch Misinformation Rev 2021 Aug 10;2(4):1-37. [doi: 10.37016/mr-2020-75]
57. Zhao X, Shaw J, Ma Z. How individuals cope with anger- and sadness-induced narrative misinformation on social media: roles of transportation and correction. Online Media Glob Commun 2024;3(3):418-446. [doi: 10.1515/omgc-2024-0021]
58. Dong W, Tao J, Xia X, Ye L, Xu H, Jiang P, et al. Public emotions and rumors spread during the COVID-19 epidemic in China: web-based correlation study. J Med Internet Res 2020 Nov 25;22(11):e21933 [FREE Full text] [doi: 10.2196/21933] [Medline: 33112757]
59. Ren SY, Gao RD, Chen YL. Fear can be more harmful than the severe acute respiratory syndrome coronavirus 2 in controlling the corona virus disease 2019 epidemic. World J Clin Cases 2020 Feb 26;8(4):652-657 [FREE Full text] [doi: 10.12998/wjcc.v8.i4.652] [Medline: 32149049]

XSL•FO
RenderX

60. Dhawan D, Bekalu M, Pinnamaneni R, McCloud R, Viswanath K. COVID-19 news and misinformation: do they matter for public health prevention? J Health Commun 2021 Nov 02;26(11):799-808. [doi: 10.1080/10810730.2021.2010841] [Medline: 34865604]
61. Tandoc EC, Ling R, Westlund O, Duffy A, Goh D, Zheng Wei L. Audiences' acts of authentication in the age of fake news: a conceptual framework. New Media Soc 2017 Sep 21;20(8):2745-2763. [doi: 10.1177/1461444817731756]
62. Varanasi RA, Pal J, Vashistha A. Accost, accede, or amplify: attitudes towards COVID-19 misinformation on WhatsApp in India. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA p. 1-17 URL: https://dl.acm.org/doi/10.1145/3491102.3517588 [doi: 10.1145/3491102.3517588]
63. P Brandtzaeg PB, Lüders M, Spangenberg J, Rath-Wiggins L, Følstad A. Emerging journalistic verification practices concerning social media. J Pract 2015 Mar 13;10(3):323-342. [doi: 10.1080/17512786.2015.1020331]
64. Sundar SS. The MAIN model: a heuristic approach to understanding technology effects on credibility. In: Metzger MJ, Flanagin AJ, editors. Digital Media, Youth, and Credibility. Cambridge, MA: The MIT Press; 2008:73-100.
65. Mohamad E, Tham JS, Ayub SH, Hamzah MR, Hashim H, Azlan AA. Relationship between COVID-19 information sources and attitudes in battling the pandemic among the Malaysian public: cross-sectional survey study. J Med Internet Res 2020 Nov 12;22(11):e23922 [FREE Full text] [doi: 10.2196/23922] [Medline: 33151897]
66. Jaafar NI, Ainin S, Yeong MW. Why bother about health? A study on the factors that influence health information seeking behaviour among Malaysian healthcare consumers. Int J Med Inform 2017 Aug;104:38-44. [doi: 10.1016/j.ijmedinf.2017.05.002] [Medline: 28599815]
67. Stimpson JP, Ortega AN. Social media users' perceptions about health mis- and disinformation on social media. Health Aff Sch 2023 Oct;1(4):qxad050 [FREE Full text] [doi: 10.1093/haschl/qxad050] [Medline: 38107206]
68. Balakrishnan V. Socio-demographic predictors for misinformation sharing and authenticating amidst the COVID-19 pandemic among Malaysian young adults. Inf Dev 2022 Aug 04;40(2):319-331. [doi: 10.1177/02666669221118922]
69. Neureiter A, Stubenvoll M, Kaskeleviciute R, Matthes J. Trust in science, perceived media exaggeration about COVID-19, and social distancing behavior. Front Public Health 2021 Dec 1;9:670485 [FREE Full text] [doi: 10.3389/fpubh.2021.670485] [Medline: 34926361]
70. Robins D, Holmes J, Stansbury M. Consumer health information on the web: the relationship of visual design and perceptions of credibility. J Am Soc Inf Sci 2009 Dec 09;61(1):13-29. [doi: 10.1002/asi.21224]
71. Piil K, Pedersen P, Gyldenvang HH, Elsborg AJ, Skaarup AB, Starklint M, et al. The development of medical infographics to raise symptom awareness and promote communication to patients with cancer: a co-creation study. PEC Innov 2023 Dec;2:100146 [FREE Full text] [doi: 10.1016/j.pecinn.2023.100146] [Medline: 37214530]
72. Dalton ME, Duffy R, Quinn E, Larsen K, Peters C, Brenner D, et al. A qualitative review of social media sharing and the 2022 monkeypox outbreak: did early labelling help to curb misinformation or fuel the fire? Sex Health 2024 Feb;21:66. [doi: 10.1071/SH23158] [Medline: 38295805]
73. Awalluddin MA, Maznorbalia AS. A systematic literature review of the preferred organizational conflict management styles in Malaysia. Environ Soc Psychol 2023 Aug 03;8(2):1597. [doi: 10.54517/esp.v8i2.1597]
74. Tandoc EC, Lim D, Ling R. Diffusion of disinformation: how social media users respond to fake news and why. Journalism 2019 Aug 07;21(3):381-398. [doi: 10.1177/1464884919868325]
75. Tully M, Bode L, Vraga EK. Mobilizing users: does exposure to misinformation and its correction affect users' responses to a health misinformation post? Soc Media Soc 2020 Dec 10;6(4):205630512097837. [doi: 10.1177/2056305120978377]
76. Chadwick A, Vaccari C, Hall NA. Covid vaccines and online personal messaging: the challenge of challenging everyday misinformation. Online Civic Culture Center. URL: https://www.lboro.ac.uk/media/media/research/o3c/pdf/Chadwick-Vaccari-Hall-Covid-Vaccines-and-Online-Personal-Messaging-2022.pdf [accessed 2025-05-29]
77. Margolin DB, Hannak A, Weber I. Political fact-checking on Twitter: when do corrections have an effect? Political Communication 2017 Sep 05;35(2):196-219. [doi: 10.1080/10584609.2017.1334018]
78. Xiao X, Borah P, Lee DK, Su Y, Kim S. A story is better told with collective interests: an experimental examination of misinformation correction during the COVID-19 pandemic. Am J Health Promot 2023 Sep;37(7):915-923 [FREE Full text] [doi: 10.1177/08901171231184075] [Medline: 37309829]
79. Bautista JR, Zhang Y, Gwizdka J. US physicians' and nurses' motivations, barriers, and recommendations for correcting health misinformation on social media: qualitative interview study. JMIR Public Health Surveill 2021 Sep 01;7(9):e27715 [FREE Full text] [doi: 10.2196/27715] [Medline: 34468331]
80. Chen L, Tang H. Intention of health experts to counter health misinformation in social media: effects of perceived threat to online users, correction efficacy, and self-affirmation. Public Underst Sci 2023 Apr;32(3):284-303. [doi: 10.1177/09636625221138357] [Medline: 36515488]
81. Mohammad NH, Chinnasamy S, Faizal S, Abdul Karim Zamri NA. Fake news and misinformation: COVID-19 and challenges confronted by Malaysian's Ministry of Health. Int J Res Bus Soc Sci 2022 May 09;12(5):1276-1295. [doi: 10.6007/ijarbss/v12-i5/13010]
82. Diethelm P, McKee M. Denialism: what is it and how should scientists respond? Eur J Public Health 2009 Jan;19(1):2-4. [doi: 10.1093/eurpub/ckn139] [Medline: 19158101]

XSL•FO
RenderX

83. Luo H, Cai M, Cui Y. Spread of misinformation in social networks: analysis based on Weibo tweets. Secur Commun Netw 2021 Dec 16;2021:1-23. [doi: 10.1155/2021/7999760]
84. Anthonysamy L, Sivakumar P. A new digital literacy framework to mitigate misinformation in social media infodemic. Glob Knowl Mem Commun 2022 Nov 14;73(6/7):809-827. [doi: 10.1108/gkmc-06-2022-0142]
85. Bicher LA, Fathy SA. Infodemic and digital literacy: the role of digital literacy in combating misinformation of COVID-19 on Facebook. Arab J Media Commun Res 2024(35):100-131. [doi: 10.21608/JKOM.2021.222480]
86. Smith R, Chen K, Winner D, Friedhoff S, Wardle C. A systematic review of COVID-19 misinformation interventions: lessons learned. Health Aff (Millwood) 2023 Dec;42(12):1738-1746. [doi: 10.1377/hlthaff.2023.00717] [Medline: 37967291]
87. Hyatt A, Shelly A, Cox R, Humphries E, Lock G, Varlow M. How can we improve information for people affected by cancer? A national survey exploring gaps in current information provision, and challenges with accessing cancer information online. Patient Educ Couns 2022 Aug;105(8):2763-2770 [FREE Full text] [doi: 10.1016/j.pec.2022.04.009] [Medline: 35465976]
88. Moretti V, Brunelli L, Conte A, Valdi G, Guelfi MR, Masoni M, et al. A web tool to help counter the spread of misinformation and fake news: pre-post study among medical students to increase digital health literacy. JMIR Med Educ 2023 Apr 18;9:e38377 [FREE Full text] [doi: 10.2196/38377] [Medline: 36996010]
89. Hadley M, Patil U, Colvin KF, Sentell T, Massey PM, Gallant M, et al. Exploring the impact of digital health literacy on COVID-19 behaviors in New York state college students during the COVID-19 pandemic. Comput Methods Programs Biomed Update 2023;4:100126. [doi: 10.1016/j.cmpbup.2023.100126]
90. Li X, Liu Q. Social media use, eHealth literacy, disease knowledge, and preventive behaviors in the COVID-19 pandemic: cross-sectional study on Chinese netizens. J Med Internet Res 2020 Oct 09;22(10):e19684 [FREE Full text] [doi: 10.2196/19684] [Medline: 33006940]
91. Spracklin E, Espina CR. SIFT-ing through misinformation: reflection notebooks to teach critical digital literacy. Nurse Educ 2025;50(2):106-110. [doi: 10.1097/NNE.0000000000001760] [Medline: 39471397]
92. Confronting health misinformation: the U.S. surgeon general's advisory on building a healthy information environment. Office of the Surgeon General (OSG), US Department of Health and Human Services. URL: https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf [accessed 2025-05-29]
93. Farr C. How to spot fake health information and root out charlatans, according to health experts. CNBC. URL: https://www.cnbc.com/2020/06/04/how-to-spot-fake-health-info-according-to-experts.html [accessed 2025-05-29]
94. Alsaad E, AlDossary S. Educational video intervention to improve health misinformation identification on WhatsApp among Saudi Arabian population: pre-post intervention study. JMIR Form Res 2024 Jan 17;8:e50211 [FREE Full text] [doi: 10.2196/50211] [Medline: 38231563]
95. Wu Y. Predicting fact-checking health information before sharing among people with different levels of altruism: based on the influence of presumed media influence. Psychol Res Behav Manag 2023;16:1495-1508 [FREE Full text] [doi: 10.2147/PRBM.S404911] [Medline: 37138700]
96. Seo H, Blomberg M, Altschwager D, Vu HT. Vulnerable populations and misinformation: a mixed-methods approach to underserved older adults' online information assessment. New Media Soc 2020 May 27;23(7):2012-2033 [FREE Full text] [doi: 10.1177/1461444820925041]
97. Jain S, Dhaon SR, Majmudar S, Zimmermann LJ, Mordell L, Walker G, et al. Empowering health care workers on social media to bolster trust in science and vaccination during the pandemic: making IMPACT using a place-based approach. J Med Internet Res 2022 Oct 17;24(10):e38949 [FREE Full text] [doi: 10.2196/38949] [Medline: 35917489]
98. Ismail N, Kbaier D, Farrell T, Kane A. The experience of health professionals with misinformation and its impact on their job practice: qualitative interview study. JMIR Form Res 2022 Nov 02;6(11):e38794 [FREE Full text] [doi: 10.2196/38794] [Medline: 36252133]
99. Chou WY, Gaysynsky A, Vanderpool RC. The COVID-19 misinfodemic: moving beyond fact-checking. Health Educ Behav 2021 Feb;48(1):9-13 [FREE Full text] [doi: 10.1177/1090198120980675] [Medline: 33322939]
100. Basol M, Roozenbeek J, Berriche M, Uenal F, McClanahan WP, Linden SV. Towards psychological herd immunity: cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. Big Data Soc 2021 May 11;8(1):205395172110138. [doi: 10.1177/20539517211013868]
101. Jolley D, Douglas KM. Prevention is better than cure: addressing anti‐vaccine conspiracy theories. J Applied Social Pyschol 2017 Jun 28;47(8):459-469. [doi: 10.1111/jasp.12453]
102. Bruns H, Dessart FJ, Krawczyk M, Lewandowsky S, Pantazi M, Pennycook G, et al. Investigating the role of source and source trust in prebunks and debunks of misinformation in online experiments across four EU countries. Sci Rep 2024 Sep 05;14(1):20723 [FREE Full text] [doi: 10.1038/s41598-024-71599-6] [Medline: 39237648]
103. Multilingualism. World Health Organization. URL: https://www.who.int/about/policies/multilingualism [accessed 2025-05-29]

Abbreviations

TMIM: theory of motivated information management
WHO: World Health Organization

XSL•FO
RenderX

Edited by T Purnat; submitted 12.10.24; peer-reviewed by D Kbaier, M Vivion; comments to author 30.11.24; revised version received 24.01.25; accepted 28.02.25; published 24.06.25.

Please cite as:
Sathianathan S, Mhd Ali A, Chong WW
How the General Public Navigates Health Misinformation on Social Media: Qualitative Study of Identification and Response Approaches
JMIR Infodemiology 2025;5:e67464
URL: https://infodemiology.jmir.org/2025/1/e67464
doi:10.2196/67464
PMID:40554777

©Sharmila Sathianathan, Adliah Mhd Ali, Wei Wen Chong. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 24.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Quality and Reliability of Online Videos as an Information Source of Public Health Education for Stroke Prevention in Mainland China: Electronic Media–Based Cross-Sectional Study

Rongguang Ge, MD; Haoyi Dai, MD; Chicheng Gong, MD; Yuhong Xia, MD; Rui Wang, MD; Jiaping Xu, MD; Shoujiang You, MD, PhD; Yongjun Cao, MD, PhD

Department of Neurology and Clinical Research Center of Neurological Disease, The Second Affiliated Hospital of Soochow University, 1055 Sanxiang Road, Suzhou, China

Corresponding Author:
Yongjun Cao, MD, PhD
Department of Neurology and Clinical Research Center of Neurological Disease, The Second Affiliated Hospital of Soochow University, 1055 Sanxiang Road, Suzhou, China

Abstract

Background: Stroke has become a leading cause of death and disability worldwide, resulting in a significant loss of healthy life years and imposing a considerable economic burden on patients, their families, and caregivers. However, despite the growing role of online videos as an emerging source of health information, the credibility and quality of stroke prevention education videos, especially those in Chinese, remain unclear.

Objective: This study aims to assess the basic characteristics, overall quality, and reliability of Chinese-language online videos related to public health education on stroke prevention.

Methods: We systematically searched and screened stroke prevention–related video resources from 4 popular Chinese domestic video platforms (Bilibili, Douyin, Haokan, and Xigua). General information, including upload date, duration, views, likes, comments, and shares, was extracted and recorded. Two validated evaluation tools were used: the modified DISCERN questionnaire to assess content reliability and the Global Quality Scale (GQS) to evaluate overall quality. Finally, Spearman correlation analysis was conducted to examine potential associations between general video metrics and their quality and reliability.

Results: After searching and screening, a total of 313 eligible videos were included for analysis: 68 from Bilibili, 74 from Douyin, 86 from Haokan, and 85 from Xigua. Among these, 113 (36.1%) were created by health care professionals, followed by news agencies (n=95, 30.4%) and general individual users (n=40, 12.8%). The median scores for the modified DISCERN and GQS were 2 and 3, respectively, suggesting that the included stroke prevention–related videos were relatively unreliable and of moderate quality. Most videos focused on primary stroke prevention and commonly recommended adopting a healthy diet; engaging in physical activity; and managing blood pressure, glucose, and lipid levels. Additionally, videos with longer durations and more comments tended to be more reliable and of higher quality. A positive association was also observed between video quality and reliability.

Conclusions: Overall, the quality and reliability of Chinese-language online videos as a source of stroke prevention information remain unsatisfactory and should be approached with caution by viewers. To address this issue, several measures should be implemented, including establishing an online monitoring and correction system, strengthening the video review process through collaboration with health care professionals, and encouraging more selective and cautious sharing of controversial content. These steps are essential to help curb the spread of online misinformation and minimize its ongoing impact.

(JMIR Infodemiology 2025;5:e64891) doi:10.2196/64891

KEYWORDS

credibility; quality; online video; stroke; prevention; public health education

Introduction

Stroke, which is primarily classified into ischemic and hemorrhagic types, is an extremely lethal and disabling neurological disease, accounting for approximately 7.3 million deaths and 160.5 million years of healthy life lost worldwide in 2021 alone [1]. China, as one of the largest developing countries, faces an increasingly severe challenge from the burden of stroke due to its large population and rapid aging [2].

In the pursuit of better combating this growing threat, the primary prevention of first stroke and secondary prevention of recurrent stroke have emerged as critical strategies and are now recognized as essential components of global stroke control [3,4]. As a result, prevention-oriented public and patient

XSL•FO
RenderX

education is increasingly emphasized, with various specific educational methods, such as standardized courses [5], peer education [6], and community-based programs [7], proven to significantly improve stroke management quality and help mitigate rising medical expenditures.

Over the past few decades, the continuous development of the internet and the growing ownership of mobile devices have inevitably created new opportunities for education on cerebrovascular disease. In particular, online video-sharing platforms, known for their easy accessibility and user-friendly interfaces, have significantly transformed how health information is disseminated and accessed by the public [8,9]. Compared with traditional educational methods involving 1-on-1 teaching and print materials, online video-based disease education may offer a distinct advantage by presenting graphical content that is easier for audiences to absorb and retain. For example, YouTube (Google LLC/Alphabet Inc) has been found to provide useful video resources that compile information about stroke treatment options for patients and their families [10].

However, despite these benefits, the drawbacks of online video-based patient education cannot be ignored. Videos created by amateurs may contain numerous rumors, hoaxes, and misinformation that lack a solid scientific basis [11,12]. Consequently, it remains an arduous and complicated task for most patients with stroke and general users to fully evaluate the overall quality and reliability of stroke-related videos automatically recommended by search engines [13]. Given that Google (Alphabet Inc) has withdrawn its operations from Mainland China, YouTube has been unavailable to the majority of the Chinese population since 2010. Nevertheless, some Chinese domestic video-sharing platforms, such as TikTok (ByteDance Ltd), have gained tremendous popularity across the nation with their engaging content and user interaction.

Although several previous studies have investigated the quality of stroke-related materials on YouTube, TikTok, and similar platforms, showing variable reliability and adherence to evidence-based information, most of this literature originates from English-speaking contexts [10,14-17]. For instance, a systematic review by Garg et al [18] assessed social media's role in stroke education globally, while other studies have evaluated video content acquired in countries such as India [14] and Turkey [17]. However, research focusing on Chinese-language video content remains limited, despite the heavy burden of stroke and the widespread use of social media in China.

Therefore, our research team aimed to systematically assess the overall quality and reliability of online videos related to stroke prevention in Mainland China by searching for and extracting relevant Chinese-language videos from the 4 most popular online video platforms. This study seeks to address the existing academic gap and help future viewers more appropriately select stroke-related video resources.

Methods

Search Strategy

In this cross-sectional study based on electronic media, the following search terms were used: "中风" "脑卒中" (Chinese for "stroke"), "脑梗死" (Chinese for "ischemic stroke"), "脑出血" (Chinese for "hemorrhagic stroke"), and "预防" (Chinese for "prevention"). These terms were used to retrieve the top 100 videos automatically recommended by the default overall ranking on 4 popular Chinese-language online video-sharing platforms: Bilibili (Bilibili Inc/Shanghai Hode Information Technology Co, Ltd/Sony Group Corporation) [19], Douyin (ByteDance Ltd) [20] (the Chinese counterpart of TikTok), Haokan Video (Baidu, Inc) [21], and Xigua Video (ByteDance Ltd) [22]. A brief introduction to these 4 Chinese-language video platforms is provided in Table 1.

Table . A brief introduction to Bilibili, Douyin, Haokan Video, and Xigua Video.

Online video platform	History and introduction
Bilibili	Founded by Shanghai Kuanyu Digital Technology on June 26, 2009, Bilibili has become a cultural community and video platform known for its bullet comment feature and strong popularity among China's younger generation.
Douyin (Chinese counterpart of TikTok)	Founded by ByteDance in 2016, Douyin (the Chinese counterpart of TikTok) is a short video platform aimed at young audiences in China. It encourages users to record and share their daily lives using smartphones or other available filming equipment.
Haokan Video	Founded by Baidu in December 2017, Haokan Video is currently positioned as the "Comprehensive Short Video Flagship" among Baidu's products. The platform is committed to developing a broad short video ecosystem that spans knowledge, lifestyle, health, culture, history, science popularization, technology, emotions, news, film, television, and other domains.
Xigua Video	Founded by ByteDance in 2017, Xigua Video has become a leading domestic professional + user-generated content platform in China. It delivers high-quality content to diverse audiences through personalized recommendations, while promoting creative diversity and making it easy for users to share their videos.

XSL•FO
RenderX

The selection of search keywords was based on both medical terminology and commonly used lay terms for stroke within the Chinese context to ensure relevance and comprehensiveness. This approach aimed to capture a broad spectrum of videos that users might realistically encounter. For example, we included both formal terms such as "脑卒中" (formal medical terminology for "stroke" in the Chinese language) and more colloquial expressions such as "中风" (colloquial description of "stroke" in the Chinese language), which are widely used in public discourse and online media. In addition, we referred to trending tags and frequently searched phrases related to stroke on Chinese video-sharing platforms to enhance the validity of the search keyword set.

With regard to the automatic recommendation and ranking algorithm, consultation with the customer service agents from the selected platforms indicated that the algorithms are primarily driven by video relevance, followed by a comprehensive evaluation based on popularity metrics such as views and likes. However, more detailed information about the algorithm's formulas and methodologies is unavailable due to concerns related to commercial confidentiality.

Video retrieval was conducted on a Windows 10 (Microsoft Corporation) personal computer using a newly installed Mozilla Firefox browser (version 121.0.1; Mozilla Foundation and Mozilla Corporation) in Suzhou, Jiangsu, China, within a single day (December 28, 2023) to minimize potential bias from newly uploaded content. To reduce the influence of personalized recommendation algorithms, all cookies, browsing history, and temporary internet files were cleared, and platform accounts were logged out before and between search queries, which were performed in incognito mode.

Video Selection Criteria

After conducting the searches, the top 100 videos automatically recommended based on the default overall ranking on each video platform were selected and prepared for initial screening. This decision was informed by previous literature indicating that most amateur health seekers rarely browse beyond the first 2 pages of search results [23]. Additionally, most Chinese-language video-sharing platforms limit the number of videos displayed per page to no more than 30. Specifically, Bilibili displays 30 videos per page, Haokan Video 10, Xigua Video 20, and Douyin 20. Therefore, we believe that our sample size (n=100) exceeds the number of videos typically displayed on the first 2 pages (Bilibili, n=60; Haokan Video, n=20; Xigua Video, n=40; and Douyin, n=40), which audiences are most likely to engage with, thereby ensuring representativeness. Furthermore, a sample size of 100 has been widely adopted in similar previous studies [24-26], further supporting the rationale for selecting the top 100 automatically recommended videos for subsequent filtering and analysis.

The exclusion criteria were defined as follows: (1) videos not directly related to stroke prevention; (2) duplicated videos repeatedly uploaded to the same platform; (3) videos not in Chinese (including Mandarin, Cantonese, and other dialects); and (4) commercial advertisements. In this study, 78 off-topic videos and 9 duplicate videos uploaded to the same platform were identified and excluded, resulting in a final sample of 313 eligible videos for further analysis (Figure 1). All eligible videos were produced in Mandarin Chinese, and no highly similar or fully duplicate videos were observed across different platforms.

XSL•FO
RenderX

Figure 1. Flowchart of the retrieval and screening process for Chinese-language online videos related to stroke prevention. GQS: Global Quality Scale.

XSL•FO
RenderX

General Information Extraction

For the included videos, general information, including URLs; authorship; number of views, likes, shares, and comments; upload date; and video length (in seconds), was collected and recorded in an Excel spreadsheet (Microsoft Corporation) by an independent researcher (RGG).

Classification of Videos

The included videos were categorized based on the identity of their producers as follows: (1) general users (eg, amateurs); (2) health professionals (eg, medical practitioners); (3) science communicators (eg, citizen science storytellers); (4) news agencies (eg, CCTV [China Central Television]); (5) medical organizations (eg, hospital); (6) for-profit organizations (eg, commercial health care–related enterprise); and (7) nonprofit organizations (eg, voluntary patient communities).

For videos classified under health professionals, a further subcategorization was conducted as follows: (1) doctors specializing in neurology within modern evidence-based medicine; (2) doctors specializing in other areas of modern evidence-based medicine; (3) doctors of traditional Chinese medicine (TCM); and (4) other health-related professionals, further specified according to their specialty. Detailed definitions of the authorship categories are provided in Table S1 in Multimedia Appendix 1.

Video Evaluation Framework

Quality and Reliability Assessment

After completing the video retrieval, exclusion, and classification, 2 separate questionnaires were used to quantitatively assess the overall quality and reliability of the included online videos on stroke prevention.

The DISCERN instrument, first introduced in 1999 to evaluate the quality of online information on treatment choices [27], has since been widely validated and applied to assess health-related content on video-sharing platforms [28,29]. In 2012, Singh et al [30] developed a modified version of DISCERN, streamlining the original 16-question format into a 5-question version that focuses on clarity, reliability, bias/balance, provision of additional information sources, and acknowledgment of areas of uncertainty. This modified version has also been extensively tested and is commonly used in previous studies [31-35]. Therefore, in this study, we used the modified DISCERN instrument to assess the reliability of video content based on its 5 items (Table 2). The total score of the modified DISCERN tool ranges from 0 to 5, with higher scores indicating greater reliability (from unreliable to reliable). Furthermore, the GQS, a commonly used 5-point scale ranging from 1 (poor quality) to 5 (excellent quality) for evaluating health-related content on the internet [25,36,37], was also applied to assess the overall quality of the included videos (Table 3).

Table . Modified DISCERN questionnaire (1 point for every yes and 0 points for no) items and their descriptions.

Item	Description
1	Are the aims clear and achieved?
2	Are reliable sources of information used?
3	Is the information presented balanced and unbiased?
4	Are additional sources of information listed for patient reference?
5	Are areas of uncertainty mentioned?

Table . Global Quality Scale grades and their descriptions.

Grade	Description
1	Poor quality, poor flow, most information missing; not helpful for patients.
2	Generally poor, some information given; limited use for patients.
3	Moderate quality, some information is adequately discussed, but important topics are missing; somewhat useful to patients.
4	Good quality and flow, most relevant information is covered; useful for patients.
5	Excellent quality and flow; very useful for patients.

Two qualified neurologists (HYD and YHX), both with sufficient medical education and clinical experience, were assigned to independently assess the included online videos using the modified DISCERN and GQS tools. Before formal scoring, both raters thoroughly reviewed and familiarized themselves with the questionnaires to minimize potential biases caused by any misunderstanding of the scoring instruments.

Moreover, the Cohen κ test was used to assess interrater agreement. A value greater than 0.8 was considered indicative of excellent consistency; a value between 0.8 and 0.6 represented substantial consistency; a value between 0.6 and 0.4 indicated moderate consistency; and a value below 0.4 was considered poor consistency [38]. Disagreements between the 2 reviewers were resolved through discussion; if a consensus could not be reached, a senior researcher (YJC) was appointed to make the final decision.

XSL•FO
RenderX

Video Content Evaluation

With regard to the video content related to stroke prevention, 3 key aspects were considered, as outlined below: (1) Which stage of stroke prevention was mentioned (primary, secondary, or both)? (2) What specific preventive measures were recommended (including both pharmacological and nonpharmacological approaches)? (3) What were the frequency and proportion of specific stroke prevention measures recommended in the videos?

As for the rationale behind using this 3-item combined evaluation strategy: Identifying the stage of prevention helps determine whether a video emphasizes preventing the initial onset of stroke or preventing recurrence, an area often overlooked in public education. Examining the specific preventive strategies provides insight into whether the content offers accurate and actionable guidance aligned with current clinical guidelines. Lastly, quantifying the frequency and distribution of these strategies allows for an assessment of the consistency and emphasis of public health messaging across different videos and platforms. Together, these 3 items offer a structured lens through which to assess both the breadth and clinical relevance of video content, aligning with the overall goal of evaluating the educational quality and reliability of Chinese-language videos on stroke prevention.

This procedure was performed by another independent reviewer (CCG), and any uncertainties were resolved through discussion with the research team.

Statistical Analysis

In this study, the Shapiro-Wilk test was used to assess the normality of the quantitative data extracted from the included videos. For nonnormally distributed data, the median and IQR were used for descriptive analysis, whereas for normally distributed data, the mean and SD were reported. Categorical data were presented as absolute numbers and percentages.

Comparisons between 2 groups were performed using the nonparametric Mann-Whitney *U* test, while comparisons among 3 or more groups were conducted using the Kruskal-Wallis *H* test. The Spearman correlation coefficient was then used to examine associations between video variables and quality and reliability scores. For interpretation, a correlation coefficient less than 0.10 was considered negligible; between 0.10 and 0.39, weak; between 0.40 and 0.69, moderate; between 0.70 and 0.89, strong; and greater than 0.90, very strong [39]. Finally, a linear regression model was applied to explore potential temporal trends in the quality and reliability of stroke prevention–related videos over the upload period. A *P* value of <.05 was considered statistically significant. All analyses were performed with the GraphPad Prism software (version 9.5.1; GraphPad Software, Inc).

Ethical Considerations

This study did not involve human participants or experimental animals; therefore, ethics approval and consent to participate were not applicable. Our study involved the analysis of publicly available data or the use of secondary data that do not contain any personal identifiers, thus ensuring the anonymity and confidentiality of the individuals concerned. Compliance with ethical standards was maintained throughout the research process.

Results

Overview

As shown in Figure 1, after the search and screening process, a total of 313 eligible online videos on stroke prevention education were included for further analysis. Of these, 68 were from Bilibili, 74 from Douyin (the Chinese counterpart of TikTok), 86 from Haokan Video, and 85 from Xigua Video.

Overall, as shown in Table 4, the median values among the included videos were as follows: video length, 137 seconds; views, 779; likes, 18; shares, 67; and comments, 1. Regarding the year of video upload, the number of uploaded videos has steadily increased over time, with the majority uploaded in 2023, accounting for 106 out of 313 (33.9%) total videos uploaded. When analyzed by video authorship, health professionals, including doctors specializing in neurology and other fields of modern medicine, doctors of TCM, and pharmacists, collectively created and uploaded 113 stroke prevention–related videos (n=113, 36.1%), followed by news agencies (n=95, 30.4%) and general users (n=40, 12.8%). More detailed descriptive results by video platforms are presented in Table 4.

XSL•FO
RenderX

Table . The general information, producer identity, quality and reliability scores, content, and recommended stroke prevention strategy of online videos uploaded on Bilibili, Douyin, Haokan Video, and Xigua Video.

Video-sharing platforms	Total	Bilibili	Douyin	Haokan	Xigua
General information					
Number, n	313	68	74	86	85
Length (seconds), median (IQR)	137 (81-213.5)	156 (89-279)	88 (55-138.3)	151.5 (94-213.5)	142 (100-253.5)
Views (n), median (IQR)	779 (336-4565)	352.5 (123-1103)	N/A[a]	1299 (577.3-6996)	922 (355-4369)
Likes (n), median (IQR)	18 (2.5-306.5)	3.5 (1-22.25)	662 (281-4477)	10 (3-64.5)	6 (1-44.5)
Shares (n), median (IQR)	67 (5-327.8)	8 (1-38)	172.5 (69.25-906.8)	N/A[b]	N/A[c]
Comments (n), median (IQR)	1 (0-10)	0 (0-2)	24.5 (7.75-106.3)	0 (0-2)	0 (0-2)
Upload year, n (%)					
2023	106 (33.9)	35 (51.5)	38 (51.4)	21 (24.4)	12 (14.1)
2022	68 (21.7)	15 (22.1)	17 (23.0)	17 (19.8)	19 (22.4)
2021	64 (20.4)	11 (16.2)	13 (17.6)	18 (20.9)	22 (25.9)
2020	43 (13.7)	5 (7.4)	5 (6.8)	17 (19.8)	16 (18.8)
2019	23 (7.3)	2 (2.9)	1 (1.4)	12 (14.0)	8 (9.4)
2018	6 (1.9)	0 (0.0)	0 (0.0)	1 (1.2)	5 (5.9)
2017	3 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (3.5)
Authorship, n (%)					
General users	40 (12.8)	16 (23.5)	0 (0.0)	12 (14.0)	12 (14.1)
Science communicators	11 (3.5)	10 (14.7)	1 (1.4)	0 (0.0)	0 (0.0)
Health professionals	113 (36.1)	14 (20.6)	60 (81.1)	12 (14.0)	27 (31.8)
Doctors specializing in neurology of modern medicine	32 (10.2)	3 (4.4)	23 (31.1)	0 (0.0)	6 (7.1)
Doctors specializing in other areas of modern medicine	59 (18.8)	9 (13.2)	33 (44.6)	5 (5.8)	12 (14.1)
Doctors of traditional Chinese medicine	19 (6.1)	2 (2.9)	4 (5.4)	7 (8.1)	6 (7.1)
Pharmacists	3 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (3.5)
Medical organizations	18 (5.8)	7 (10.3)	9 (12.2)	0 (0.0)	2 (2.4)
For-profit organizations	18 (5.8)	6 (8.8)	0 (0.0)	3 (3.5)	9 (10.6)
Nonprofit organizations	18 (5.8)	11 (16.2)	2 (2.7)	1 (1.2)	4 (4.7)
News agencies	95 (30.4)	4 (5.9)	2 (2.7)	58 (67.4)	31 (36.5)
Quality and reliability					
Global Quality Scale score, mean (SD)	3.12 (0.94)	3.41 (0.95)	2.93 (0.83)	2.92 (1.00)	3.26 (0.89)
Global Quality Scale score, median (IQR)	3 (3-4)	3 (3-4)	3 (3-3)	3 (2-3)	3 (3-4)

XSL•FO
RenderX

Video-sharing platforms	Total	Bilibili	Douyin	Haokan	Xigua
DISCERN score, mean (SD)	2.47 (0.92)	2.60 (0.87)	2.70 (0.95)	2.11 (0.85)	2.52 (0.91)
DISCERN score, median (IQR)	2 (2-3)	2.5 (2-3)	3 (2-3)	2 (2-3)	2 (2-3)
Content analysis					
Primary stroke prevention, n (%)					
Physical activity	144 (46.0)	36 (52.9)	32 (43.2)	38 (44.2)	38 (44.7)
Healthy diet	145 (46.3)	39 (57.4)	28 (37.8)	38 (44.2)	40 (47.1)
Control blood pressure	158 (50.5)	31 (45.6)	38 (51.4)	40 (46.5)	49 (57.6)
Control blood glucose	127 (40.6)	27 (39.7)	29 (39.2)	34 (39.5)	37 (43.5)
Control blood lipid	129 (41.2)	29 (42.6)	32 (43.2)	34 (39.5)	34 (40.0)
Reduce weight	73 (23.3)	14 (20.6)	21 (28.4)	17 (19.8)	21 (24.7)
Control blood homocysteine	21 (6.7)	6 (8.8)	4 (5.4)	5 (5.8)	6 (7.1)
Quit smoking	120 (38.3)	34 (50.0)	28 (37.8)	23 (26.7)	35 (41.2)
Quit drinking	99 (31.6)	31 (45.6)	15 (20.3)	22 (25.6)	31 (36.5)
Treat atrial fibrillation	33 (10.5)	8 (11.8)	10 (13.5)	6 (7.0)	9 (10.6)
Treat other heart diseases	34 (10.9)	9 (13.2)	13 (17.6)	6 (7.0)	6 (7.1)
Monitor carotid artery stenosis	19 (6.1)	1 (1.5)	11 (14.9)	1 (1.2)	6 (7.1)
Treat migraine	2 (0.6)	1 (1.5)	1 (1.4)	0 (0.0)	0 (0.0)
Treat obstructive sleep apnea-hypopnea syndrome	4 (1.3)	0 (0.0)	3 (4.1)	0 (0.0)	1 (1.2)
Control hypercoagulable state	5 (1.6)	2 (2.9)	1 (1.4)	0 (0.0)	2 (2.4)
Control inflammation or infection	5 (1.6)	0 (0.0)	3 (4.1)	0 (0.0)	2 (2.4)
Oral antiplatelet drugs	22 (7.0)	4 (5.9)	7 (9.5)	5 (5.8)	6 (7.1)
Traditional Chinese herbs	17 (5.4)	2 (2.9)	2 (2.7)	10 (11.6)	3 (3.5)
Acupuncture and moxibustion	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.2)
Massage	1 (0.3)	0 (0.0)	0 (0.0)	1 (1.2)	0 (0.0)
Avoid sudden changes in temperature	28 (8.9)	12 (17.6)	2 (2.7)	3 (3.5)	11 (12.9)
Remain emotionally stable	68 (21.7)	20 (29.4)	14 (18.9)	14 (16.3)	20 (23.5)
Secondary stroke prevention, n (%)					
Persistent clinical follow-up	2 (0.6)	0 (0.0)	2 (2.7)	0 (0.0)	0 (0.0)
Healthy diet	13 (4.2)	6 (8.8)	2 (2.7)	3 (3.5)	2 (2.4)
Physical activity	14 (4.5)	6 (8.8)	4 (5.4)	3 (3.5)	1 (1.2)

XSL•FO
RenderX

Video-sharing platforms	Total	Bilibili	Douyin	Haokan	Xigua
Quit smoking	13 (4.2)	6 (8.8)	4 (5.4)	2 (2.3)	1 (1.2)
Quit drinking	14 (4.5)	6 (8.8)	5 (6.8)	2 (2.3)	1 (1.2)
Control blood pressure	24 (7.7)	8 (11.8)	7 (9.5)	6 (7.0)	3 (3.5)
Control blood glucose	21 (6.7)	7 (10.3)	7 (9.5)	6 (7.0)	1 (1.2)
Control blood lipid	21 (6.7)	6 (8.8)	7 (9.5)	5 (5.8)	3 (3.5)
Reduce weight	7 (2.2)	2 (2.9)	4 (5.4)	1 (1.2)	0 (0.0)
Control blood homocysteine	1 (0.3)	1 (1.5)	0 (0.0)	0 (0.0)	0 (0.0)
Oral antiplatelet drugs	21 (6.7)	6 (8.8)	6 (8.1)	4 (4.7)	5 (5.9)
Treat atrial fibrillation	7 (2.2)	5 (7.4)	1 (1.4)	1 (1.2)	0 (0.0)
Treat other heart diseases	6 (1.9)	3 (4.4)	2 (2.7)	1 (1.2)	0 (0.0)
Monitor carotid artery stenosis	1 (0.3)	0 (0.0)	1 (1.4)	0 (0.0)	0 (0.0)
Carotid artery stenting	5 (1.6)	3 (4.4)	2 (2.7)	0 (0.0)	0 (0.0)
Avoid sudden change in temperature	1 (0.3)	1 (1.5)	0 (0.0)	0 (0.0)	0 (0.0)
Remain emotionally stable	5 (1.6)	1 (1.5)	1 (1.4)	2 (2.3)	1 (1.2)

[a]The number of views was not available for the Douyin platform.

[b]The number of shares was not available for the Haokan platform.

[c]The number of shares was not available for the Xigua platform.

With regard to the overall quality and reliability of the included online videos, the kappa coefficients for the modified DISCERN and GQS scores, as assessed by 2 independent raters, were 0.71 and 0.73, respectively, indicating substantial interrater agreement. The observed discrepancies were primarily due to 1 rater slightly overestimating the quality and reliability of certain stroke prevention–related videos. However, most of these minor disagreements were resolved through discussion between the 2 raters. In cases where disagreement persisted, a senior researcher was assigned to make the final decision. Taken together, with the median scores for the modified DISCERN and GQS being 2 and 3, respectively, our results suggest that the included videos exhibited relatively poor reliability but moderate quality.

Video Upload Timeline

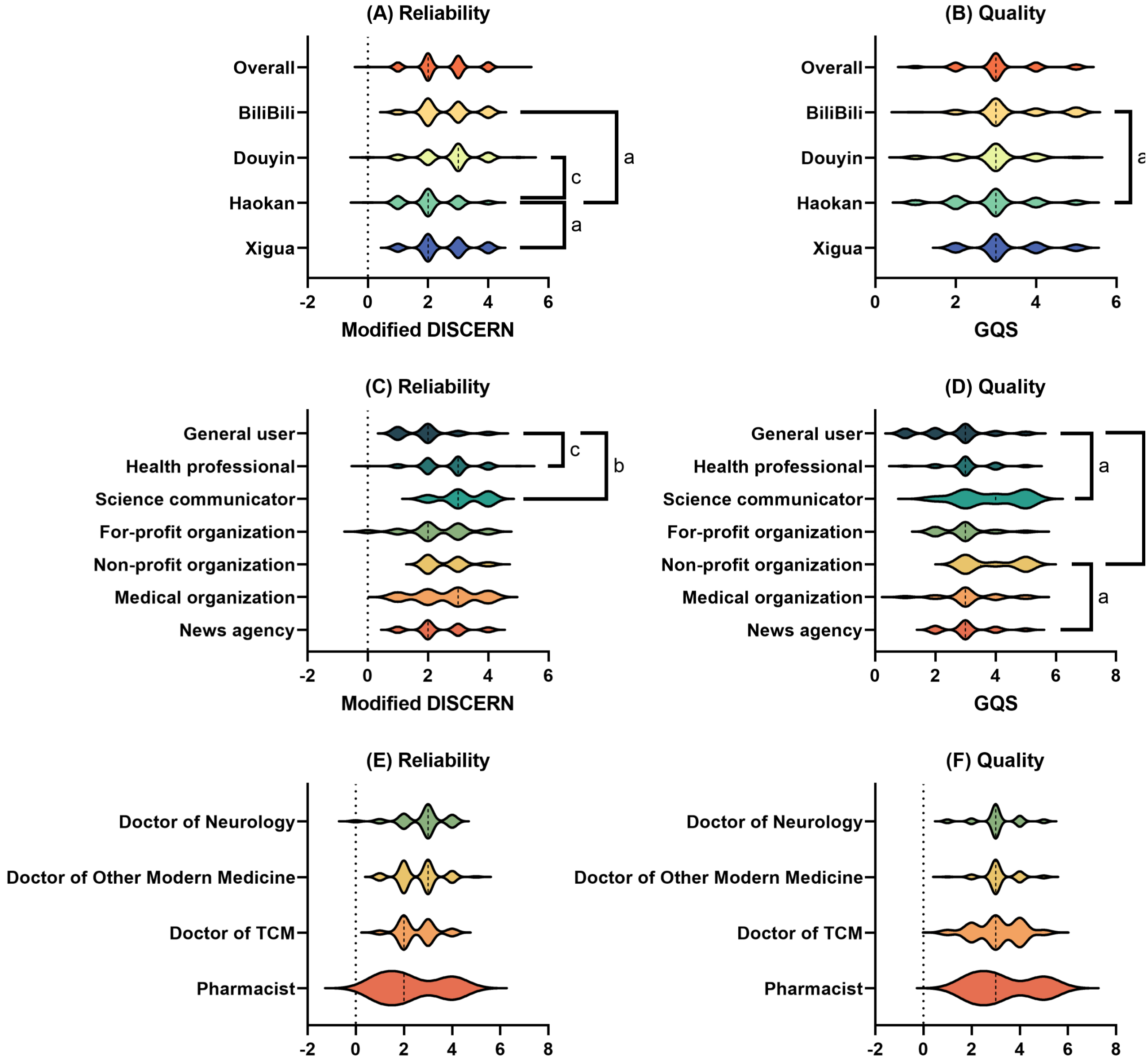
The distribution of video uploads over time revealed a clear upward trend. As shown in Table 4, the number of stroke prevention–related videos increased steadily each year. Only a small proportion of videos were uploaded before 2020, with a marked rise beginning in 2021. Notably, the highest number of videos (n=106, 33.9%) was uploaded in 2023 alone, suggesting growing public and institutional engagement with stroke-related health education on video platforms in recent years. This increase may reflect heightened awareness of cerebrovascular disease prevention, evolving content production practices, or rising user engagement with online platforms across China. These temporal patterns underscore the relevance and urgency of evaluating video-based health information, as both the volume and impact of such content continue to grow.

Quality and Reliability by Video Platform

When examined by specific online video platforms, as shown in Figure 2A and B, our results suggest that videos from Douyin (median modified DISCERN score of 3) exhibited higher median reliability score compared with those from other platforms (median modified DISCERN scores of 2.5, 2, and 2 for Bilibili, Haokan Video, and Xigua Video, respectively). However, this difference was statistically significant only between Douyin and Haokan Video ($P<.001$). Meanwhile, in terms of video quality, the median GQS scores were highly consistent across the 4 video platforms (GQS=3 for all 4 platforms). A statistically significant difference was observed only between Bilibili and Haokan Video ($P=.03$).

XSL•FO
RenderX

Figure 2. The modified DISCERN and Global Quality Scale (GQS) scores by (A, B) specific video platform, (C, D) video creators, and (E, F) health care professionals. [a]P<.05, [b]P<.01, [c]P<.001. TCM: traditional Chinese medicine.



Quality and Reliability by Video Authorship

When analyzed by specific video creator categories, as shown in Figure 2C and D, videos produced by health care professionals (median modified DISCERN score of 3) and science communicators (median modified DISCERN score of 3) demonstrated greater reliability compared with those created by amateur individual users (median modified DISCERN score of 2). However, no statistically significant difference was observed among videos created by different organizational authors (*P*>.99 in all cases). In terms of video quality, the findings were consistent with those for video reliability. Stroke prevention–related videos created by science communicators received higher GQS scores (median GQS score of 4) than those by amateur individual users (median GQS score of 3). No significant difference was observed between videos created by health care professionals and those by amateur users (*P*=.06). Among organizational video creators, videos from nonprofit organizations (median GQS score of 3.5) demonstrated higher quality compared with those from amateur users and news agencies (median GQS score of 3 for both cases).

When the health care professional authorship category was further subdivided (Figure 2E and F), no statistically significant differences were found among videos produced by neurologists (*P*>.99), doctors from other fields of modern medicine (*P*>.99), doctors of TCM (*P*>.99), and pharmacists (*P*>.99), regardless of whether the assessment was based on reliability or quality. However, despite the lack of statistical significance (*P*>.99), videos created by doctors specializing in modern medicine received higher reliability scores than those produced by TCM doctors (median modified DISCERN scores of 3 for both neurologists and other modern medical specialists vs 2 for TCM doctors). Interestingly, pharmacists, though contributing a relatively small proportion of the videos, produced content with moderate quality (median GQS score of 3), possibly reflecting their specialized knowledge in pharmacological prevention

XSL•FO
RenderX

strategies, such as the use of antiplatelet agents for stroke risk reduction.

These findings suggest that professional background may influence the rigor of content, even if the differences are not substantial, and underscore the potential value of encouraging more contributions from neurologists and other clinical specialists in stroke-related video creation.

Video Content Analysis

Regarding the stroke prevention strategies recommended in the included online videos, the vast majority focused on primary prevention (n=262), targeting the first onset of stroke. By contrast, relatively few addressed secondary prevention aimed at preventing stroke recurrence (n=28), and only a small number of videos covered both primary and secondary prevention simultaneously (n=23).

To be more specific, among primary stroke prevention measures, the most frequently mentioned strategy was controlling abnormal blood pressure (158/313, 50.5%), followed by adopting healthy dietary habits (145/313, 46.3%) and engaging in physical activity (144/313, 46.0%). These lifestyle-oriented recommendations reflect prevailing public health messaging trends and align with international stroke prevention guidelines. By contrast, for secondary prevention, the most frequently recommended measure was also managing abnormal blood pressure (24/313, 7.7%), followed by controlling abnormal blood glucose (21/313, 6.7%), regulating abnormal blood lipid levels (21/313, 6.7%), and using oral antiplatelet agents (21/313, 6.7%). This distribution suggests that, where addressed, secondary prevention videos tended to emphasize pharmacological and laboratory-monitored interventions rather than lifestyle-based recommendations.

Collectively, these findings suggest that while primary prevention dominates the landscape of online stroke-related education in China, a considerable gap exists in public-facing content addressing secondary prevention. Enhancing the visibility and clarity of such content, particularly among stroke survivors, may represent a valuable opportunity for public health intervention.

Correlation of Video General Information With Quality and Reliability

As shown in Table 5, significant correlations were observed among several general video information metrics. In particular, the number of views was positively associated with video length, number of likes, number of shares, and number of comments (r=0.142, P=.03; r=0.838, P<.001; r=0.897, P<.001; and r=0.593, P<.001, respectively). Similarly, videos that received more likes were more likely to have higher numbers of shares and comments (r=0.900, P<.001 and r=0.816, P<.001, respectively).

Table . The Spearman correlation coefficient between general video information and video quality and reliability.

Correlation	Video Length (s)	Views	Likes	Comments	Shares	Modified DISCERN score	Global Quality Scale score
Video Length (seconds)	1.000	—[a]	—	—	—	—	—
Views (n)	0.142[b]	1.000	—	—	—	—	—
Likes (n)	−0.060	0.838[b]	1.000	—	—	—	—
Comments (n)	−0.045	0.593[b]	0.816[b]	1.000	—	—	—
Shares (n)	−0.017	0.897[b]	0.900[b]	0.826[b]	1.000	—	—
Modified DISCERN scores	0.151[b]	−0.105	0.027	0.137[b]	0.019	1.000	—
Global Quality Scale score	0.441[b]	−0.042	−0.077	−0.017	−0.110	0.559[b]	1.000

[a]Not applicable.

[b]Statistical significance (P<.05).

Nevertheless, when considering video quality and reliability, video length was found to be positively correlated with both the modified DISCERN and GQS scores (r=0.151, P=.007 and r=0.441, P<.001, respectively). Additionally, a greater number of comments was associated with higher video reliability, although this correlation was relatively weak (r=0.137, P=.01).

Temporal Trend in Video Quality and Reliability

According to Figure S1 in Multimedia Appendix 1, an analysis of video upload time revealed a slight but statistically significant positive correlation between more recent upload dates and both video quality (P=.002) and reliability (P=.01). Specifically, Pearson correlation coefficients indicated that videos uploaded more recently tended to have higher GQS scores (r=0.113, P=.002) and higher modified DISCERN scores (r=0.086, P=.01). Although the strength of these associations was relatively weak, the findings suggest a gradual improvement in the overall quality and reliability of stroke-related video content over time. This trend may reflect growing awareness among content creators regarding evidence-based health communication, as well as improved platform policies aimed at enhancing the accuracy of health information.

XSL•FO
RenderX

Discussion

Principal Findings

In this study, by searching and screening online videos related to stroke prevention across 4 popular Chinese-language video platforms, we systematically described and analyzed the general characteristics, as well as the overall quality and reliability, of the included health-related online videos. To the best of our knowledge, this may be one of the first studies to specifically evaluate and report on the quality and reliability of Chinese-language online video platforms as information sources for stroke prevention education in Mainland China. We hope that our findings will provide valuable data for real-world policy makers and medical practitioners, ultimately contributing to improved stroke prevention and management efforts in the future.

Overall, using the modified DISCERN and GQS instruments, the included online videos related to stroke prevention were evaluated as relatively unreliable and of moderate quality. This suggests that, although video creators may offer some useful advice on primary and secondary stroke prevention measures, they often fail to provide supporting references and frequently lack fairness and objectivity in presenting information. These findings are consistent with previous research in other medical domains, which has similarly shown that health-related content on new media, particularly online video-sharing platforms, is often unsatisfactory and, at times, misleading [40-43]. This phenomenon may be largely attributed to the competition for audience attention in the digital marketplace, where accurate scientific information, often complex and less engaging for the general public, is easily overshadowed by sensationalized or simplified claims [44]. Consequently, entertainment-oriented online platforms are more likely to recommend light, engaging content over scientifically rigorous materials, which may be more difficult for general audiences to understand. Nevertheless, a growing number of stroke-related educational videos are being created and uploaded each year, suggesting that stroke prevention is becoming an increasingly visible topic on Chinese social media, particularly on video-sharing platforms. Notably, a positive temporal trend was observed in the quality and reliability of stroke-related videos uploaded between 2017 and 2023. We speculate that this improvement may be driven by increased participation from health care professionals in content creation, thereby directly contributing to the enhanced quality of stroke prevention information available online.

When analyzed by specific video platforms, videos from Douyin appeared to be more favorable, exhibiting higher content quality and reliability compared with those from the other 3 platforms, although statistical significance was not consistently observed. This finding may be explained by the fact that approximately 81% of eligible videos on Douyin were created and uploaded by qualified health care professionals, who generally receive more comprehensive and systematic evidence-based medical education and training than general users. Therefore, this phenomenon is relatively understandable. However, when the data for health care professionals were further stratified into more specific subgroups, videos created by doctors specializing in modern medicine—regardless of their field—tended to be more reliable than those produced by doctors of TCM, despite the absence of statistically significant differences.

Unfortunately, alongside the potentially beneficial health-related recommendations found in stroke prevention videos, concerns have also been raised about the prevalence of misinformation in the included study videos. Notably, some general individual users without a medical education background may share personal opinions in their videos that lack solid scientific evidence. For example, one video created by amateurs claimed that "regularly receiving intravenous injections can prevent stroke onset because it may help clear blood vessels." This statement has been identified as a widely circulated misconception among the Chinese population, particularly among older adults. Additionally, some video creators were observed to use emotionally charged and sensationalized titles and content, rather than objective and evidence-based information, in an effort to attract public attention and maximize viewership. However, this approach may lead to severe misunderstandings that could potentially jeopardize public health at the community level. Furthermore, the risk of misinformation is not limited to nonprofessionals, as certain content produced by health care professionals also warrants scrutiny. Specifically, current stroke prevention recommendations and suggestions from TCM doctors are highly heterogeneous, with some approaches not yet completely validated by formal clinical trials and largely based on empirical knowledge. For instance, some TCM practitioners propose that massage on specific body points (commonly referred to as "acupoints") may help prevent cerebral stroke. While this perspective is rooted in traditional practice, there is currently insufficient evidence from randomized controlled trials to substantiate its efficacy. Nevertheless, despite TCM's historically experience-driven approach, it remains an integral component of China's health care system and continues to serve as a complementary and alternative option in stroke prevention and management. In recent decades, TCM practitioners have made concerted efforts to integrate traditional practices with modern scientific research, contributing to its ongoing development and potential role in future stroke care [45].

To effectively address these issues, several targeted countermeasures can be implemented to enhance the quality and reliability of health-related online videos and information [46-48]. First and foremost, the review process for medical information videos uploaded to online platforms should be reinforced and supervised in collaboration with qualified medical professionals. Additionally, establishing and strengthening an online monitoring and correction system could help limit the spread of misinformation and reduce its long-term impact. Finally, general social media users should be encouraged to share controversial health information more selectively and cautiously, in order to curb the dissemination of false content and restrict accounts that produce or promote it.

Recently, Chinese governmental departments and relevant enterprises have begun taking active measures to combat the widespread dissemination of health-related misinformation online. For example, in collaboration with the Ministry of Industry and Information Technology, the Chinese National

XSL•FO
RenderX

Health Commission issued a policy on May 27, 2024, aimed at urgently strengthening the regulation of health-related webcasting, video content, and promotional activities on digital platforms. Simultaneously, ByteDance (TikTok's parent company) launched a centralized governance initiative in 2024 focused on identifying and removing poor-quality or even fraudulent sources of medical information. Beyond China, the World Health Organization (WHO) has proposed several countermeasures to limit online misinformation [49]. These include awareness campaigns targeting patients and health care professionals, the promotion of platforms that provide evidence-based data, the integration of scientific evidence into health-related content in mass media, and efforts to improve media and health literacy. Additionally, the WHO encouraged the active involvement of experts and health care professionals in debunking misinformation and guiding internet users toward credible, evidence-based information sources. However, the persistence of misinformation in the digital space, particularly in health care, continues to pose a significant threat to public well-being worldwide. This ongoing challenge underscores the urgent need for stronger guidelines and innovative strategies to mitigate its impact effectively.

In addition to evaluating the quality and characteristics of stroke-related videos, it is important to consider the platform-specific policies that may influence the dissemination of health content. Major Chinese video-sharing platforms, such as Douyin, have publicly issued guidelines encouraging the distribution of accurate and authoritative health information. For instance, these platforms typically require health-related content creators to verify their professional credentials and may implement content moderation mechanisms to flag or remove misleading medical information. However, the enforcement and transparency of these policies vary across platforms, and the extent to which they affect the quality and reach of health-related videos remains largely understudied. Future research could benefit from a more systematic analysis of how platform-level governance affects the credibility and visibility of online health communication.

Limitations

Despite these findings, several limitations of this study should be acknowledged. First and most importantly, our research team retrieved and analyzed only Chinese-language videos targeting Chinese audiences, using a limited set of search keywords. Videos in other languages, such as English, Japanese, and Korean, were not included, indicating a need for broader investigation in future research. Second, although measures were taken to minimize potential bias introduced by platform algorithms during the initial video search, we cannot fully exclude the possibility that our results were influenced by the artificial intelligence algorithms used by commercial video platforms. To be more specific, video search engines tend to prioritize and recommend content with higher engagement metrics. Although we logged out of our accounts, deleted temporary files and cookies, and conducted searches using a newly installed internet browser, the inherent nature of these automatic algorithms amplifies disparities between videos—allowing popular content to gain increasing visibility, while less popular videos receive minimal exposure. Consequently, this unavoidable limitation may have influenced the collection of general video-related metrics, particularly likes and shares. Third, this study was based on a limited sample size (from only 4 Chinese domestic video platforms). Therefore, further research involving a larger number of videos and platforms with broader representativeness is urgently needed. Fourth, data collection was conducted on a single day, which may not fully capture the dynamic and evolving nature of content rankings on video platforms. Videos that consistently maintain high visibility over time might have been overlooked, and our findings may therefore represent only a snapshot of what users were most likely to encounter at that specific moment. Future studies could incorporate longitudinal sampling strategies to account for temporal variations in content prominence. Finally, as a cross-sectional study using a relatively subjective evaluation strategy, this study primarily focused on a descriptive assessment of the current state of online video-based patient education in Mainland China. Therefore, future prospective studies are warranted to evaluate the real-world effectiveness of physician-created and peer-reviewed video materials for patient education.

Conclusion

This study examined the overall quality and reliability of Chinese-language videos related to public patient education on stroke prevention. The findings suggest that the quality and reliability of health-related videos, as a source of information for stroke management, remain relatively unsatisfactory and should be approached with caution by Chinese audiences. Therefore, targeted countermeasures, such as reinforcing the video content review process, establishing online correction mechanisms, and encouraging users to selectively and carefully share controversial information, are urgently needed to enhance the credibility and impact of stroke-related educational content in the future.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant 82171296), the Jiangsu Provincial Medical Key Discipline (grant ZDXK202217), and the 6th Jiangsu Province "333 High-Level Talents Training Project."

Data Availability

The data that support the findings of this study are available from the corresponding author (YJC) upon reasonable request.

XSL•FO
RenderX

Authors' Contributions

RGG and HYD were responsible for the concept and design of the study. Data collection and analysis were carried out by RGG, HYD, CCG, and YHX, with extensive contributions from RGG and HYD. The initial draft of the article was prepared by RGG. SJY, RW, and JPX critically revised the article for important intellectual content. YJC and SJY provided study supervision and secured funding. All authors have read and approved the final article.

RGG and HYD contributed equally to this work and should be recognized as co-first authors.

SJY (0319503013@163.com) and YJC (yongjuncao@126.com) jointly supervised this study and should be recognized as co-corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of author identity of included online videos and the temporal trends of modified DISCERN and Global Quality Scale scores of Chinese language stroke prevention–related videos.

[DOCX File, 313 KB - infodemiology_v5i1e64891_app1.docx]

References

1. Feigin VL, Abate MD, Abate YH, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Neurol 2024 Oct;23(10):973-1003. [doi: 10.1016/S1474-4422(24)00369-7]
2. Wu S, Wu B, Liu M, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. Lancet Neurol 2019 Apr;18(4):394-405. [doi: 10.1016/S1474-4422(18)30500-3] [Medline: 30878104]
3. Owolabi MO, Thrift AG, Mahal A, et al. Primary stroke prevention worldwide: translating evidence into action. Lancet Public Health 2022 Jan;7(1):e74-e85. [doi: 10.1016/S2468-2667(21)00230-9] [Medline: 34756176]
4. Pandian JD, Gall SL, Kate MP, et al. Prevention of stroke: a global perspective. Lancet 2018 Oct 6;392(10154):1269-1278. [doi: 10.1016/S0140-6736(18)31269-8] [Medline: 30319114]
5. Jeon MY, Jeong H. Effects of a stroke primary prevention program on risk factors for at-home elderly. Med Sci Monit 2015 Nov 28;21:3696-3703. [doi: 10.12659/msm.895519] [Medline: 26615035]
6. Kronish IM, Goldfinger JZ, Negron R, et al. Effect of peer education on stroke prevention: the prevent recurrence of all inner-city strokes through education randomized controlled trial. Stroke 2014 Nov;45(11):3330-3336. [doi: 10.1161/STROKEAHA.114.006623] [Medline: 25248910]
7. Yamagishi K, Sankai T, Muraki I, et al. Trends in stroke, cardiovascular disease, and medical expenditure under a community-based long-term stroke prevention program. J Hypertens 2023 Mar 1;41(3):429-436. [doi: 10.1097/HJH.0000000000003351] [Medline: 36728733]
8. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
9. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85. [doi: 10.2196/jmir.1933] [Medline: 23615206]
10. Szmuda T, Alkhater A, Albrahim M, et al. YouTube as a source of patient information for stroke: a content-quality and an audience engagement analysis. J Stroke Cerebrovasc Dis 2020 Sep;29(9):105065. [doi: 10.1016/j.jstrokecerebrovasdis.2020.105065] [Medline: 32807469]
11. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. J Med Internet Res 2022 Aug 4;24(8):e37367. [doi: 10.2196/37367] [Medline: 35816685]
12. Kessler SH, Humprecht E. COVID-19 misinformation on YouTube: an analysis of its impact and subsequent online information searches for verification. Digit Health 2023;9:20552076231177131. [doi: 10.1177/20552076231177131] [Medline: 37256011]
13. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019 May 2;21(5):e12522. [doi: 10.2196/12522] [Medline: 31045507]
14. Nanda A, M N M, Jammula G, et al. Study on the quality, relevance, and comparability of YouTube videos expressing stroke and its consequences from various sources. Cureus 2023 Aug;15(8):e43277. [doi: 10.7759/cureus.43277] [Medline: 37692625]
15. Denham AM, Baker AL, Spratt NJ, et al. YouTube as a resource for evaluating the unmet needs of caregivers of stroke survivors. Health Informatics J 2020 Sep;26(3):1599-1616. [doi: 10.1177/1460458219873538] [Medline: 31722610]

XSL•FO
RenderX

16. Assylbek MI, Zimba O, Akyol A, Yessirkepov M, Kocyigit BF. YouTube as a source of information for stroke rehabilitation: a cross-sectional analysis of quality and reliability of videos. Rheumatol Int 2025 Mar 22;45(4):77. [doi: 10.1007/s00296-025-05832-4] [Medline: 40119932]
17. Özdemir Z, Acar E. YouTube as a source of recognizing acute stroke; progress in 2 years. BMC Public Health 2024 Aug 13;24(1):2208. [doi: 10.1186/s12889-024-19710-4] [Medline: 39138572]
18. Garg D, Agarwal A, Srivastava MP, Vishnu VY. Use of social media in stroke: a systematic review. Ann Indian Acad Neurol 2023;26(3):206-212. [doi: 10.4103/aian.aian_58_23] [Medline: 37538420]
19. Bilibili. URL: http://www.bilibili.com [accessed 2025-07-09]
20. Douyin. URL: http://www.douyin.com [accessed 2025-07-09]
21. Haokan Video. URL: https://haokan.baidu.com [accessed 2025-07-09]
22. Xigua Video. URL: https://www.ixigua.com [accessed 2025-07-09]
23. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004 Oct;7(5):497-510. [doi: 10.1089/cpb.2004.7.497] [Medline: 15667044]
24. Mueller SM, Hongler VNS, Jungo P, et al. Fiction, falsehoods, and few facts: cross-sectional study on the content-related quality of atopic eczema-related videos on YouTube. J Med Internet Res 2020 Apr 24;22(4):e15599. [doi: 10.2196/15599] [Medline: 32329744]
25. Sun F, Zheng S, Wu J. Quality of information in gallstone disease videos on TikTok: cross-sectional study. J Med Internet Res 2023 Feb 8;25:e39162. [doi: 10.2196/39162] [Medline: 36753307]
26. Zheng S, Tong X, Wan D, Hu C, Hu Q, Ke Q. Quality and reliability of liver cancer-related short Chinese videos on TikTok and Bilibili: cross-sectional content analysis study. J Med Internet Res 2023 Jul 5;25:e47210. [doi: 10.2196/47210] [Medline: 37405825]
27. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111. [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
28. Loeb S, Reines K, Abu-Salha Y, et al. Quality of bladder cancer information on YouTube. Eur Urol 2021 Jan;79(1):56-59. [doi: 10.1016/j.eururo.2020.09.014]
29. Gorrepati PL, Smith GP. DISCERN scores of YouTube information on eczema treatments. J Am Acad Dermatol 2021 Nov;85(5):1354-1355. [doi: 10.1016/j.jaad.2020.11.007] [Medline: 33181240]
30. Singh AG, Singh S, Singh PP. YouTube for information on rheumatoid arthritis—a wakeup call? J Rheumatol 2012 May;39(5):899-903. [doi: 10.3899/jrheum.111114] [Medline: 22467934]
31. Cassidy JT, Fitzgerald E, Cassidy ES, et al. YouTube provides poor information regarding anterior cruciate ligament injury and reconstruction. Knee Surg Sports Traumatol Arthrosc 2018 Mar;26(3):840-845. [doi: 10.1007/s00167-017-4514-x] [Medline: 28314888]
32. Li HOY, Pastukhova E, Brandts-Longtin O, Tan MG, Kirchhof MG. YouTube as a source of misinformation on COVID-19 vaccination: a systematic analysis. BMJ Glob Health 2022 Mar;7(3):e008334. [doi: 10.1136/bmjgh-2021-008334] [Medline: 35264318]
33. Radonjic A, Fat Hing NN, Harlock J, Naji F. YouTube as a source of patient information for abdominal aortic aneurysms. J Vasc Surg 2020 Feb;71(2):637-644. [doi: 10.1016/j.jvs.2019.08.230] [Medline: 31611104]
34. Yoo M, Hong J, Jang CW. Suitability of YouTube videos for learning knee stability tests: a cross-sectional review. Arch Phys Med Rehabil 2020 Dec;101(12):2087-2092. [doi: 10.1016/j.apmr.2020.05.024] [Medline: 32593549]
35. Moon H, Lee GH. Evaluation of Korean-language COVID-19-related medical information on YouTube: cross-sectional infodemiology study. J Med Internet Res 2020 Aug 12;22(8):e20775. [doi: 10.2196/20775] [Medline: 32730221]
36. Ding R, Kong Q, Sun L, et al. Health information in short videos about metabolic dysfunction-associated steatotic liver disease: analysing quality and reliability. Liver Int 2024 Jun;44(6):1373-1382. [doi: 10.1111/liv.15871] [Medline: 38441405]
37. Barlas T, Ecem Avci D, Cinici B, Ozkilicaslan H, Muhittin Yalcin M, Eroglu Altinova A. The quality and reliability analysis of YouTube videos about insulin resistance. Int J Med Inform 2023 Feb;170:104960. [doi: 10.1016/j.ijmedinf.2022.104960] [Medline: 36525801]
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310] [Medline: 843571]
39. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg 2018 May;126(5):1763-1768. [doi: 10.1213/ANE.0000000000002864] [Medline: 29481436]
40. Libretti A, Vitale SG, Saponara S, et al. Hysteroscopy in the new media: quality and reliability analysis of hysteroscopy procedures on YouTube. Arch Gynecol Obstet 2023 Nov;308(5):1515-1524. [doi: 10.1007/s00404-023-07172-9] [Medline: 37568070]
41. Ge R, Xu X, Chen Y, et al. The quality and reliability of online video resources concerning patient education for migraine treatment in China Mainland: an electronic media-based study. BMC Public Health 2025 Feb 18;25(1):656. [doi: 10.1186/s12889-025-21861-x] [Medline: 39966862]

XSL•FO
RenderX

42. Vitale SG, Angioni S, Saponara S, et al. Hysteroscopic metroplasty and its reproductive impact among the social networks: a cross-sectional analysis on video quality, reliability and creators' opinions on YouTube, TikTok and Instagram. Int J Med Inform 2025 Mar;195:105776. [doi: 10.1016/j.ijmedinf.2024.105776] [Medline: 39729920]
43. Vitale SG, Angioni S, Saponara S, et al. TikTok as a platform for hysteroscopy information: an analytical video-based cross-sectional study to assess quality, reliability, and accuracy. Int J Gynaecol Obstet 2025 Jan;168(1):353-361. [doi: 10.1002/ijgo.15846] [Medline: 39118465]
44. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552. [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
45. Guo DA, Lu A, Liu L. Modernization of traditional Chinese medicine. J Ethnopharmacol 2012 Jun 1;141(2):547-548. [doi: 10.1016/j.jep.2012.05.001] [Medline: 22569212]
46. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 2;41:433-451. [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
47. Kurowski JA, Bewtra M, Kodish E, Lashner B. Reducing the spread of internet misinformation in IBD: ethics and responsibility. Inflamm Bowel Dis 2021 May 17;27(6):960-962. [doi: 10.1093/ibd/izaa335] [Medline: 33399199]
48. Bak-Coleman JB, Kennedy I, Wack M, et al. Combining interventions to reduce the spread of viral misinformation. Nat Hum Behav 2022 Oct;6(10):1372-1380. [doi: 10.1038/s41562-022-01388-6] [Medline: 35739250]
49. Infodemics and misinformation negatively affect people's health behaviours, new WHO review finds. World Health Organization. URL: https://www.who.int/europe/news-room/01-09-2022-infodemics-and-misinformation-negatively-affect-people-s-health-behaviours--new-who-review-finds [accessed 2025-06-28]

Abbreviations

CCTV: China Central Television
GQS: Global Quality Scale
TCM: traditional Chinese medicine
WHO: World Health Organization

Edited by T Mackey; submitted 30.07.24; peer-reviewed by J Hao, S Saponara; revised version received 01.06.25; accepted 04.06.25; published 21.07.25.

Please cite as:
Ge R, Dai H, Gong C, Xia Y, Wang R, Xu J, You S, Cao Y
The Quality and Reliability of Online Videos as an Information Source of Public Health Education for Stroke Prevention in Mainland China: Electronic Media–Based Cross-Sectional Study
JMIR Infodemiology 2025;5:e64891
URL: https://infodemiology.jmir.org/2025/1/e64891
doi:10.2196/64891

© Rongguang Ge, Haoyi Dai, Chicheng Gong, Yuhong Xia, Rui Wang, Jiaping Xu, Shoujiang You, Yongjun Cao. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 21.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of Influencers and Echo Chambers in the Diffusion of Vaccine Misinformation: Opinion Mining in a Taiwanese Online Community

Jason Dean-Chen Yin[1], PhD; Tzu-Chin Wu[2], MA; Chia-Yun Chen[2], MA; Fen Lin[3], PhD; Xiaohui Wang[3], PhD

[1]Li Ka Shing Faculty of Medicine, School of Public Health, University of Hong Kong, Hong Kong, China

[2]College of Public Health, Institute of Health Behaviors and Community Sciences, National Taiwan University, Taiwan, Taiwan

[3]Department of Media and Communication, City University of Hong Kong, Room 5086, Creative Media Center, Hong Kong, China

Corresponding Author:
Fen Lin, PhD
Department of Media and Communication, City University of Hong Kong, Room 5086, Creative Media Center, Hong Kong, China

Abstract

Background: Prevalence and spread of misinformation are a concern for the exacerbation of vaccine hesitancy and a resulting reduction in vaccine intent. However, few studies have focused on how vaccine misinformation diffuses online, who is responsible for the diffusion, and the mechanisms by which that happens. In addition, researchers have rarely investigated this in non-Western contexts particularly vulnerable to misinformation.

Objective: This study aims to identify COVID-19 vaccine misinformation, map its diffusion, and identify the effect of echo chamber users on misinformation diffusion on a Taiwanese online forum.

Methods: The study uses data from a popular forum in Taiwan called PTT. A crawler scraped all threads on the most popular subforum from January 2021 until December 2022. Vaccine-related threads were identified through keyword searching (n=5818). Types of misinformation, including misleading, disinformation, conspiracy, propaganda, and fabricated content, were coded by 2 researchers. Polarity was proposed as a proxy for measuring an individual's level of involvement in the echo chamber, one of the mechanisms responsible for the viral misinformation on social media. Factors related to information diffusion, including misinformation type and polarity, were then assessed with negative binomial regression.

Results: Of 5818 threads, 3830 (65.8%) were identified as true information, and 1601 (27.5%) contained misinformation, yielding 5431 boards for analysis. Misinformation content did not vary much from other contexts. Propaganda-related information was most likely to be reposted (relative risk: 2.07; P<.001) when comparing to true information. However, the more polarized a user was, the less likely his or her content was to be reposted (relative risk: 0.22; P<.001). By removing the nodes with a high level of indegree, outdegree, and betweenness centrality, we found that the core network and the entire network demonstrated a decreasing trend in average polarity score, which showed that influential users contributed to the polarization in misinformation consumption.

Conclusions: Although the forum exhibits a resilience to echo chambering, active users and brokers contribute significantly to the polarization of the community, particularly through propaganda-style misinformation. This popularity of propaganda-style misinformation may be linked to the political nature of the forum, where public opinion follows "elite cues" on issues, as observed in the United States. The work in this study corroborates this finding and contributes a data point in a non-Western context. To manage the echo chambering of misinformation, more effort can be put into moderating these users to prevent polarization and the spread of misinformation to prevent growing vaccine hesitancy.

(JMIR Infodemiology 2025;5:e57951) doi:10.2196/57951

KEYWORDS

misinformation; vaccine; online community; influencer; echo chamber; Taiwan

Introduction

Background

As individuals increasingly turn to social media as a primary source of information, the prevalence and spread of unverified or misinformed scientific claims is concerning [1]. Social media users gravitate toward information that validates their belief systems, forming echo chambers that validate their shared narrative [2]. These echo chambers can be the launching pads for misinformation that goes viral [3]. Worse, they can influence opinions on issues of public concern, such as vaccination [4].

XSL•FO
RenderX

As early as 2001, studies identified the rise of vaccine hesitancy online [5,6] and cataloged the techniques in antivaccination misinformation transmission [7,8]. These studies paralleled the early days of the World Wide Web and focused solely on analyzing web pages. An early study by Zimmerman et al [6] found that misrepresentation (twisting of content) was a method for conveying vaccine misinformation. Kata [8] elaborated on misrepresentation in her study in 2010 on childhood vaccination. In her study, misinformation, a sort of misrepresentation, was a main antivaccination theme that arose in classification. Under the misinformation umbrella, she found that using outdated sources; misrepresenting facts; self-referencing to “experts”; not referencing statistics or citations; and making unsupported statements were all ways of passing misinformation. The use of negative tones is also a method of strengthening antivaccination methods [9].

As misinformation became more prevalent on social media, exacerbated by the terming of “fake news” and a global pandemic, more research was conducted on clarifying the definition of misinformation and classifying the different kinds of misinformation. The term “misinformation” is often used interchangeably with related concepts such as spam, rumors, fake news, and disinformation. In this study, we use “misinformation” as an umbrella term to encompass all false or inaccurate information disseminated through social media. Under these umbrellas, there are different types of misinformation. Some information can be fabricated content, while some can be manipulated to be clickbait. Some can be satire or parody being passed off as true, and some can be propaganda [10]. Information can be passed off as true via sponsors and be partially true or totally false. New modes of producing synthetic media through artificial intelligence such as “deep fakes” distort reality by passing as true. Although there is nuance between the different types of misinformation, all types are distinct from true information, which is not deliberately fabricated with malicious intent and does not contain false (scientific) information.

There are several studies on classifying types of vaccine misinformation. Wu et al [11], in their study on general misinformation in social media, tentatively organized misinformation into nonexclusive categories such as unintentionally and intentionally spread misinformation, urban legend, fake news, unverified information, rumor, crowdturfing, spam, and trolling. Zhao et al’s [12] study does not deviate far from this. They identified that misinformation could include conspiracies; concerns about vaccine safety and efficacy; a flat rejection of all vaccines; morality (including religion and human experimenting); and a violation of civil liberties. Classification studies also identified new modes of transmission. Basch et al [13] studied the types of misinformation on TikTok, a medium mixing audiovisual and textual cues, and found that parodying or the overemphasis of false consequences of available vaccines were all methods of strengthening an antivaccination narrative.

In addition to modes of transmission and classification, some studies have focused on linking misinformation to vaccination intent, although these are sparse. One study used a questionnaire to identify that COVID-19 vaccine hesitancy mediated the relationship between vaccine knowledge and vaccination intention, with misinformation on vaccines being associated with higher vaccine hesitancy. The same study also claimed that most respondents were exposed to COVID-19 misinformation [14]. Another study conducted a randomized controlled trial, finding that exposure to recent misinformation induced a decline in intent of vaccination. This is especially true when the misinformation is transmitted as misrepresenting facts of a scientific nature [15].

In the studies on vaccine misinformation, there are unexplored avenues. Few studies have examined how vaccine misinformation diffuses online, both in time and in quantity. Diffusion is a communication process whereby information is communicated over time “among the members of a social system” [16]. Often, this process is measured in terms of depth, breadth, and speed [17,18]. Diffusion depth refers to the length of the longest transmission chain, breadth refers to the capacity to generate offspring posts, and speed refers to the efficiency of the information diffusion process [19]. Understanding the diffusion process between true information and misinformation is important because, should approaches such as psychological inoculation [20] or other vaccination community strategies work [21,22], understanding the scale of spread and when to intervene diminishes the spread.

In addition, no studies have studied key users’ roles in spreading vaccine misinformation. This can be broken down into 2 symbiotic perspectives. One perspective is looking at how individuals’ engagements online may be polarizing or divisive. Previous research on polarity focused on this mostly in politics, using different metrics of measurement for polarity. One study assessed political polarization using interactional, positional, and affective polarity across various platforms to find that polarization is platform dependent [23]. Another study looked at affective polarization against the feminist cause following protests [24]. In the vaccination literature, polarity mostly focuses on differences in vaccine sentiment [4,25,26], aligning it with affective polarization. However, there are a paucity of studies analyzing polarity in engagement with misinformation. This is important because the deliberate exclusion of true perspectives on vaccination may affect the downstream intention to vaccinate. Echo chambers that form around misinformation are thus one avenue worth exploring. Previous research has found that echo chambers are related closely to misinformation diffusion [27] and that misinformation disseminated by echo chambers spread more virally than misinformation not distributed by them [28].

The second perspective is identifying who the key users are in this process. One can see how the polarity in the endorsement of misinformation can be exponentially more problematic when the users doing the exclusion are both laypeople [29] and central in the network. Two studies found that information diffused by “brokers”—those who have high betweenness centrality—affects the final size of information cascades [30,31]. For health information, the diffusion size of posts by the US Centers for Disease Control and Prevention (CDC) were related to broker involvement [32]. More work can thus be done on understanding how key users aid in spreading vaccine misinformation through the deliberate exclusion of “true news,” thereby enhancing its dissemination.

XSL•FO
RenderX

The consequence of having these key users spread information lies in the fact that individual beliefs are shaped by their immediate social network, both online and offline [33]. These worlds feed back to each other, entrenching belief. Scholars have found that dissemination of misinformation is driven by social reinforcement in an individual's digital circles [34]. If a person's network consists of mostly rumormongers, that person in turn likely propagates the same misinformation. This propagation often spills into the offline space, influencing judgment through reinforcing the legitimacy of the information online [35], eventually creating fissures along sociodemographic lines, further polarizing the offline world [36]. Should this process occur for vaccine information, it would exacerbate the issue of vaccine hesitancy and have negative implications for vaccination. Few studies have examined the upstream part of this process, including in non-Western contexts [36]. This study focuses on vaccine misinformation on a local forum in Taiwan (PTT) from 3 aspects: identifying the types of misinformation, describing how it diffuses relative to regular news, and assessing how influential users fall into chambers of misinformation, affecting its spread.

Research Questions

This study uses the same PTT data to address the following questions about misinformation. First, we will determine what types of misinformation are present in a Taiwanese online community. This identification and categorization archives how misinformation topics differ to other contexts considering the sociopolitical context of Taiwan. Second, the study will determine how misinformation cascades differ from true information by dichotomizing the topics, comparing their breadth [19]. Finally, we will assess to what extent the echo chamber phenomenon exists in the diffusion of misinformation and how influential users, as a proxy for understanding key spreaders of misinformation, affect the echo chamber in the network. The findings of this study can inform how Taiwan can combat growing misinformation in its diverse online ecology.

Methods

Data Procurement

PTT is a terminal-based bulletin board system developed in Taiwan in 1995. Functioning like an internet forum, PTT is often termed Taiwan's "Reddit" and is one of the most active forums in Taiwan. From July 2022 to July 2023, the average number of users per day was 56,000 [37]. The web-based version of PTT is structured into different boards, which are equivalent to thematic groups, on a forum. Within each board, threads or topics can be started. On each thread, the poster's metadata such as the username, time of post, and IP address are available for scraping. In addition, within the threads there is a comment section for which the corresponding metadata is also available.

The data for this study were collected from the "Gossiping" board on PTT using a web scraper developed in Python to extract HTML data. The scraping targeted posts made between January 2021 and December 2022. We filtered the dataset using the Chinese term for "vaccine" to ensure we captured vaccine-related threads. The language of all collected data was Mandarin Chinese, the primary language used on PTT. The "Gossiping" board was specifically chosen due to its high activity levels and its reputation for users engaging in a wide range of discussions, including those related to public health, politics, and societal issues. This board is particularly active during times of political or social crises, making it a rich source for analyzing vaccine misinformation during the COVID-19 pandemic. Other thematic boards were excluded to maintain focus on a general-purpose discussion forum where misinformation may more organically spread across a wider audience. In total, 5818 boards were pulled.

Identifying Misinformation

The coding process for classifying vaccine misinformation followed an established framework based on existing literature. We used a broad classification scheme (Table 1) derived from studies on misinformation in digital spaces, which categorize misinformation into types such as conspiracies, fabricated content, fake news, and political propaganda. The classification of posts into misinformation categories was conducted in Mandarin by 2 trained coders, who worked independently to classify an initial sample of 500 posts. The initial codebook was developed using an inductive coding approach, where new categories emerged from the data during this pilot phase. The coders reached a high level of intercoder reliability, with a Cohen κ of 0.91, indicating strong agreement. Discrepancies were resolved through discussion, and the finalized codebook was used to classify the remaining 5318 posts.

XSL•FO
RenderX

Table . Labeling scheme for true information and false information.

Category and label	Classification criteria
True information	
Factual news	Reporting of news events or facts without interpretation or analysis
Scientifically accurate analytical content	Interpreting or analyzing facts or data to provide deeper understanding
Misinformation	
Misleading information	False or inaccurate information, regardless of intention
Disinformation	Deliberately created and shared with the intent to mislead or deceive
Propaganda	Biased or misleading information used to promote a political cause or push a certain point of view
Fabricated content	Entirely false content designed to deceive and mislead (create a fake news source, fake quotes, or nonexistent events)
Conspiracy	Belief or explanation that something is a result of a secret plot by a group or organization
Mixed misinformation	Two or more misinformation categories
Religious beliefs	Any discussion of religion in relation to vaccination
Unrelated	Boards containing “vaccine” keyword but unrelated to vaccine information or misinformation

Any disagreements among coders were resolved through discussion among themselves. The coders informed the principal investigator of these discrepancies, who further approved or reassessed their decision, informing the coders. As a standard reference, we consolidated a library of misinformation on COVID-19 vaccines in Taiwan from three fact-checking organizations in Taiwan: MyGoPen [38], Cofacts [39], and Taiwan FactCheck Center [40]. These organizations, led by civil society movements combatting misinformation, consolidated potentially misinformative news on a variety of topics, including those related to COVID-19. Boards related to the COVID-19 vaccine were filtered by using “vaccine” as a keyword in Chinese. In the event of disagreements on classification, the vaccine-related misinformation on these sites was used as a final check. Following the initial calibration, the remaining 5318 boards were split into 2 datasets (n=2659) and independently coded. Any questions or concerns were brought up to all coders and the principal investigator; the final classification was determined by a majority vote or mutual agreement. For comparing diffusion of misinformation and users’ polarized engagement with misinformation, the categories in Table 1 are further collapsed into “true” and “false” information. Posts unrelated to vaccines or involving purely religious content were excluded from the final analysis. Thematic examples of excluded posts include those focused on general prayer requests, religious teachings not tied to vaccination, or unrelated social and political discussions. For example, posts discussing the afterlife or religious rituals were categorized as unrelated, as they did not directly engage with the vaccine misinformation discourse.

For subsequent analyses, we wanted to remove users who occasionally shared misinformation and focus on users who shared misinformation frequently (ie, we distilled a “core network” by extracting users with over 5 posts; n≥5, n=number of posts). This is also consistent with previous studies that used multiple filtering techniques to remove randomness in a large network [41,42]. To justify the choice of cutoff, robustness tests were conducted, as shown in Multimedia Appendix 1. In short, different threshold values (from 1 to 10) and the average polarity scores of the selected core network were calculated to see any significant changes. Although the average polarity increased as the cutoff increased, the trend of decreasing polarity by cutting out key users was the same across all values. To balance between cutting off too many or too few nodes for a forum, the choice of n≥5 was made.

Diffusion of Misinformation

The diffusion of information is traditionally measured in 3 ways: breadth, depth, and speed. These 3 aspects follow previously defined measurements on diffusion on social media platforms [18]. However, given the structure of the PTT forum, depth is unascertainable. The reason for this is because, in PTT, replies of replied posts always refer to the original post, and not the replied post. This metadata obscures the length of the diffusion chain. Compare this to platforms like X (formerly Twitter) that link the diffusion chain through explicitly stating a “retweet” of a “retweet” is “retweeted” from the “retweet,” and not the original tweet. Due to the structure in PTT, breadth will be used to capture the size of the information diffusion. Breadth is the number of first-degree child nodes that repost it. If we denote a message as m and the set of first-degree child nodes as $N(m)$, the breadth $B(m)$ is equivalent to $|N(m)|$.

In addition, a negative binomial regression model will estimate the predictors of breadth. A negative binomial regression is used since overdispersion is expected for the distribution of breadth of information diffusion. To build the model, the categorical misinformation types will be input as categorical variables, with “true information” as the baseline. Besides, the polarity score generated in the next section will be used. Polarity is included because it is possible that users who tend to be extreme on either spectrum are likely to have more engagement

XSL•FO
RenderX

in the network. In addition, 2 control variables will be used, including the word count of the post and the activeness of the users. Previous research has suggested that longer posts have a higher likelihood of being transmitted [43]. The activeness of a user is operationalized as the number of historical posts. Although activity on a media platform does not necessarily mean higher engagement [44], it is an important confounding factor to be included in the model. The regression model estimate is: $\log(\mu_i) = \beta_0 + \beta_1(Misinformation) + \beta_2(Disinformation) + \beta_3(Propaganda) + \beta_4(Fabricated) + \beta_5(Propaganda) + \beta_6(WordCount) + \beta_7(UserActiveness) + \beta_8(PolarityScore) + \varepsilon$.

where μ_i is the count of the dependent variable for the i^{th} observation of breadth of misinformation spread. The log link function relates the mean of the response variable to the linear predictors. Variables are tested for multicollinearity prior to inclusion into the model. For a deeper analysis into the variable correlations, see Multimedia Appendix 2. Given that there is no offset term in the regression, the results of the negative binomial regression produce a relative risk (RR) indicating the increased risk of the outcome variable.

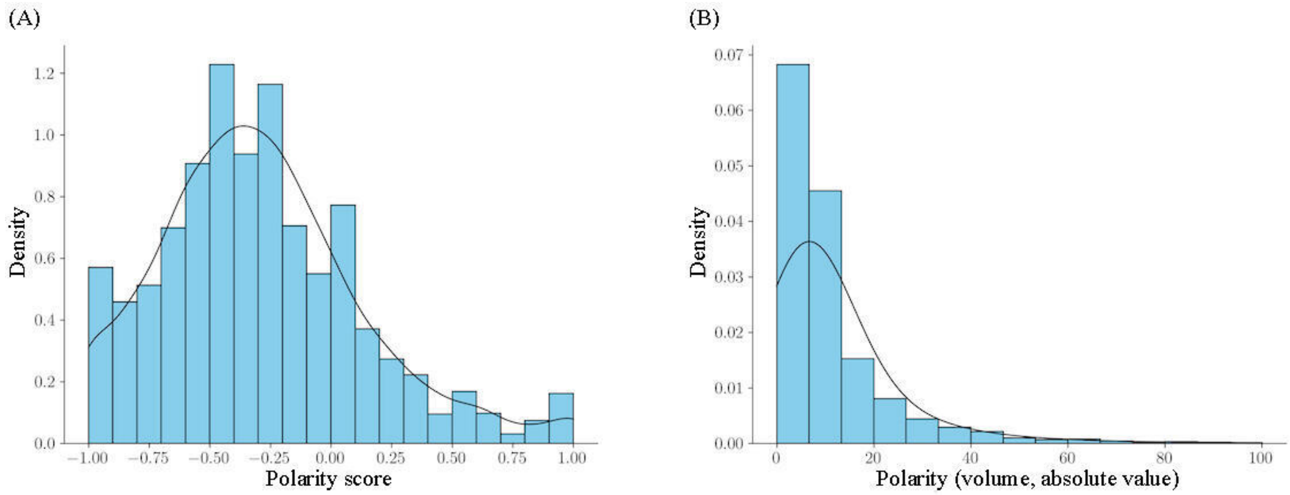
Measuring Polarity

Polarization in public opinion refers to the extent to which the views of a population (in this case, support for true or false information) are extreme and distinctly divided [45]. In the context of this study, polarity represents the predominant commenting activity on either true or false information (ie, misinformation), suggesting that users are chambered in their engagement with certain types of information. This measure aligns with the "affective polarization" measures previously used in the politics [23,24] and vaccine sentiment literature; however, they are extended to this study to measure endorsement of either true or false information. Further, polarization can be measured at the individual level or aggregated at the community level. In this study, we measure individuals' level of polarity, proposing 2 metrics for analysis.

The first is polarity measured by the difference in proportion of comments on true information and misinformation ("proportion polarity"). To measure this, we collected the commenting behavior for each node v in the network. With their commenting history, $C(v)$, we calculated the number of comments on true and false information, denoted as $C_{pos}(v)$ and $C_{neg}(v)$, respectively. The proportions of positive and negative comments were then Pposv=Cpos(v)C(v) and Pnegv=CnegvC(v), respectively. We subtracted the proportion of negative comments from the proportion of positive comments to get the polarity score, $\pi_{prop}(v)$, per node using the equation $\pi_{prop}(v) = P_{pos}(v) – P_{neg}(v)$. The range of $\pi_{prop}(v)$ is $–1 \leq \pi_{prop}(v) \leq 1$, with a score of −1 representing a polarity in commenting only on false information, 0 representing equal commenting on both, and 1 representing entirely commenting on true information.

The second is polarity measured by the absolute value of the difference in volume of comments on true and false information ("volume polarity"). Measuring by absolute value of the difference removes the true-false dichotomy and allows a straightforward interpretation of any polarity in the network, permitting a clearer interpretation of echo chambering. To calculate volume polarity, $\pi_{vol}(v)$, we took the absolute value of the difference of the number of negative to positive comments, $|C_{neg}(v) – C_{pos}(v)$. The range of $\pi_{vol}(v)$ is then $0 \leq \pi_{vol}(v) < inf$, with higher values indicating higher polarity. Figure 1 illustrates the distribution in polarity scores calculated by proportion (Figure 1A) and volume (Figure 1B) for the core network. For ease of interpretation in the negative binomial regression, volume polarity value is min-max normalized to create a value between 0 and 1 due to the expected heavier left skew. To preserve the weight of posting in information diffusion, we use the polarity by volume in our following analyses.

Figure 1. Distribution of polarity scores (n=2422) calculated by (A) proportion and (B) volume for the core network.



Node-Level Predictors of Polarity

Influential users generally have disproportionate impact on the flow of network information, also shown in the previous section. Identifying them and their relation to polarity in the network is therefore an important step in diminishing the spread of misinformation. Measures of centrality are usually used to identify node importance in the network. We used 2 centrality

XSL•FO
RenderX

metrics to identify influential nodes, with each representing a different concept of influence.

The first measure of centrality is degree centrality, which directly measures the number of connections a node has. The more connections a node has, the more central and influential it is within the network. In PTT, since forum data is directional, the indegree and outdegree of a node will be calculated for each node separately. Decomposing degrees into indegrees and outdegrees helps distinguish whether nodes on either polar end are more an authority (indegree) or a broadcaster (outdegree) of misinformation.

The second measure of centrality is betweenness centrality, which measure the extent to which a node lies on paths between other nodes; specifically, capturing the frequency with which a node appears on the shortest paths between pairs of nodes. Nodes with high betweenness centrality have control over information flow in the network, acting as "bridges" or "brokers" across the network [46]. These nodes are critical for the flow of information in the network because they function as "switches" for facilitating or inhibiting information flow. For the network, let σ_{st} denote the total number of shortest paths from node *s* to node *t*, and $\sigma_{st}(v)$ denote the number of those paths that pass through a node *v*. The betweenness centrality *b*(*v*) can be defined as:

$$\sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

which sums over all pairs of nodes (*s*,*t*) in the network, excluding those for which the pair is *v*, and for each pair, calculates the fraction of shortest paths between *s* and *t* that pass through *v*. Said another way, it calculates the proportion of times *v* is a bridge along the shortest path between 2 other nodes.

To analyze how polarity in the network changes in response to these influencers in misinformation diffusion, we removed a subset of the percentage of top influential users—by increasing increments of 5%—and calculated and plotted the average absolute value polarity score in the network. The resulting graphs of changing polarity indicate the change in overall posting behavior on either true or misinformative threads by influential users.

Ethical Considerations

This study involved the analysis of secondary data collected from publicly available social media platforms. The data were obtained from PTT, an open-access online forum, and all posts were publicly accessible at the time of data collection. No private or sensitive information was collected, and all user data were anonymized before analysis. The study did not involve any interaction with the individuals behind the social media accounts, and no identifiable information is included in the results. All findings are presented in aggregate to ensure anonymity. As the research only analyzed publicly available data and did not involve human subjects directly, formal ethics approval was not required according to institutional guidelines. Nonetheless, the study adhered to ethical principles of data privacy and responsible handling of social media data.

Results

Identified Misinformation

The results in Table 2 show the number of boards by each misinformation type in Table 1, in addition to an elaboration of several of the most common thread types. Out of 5818 threads, most threads (n=2227, 38.3%) involved netizens asking questions for further clarification on vaccines. The next most common was reporting of official news reports or press conference news (n=1603, 27.8%). For all threads containing misinformation, the most common was the "propaganda" type (n=858, 14.7%), indicating the relatively political nature of this forum. After disregarding the unrelated and religious threads, and recategorizing the "mixed" threads into the conspiracy category, there were 3830 true threads and 1601 threads containing misinformation, for a total of 5431 boards for analysis moving forward. For these boards, the number of comments and reposts for true and false information is illustrated in Table 3.

XSL•FO
RenderX

Table . Counts and examples of common misinformation types on PTT.

Label	Threads, n (%)	Common types found	Example
True information (n=3830, 65.8%)			
Factual news	1603 (27.6)	Official reports of vaccine side effects from government sites, often containing a "News" tag, with text fully reported. Most report official announcements or press conference news with no expression of opinion.	Title: "[News] Free registration now open! Vaccination booth stationed at the PX Mart near Christmas Land"
Scientifically accurate analytical content	2227 (38.3)	Nonbiased question-asking by "villagers"[a] that do not carry any slant in content, or a tone that instigates comments.	"Recently, there was a debate on vaccination, which we only heard the opposite side to not vaccinate. Yes, there are always safety concerns for any medical technology, and it is not 100% preventive of reinfection, but are there other reasons?"
Misinformation (n=1601, 27.5%)			
Misleading information	452 (7.8)	Appears as a question (or comment) but has negative intentions of misleading the public (without understanding the origin of the intention).	"We are now allowing children to get Moderna vaccines, but the side effects are large. According to Murphy's Law, even if the probability is low, it is still possible. If a child does die, after vaccination, who is responsible? … looking at a dazzling object, it's just tin foil. Behind cute makeup is just powder; even polished nails have dark edges."
Disinformation	97 (1.7)	Linking the vaccine as a direct cause of other diseases or ailments (death, cyborgism, myocarditis, balding, etc), and deriving these conclusions from personal experience.	"A lot of people are experiencing side effects from the vaccine, and if something happens, no one takes responsibility. Young people do not need vaccination since their ability to recover is incredibly strong, right? QQ"
Propaganda	858 (14.7)	Linking the reasons for vaccination to specific political attitudes. Often have the characteristic of replying to news posts, guided by personal opinions to specific political positions.	"The 'Taliban'[b] really is ahead internationally, even half of their supporters refuse to roll up their sleeves for Medigen [Taiwan's domestic vaccine]. But maybe they can use it for a bath? The party can enjoy a good bath in their benevolent fluid"
Fabricated content	195 (3.4)	Making nonfactual statements or describing stories related to vaccines based on unwarranted assumptions. Often begins with statements such as "I dreamed that..." or "My friend did..." It is difficult to verify the level of fabrication through text.	"Recently, a female college went to her OBGYN after her Moderna injection. She said her menstruation came twice a month. Some people also said after BNT, the heart was uncomfortable, and took a few days to pass. Why are there still people taking the third dose of MRNA? What's the gossip?"

XSL•FO
RenderX

Label	Threads, n (%)	Common types found	Example
Conspiracy	179 (3.1)	Linking vaccine administration or distribution with the hidden interests of the government and pharmaceutical companies. Mostly done along the lines of them achieving the "benefits" of controlling people if they control vaccine distribution.	"The original vaccines are developed using the original virus, but the virus is constantly mutating. Vaccine factories are not quick to develop new ones, and just encourage us to take booster doses. Isn't this just being done to cheat us of our money? Just like Intel, slowly squeezing out toothpaste, then launching the next generation of products when they can no longer make money."
Mixed misinformation	29 (0.5)	Most cases of 2 or more involved some analytical component, used in a propaganda way.	—[c]
Religious beliefs	23 (0.4)	Adding religious slogans or arguments after raising vaccine-related views. In some texts, elaborated arguments include the viewpoints of a "false world" or "unknown forces." Could also be categorized as broadly conspiracy, with religious grounding.	"I believe [username] now. He said that vaccines are all planned to change human beings. Viruses are man-made and everything is planned. He said the heart used to be on the left side of the body, and now it's in the center according to Google. I'm convinced. I do not want the vaccine. Do you believe [username]?"
Unrelated	155 (2.7)	—	—

[a]Villagers is a common term used on PTT that roughly translates to netizens.

[b]"Taliban" is a derogatory internet slang term that refers to the Democratic Progressive Party or "green" party in Taiwan. The Mandarin Chinese term substitutes the "li" in "Taliban" for the homophonic sound for "green."

[c]Not applicable.

Table . Number of comments and reposts by true information and misinformation.

	Average number of comments	Average number of reposts
True information	74.01	0.40
Misinformation	52.63	0.20
Misleading information	52.63	0.20
Disinformation	75.16	0.16
Propaganda	66.63	0.39
Fabricated content	50.21	0.20
Conspiracy theories	51.41	0.26

Based on our coding and descriptive analysis of the different types of misinformation, we found that the narratives in Taiwan closely resemble those observed in Western contexts. One example is the conspiracy theory that vaccination is used as a means of control by governments or corporations [47,48]. Broadly speaking, these are symptoms of an overarching lack of trust in authorities on health, and part of the larger antiscience trend [48-52]. This point is potentially exacerbated in geopolitically tense contexts such as Taiwan. One such example was the many narratives of faulty quality control prevalent for the Pfizer BioNTech vaccine, since its distributor for Greater China, Fosun, was Shanghai-based, and vaccines were refused as a result. The frequent discussion about and linking of vaccine decisions and political parties (66.63 comments per board, 0.39 average reposts, Table 3) also corroborates this point, and it is also a trend found in the United States. As a side note, the United States and Taiwan share similar forms of government systems, as well as bipartisan polarization.

The following analyses use the core network. In the entire network, there were 23,004 unique users, with the average polarity (by volume) being 3.27. The core network is relatively smaller and more polarized than the entire network, with 6035 unique users with an average polarity of 8.70.

Diffusion of Misinformation

Overall, the average word count of posts was 525.5 words while the median was 326.0, indicating a heavy right tail. The same trend was observed for the activeness of the user, with an average of 279.3 engagements—any posts or comments—compared to a median of 129.0. The polarity score

XSL•FO
RenderX

of the core network is long-tail distributed, and it has an average of 38.2 but a median of 4 (Figure 1). For breadth, among threads identified containing true information, 608 (15.9%) were shared at least once. Conversely, 165 (10.3%) of misinformative threads were reposted. There is no statistical difference between the average repost counts across true information and misinformation, with means of 2.39 and 2.33 reposts for true information and misinformation, respectively (t=0.250; P=.803). The 95% confidence interval for the 2 values ranged from −2.2 to 29.1 for true information threads and −1.1 to 10.3 for misinformation threads.

Table 4 presents negative binomial regression results of predicting breadth of information diffusion. As can be seen, the RR of repost for propaganda information is double that of true information (RR=2.07; P<.001). The risk of reposting disinformation is half that of true information (RR=0.48; P=.001). Moreover, posts from more polarized users (RR=0.22; P<.001) do not arouse as much discussion, suggesting the forum is relatively averse to echo chambering in relation to misinformation.

Table . Predictors of breadth (n=2422).

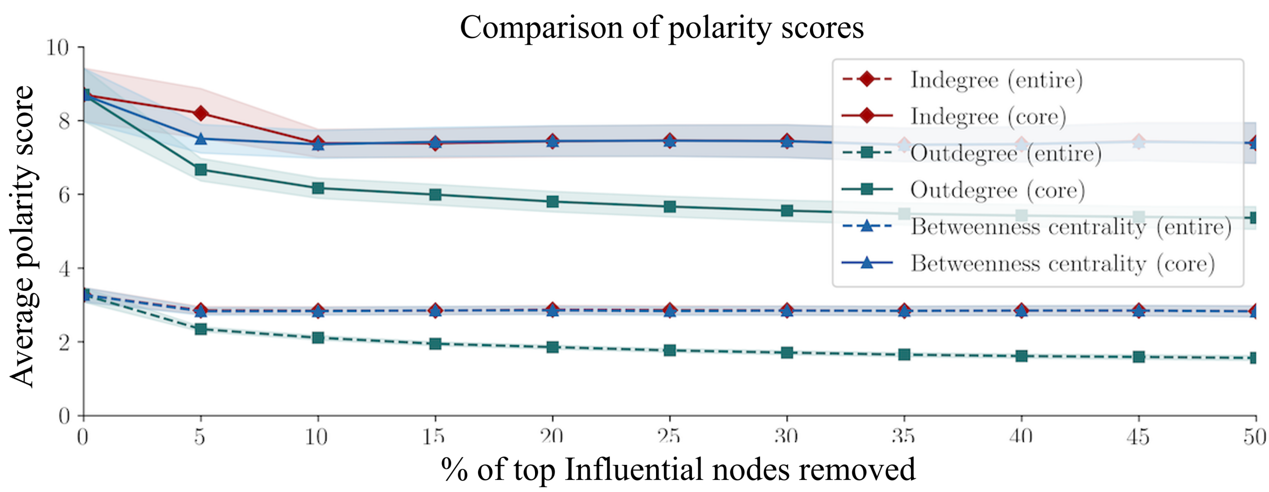
Variables	Breadth		
	Exp(coef)	SE	*P* value
Information type			
True information (baseline)	N/A[a]		
Misleading information	1.10	0.085	.28
Disinformation	0.48	0.227	.001
Propaganda	2.07	0.063	<.001
Fabricate information	0.79	0.156	.12
Conspiracy theories	1.23	0.132	.12
Polarity of user	0.22	0.202	<.001
Control variables			
Word count (of post)	1.0002	0.25×10^{-4}	<.001
User activeness (number of historical posts)	0.998	0.53×10^{-4}	.001

[a]Not applicable.

Node-Level Predictors of Polarity

We also examined whether removing influential users would contribute to decreasing polarization in the misinformation diffusion network. Figure 2 shows the changes of polarity scores for the entire network and the core network of more than 5 posts. The first finding is that, overall, the entire network is less polarized than the core network, as exemplified by the lower average polarity score. This means most of the active users are polarized; they will post predominantly on exclusively true information or misinformation threads. The second finding is that the average polarity score of the network decreased as the most influential users of the network were removed (the top 5% in the entire network and the top 10% in the core network). The 95% confidence intervals are generated from t tests testing the difference in means, with no overlap suggesting significance.

Figure 2. Impact of removing top percent of influential nodes on network (polarity score calculation: absolute value of volume).



XSL•FO
RenderX

These results suggest several trends. The first is that the core network exhibits more echo chambering in its behavior. Thus, those that are more influential on PTT are more likely to be more polarized, gravitating toward true or false information. However, when interpreted in conjunction with the regression results, posts from polarized users were likely to arouse less attention, suggesting that polarized users are more active and influential in commenting instead of posting. The second is that polarization happens with a few key users. In the core network, when the top 10% of nodes using any metric are cut, the polarity score decreases sharply. After 10%, the rate of decrease tapers off. The steeper declines for outdegree and betweenness centrality suggest that identifying active commenters and brokers in the network may be more useful than identifying those that receive many comments.

Discussion

Principal Findings

This study aims to explore vaccine misinformation in Taiwan in 3 ways. The first is by cataloging the types of vaccine misinformation encountered during the pandemic. The second is comparing the diffusion of different types of misinformation. The third is measuring how influential users contribute to the level of polarity of the discussion forum. Overall, the results show that propaganda-style misinformation (RR=2.07; $P<.05$) has wider diffusion, whereas polarized users (RR=0.22; $P<.001$) received less attention for their posts. In addition, by removing the most influential users in the misinformation diffusion network, measured by indegree, outdegree, and betweenness centrality, the average polarity of the entire network decreased. This suggests a potential strategy to combating the echo chamber effect in misinformation by targeting the influential users.

The results from the regression supplement the literature on misinformation virality [53] that suggests that misinformation type may contribute to differences in diffusion. Misinformation propagated with political propaganda intentions has a higher rate of reposting compared to true information. This, in part, could be due to the nature of the PTT forum, which is a heavily politicized forum. In the broader echo chamber literature on science, there are growing concerns of politically aligned selective exposure in science. A study by Nisbet et al [54] focusing on the United States found that, regardless of political affiliation, there are negative and emotionally loaded reactions to dissonant science communication. The downstream effects of these trends are drastic in that overall trust of the scientific community diminishes. Some studies posit that this is due to a phenomenon of "elite cues" in which public opinion follows the elites' partisan battles on the specific issue, such as shown in the climate change debate in the United States [55,56]. Another closer example to vaccination is Hamilton and Safford's study [57] on trust in the US CDC among Republicans following Donald Trump's changing views toward the CDC. Although this study does not analyze elite cues, the same trend may also be true for Taiwan, which has political systems very similar to the United States in terms of bipartisanship and polarization. A most recent cross-national study further suggests the close relationship between individuals' voting choices and their conspiracy beliefs during the COVID-19 pandemic [58]. A further study on the narratives in the propaganda stream is one next step toward corroborating these trends in the literature, and enriching it by providing analysis from a different perspective.

Other than propaganda-based misinformation, the finding that disinformation was reposted less also reinforces that certain streams of misinformation may have more viral potential than others. On content, the findings corroborate other literature findings of longer posts being more transmissible. Posts with longer text are more likely to be spread [43]. However, more active users generally have proportionally less of their boards arouse attention, a finding different from other platforms like X [44]. This could be a natural trend for users whose post volumes are high.

Other findings similar to previous studies include the more prevalent engagement with misinformation compared to true information [59-62], as suggested by the skew of the graph in Figure 1A. This is true despite most boards containing true information. However, the diffusion of true information and misinformation is the same, as seen in Table 4. In addition, this study finds that most misinformation streams also tend to be in either vaccine hesitant or antivaccination stance boards [63-65], a finding not surprising given the affinity of misinformation to an antivaccination agenda [66]. Another new finding is that "brokers" of forum information—influential users as measured by betweenness—disproportionately engage with true information or misinformation across the entire network, suggesting that they are the polarizing forces in the echo chambering of misinformation. These results highlight the importance of measuring polarity in endorsement of misinformation for vaccine hesitancy, an aspect previously underexplored in the literature. By analyzing the polarized endorsement of misinformation, we gain insights into the formation of echo chambers of misinformation arising in a forum and can devise strategies for managing them. The findings for PTT suggest 2 main modes of moving forward in vaccine misinformation management for Taiwan.

The first is the implementation of a consolidated, automatic early detection system of online media potentially containing vaccine misinformation. Much like a disease surveillance system, a misinformation surveillance system should be able to detect and flag potential misinformation early such that its transmission is inhibited if necessary. The reason this is useful is because of the seeming affinity that misinformation has in attracting users, which is more dangerous when those users are influential. Once detected, this information can be used to inform the public as a means of "psychological inoculation" to help internet users better discern misinformation [67]. Given that the narratives of misinformation are not novel in Taiwan, this system can be trained using global reports of vaccine-related misinformation in addition to the civilian-led misinformation clarification platforms already present in Taiwan (MyGoPen, CoFacts, etc). This management system would be particularly important during outbreaks or other flashpoints as online information often peaks as a response to events [68-71]. As a step in infodemic management, it would help improve health literacy.

XSL•FO
RenderX

The second is to create a system that can identify then neutralize influential brokers (ie, high betweenness centrality) and commenters (ie, high outdegree) who have high interactions with misinformation posts. This can reduce the polarity in the network away from negative polarity. In this study, their connectivity means that neutralization may reduce the polarity of the network. Identifying users that are either antivaccination or spreaders of misinformation using these metrics represents an untapped potential for positive engagement that continually breaks the echo chamber effect for both vaccine stance and misinformation spread.

Although the technical aspects of such a system are relatively straightforward, the challenge extends beyond just technical regulation and public health, requiring the balancing of ethical considerations in moderating harmful information without impinging on free speech. This moral dilemma has been studied in other contexts, such as politics and culture [72,73], suggesting that the public supports misinformation management if it causes harm, defined as something undermining people's ability to make informed decisions (particularly around public health and elections) [74]. Studies on the association of misinformation and vaccine intention suggest that they are negatively correlated [15,75]. Other considerations are the scope of required management, such as removal of posts or the temporary or permanent suspension of users. During the COVID-19 pandemic, many social media platforms (eg, Facebook, Instagram) assumed this regulatory role and intervened to prevent the spread of misinformation and conspiracies around vaccines. However, in Taiwan, no such action was taken to actively manage misinformation on local platforms. Rather, the government provided information platforms as secondary references for those already exposed. Moving forward, misinformation management should constitute part of the overall architecture for epidemic management in Taiwan, a process likely to involve fierce democratic discussion or debate on the moral dilemma of speech regulation.

Limitations

A significant limitation of this study is its focus on a single platform, PTT, which may not be representative of the broader online landscape in Taiwan. PTT's user base is distinct in its demographics and its heavily politicized nature, which could skew the findings on the spread of vaccine misinformation and its interaction with user behavior. As such, the results may not fully capture how misinformation diffuses across other popular platforms like Facebook, LINE, or Instagram, which have different user profiles and communication dynamics. Future research could address this limitation by conducting comparative studies across multiple social media platforms to gain a more comprehensive understanding of misinformation diffusion.

Another limitation stems from the use of social media data, which inherently introduces challenges related to user anonymity and the authenticity of user identities. Social media platforms often allow for pseudonymous accounts, making it difficult to verify whether the key users identified in this study are representative of broader population segments or if they include bots or automated accounts, which are known to play a role in the spread of misinformation. Additionally, social media data are typically unstructured, which presents challenges in accurately capturing the context, sentiment, and intent behind posts. This study relied on manual coding, which, despite high intercoder reliability, is subject to human bias. Incorporating more automated methods, such as natural language processing, could reduce bias and improve the scalability of the analysis. Moreover, the rapidly evolving nature of social media platforms, algorithms, and user behavior means that findings based on past data may not hold in future contexts, making longitudinal studies crucial for a more dynamic understanding of misinformation patterns over time.

Conclusion

This study provides key insights into the dynamics of vaccine misinformation in Taiwan, highlighting the influence of core users in spreading polarizing content and the different diffusion patterns between true information and misinformation. The findings underscore the potential for targeted interventions, such as identifying and neutralizing influential users who propagate misinformation, as a means of reducing network polarization and improving public health outcomes. By addressing the mechanisms of misinformation diffusion and understanding the role of user behavior in its propagation, this research contributes to the broader effort to combat misinformation in digital spaces and enhance public trust in scientific communication. Future studies can build on these findings by extending the analysis to other platforms and exploring cross-national comparisons to generalize the results.

Acknowledgments

This study was supported by the CityUHK Strategic Research Grant (7005826 and 7005823).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Testing different thresholds for the cutoff of the core group.

[DOCX File, 729 KB - infodemiology_v5i1e57951_app1.docx]

Multimedia Appendix 2

Correlation analysis for quantitative variables in negative binomial regression.

XSL•FO
RenderX

[DOCX File, 57 KB - infodemiology_v5i1e57951_app2.docx]

References

1. Bessi A, Coletto M, Davidescu GA, Scala A, Caldarelli G, Quattrociocchi W. Science vs conspiracy: collective narratives in the age of misinformation. PLoS One 2015;10(2):e0118093. [doi: 10.1371/journal.pone.0118093] [Medline: 25706981]
2. Colleoni E, Rozza A, Arvidsson A. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. J Commun 2014 Apr;64(2):317-332. [doi: 10.1111/jcom.12084]
3. Bessi A, Petroni F, Del Vicario M, et al. Viral misinformation: the role of homophily and polarization. Presented at: WWW '15: 24th International World Wide Web Conference; May 18-22, 2015; Florence Italy p. 355-356. [doi: 10.1145/2740908.2745939]
4. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. Vaccine (Auckl) 2018 Jun 14;36(25):3606-3612. [doi: 10.1016/j.vaccine.2018.05.040] [Medline: 29773322]
5. Davies P, Chapman S, Leask J. Antivaccination activists on the world wide web. Arch Dis Child 2002 Jul;87(1):22-25. [doi: 10.1136/adc.87.1.22] [Medline: 12089115]
6. Zimmerman RK, Wolfe RM, Fox DE, et al. Vaccine criticism on the World Wide Web. J Med Internet Res 2005 Jun 29;7(2):e17. [doi: 10.2196/jmir.7.2.e17] [Medline: 15998608]
7. Friedlander ER. Opposition to immunization: a pattern of deception. Sci Rev Altern Med 2001;5(1):18-23.
8. Kata A. A postmodern Pandora's box: anti-vaccination misinformation on the Internet. Vaccine (Auckl) 2010 Feb 17;28(7):1709-1716. [doi: 10.1016/j.vaccine.2009.12.022] [Medline: 20045099]
9. Donzelli G, Palomba G, Federigi I, et al. Misinformation on vaccination: a quantitative analysis of YouTube videos. Hum Vaccin Immunother 2018 Jul 3;14(7):1654-1659. [doi: 10.1080/21645515.2018.1454572] [Medline: 29553872]
10. Tandoc EC, Lim ZW, Ling R. Defining 'fake news'. Digital Journalism 2018 Feb 7;6(2):137-153. [doi: 10.1080/21670811.2017.1360143]
11. Wu L, Morstatter F, Carley KM, et al. Misinformation in social media: definition, manipulation, and detection. ACM SIGKDD Explorations Newsletter 2019;21(2):80-90. [doi: 10.1145/3373464.3373475]
12. Zhao S, Hu S, Zhou X, et al. The prevalence, features, influencing factors, and solutions for COVID-19 vaccine misinformation: systematic review. JMIR Public Health Surveill 2023 Jan 11;9(1):e40201. [doi: 10.2196/40201] [Medline: 36469911]
13. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. Hum Vaccin Immunother 2021 Aug 3;17(8):2373-2377. [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
14. Lee SK, Sun J, Jang S, Connelly S. Misinformation of COVID-19 vaccines and vaccine hesitancy. Sci Rep 2022 Aug 11;12(1):13681. [doi: 10.1038/s41598-022-17430-6] [Medline: 35953500]
15. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021 Mar;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
16. Rogers EM. Diffusion of innovations: modifications of a model for telecommunications. Die Diffus von Innov der Telekommunikation 1995:25-38. [doi: 10.1007/978-3-642-79868-9_2]
17. Van den Bulte C. New product diffusion acceleration: measurement and analysis. Marketing Science 2000 Nov;19(4):366-380. [doi: 10.1287/mksc.19.4.366.11795]
18. Yang J, Counts S. Predicting the speed, scale, and range of information diffusion in Twitter. ICWSM 2010;4(1):355-358. [doi: 10.1609/icwsm.v4i1.14039]
19. Zhang L, Peng TQ. Breadth, depth, and speed: diffusion of advertising messages on microblogging sites. Internet Research 2015 Jun 1;25(3):453-470. [doi: 10.1108/IntR-01-2014-0021]
20. van der Linden S, Dixon G, Clarke C, Cook J. Inoculating against COVID-19 vaccine misinformation. EClinicalMedicine 2021 Mar;33:100772. [doi: 10.1016/j.eclinm.2021.100772] [Medline: 33655205]
21. Whitehead HS, French CE, Caldwell DM, Letley L, Mounier-Jack S. A systematic review of communication interventions for countering vaccine misinformation. Vaccine (Auckl) 2023 Jan 27;41(5):1018-1034. [doi: 10.1016/j.vaccine.2022.12.059] [Medline: 36628653]
22. Shelby A, Ernst K. Story and science. Hum Vaccin Immunother 2013 Aug 8;9(8):1795-1801. [doi: 10.4161/hv.24828]
23. Yarchi M, Baden C, Kligler-Vilenchik N. Political polarization on the digital sphere: a cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. Polit Commun 2021 Mar 15;38(1-2):98-139. [doi: 10.1080/10584609.2020.1785067]
24. Suarez Estrada M, Juarez Y, Piña-García CA. Toxic social media: affective polarization after feminist protests. Social Media + Society 2022 Apr;8(2). [doi: 10.1177/20563051221098343]
25. Cossard A, De Francisci Morales G, Kalimeri K, Mejova Y, Paolotti D, Starnini M. Falling into the echo chamber: the Italian vaccination debate on Twitter. ICWSM 2020;14:130-140. [doi: 10.1609/icwsm.v14i1.7285]

XSL•FO
RenderX

26. Dunn AG, Leask J, Zhou X, Mandl KD, Coiera E. Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: an observational study. J Med Internet Res 2015 Jun 10;17(6):e144. [doi: 10.2196/jmir.4343] [Medline: 26063290]
27. Törnberg P. Echo chambers and viral misinformation: modeling fake news as complex contagion. PLoS One 2018;13(9):e0203958. [doi: 10.1371/journal.pone.0203958] [Medline: 30235239]
28. Choi D, Chun S, Oh H, Han J, Kwon TT. Rumor propagation is amplified by echo chambers in social media. Sci Rep 2020 Jan 15;10(1):310. [doi: 10.1038/s41598-019-57272-3] [Medline: 31941980]
29. Kim I, Valente TW. COVID-19 health communication networks on Twitter: identifying sources, disseminators, and brokers. Connect (Tor) 2020 Jan 1;40(1):129-142. [doi: 10.21307/connections-2019.018]
30. Bakshy E, Rosenn I, Marlow C, Adamic L. The role of social networks in information diffusion. 2012 Apr 16 Presented at: WWW 2012; Apr 16-20, 2012; Lyon, France p. 519-528. [doi: 10.1145/2187836.2187907]
31. Weng L, Menczer F, Ahn YY. Virality prediction and community structure in social networks. Sci Rep 2013;3(1):2522. [doi: 10.1038/srep02522] [Medline: 23982106]
32. Meng J, Peng W, Tan PN, Liu W, Cheng Y, Bae A. Diffusion size and structural virality: the effects of message and network features on spreading health information on Twitter. Comput Human Behav 2018 Dec;89:111-120. [doi: 10.1016/j.chb.2018.07.039] [Medline: 32288177]
33. DiFonzo N, Suls J, Beckstead JW, et al. Network structure moderates intergroup differentiation of stereotyped rumors. Soc Cogn 2014 Oct;32(5):409-448. [doi: 10.1521/soco.2014.32.5.409]
34. De Domenico M, Lima A, Mougel P, Musolesi M. The anatomy of a scientific rumor. Sci Rep 2013 Oct 18;3(1):2980. [doi: 10.1038/srep02980] [Medline: 24135961]
35. Difonzo N. Conspiracy rumor psychology. In: Conspiracy Theories and the People Who Believe Them 2018:257-268. [doi: 10.1093/oso/9780190844073.003.0017]
36. Fine GA. Rumor, trust and civil society: collective memory and cultures of judgment. Diogenes (Engl ed) 2007 Feb;54(1):5-18. [doi: 10.1177/0392192107073432]
37. PTT online user statistics [Website in Chinese]. PTT. URL: https://www.ptt.cc/bbs/Record/search?q=上站人次統計 [accessed 2023-07-26]
38. Search results for vaccines. MyGoPen. URL: https://www.mygopen.com/search?q=疫苗 [accessed 2023-11-01]
39. Search “vaccine” (Chinese: 真的假的 “疫苗”). Cofacts. URL: https://cofacts.tw/search?type=messages&q=疫苗 [accessed 2023-11-01]
40. Taiwan FactCheck Center (台灣事實查核中心). URL: https://tfc-taiwan.org.tw/ [accessed 2023-11-01]
41. Zhao X, Wang X, Ma Z, Ma R. Primacy effect of emotions in social stories: user engagement behaviors with breast cancer narratives on Facebook. Comput Human Behav 2022 Dec;137:107405. [doi: 10.1016/j.chb.2022.107405]
42. Wang X, Song Y. Viral misinformation and echo chambers: the diffusion of rumors about genetically modified organisms on social media. INTR 2020 Jun 22;30(5):1547-1564. [doi: 10.1108/INTR-11-2019-0491]
43. Berger J, Milkman KL. What makes online content viral? Journal of Marketing Research 2012 Apr;49(2):192-205. [doi: 10.1509/jmr.10.0353]
44. Suh B, Hong L, Pirolli P, Chi EH. Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. Presented at: 2010 IEEE Second International Conference on Social Computing; Aug 20-22, 2010; Minneapolis, MN. [doi: 10.1109/SocialCom.2010.33]
45. Del Vicario M, Bessi A, Zollo F, et al. The spreading of misinformation online. Proc Natl Acad Sci U S A 2016 Jan 19;113(3):554-559. [doi: 10.1073/pnas.1517441113] [Medline: 26729863]
46. Everett MG, Valente TW. Bridging, brokerage and betweenness. Soc Networks 2016 Jan;44:202-208. [doi: 10.1016/j.socnet.2015.09.001] [Medline: 28239229]
47. Jiang X, Su MH, Hwang J, et al. Polarization over vaccination: ideological differences in Twitter expression about COVID-19 vaccine favorability and specific hesitancy concerns. Social Media + Society 2021 Jul;7(3). [doi: 10.1177/20563051211048413]
48. Wawrzuta D, Jaworski M, Gotlib J, Panczyk M. What arguments against COVID-19 vaccines run on Facebook in Poland: content analysis of comments. Vaccines (Basel) 2021 May 10;9(5):481. [doi: 10.3390/vaccines9050481] [Medline: 34068500]
49. Tustin JL, Crowcroft NS, Gesink D, Johnson I, Keelan J, Lachapelle B. User-driven comments on a Facebook advertisement recruiting Canadian parents in a atudy on immunization: content analysis. JMIR Public Health Surveill 2018 Sep 20;4(3):e10090. [doi: 10.2196/10090] [Medline: 30249585]
50. Okuhara T, Ishikawa H, Okada M, Kato M, Kiuchi T. Japanese anti- versus pro-influenza vaccination websites: a text-mining analysis. Health Promot Int 2019 Jun 1;34(3):552-566. [doi: 10.1093/heapro/day015] [Medline: 29584863]
51. Baines A, Ittefaq M, Abwao M. #Scamdemic, #plandemic, or #scaredemic: what Parler social media platform tells us about COVID-19 vaccine. Vaccines (Basel) 2021 Apr 22;9(5):421. [doi: 10.3390/vaccines9050421] [Medline: 33922343]
52. Lahouati M, De Coucy A, Sarlangue J, Cazanave C. Spread of vaccine hesitancy in France: what about YouTube. Vaccine (Auckl) 2020 Aug 10;38(36):5779-5782. [doi: 10.1016/j.vaccine.2020.07.002] [Medline: 32682617]
53. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 9;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]

XSL•FO
RenderX

54. Nisbet EC, Cooper KE, Garrett RK. The partisan brain: how dissonant science messages lead conservatives and liberals to (dis)trust science. Ann Am Acad Pol Soc Sci 2015;658(1):36-66. [doi: 10.1177/0002716214555474]
55. Merkley E, Stecula DA. Party elites or manufactured doubt? The informational context of climate change polarization. Sci Commun 2018 Apr;40(2):258-274. [doi: 10.1177/1075547018760334]
56. Brulle RJ, Carmichael J, Jenkins JC. Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the U.S., 2002–2010. Clim Change 2012 Sep;114(2):169-188. [doi: 10.1007/s10584-012-0403-y]
57. Hamilton LC, Safford TG. Elite cues and the rapid decline in trust in science agencies on COVID-19. Sociol Perspect 2021 Oct;64(5):988-1011. [doi: 10.1177/07311214211022391]
58. Lin F, Meng X, Zhi P. Are COVID-19 conspiracy theories for losers? Probing the interactive effect of voting choice and emotional distress on anti-vaccine conspiracy beliefs. Humanit Soc Sci Commun 2025;12(1):470 [FREE Full text] [doi: 10.1057/s41599-025-04774-3]
59. Hou Z, Tong Y, Du F, et al. Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. J Med Internet Res 2021 Jun 11;23(6):e27632. [doi: 10.2196/27632] [Medline: 34061757]
60. Yiannakoulias N, Slavik CE, Chase M. Expressions of pro- and anti-vaccine sentiment on YouTube. Vaccine (Auckl) 2019 Apr 3;37(15):2057-2064. [doi: 10.1016/j.vaccine.2019.03.001] [Medline: 30862365]
61. Gori D, Durazzi F, Montalti M, et al. Mis-tweeting communication: a vaccine hesitancy analysis among Twitter users in Italy. Acta Biomed 2021 Oct 5;92(S6):e2021416. [doi: 10.23750/abm.v92iS6.12251] [Medline: 34739459]
62. Karapetiantz P, Audeh B, Lillo-Le Louët A, Bousquet C. Discrepancy between personal experience and negative opinion with human papillomavirus vaccine in web forums. Stud Health Technol Inform 2020 Jun 26;272:417-420. [doi: 10.3233/SHTI200584] [Medline: 32604691]
63. Bradshaw AS, Treise D, Shelton SS, et al. Propagandizing anti-vaccination: analysis of Vaccines Revealed documentary series. Vaccine (Auckl) 2020 Feb 18;38(8):2058-2069. [doi: 10.1016/j.vaccine.2019.12.027] [Medline: 31980194]
64. Basch CH, Zybert P, Reeves R, Basch CE. What do popular YouTubeTM videos say about vaccines. Child Care Health Dev 2017;43(4):499-503. [doi: 10.1111/cch.12442]
65. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis 2021 Jul;108:256-262. [doi: 10.1016/j.ijid.2021.05.059] [Medline: 34052407]
66. Yin JDC. Media data and vaccine hesitancy: scoping review. JMIR Infodemiology 2022;2(2):e37300. [doi: 10.2196/37300] [Medline: 37113443]
67. Roozenbeek J, van der Linden S, Goldberg B, Rathje S, Lewandowsky S. Psychological inoculation improves resilience against misinformation on social media. Sci Adv 2022 Aug 26;8(34):eabo6254. [doi: 10.1126/sciadv.abo6254] [Medline: 36001675]
68. Furini M. Identifying the features of ProVax and NoVax groups from social media conversations. Comput Human Behav 2021 Jul;120:106751. [doi: 10.1016/j.chb.2021.106751]
69. Diaz P, Reddy P, Ramasahayam R, Kuchakulla M, Ramasamy R. COVID-19 vaccine hesitancy linked to increased internet search queries for side effects on fertility potential in the initial rollout phase following Emergency Use Authorization. Andrologia 2021 Oct;53(9):e14156. [doi: 10.1111/and.14156] [Medline: 34181273]
70. Odone A, Tramutola V, Morgado M, Signorelli C. Immunization and media coverage in Italy: an eleven-year analysis (2007-17). Hum Vaccin Immunother 2018;14(10):2533-2536. [doi: 10.1080/21645515.2018.1486156] [Medline: 29993316]
71. Deiner MS, Fathy C, Kim J, et al. Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Informatics J 2019 Sep;25(3):1116-1132. [doi: 10.1177/1460458217740723] [Medline: 29148313]
72. Kozyreva A, Herzog SM, Lewandowsky S, et al. Resolving content moderation dilemmas between free speech and harmful misinformation. Proc Natl Acad Sci U S A 2023 Feb 14;120(7):e2210666120. [doi: 10.1073/pnas.2210666120] [Medline: 36749721]
73. Lin F, Chen X, Cheng EW. Contextualized impacts of an infodemic on vaccine hesitancy: The moderating role of socioeconomic and cultural factors. Inf Process Manag 2022;59(5):103013.
74. Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a single market for digital services (digital services act) AND amending directive 2000/31/EC. European Union. URL: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A52020PC0825 [accessed 2024-01-09]
75. Chen YP, Chen YY, Yang KC, et al. The prevalence and impact of fake news on COVID-19 vaccination in Taiwan: retrospective study of digital media. J Med Internet Res 2022 Apr 26;24(4):e36830. [doi: 10.2196/36830] [Medline: 35380546]

Abbreviations

CDC: Centers for Disease Control and Prevention
RR: relative risk

XSL•FO
RenderX

Edited by T Mackey; submitted 01.03.24; peer-reviewed by S Zhang, T Cerqueira-Silva; revised version received 14.10.24; accepted 24.10.24; published 18.08.25.

Please cite as:
Yin JDC, Wu TC, Chen CY, Lin F, Wang X
The Role of Influencers and Echo Chambers in the Diffusion of Vaccine Misinformation: Opinion Mining in a Taiwanese Online Community
JMIR Infodemiology 2025;5:e57951
URL: https://infodemiology.jmir.org/2025/1/e57951
doi:10.2196/57951

© Jason Dean-Chen Yin, Tzu-Chin Wu, Chia-Yun Chen, Fen Lin, Xiaohui Wang. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 18.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Unraveling the Use of Disinformation Hashtags by Social Bots During the COVID-19 Pandemic: Social Networks Analysis

Victor Suarez-Lledo[1,2*], PhD; Esther Ortega-Martin[1,3*], MSc; Jesus Carretero-Bravo[1,4*], PhD; Begoña Ramos-Fiol[1*], MSc; Javier Alvarez-Galvez[1,3*], PhD

[1]Computational Social Science DataLab, University Institute of Research for Sustainable Social Development (INDESS), University of Cadiz, Jerez de la Frontera, Spain

[2]Department of Sociology, University of Granada, Granada, Spain

[3]Department of General Economy (Sociology Area), Faculty of Nursing and Physiotherapy, University of Cadiz, Cadiz, Spain

[4]Department of Quantitative Methods, Universidad Loyola Andalucía, Seville, Spain

[*]all authors contributed equally

Corresponding Author:
Victor Suarez-Lledo, PhD
Computational Social Science DataLab
University Institute of Research for Sustainable Social Development (INDESS)
University of Cadiz
Avda. de la Universidad, 4
Jerez de la Frontera, 11406
Spain
Phone: 34 956167216
Email: victor.sanz@uca.es

Abstract

Background: During the COVID-19 pandemic, social media platforms have been a venue for the exchange of messages, including those related to fake news. There are also accounts programmed to disseminate and amplify specific messages, which can affect individual decision-making and present new challenges for public health.

Objective: This study aimed to analyze how social bots use hashtags compared to human users on topics related to misinformation during the outbreak of the COVID-19 pandemic.

Methods: We selected posts on specific topics related to infodemics such as vaccines, hydroxychloroquine, military, conspiracy, laboratory, Bill Gates, 5G, and UV. We built a network based on the co-occurrence of hashtags and classified the posts based on their source. Using network analysis and community detection algorithms, we identified hashtags that tend to appear together in messages. For each topic, we extracted the most relevant subtopic communities, which are groups of interconnected hashtags.

Results: The distribution of bots and nonbots in each of these communities was uneven, with some sets of hashtags being more common among accounts classified as bots or nonbots. Hashtags related to the Trump and QAnon social movements were common among bots, and specific hashtags with anti-Asian sentiments were also identified. In the subcommunities most populated by bots in the case of vaccines, the group of hashtags including #billgates, #pandemic, and #china was among the most common.

Conclusions: The use of certain hashtags varies depending on the source, and some hashtags are used for different purposes. Understanding these patterns may help address the spread of health misinformation on social media networks.

(JMIR Infodemiology 2025;5:e50021) doi:10.2196/50021

KEYWORDS

social media; misinformation; COVID-19; bot; hashtags; disinformation; network analysis; community detection; dissemination; decision-making; social bot; infodemics; tweets; social media network

Introduction

From the swine influenza (H1N1) pandemic in 2009 to the subsequent outbreak of the H7N9 virus, also known as bird flu, in 2013, Twitter (subsequently rebranded as X) has increasingly become a popular platform for sharing health information [1,2]. Using posts, users can express their thoughts and opinions on many health topics. That is why specific interaction tasks have attracted the attention of researchers. This research can inform public policy by encouraging governments and health care professionals to allocate necessary resources, act, and plan accordingly [3,4]. These social media platforms have played a crucial role in providing information to the public during the COVID-19 pandemic. However, there was an increase in low-quality information, as well as the infodemic phenomenon. The infodemic, defined as an excess of information that makes it difficult for people to find reliable sources [5], can have harmful consequences [6].

The COVID-19 pandemic triggered mandatory lockdowns, social distancing, quarantines, and SARS-CoV-2–protective measures that would give rise to all sorts of opinions and behaviors [7]. During the COVID-19 pandemic, mandatory lockouts drastically altered people's daily routines (work, travel, and leisure activities) to levels never before experienced by the populations of the different countries affected by the new disease [8]. The state of uncertainty in the face of an invisible threat would transform previously normal situations into situations of risk. Direct social interaction with people outside the nuclear family, attending a concert, meeting for dinner with friends and family, shaking hands with someone, and even hugging or kissing became exceptional situations during the most uncertain periods of the pandemic—situations that, as has been observed retrospectively, would have a significant impact on the mental health of the population [9]. Likewise, the health crisis gave rise to the infodemic that, through social media platforms, opened the door to fake news, misconceptions, hoaxes, and anecdotal evidence about the origin of the pandemic, the social agents to blame for the situation, and the possible measures to be taken at a time of maximum uncertainty [10].

To understand how during the new context of health emergency misinformation spreads on these platforms, studies analyzed different elements, including the quality of information sources through URL analysis; identification of topics that generate misinformation; and analysis of online communities that spread misinformation, such as the antivaccine movement [11-14]. Others focused on the use of hashtags to describe the organization of the debate around the COVID-19–related topics. Researchers examined the frequency of use and the topic analysis of hashtags, and emphasized their main role in certain conversations [15,16]. By analyzing specific hashtags, studies also demonstrated how antivaccine communities, the proliferation of racist sentiments, or the spread of conspiracy theories are articulated on social media [17-19]. Some studies paid particular attention to how hashtags were used or combined in online conversations about the COVID-19 pandemic, using clustering techniques to describe the themes and combining hashtags with semantic text analysis and natural language processing (NLP) methods to improve topic modeling [20-22]. In addition, social network analysis (SNA) became useful to examine the co-occurrence of hashtags [23]. These studies demonstrate how the combination of different approach is useful to analyze online conversations more thoroughly.

Recently, the role of social bots has contributed to the spread of misinformation on social media platforms in various ways [24]. This issue garnered more attention as fake news and misinformation were significant factors during the COVID-19 pandemic. In this sense, some studies analyzed the role of bots regarding the spread of misinformation in general, while others have focused specifically on topics such as vaccines, conspiracy theories, hate speech, or reactions to other political actions [25-31]. However, a small amount of research compared the behavior of bots and humans [32,33].

To better understand the influence of bots on social media conversations, a previous study used topic modeling to segment the Twitter conversation and compare differences between accounts [34]. Nevertheless, the analysis did not focus on the usage of hashtags, which is the primary focus of this study. We aim to identify patterns and trends in hashtag usage to describe how bots and nonbots differ in their use of hashtags.

Only a few studies analyzed how social media bots use hashtags compared to humans. Most studies in this field examine specific hashtags [17-19,35-37]. To address this knowledge gap, we explore how social bots use hashtags specifically in connection with certain infodemic topics, issues that contribute to the generation or spread of fake news, misinformation, or discriminatory narratives. By analyzing how frequently hashtags co-occur, we aim to understand how they appear in the conversation and how they are combined. Besides, we also considered the context in which hashtags are used. They can be used ironically or convey disagreement. Our goal is to address three key questions: (1) What are the most common hashtag co-occurrences? (2) What are the differences in hashtag usage between bots and nonbots? and (3) Do bots and nonbots use certain hashtags in different ways?

Methods

Data Collection

Data collection for this study took place from March 16 to June 15, 2020, using the Twitter Streaming application programming interface (API). The hashtags #covid_19, #covid19, #covid, and #coronavirus were used to capture conversations about the first wave of COVID-19 pandemic, and only English-language posts were selected.

Based on previous research, we created a list of topics that were commonly associated with fake news or misinformation. This list includes ozone, laboratory, 5G, conspiracy, Bill Gates, milk, military, and UV. Vaccines were also identified as a controversial topic in multiple studies, so we added them to the list [38-40].

Ethical Considerations

The present study was approved by the Ethics Committee of the University of Cadiz (005_2024).

Bot Classification

To identify whether accounts on Twitter were bots or not, we used Botometer by OsoMe (formerly known as BotOrNot) [41]. This publicly available application uses over a thousand criteria to determine how closely a Twitter account matches the typical characteristics of social bots.

To create a binary classification (bot or nonbot) and prioritize identifying true positives over true negatives, we set a threshold

XSL•FO
RenderX

value of 0.8 [34]. Using this threshold, we classified approximately 14.8% of the accounts as bots, which is in line with the findings of other research that found bot levels to be between 9% and 15% of the total number of Twitter accounts [42].

Botometer also provides rankings for 6 main types of bots, including echo-chamber, fake follower, financial, self-declared, spammer, and others, in addition to the overall likelihood of being a bot. In this study, we focused on analyzing the behavior of social bot accounts, particularly those that were not identified as automated accounts. These types of accounts are often associated with press agencies, companies, newspapers, or journals, and their primary purpose is to automatically publish information about a specific topic. These accounts may indicate that they are automated, for example, by including the word "bot" in their screen name or being identified as bots on Botwiki [41]. Therefore, we chose to exclude self-declared bots from our analysis due to their different characteristics compared with other social bots [41].

For this study, we classified accounts as nonbots if their probability of being a bot was less than 0.8, as self-declared bots if their probability of being a self-declared bot was greater than 0.8, and as bots if their probability of being a bot was greater than 0.8 and their probability of being a self-declared bot was less than 0.8. We then filtered out self-declared bots and considered both bots and nonbots for analysis.

Network Analysis

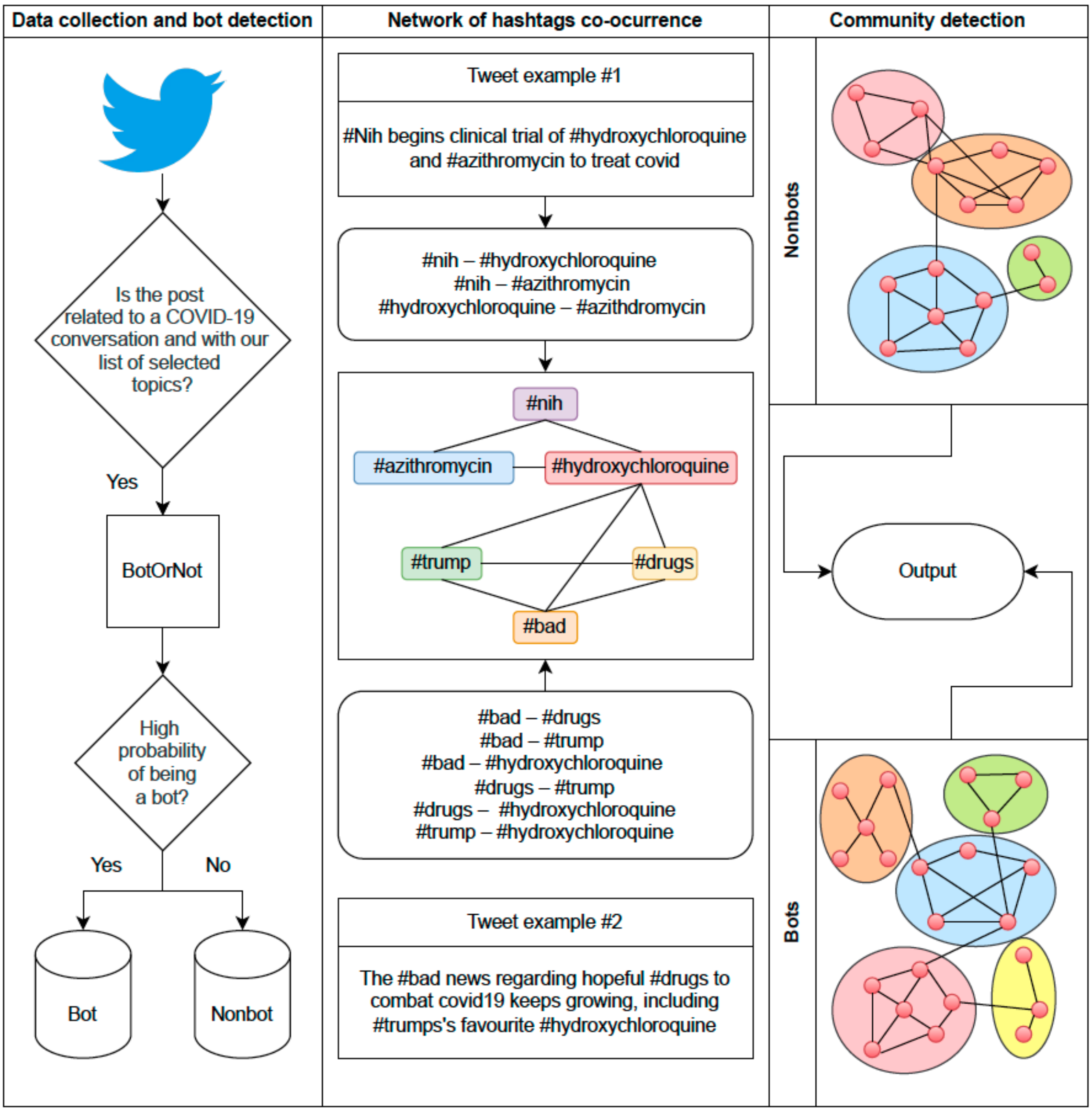
To identify patterns in the usage of hashtags, we applied network analysis. We constructed a network by analyzing the co-occurrence of hashtags in posts and comparing the use of hashtags by bots and nonbots. In the network, hashtags were represented as nodes, and they were connected if they appeared in the same post. The weight of the connection between 2 hashtags was determined by the number of times they co-occurred.

We also calculated various metrics of connection, distribution, and segmentation of the hashtag network. We used the PageRank algorithm to identify the most important nodes in the network and the degree value, which represents the number of connections each hashtag has [43]. We also used the betweenness metric, which measures centrality [44]. In addition, we used the Louvain algorithm to detect the most important communities in the network. This algorithm maximizes a modularity score for each community, where the modularity measures the quality of the assignment of nodes to communities. This allowed us to identify hashtags that often co-occur together. We computed each metric separately considering whether the hashtags appear in posts posted by a bot or a nonbot. Figure 1 contains a flow diagram for the entire process.

In the following section, we first present the results for the entire network. In the following subsections, 1 for each topic, we segment the overall network of hashtag co-occurrences by extracting posts that specifically mention words related to each topic. For example, the network for vaccines will show the co-occurrences of all hashtags that appeared in posts about vaccines.

XSL•FO
RenderX

Figure 1. Flowchart from data collection to analysis.



Results

Overview

In total, we extracted around 107,173 posts from March to July 2020 that were related to the topics on our list. Most of these posts were about vaccines (59,090/107,173, 55.1%), hydroxychloroquine (17,731/107,173, 16.5%), or the military (12,548/107,173, 11.5%). Out of all the accounts analyzed, 85.2% (91,311/107,173) were identified with a low likely of being bots, that is, nonbots. Approximately 14.8% (15,862/107,173) of the posts were classified as likely being from bot accounts. As shown in Figure 2, the number of posts related to vaccines was consistently higher throughout the period, except for 2 specific moments. The first of these coincides with a message from US President Donald Trump recommending the use of hydroxychloroquine, an unproven drug. The second date also coincides with a message from Trump suggesting the injection of disinfectant to beat COVID-19 pandemic.

We created a graph of the full network of hashtags. For clarity, we selected a random sample from the entire collection of posts and depicted it in Figure 3. We also applied color to the Louvain communities and highlighted some hashtags that represent the topics analyzed in the study. This process is like the one we used for each topic in the list.

XSL•FO
RenderX

Figure 2. Bot and nonbot distribution by topic and date.

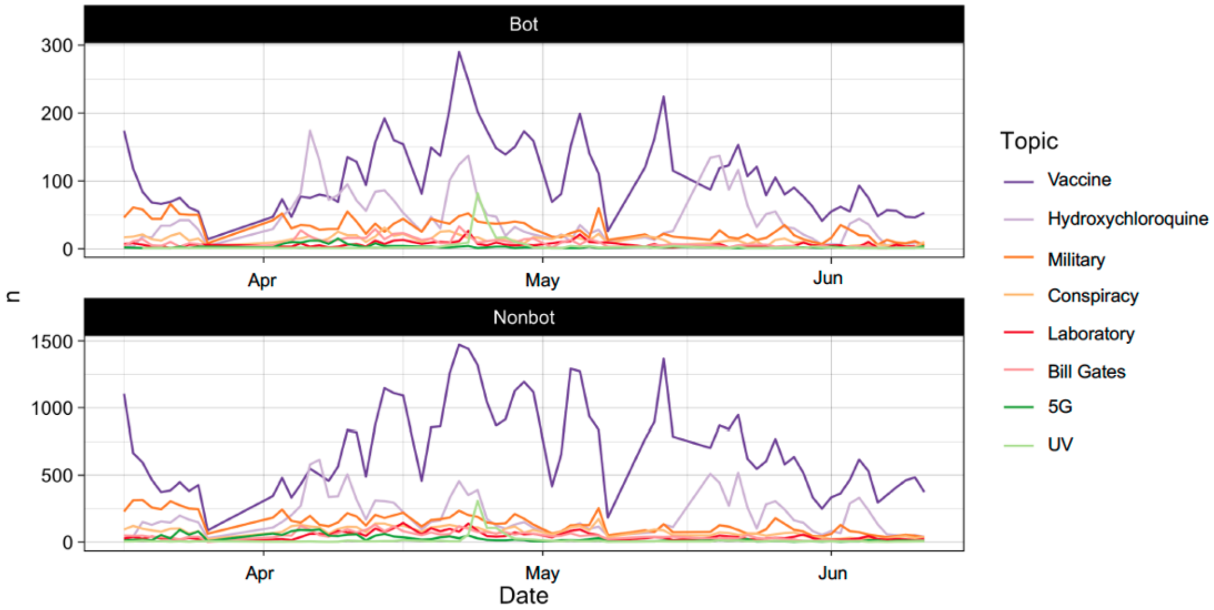
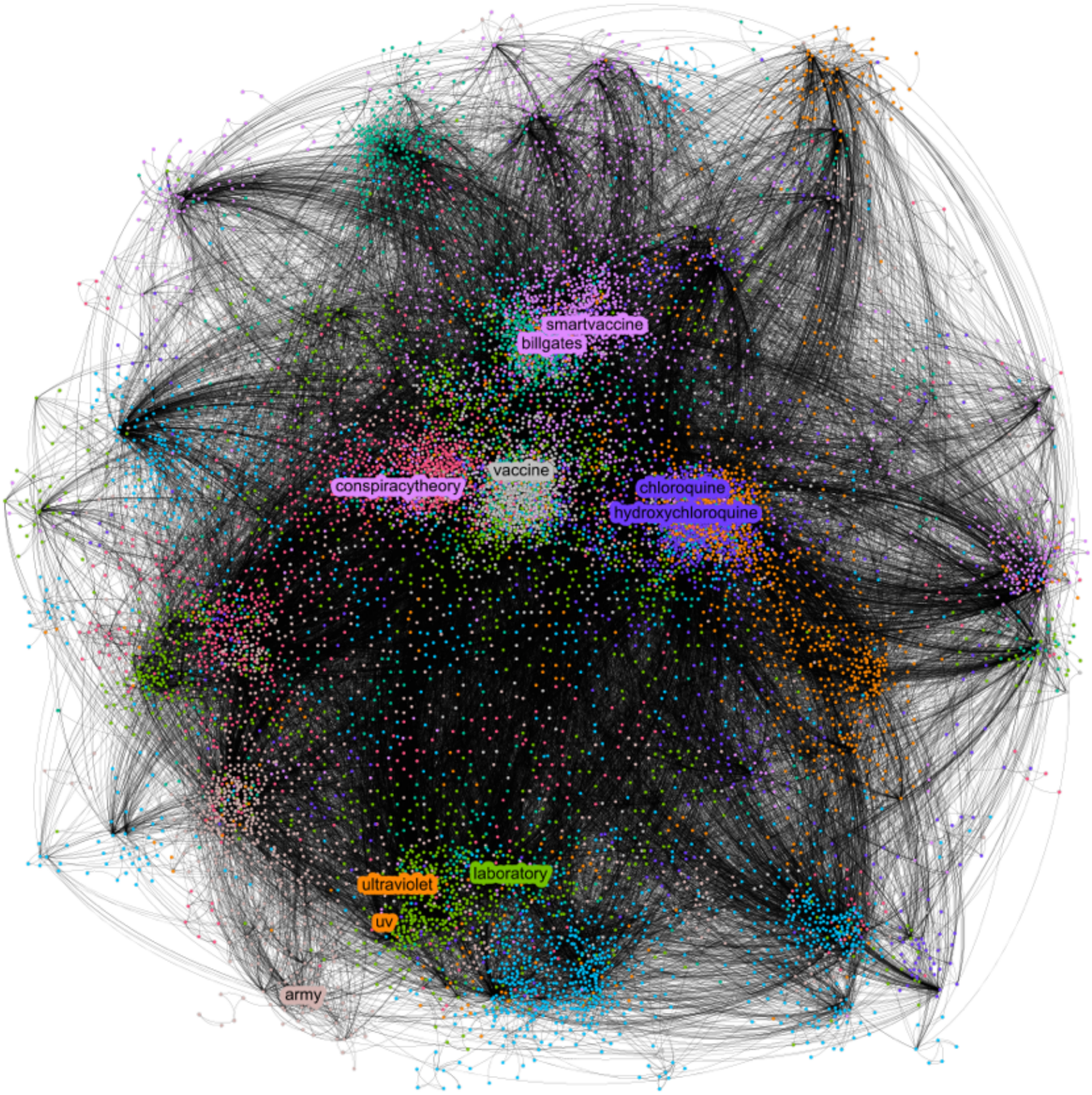


Figure 3. Hashtag network.



XSL•FO
RenderX

In Table 1, we present statistics for the overall hashtags network to provide a broad overview. As mentioned earlier, we calculated the metrics separately for each type of account. There are some differences in the most used hashtags between the 2 groups. For example, hashtags such as #Trump, #China, and #BillGates appear in both groups. However, the hashtag #vaccineswork is one of the most used by nonbots, while the hashtag #lka (which is the country code for Sri Lanka) is more frequently used by bots.

Table 1. Most common co-occurrences by bot and nonbot.

Hashtags	Posts, n (%)
Bots (n=3459)	
#chloroquine - #hydroxychloroquine	537 (15.52)
#hydroxychloroquine - #trump	490 (14.17)
#africaisnotalaboratory - #changeyourworld	437 (12.63)
#azithromycin - #hydroxychloroquine	345 (9.97)
#coronavirushoax - #prisonearth	280 (8.09)
#digitalvirus - #policestate	280 (8.09)
#digitalvirus - #prisonearth	280 (8.09)
#policestate - #prisonearth	280 (8.09)
#coronaviruslockdown - #lockdownextension	267 (7.72)
#changeyourworld - #coronacrisisuk	263 (7.6)
Nonbots (n=665)	
#hydroxychloroquine - #trump	133 (20)
#climatechange - #sustainability	106 (15.94)
#lka - #srilanka	86 (12.93)
#chloroquine - #hydroxychloroquine	84 (12.63)
#azithromycin - #hydroxychloroquine	72 (10.83)
#kag - #maga	53 (7.97)
#pandemic - #vaccine	35 (5.26)
#billgates - #vaccines	33 (4.96)
#kag - #qanon	33 (4.96)
#china - #vaccine	30 (4.51)

There are also some similarities in the co-occurrence of hashtags between the 2 groups. For example, hashtags #hydroxychloroquine and #trump appear in the same posts with higher frequency in both cases, at 14.17% (490/3459) and 20% (133/665), respectively. However, other hashtag pairs such as #kag-#maga, #billgates-#vaccines, or #kag-#qanon are common among bots. “KAG” stands for “Keep America Great,” which was President Trump’s campaign slogan in 2020, while “MAGA” stands for “Make America Great Again,” which was his campaign slogan in 2016. Both slogans have been associated with American nationalism, and the hashtag #MAGA has sometimes been used by white supremacist groups and Trump supporters.

On the other hand, nonbots tend to use other hashtag pairs such as #coronavirushoax-#prisionearth, #digitalvirus-#policestate, and #digitalvirus-#prisionearth. These hashtags, especially “#prisionearth,” were often used ironically to mock false rumors or exaggerations that were circulated online.

Vaccines

The most common co-occurrent hashtags used by nonbots regarding vaccines are #uk-#usa, #research-#science, #vaccineswork-#worldimmunizationweek. However, the most common hashtags in those posts posted by bots are #trump-#votebluetosaveamerica, #healthcare-#ppe, or even #healthcare-#ventilators. In addition, these last mentioned are exclusive of bots. That is, they only co-occur in posts from accounts classified as bots. Besides, it is worth mentioning that #billgates, along with #pandemic or #china, are the hashtags with the highest degree of connections, as seen in Table 2.

XSL•FO
RenderX

Table 2. Most important hashtags by topic.

Hashtags	Degree	PageRank	Betweenness
Vaccine			
billgates	44	0.025	22,728
pandemic	39	0.019	26,196
china	35	0.019	12,380
usa	30	0.013	7,375
vaccineswork	28	0.019	8,833
trump	28	0.015	15,704
stayhome	22	0.011	4,583
uk	21	0.010	2,703
science	21	0.011	5,048
france	19	0.008	2,064
Military			
trump	34	0.042	8,032
china	27	0.030	3,733
usa	22	0.026	5,561
italy	16	0.023	4,219
us	16	0.019	1,667
iran	15	0.020	1,938
russia	11	0.015	1,353
maga	10	0.012	620
wuhan	10	0.012	497
breaking	9	0.012	2,372
Laboratory			
wuhan	36	0.045	8,422
laboratory	26	0.033	11,660
africaisnotalaboratory	21	0.041	4,641
china	20	0.023	3,470
staysafe	11	0.017	7,566
stayhome	10	0.013	9,242
us	8	0.009	476
pandemic	8	0.009	8,614
coronaviruslockdown	7	0.011	1,676
healthcare	7	0.009	1,331
5G			
china	42	0.020	31,413
pandemic	27	0.012	25,136
wuhan	19	0.009	13,463
iot	18	0.008	11,045
qanon	17	0.008	6,437
bigdata	17	0.007	7,446
technology	17	0.008	8,731
ai	14	0.007	4,819

XSL•FO
RenderX

Hashtags	Degree	PageRank	Betweenness
tech	14	0.006	4,455
fakenews	14	0.007	8,353
Hydroxychloroquine			
trump	54	0.074	10,106
chloroquine	20	0.028	2,538
coronaviruspandemic	15	0.020	1,515
kag	14	0.017	897
maga	13	0.017	2,197
coronavirusoutbreak	12	0.016	1,089
india	12	0.017	855
hcq	12	0.020	1,468
usa	12	0.015	2,095
gop	11	0.014	636
Conspiracy			
conspiracy	35	0.084	1,872
conspiracytheory	25	0.054	2,111
conspiracytheories	16	0.037	686
pandemic	16	0.033	878
china	15	0.032	785
trump	12	0.030	732
disinformation	10	0.022	77
fakenews	10	0.023	321
usa	10	0.024	778
us	9	0.020	213
Bill Gates			
billgates	68	0.056	17,637
qanon	29	0.023	4,043
pandemic	27	0.024	7,341
maga	23	0.017	1,650
vaccines	19	0.016	5,232
stopbillgates	15	0.011	862
kag	13	0.009	104
trump	13	0.011	1,049
microsoft	13	0.010	1,978
usa	13	0.010	1,173
UV			
ai	14	0.041	839
trump	11	0.044	1,427
health	8	0.025	491
innovation	8	0.024	171
pandemic	8	0.029	428
uvlight	8	0.028	1,617
robots	7	0.023	754

XSL•FO
RenderX

Hashtags	Degree	PageRank	Betweenness
artificialintelligence	6	0.018	112
lysol	5	0.018	122
machinelearning	5	0.016	255

The algorithm extracted 5 different communities (Multimedia Appendix 1). We found significant differences in the hashtags that made up the Louvain communities. The first community contains hashtags related to news (#breaking, #usnews, and #breakingnews); countries (#canada, #france, #japan, #spain, and #africa); and others related to fake news like #wuhanvirus, #ccpvirus, #bioweapon, #hiddenhand, #psychopaths, #chinaisassho, and #madeinchina. This community is the most populated by bots, and the difference between bots and nonbots is the highest.

The second community contains hashtags related to famous people (#billgates, #anthonyfauci, and #georgesoros). These include people like Bill Gates and Anthony Fauci who played a leading role by holding provaccine positions. As in the previous case, we also found some hashtags related to fake news or conspiracy theories such as #billgatesisevil, #billgatesvaccine, #vaccinemafia, or #newworldorder. In this community, the quantity of nonbots is slightly higher than the number of bots.

On the other hand, the number of bots is also higher in the third community. In this case, the hashtags mention politics, such as #trump, #biden, and #borisjohnson. In addition, there were also some hashtags related to measures to curb the pandemic, such as #stayhome, #socialdistancing, or #lockdown. Only a few infodemic-related hashtags were found: #methanemouth, #pussygrabber, or #bananarepublic. The number of nonbots is higher in the other 2 communities. The fourth and fifth communities contain hashtags related to research and vaccines (#research, #health, and #medicine) or diseases and public health campaigns (#vaccineswork, #measles, #endpolio, and #healthforall), respectively. In particular, #vaccineswork is a hashtag used by health institutions such as the World Health Organization. Conversations on these hashtags were related to second waves and the importance of vaccines to fight against the COVID-19 pandemic.

Hydroxychloroquine

Hashtags related to Trump and the Republican movement were common in the case of hydroxychloroquine. These hashtags, such as #kag, #maga, #gop, #qanon, and #tcot, were more common in bot posts. Although #trump also appears in the case of nonbots, there were other hashtags related to news: #breaking-#breakingnew and #chinavirus-#wuhanvirus. Consequently, #trump has the highest degree of connection and the one with the highest betweenness. This hashtag, along with #chloroquine or #coronaviruspandemic, is the hashtag with the highest number of connections. There is a big difference between the first and the rest of the hashtags shown in Table 2. This difference indicates the leading role that #trump plays in the conversation about hydroxychloroquine.

We identified 8 different communities (Multimedia Appendix 1). Regarding the composition of the communities, it is worth mentioning the difference between the 2 most important ones. On the one hand, the first contains hashtags related to drugs, vaccines, or the pharmaceutical industry: #azithromycin, #biotech, #chloroquine, #lupus, #malaria, #cdc, or #hcq. In the same line, in the fourth community, the predominance of nonbots is noticeable. This time the hashtags mention countries (#uk, #us, #coronavirusuk, #france, #italy, and #germany), news (#worldnews and #usnews), TV series (#greysanatomy and #littlefireseverywhere), and supporting hashtags (#inthistogether).

On the other hand, in the second community, most of the hashtags are related to Trump or social movements related to him (#trump, #gop, #maga, and #donaldtrump). Nonetheless, some are against him (#notaleader, #worstpresidentinhistory, and #putinpuppet). In addition, the number of bots is higher than the number of nonbots, contrary to what happens in the first one.

Military

In this case, hashtags are related to specific countries that were mentioned during the pandemic. For nonbots, those most mentioned are #china-#us, #italy-#russia, and #lka-#srilanka. The latter is the most common among bots, followed in fourth place by #italy-#russia. Among the sets that do not mention countries, we find hashtags related to Trump (#gop-#trump, #kag-#maga, and #kag-#qanon).

These hashtags have similarities to those of hydroxychloroquine. The bots' unique hashtags are related to the Trump movement or Republican movements (#gop, #kag, and #qanon). In addition, #trump has the highest degree of connectivity and betweenness. This situation is also present in the communities (Multimedia Appendix 1). The first community detected contains hashtags related to Trump, and the second is related to military and veterans (#usmc, #veterans, or #usairforce). In both cases, these relationships take place in posts posted by bots.

Conspiracy

In this group, we found some hashtags related to conspiracy theories (or misinformation) and others related to countries. Regarding bots, the most common hashtags are #fakenews-#technology, #conspiracytheories-#socialmedia, and #donthecon-#trumplies. In line with this, for the nonbots, the most common hashtags are #conspiracytheory-#woke. The hashtags used only by bots are also related to racism (#racism-#sinophobia) or the economic system (#capitalismfails-#socialismworks).

Of the 6 most prominent communities (Multimedia Appendix 1), 3 of them have only nonbots. Topics in these communities are about minority groups (#blackpeople, #lgbt, and #amerikkka), about Trump (#maga, #bananarepublic, and #qanon), and about the pandemic (#coronavirusoutbreak,

XSL•FO
RenderX

#coronaviruspandemic, and #pandemictech). Of the other 3, in the first one, the number of nonbots is slightly higher than the number of bots. Some of the hashtags have to do with conspiracy theories (#conspiracytheory, #disinformation, and #propaganda), media (#qanonnfoxnews, #propaganda, and #fakenews), and others in a derogatory tone (#covidiot, #plandemic, and #plandemicdocumentary). On the other hand, in the second and fifth communities, the numbers of bots are higher. In this case, the most common hashtags are related to countries (#china, #us, and #iran), Iran specifically (#irancovidtruth and #iranregimechange), or against right-wing political parties (#rightwingignorance).

Laboratory

In this case, there are apparent differences in the geographical areas of the most used hashtags. On the one hand, nonbots mostly use #africaisnotalaboratory, while bots use #srilanka and #lka (country code for Sri Lanka). The hashtag #indiafightscorona is also common for bots. The hashtags #china-#wuhan are very common in both cases. This explains why #wuhan is the hashtag with the highest PageRank value and the highest degree of connection (Table 2), followed by #laboratory in second place and #africaisnotalaboratory in third place.

The differences between hashtags and the type of account that wrote the message were very clear in this case. On the one hand, in the first and fourth communities, the presence of bots is higher than nonbots (Multimedia Appendix 1). The first is focused on China, with some examples such as #ccpvirus, #chinamustexplain, or #chinaliedpeopedied, and the second is focused on Southeast Asia, such as #armenia, #abudhabi, or #masdarcity.

Bill Gates

The data from the Bill Gates conversation are similar to those obtained in the case of hydroxychloroquine. Trump-related hashtags were very common (#kag, #maga, and #qanon) in both bots and nonbots. The centrality and degree values are among the highest, as can be seen in Table 2. There were also new hashtags related to this type of political movement that only appears in this conversation, such as #crimesagainsthumanity, #gatesofhell, or #greatawakening. In addition, hashtags disparaging the figure of Bill Gates are also common, such as #saynotobillgates or #billgatesisevil.

We identified 5 communities of hashtags (Multimedia Appendix 1). Among the 3 largest communities, the number of bots is higher than the number of nonbots in the second one. In this community, the most frequent hashtags are #trump, #depopulationagenda, #eugenetics, #repubicans, #auspol, #qanon, and #americafirst. The hashtags, as mentioned above, are related to Trump or against some figures who have publicly supported vaccines. Examples are #trump, #americafirst, or #faucifraud. These hashtags can also be found in the first community, where the percentage of both account types is similar. However, in this community, the number of bots is not higher than that of nonbots. In the third community, the number of nonbots is higher than bots. Most hashtags in this community mention COVID-19 (#coronaviruschallenge, #coronavirusbill, #coronaviruschina, and #coronavirusnewyork), but other hashtags such as #hoaxvirus, #tedconnnect, #freedomovefear, or #trumpisevil also appear.

5G

Regarding 5G, hashtags related to technology or news were the predominant ones in the case of nonbots, such as #techwar-#tradewar or #bbcaq-#itvnews. On the other hand, in the case of bots, the hashtags continue to mention geographical areas: #america-#china and #america-#lka. There are other hashtags with higher intensity, for example, #chinesecoronavirus-#democratshateamerica or #conspiracytheories-#technology. As can be seen in Table 2, the #china hashtag gets the highest PageRank value, followed by #pandemic and #wuhan. In addition, #china has 42 degrees of connectivity, doubling the value of the second, which is #pandemic with 27 connections. But above all, these values indicate the central place these hashtags take in the conversation. On the one hand, the high degree indicates they co-occur with many different hashtags. On the other hand, a high betweenness value indicates a central place in the network.

This time, the algorithm found 5 different communities of hashtags (Multimedia Appendix 1). The presence of bots is higher than nonbots in the first 3. The first is related to #tech, #bigdata, #cibersecurity, and so on. The second one is focused on #conspiracytheories, #digitalskynet, and #misinformation. The third is focused on China, with hashtags such as #batflu, #chinesevirus, and #huaweithis. The last 2 communities, where the level of nonbots is higher, are formed by varied hashtags. The fourth community is formed by hashtags such as #kag or #maga. The fifth one contains hashtags mentioning rumors or disinformation: #fakenews, #disinformation, and #democrathoax. In this community, it is worth mentioning the appearance of hashtags related to #blacklivesmatter, such as #racism, #blacklivesmatteraustralia, or #policebrutality.

UV

In this case, the appearance of technology-related hashtags (#ai and #healthtech) is even more noticeable, especially in the case of bots (Table 2). On the other hand, the most common hashtags are #batflu-#quarantine in the case of nonbots. Concerning the 6 communities we found (Multimedia Appendix 1), in the first 3, the number of nonbots is higher. The subject matter of these communities is related to politicians (#trump, #joebiden, and #berniesanders), technology (#artificialintelligence, #bioinformatics, and #machinelearning), or more specifically to technological innovation (#health, #innovation, #coronavirusnewyorkty, and #smartcities).

Discussion

Principal Findings

This study examined the use of hashtags by social bots on Twitter during the early stages of the COVID-19 pandemic. By analyzing the co-occurrence of hashtags, we were able to identify differences between accounts classified as bots and nonbots. We used Louvain communities to further classify these co-occurrences and found consistent differences in hashtag usage between the 2 groups. We used social network analysis

XSL•FO
RenderX

based on the co-occurrence of hashtags to take advantage of hashtags as key elements of online texts and understand how different users tag posts.

The analysis of hashtags provided several key insights into attitudes toward the COVID-19 pandemic and related behaviors. We consistently observed differences between bots and nonbots. In the case of bots, it was more common to find co-occurrences of hashtags related to political movements, particularly those on the right wing and related to Trump. This is consistent with findings in the literature showing a higher presence of conservatives in topics related to misinformation about COVID-19 pandemic [45].

In the conversation about vaccines, we observed that bots used hashtags related to fake news, such as #billgates and #china, more frequently. This analysis also identified specific uninformative hashtags (#ccpvirus and #chinesevirus) associated with anti-Asian sentiment [18]. Other hashtags expressed different opinions, such as criticism (#billgateisevil) or hate (#chinaliedpeopledied). It is worth noting that most of the tweets posted by nonbot users came from official accounts of institutions such as the World Health Organization, ministries of health, or entities related to public health. These messages focused on reporting on the evolution of the pandemic; the number of deaths; infection rates; and the health measures implemented, such as lockdowns and vaccination campaigns to contain the spread of the virus.

In our analysis of the conversation related to hydroxychloroquine, we identified 2 distinct communities of hashtags. One group was related to public health or medicine, while the other group was related to political movements and associated with Trump. Other studies have also found that Trump was involved in this conversation [46,47]. However, we also found that some of the hashtags in the conversation about hydroxychloroquine related to scientific facts. These differences suggest a highly polarized conversation with scientific arguments pitted against controversial political campaigns.

According to one of these studies [47], accounts with a higher impact on topics related to hydroxychloroquine disinformation were more likely to support President Trump. In addition, these types of content had a higher volume of tweets, longer duration in time, and greater echo. Our findings on the number of bots in these communities with politicized hashtags would partly explain the permanence over time and high echo values. Bots amplify these debates and increase the impact of the messages they disseminate [29,48,49]. However, our results also identify communities with anti–President Trump hashtags and higher numbers of bots. Liberals also engage in these conversations, although to a lesser extent than Conservatives [45].

These findings are extensible to topics such as the military or Bill Gates, where the conversation has been highly politicized and permeated with fake news. According to the results obtained, Trump occupied a leading role in the Twitter conversations during the period analyzed. This fact has also been noted in other previous works. Trump publicly supported the use of hydroxychloroquine and other drugs to combat the advance of the COVID-19 pandemic, with its corresponding impact on increased searches [50]. In addition, Bill Gates is often the protagonist in conspiracy theories [51].

Limitations and Strengths

There are several factors to consider when categorizing accounts as nonbot or bot. Botometer is backed by a large volume of research, but its effectiveness has been debated. It is important to remember that Botometer only provides a probability that an account is a bot, not a definitive classification. To get the most accurate results, it is recommended to compare probability distribution. However, in some cases it may be necessary to establish a binary classification for research purposes. In such cases, previous research has shown that using a cutoff value and comparing the results is a successful strategy [52].

It is important to consider the language constraint of this study. Only selecting tweets written in English may limit the focus to actors and events from English-speaking countries. In addition, no geographic limitations were placed on the collection of tweets, which allows for a larger volume of data but may also make it difficult to interpret results. It is also worth noting that the tweets analyzed in this study were from the early stages of the pandemic, and conversations and topics may have evolved over time.

Conclusion

Our analysis of hashtag usage on Twitter showed that there were differences in the patterns of use between bot and nonbot accounts. By grouping hashtags based on co-occurrence, we were able to identify distinct patterns in the usage of hashtags. On controversial or highly polarized issues, the hashtags used often pertained to the campaign or movement being promoted, with a significant portion related to Trump. In some cases, hashtags opposing these movements were also identified. On less polarized topics, hashtag usage was more diverse and included references to specific geographic locations or social groups. This analysis method can be useful in detecting hashtags that may be linked to fake news or misinformation, or in tracing the spread of such content on social media platforms.

Acknowledgments

We would like to acknowledge the support of the University Research Institute for Sustainable Social Development and the University of Cádiz. The publication is part of project NETDYNAMIC (CNS2022-135907), funded by MCIN/AEI/10.13039/501100011033 and by the European Union “Next Generation EU”/PRTR. The present study has also been supported by the project DCODES (PID2020-118589RB-I00), granted by the Spanish Ministry of Science and Innovation and financed by MCIN/AEI/10.13039/501100011033.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Bot distribution by topic.

[PNG File , 104 KB - infodemiology_v5i1e50021_app1.png]

References

1. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
2. Vos SC, Buckner MM. Social media messages in an emerging health crisis: tweeting bird Flu. J Health Commun 2016;21(3):301-308. [doi: 10.1080/10810730.2015.1064495] [Medline: 26192209]
3. Page BI, Shapiro RY. Effects of public opinion on policy. Am Polit Sci Rev 2014;77(1):175-190. [doi: 10.2307/1956018]
4. Han X, Wang J, Zhang M, Wang X. Using social media to mine and analyze public opinion related to COVID-19 in China. Int J Environ Res Public Health 2020;17(8):2788 [FREE Full text] [doi: 10.3390/ijerph17082788] [Medline: 32316647]
5. EPI-WIN: WHO information network for epidemics. World Health Organization. URL: https://www.who.int/teams/epi-win [accessed 2022-09-23]
6. Sylvia Chou WY, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. Am J Public Health 2020;110(S3):S273-S275. [doi: 10.2105/AJPH.2020.305905] [Medline: 33001722]
7. Kyprianidou M, Konstantinou P, Alvarez-Galvez J, Ceccarelli L, Gruszczyńska E, Mierzejewska-Floreani D, et al. Profiling hesitancy to COVID-19 vaccinations in six European countries: behavioral, attitudinal and demographic determinants. Behav Med 2023;49(4):362-373. [doi: 10.1080/08964289.2022.2071198] [Medline: 35546258]
8. Alvarez-Galvez J, Anastasiou A, Lamnisos D, Constantinou M, Nicolaou C, Papacostas S, et al. The impact of government actions and risk perception on the promotion of self-protective behaviors during the COVID-19 pandemic. PLoS One 2023;18(4 April). [doi: 10.21203/rs.3.rs-1270359/v1]
9. Gloster AT, Lamnisos D, Lubenko J, Presti G, Squatrito V, Constantinou M, et al. Impact of COVID-19 pandemic on mental health: an international study. PLoS One 2020;15(12):e0244809 [FREE Full text] [doi: 10.1371/journal.pone.0244809] [Medline: 33382859]
10. Balakrishnan V, Ng WZ, Soo MC, Han GJ, Lee CJ. Infodemic and fake news - A comprehensive overview of its global magnitude during the COVID-19 pandemic in 2021: a scoping review. Int J Disaster Risk Reduct 2022;78:103144 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.103144] [Medline: 35791376]
11. Yang KC, Pierri F, Hui PM, Axelrod D, Torres-Lugo C, Bryden J, et al. The COVID-19 infodemic: Twitter versus Facebook. Big Data & Society 2021;8(1):205395172110138. [doi: 10.1177/20539517211013861]
12. Singh L, Bode L, Budak C, Kawintiranon K, Padden C, Vraga E. Understanding high- and low-quality URL sharing on COVID-19 Twitter streams. J Comput Soc Sci 2020;3(2):343-366. [doi: 10.1007/s42001-020-00093-6] [Medline: 33263092]
13. Rovetta A, Bhagavathula AS. Global infodemiology of COVID-19: analysis of Google web searches and Instagram hashtags. J Med Internet Res 2020;22(8):e20673 [FREE Full text] [doi: 10.2196/20673] [Medline: 32748790]
14. Blane JT, Bellutta D, Carley KM. Social-cyber maneuvers during the COVID-19 vaccine initial rollout: content analysis of tweets. J Med Internet Res 2022;24(3):e34040 [FREE Full text] [doi: 10.2196/34040] [Medline: 35044302]
15. Petersen K, Gerken JM. #COVID-19: an exploratory investigation of hashtag usage on Twitter. Health Policy 2021;125(4):541-547. [doi: 10.1016/j.healthpol.2021.01.001] [Medline: 33487479]
16. Xi W, Xu W, Zhang X, Ayalon L. A thematic analysis of Weibo topics (Chinese Twitter hashtags) regarding older adults during the COVID-19 outbreak. J Gerontol B Psychol Sci Soc Sci 2021;76(7):e306-e312. [doi: 10.1093/geronb/gbaa148] [Medline: 32882029]
17. Khadafi R, Nurmandi A, Qodir Z, Misran. Hashtag as a new weapon to resist the COVID-19 vaccination policy: a qualitative study of the anti-vaccine movement in Brazil, USA, and Indonesia. Hum Vaccin Immunother 2022;18(1):2042135 [FREE Full text] [doi: 10.1080/21645515.2022.2042135] [Medline: 35240923]
18. Hswen Y, Xu X, Hing A, Hawkins JB, Brownstein JS, Gee GC. Association of "#covid19" versus "#chinesevirus" with anti-Asian sentiments on Twitter: March 9-23, 2020. Am J Public Health 2021;111(5):956-964. [doi: 10.2105/AJPH.2021.306154] [Medline: 33734838]
19. Gruzd A, Mai P. Going viral: how a single tweet spawned a COVID-19 conspiracy theory on Twitter. Big Data & Society 2020;7(2):205395172093840. [doi: 10.1177/2053951720938405]
20. Tahamtan I, Potnis D, Mohammadi E, Miller LE, Singh V. Framing of and attention to COVID-19 on Twitter: thematic analysis of hashtags. J Med Internet Res 2021;23(9):e30800 [FREE Full text] [doi: 10.2196/30800] [Medline: 34406961]
21. Vicient C, Moreno A. Unsupervised topic discovery in micro-blogging networks. Expert Systems with Applications 2015;42(17-18):6472-6485. [doi: 10.1016/j.eswa.2015.04.014]
22. Alkhodair SA, Fung BCM, Rahman O, Hung PCK. Improving interpretations of topic modeling in microblogs. Asso for Info Science & Tech 2017;69(4):528-540. [doi: 10.1002/asi.23980]

XSL•FO
RenderX

23. Eriksson Krutrök M, Lindgren S. Continued contexts of terror: analyzing temporal patterns of hashtag co-occurrence as discursive articulations. Social Media + Society 2018;4(4):205630511881364. [doi: 10.1177/2056305118813649]
24. Caldarelli G, de Nicola R, del Vigna F, Petrocchi M, Saracco F. The role of bot squads in the political propaganda on Twitter. Commun Phys 2020;3(1):1-15. [doi: 10.1038/s42005-020-0340-4]
25. Ruiz-Núñez C, Segado-Fernández S, Jiménez-Gómez B, Hidalgo PJJ, Magdalena CSR, Pollo MDCA, et al. Bots' activity on COVID-19 pro and anti-vaccination networks: analysis of Spanish-written messages on Twitter. Vaccines (Basel) 2022;10(8):1240 [FREE Full text] [doi: 10.3390/vaccines10081240] [Medline: 36016126]
26. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis 2021;108:256-262 [FREE Full text] [doi: 10.1016/j.ijid.2021.05.059] [Medline: 34052407]
27. Scannell D, Desens L, Guadagno M, Tra Y, Acker E, Sheridan K, et al. COVID-19 vaccine discourse on Twitter: a content analysis of persuasion techniques, sentiment and mis/disinformation. J Health Commun 2021;26(7):443-459. [doi: 10.1080/10810730.2021.1955050] [Medline: 34346288]
28. Ahmed W, López Seguí F, Vidal-Alaball J, Katz MS. COVID-19 and the "Film Your Hospital" conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020;22(10):e22374 [FREE Full text] [doi: 10.2196/22374] [Medline: 32936771]
29. Uyheng J, Carley KM. Bots and online hate during the COVID-19 pandemic: case studies in the United States and the Philippines. J Comput Soc Sci 2020;3(2):445-468. [doi: 10.1007/s42001-020-00087-4] [Medline: 33102925]
30. Uyheng J, Bellutta D, Carley KM. Bots amplify and redirect hate speech in online discourse about racism during the COVID-19 pandemic. Social Media + Society 2022;8(3):205630512211047. [doi: 10.1177/20563051221104749]
31. Green M, Musi E, Rowe F, Charles D, Pollock FD, Kypridemos C, et al. Identifying how COVID-19-related misinformation reacts to the announcement of the UK national lockdown: an interrupted time-series study. Big Data & Society 2021;8(1):205395172110138. [doi: 10.1177/20539517211013869]
32. Chang HCH, Ferrara E. Comparative analysis of social bots and humans during the COVID-19 pandemic. J Comput Soc Sci 2022;5(2):1409-1425 [FREE Full text] [doi: 10.1007/s42001-022-00173-9] [Medline: 35789937]
33. Zhang M, Chen Z, Qi X, Liu J. Could social bots' sentiment engagement shape humans' sentiment on COVID-19 vaccine discussion on Twitter? Sustainability 2022;14(9):5566. [doi: 10.3390/su14095566]
34. Suarez-Lledo V, Alvarez-Galvez J. Assessing the role of social bots during the COVID-19 pandemic: infodemic, disagreement, and criticism. J Med Internet Res 2022;24(8):e36085 [FREE Full text] [doi: 10.2196/36085] [Medline: 35839385]
35. Graham T, Bruns A, Angus D, Hurcombe E, Hames S. #IStandWithDan versus #DictatorDan: the polarised dynamics of Twitter discussions about Victoria's COVID-19 restrictions. Media International Australia 2020;179(1):1329878X2098178. [doi: 10.1177/1329878X20981780]
36. Ruffer N, Knitza J, Krusche M. #Covid4Rheum: an analytical Twitter study in the time of the COVID-19 pandemic. Rheumatol Int 2020;40(12):2031-2037. [doi: 10.1007/s00296-020-04710-5] [Medline: 32995894]
37. Liu AKC, Ophir Y, Tsai SA, Walter D, Himelboim I. Hashtag activism in a politicized pandemic: framing the campaign to include Taiwan in the World Health Organization's efforts to combat COVID-19. New Media & Society 2022;26(6):146144482210991. [doi: 10.1177/14614448221099173]
38. Sharevski F, Alsaadi R, Jachim P, Pieroni E. Misinformation warnings: Twitter's soft moderation effects on COVID-19 vaccine belief echoes. Comput Secur 2022;114:102577 [FREE Full text] [doi: 10.1016/j.cose.2021.102577] [Medline: 34934255]
39. Calac AJ, Haupt MR, Li Z, Mackey T. Spread of COVID-19 vaccine misinformation in the ninth inning: retrospective observational infodemic study. JMIR Infodemiology 2022;2(1):e33587 [FREE Full text] [doi: 10.2196/33587] [Medline: 35320982]
40. Herrera-Peco I, Jiménez-Gómez B, Romero Magdalena CS, Deudero JJ, García-Puente M, Benítez de Gracia E, et al. Antivaccine movement and COVID-19 negationism: a content analysis of Spanish-written messages on Twitter. Vaccines (Basel) 2021;9(6):656 [FREE Full text] [doi: 10.3390/vaccines9060656] [Medline: 34203946]
41. Sayyadiharikandeh M, Varol O, Yang KC, Flammini A, Menczer F. Detection of novel social bots by ensembles of specialized classifiers. 2020 Presented at: CIKM '20: The 29th ACM International Conference on Information and Knowledge Management; October 19-23, 2020; Virtual event, Ireland. [doi: 10.1145/3340531.3412698]
42. Varol O, Ferrara E, Davis CA, Menczer F, Flammini A. Online human-bot interactions: detection, estimation, and characterization. 2017 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; May 15-18, 2017; Montreal, QC p. 280-289 URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14871 [doi: 10.1609/icwsm.v11i1.14871]
43. Page L, Brin S, Motwani R, Winograd T. The pageRank citation ranking: bringing order to the web. WWW Internet And Web Information Systems 1998;54(1999-66).
44. Freeman LC. A set of measures of centrality based on betweenness. Sociometry 1977;40(1):35-41. [doi: 10.2307/3033543]
45. Havey NF. Partisan public health: how does political ideology influence support for COVID-19 related misinformation? J Comput Soc Sci 2020;3(2):319-342. [doi: 10.1007/s42001-020-00089-2] [Medline: 33163686]

XSL•FO
RenderX

46. Blevins JL, Edgerton E, Jason DP, Lee JJ. Shouting into the wind: medical science versus “B.S.” in the Twitter maelstrom of politics and misinformation about hydroxychloroquine. Social Media + Society 2021;7(2):205630512110249. [doi: 10.1177/20563051211024977]
47. Haupt MR, Li J, Mackey TK. Identifying and characterizing scientific authority-related misinformation discourse about hydroxychloroquine on Twitter using unsupervised machine learning. Big Data & Society 2021;8(1). [doi: 10.1177/20539517211013843]
48. Xu W, Sasahara K. Characterizing the roles of bots on Twitter during the COVID-19 infodemic. J Comput Soc Sci 2022;5(1):591-609 [FREE Full text] [doi: 10.1007/s42001-021-00139-3] [Medline: 34485752]
49. Stella M, Ferrara E, de Domenico M. Bots increase exposure to negative and inflammatory content in online social systems. Proc Natl Acad Sci U S A 2018;115(49):12435-12440 [FREE Full text] [doi: 10.1073/pnas.1803470115] [Medline: 30459270]
50. Niburski K, Niburski O. Impact of trump's promotion of unproven COVID-19 treatments and subsequent internet trends: observational study. J Med Internet Res 2020;22(11):e20044 [FREE Full text] [doi: 10.2196/20044] [Medline: 33151895]
51. Shahsavari S, Holur P, Wang T, Tangherlini TR, Roychowdhury V. Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news. J Comput Soc Sci 2020;3(2):279-317. [doi: 10.1007/s42001-020-00086-5] [Medline: 33134595]
52. Shao C, Ciampaglia GL, Varol O, Yang KC, Flammini A, Menczer F. The spread of low-credibility content by social bots. Nat Commun 2018;9(1):4787 [FREE Full text] [doi: 10.1038/s41467-018-06930-7] [Medline: 30459415]

Abbreviations

API: application programming interface
NLP: natural language processing
SNA: social network analysis

Edited by T Purnat; submitted 16.06.23; peer-reviewed by I Herrera-Peco, A Rovetta; comments to author 21.10.23; revised version received 08.02.24; accepted 15.05.24; published 09.01.25.

Please cite as:
Suarez-Lledo V, Ortega-Martin E, Carretero-Bravo J, Ramos-Fiol B, Alvarez-Galvez J
Unraveling the Use of Disinformation Hashtags by Social Bots During the COVID-19 Pandemic: Social Networks Analysis
JMIR Infodemiology 2025;5:e50021
URL: https://infodemiology.jmir.org/2025/1/e50021
doi:10.2196/50021
PMID:

©Victor Suarez-Lledo, Esther Ortega-Martin, Jesus Carretero-Bravo, Begoña Ramos-Fiol, Javier Alvarez-Galvez. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 09.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Misinformation About Unproven Cancer Treatments on Social Media Using User-Friendly Linguistic Characteristics: Content Analysis

Ilona Fridman[1], PhD; Dahlia Boyles[2], MA; Ria Chheda[3], BS; Carrie Baldwin-SoRelle[4], MLiS; Angela B Smith[1], MD; Jennifer Elston Lafata[1,5], PhD

[1]Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC, United States
[2]Department of Communication, University of North Carolina, Chapel Hill, NC, United States
[3]Computer Science Department, University of North Carolina, Chapel Hill, NC, United States
[4]Health Sciences Library, University of North Carolina, Chapel Hill, NC, United States
[5]Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, United States

Corresponding Author:
Ilona Fridman, PhD
Lineberger Comprehensive Cancer Center
University of North Carolina
450 West Dr
Chapel Hill, NC, 27599
United States
Phone: 1 6469028137
Email: ilona_fridman@med.unc.edu

Abstract

Background: Health misinformation, prevalent in social media, poses a significant threat to individuals, particularly those dealing with serious illnesses such as cancer. The current recommendations for users on how to avoid cancer misinformation are challenging because they require users to have research skills.

Objective: This study addresses this problem by identifying user-friendly characteristics of misinformation that could be easily observed by users to help them flag misinformation on social media.

Methods: Using a structured review of the literature on algorithmic misinformation detection across political, social, and computer science, we assembled linguistic characteristics associated with misinformation. We then collected datasets by mining X (previously known as Twitter) posts using keywords related to unproven cancer therapies and cancer center usernames. This search, coupled with manual labeling, allowed us to create a dataset with misinformation and 2 control datasets. We used natural language processing to model linguistic characteristics within these datasets. Two experiments with 2 control datasets used predictive modeling and Lasso regression to evaluate the effectiveness of linguistic characteristics in identifying misinformation.

Results: User-friendly linguistic characteristics were extracted from 88 papers. The short-listed characteristics did not yield optimal results in the first experiment but predicted misinformation with an accuracy of 73% in the second experiment, in which posts with misinformation were compared with posts from health care systems. The linguistic characteristics that consistently negatively predicted misinformation included tentative language, location, URLs, and hashtags, while numbers, absolute language, and certainty expressions consistently predicted misinformation positively.

Conclusions: This analysis resulted in user-friendly recommendations, such as exercising caution when encountering social media posts featuring unwavering assurances or specific numbers lacking references. Future studies should test the efficacy of the recommendations among information users.

(JMIR Infodemiology 2025;5:e62703) doi:10.2196/62703

KEYWORDS

linguistic characteristics; linguistic features; cancer; Linguistic Inquiry and Word Count; misinformation; X; Twitter; cancer; alternative therapy; oncology; social media; natural language processing; machine learning; synthesis; review methodology; search; literature review

XSL•FO
RenderX

Introduction

Approximately 16% of people reported using social media to inform their medical decisions [1]. This percentage, based on estimates from the National Cancer Center, equates to 37 million adults in the United States. A recent systematic review estimated that up to 40% of health-related social media posts contain misinformation [2]. Misinformation could cause more harm to individuals with serious conditions such as cancer. Patients who believe in misinformation and use unproven therapies in parallel or in place of cancer treatment tend to be less adherent to evidence-based treatment [3-5]. Moreover, patients with cancer might choose to delay or reject evidence-based treatment and instead pursue unproven and potentially toxic therapies, which, for some patients, results in up to 2.5 times shorter life expectancy [6]. Approximately 30% of cancer-related social media posts on Facebook, Reddit, Pinterest, and X (previously known as Twitter) contain misinformation, and a staggering 77% of these posts have the potential to encourage patients to pursue futile and toxic therapies, resulting in physical, psychological, and logistical burdens [7]. Cancer misinformation persists across various cancer types and is more pervasive in more prevalent cancers. Across various social media platforms, two-thirds of the most shared posts about prostate cancer contain misinformation [8]. Researchers identified misinformation in 59% of posts related to breast cancer prevention and treatment [9] and 30% of posts related to gynecological cancer [10]. When surveyed, 70% of patients with cancer reported encountering misinformation about cancer on social media, with 71% believing that some of this misinformation was accurate [11].

There is a growing need to protect health information users from misinformation, especially those who are affected by serious conditions such as cancer. Multiple recommendations have been developed to assist individuals in their search for reliable health information [12-14]. However, many of the recommendations are complex, as they require individuals to possess a certain level of scientific knowledge and skills. For instance, recommendations frequently suggest taking steps such as identifying authors and their credentials, evaluating potential conflicts of interest, understanding funding sources, and assessing the original sources of scientific information. Considering the time and expertise required, expecting individuals to perform these tasks routinely is unrealistic. Moreover, these guidelines often fall short when it comes to addressing the challenges posed by social media platforms. Those who post may not disclose their real names or sources of findings, which makes some recommended steps not possible.

In this work, our goal is to identify user-friendly recommendations for addressing the high rate of misinformation on social media. We began by exploring literature on the algorithmic detection of misinformation. The algorithmic approach often involves the analysis of linguistic characteristics that differentiate between factual information and misinformation [15]. Linguistic characteristics describe a body of text in an abstract manner regardless of context and may include counts of words and word parts such as nouns, verbs, adjectives, and negations, as well as specific symbols such as URLs, hashtags, and question marks. An additional category of linguistic characteristics includes words associated with the psychological state of an author [16], which includes words related to emotions, expressions of certainty, tentativeness, insight, persuasion, and gratitude. To date, linguistic characteristics have been used by algorithms only. However, some of these characteristics are observable and could be used by individuals when they need to evaluate the credibility of the text. While individuals are unlikely to count words in social media posts regularly, they may routinely note other linguistic characteristics, such as emotions, URLs, and a strong degree of certainty. Linguistic characteristics have been shown to be effective in distinguishing misinformation from factual information across multiple contexts. However, it is unknown (1) whether the linguistic characteristics are effective in cancer-related context and (2) which subset of user-friendly linguistic characteristics could effectively distinguish misinformation. In this work, we identify the linguistic characteristics specific to the context of cancer. These characteristics will be recommended as guidelines for health information users when browsing social media.

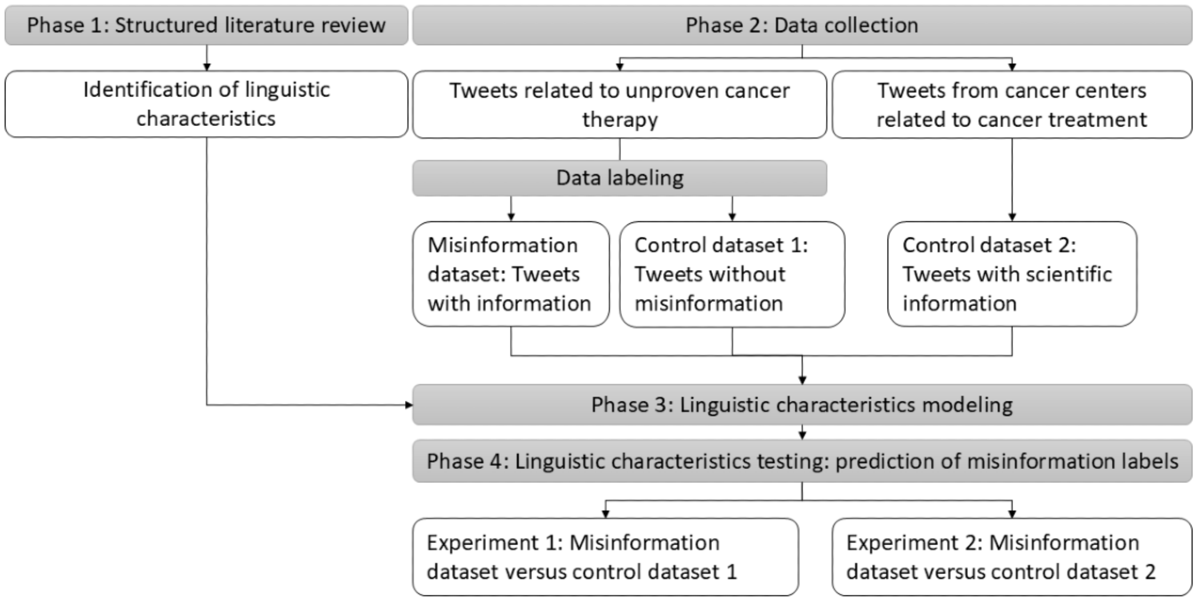
Methods

Study Design

The main sequence of study procedures is illustrated in Figure 1 and includes (1) a structured literature review, in which we assemble linguistic characteristics that were used in algorithms for distinguishing factual information and misinformation (phase 1); (2) data collection, which encompasses assembling cancer-related posts using the X application programming interface (API) and labeling them as misinformation and non-misinformation (phase 2); (3) identification of the linguistic characteristics in collected datasets using natural language processing tools (phase 3); and (4) conducting predictive modeling analysis to evaluate the effectiveness of linguistic characteristics in distinguishing social media posts with cancer misinformation (phase 4).

XSL•FO
RenderX

Figure 1. Summary of the study procedures.



Ethical Considerations

The study was institutional review board–approved by the University of North Carolina (IRB#21-2861). This was an analysis of publicly available data. As such, participants were not compensated and did not need to provide consent for the study, because the study did not involve any prospective data collection. To protect the confidentiality and anonymity of participants in this secondary data analysis, we reworded reported posts from X.

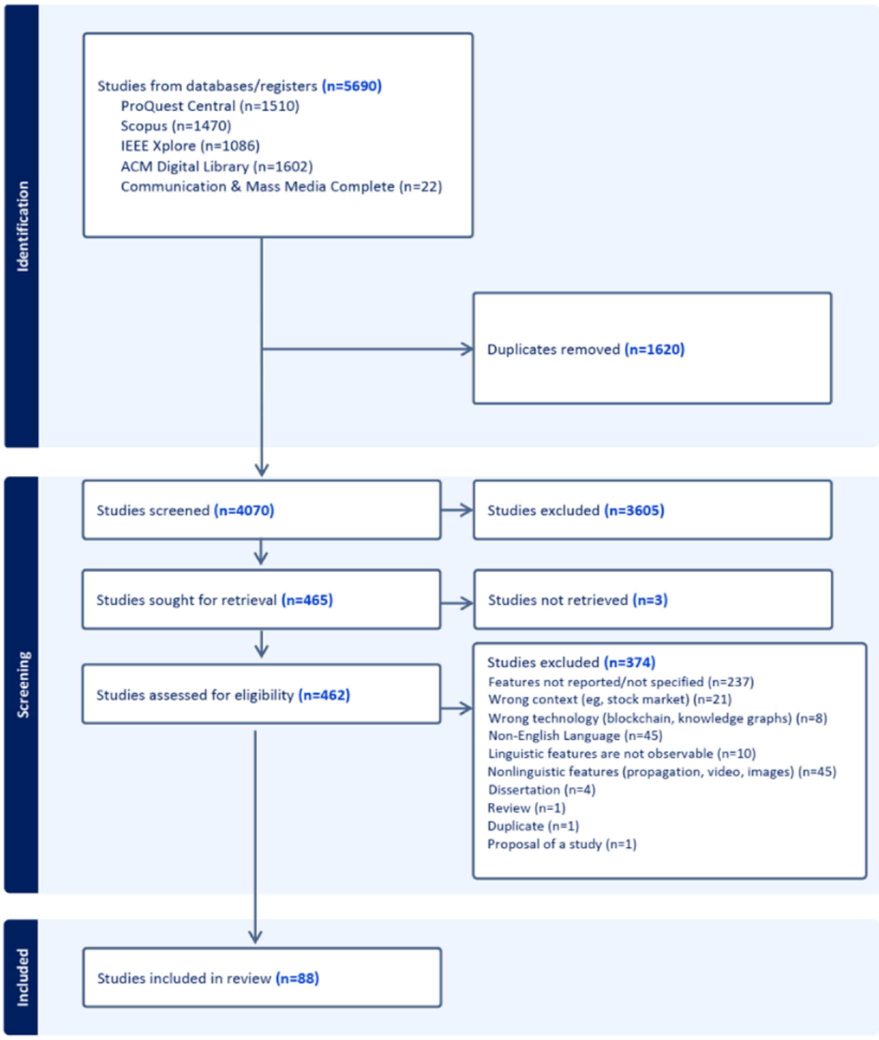
Structured Literature Review

To identify linguistic characteristics, we developed a literature review protocol that included the search strategy and keywords. This process was informed by a collaboration with a health sciences librarian (CBS), who suggested an initial set of keywords referenced in several relevant reviews [17-21]. She also created an expanded title, abstract, and keyword search strategies for each of the following concepts: (1) text as a unit of analysis, (2) misinformation, (3) algorithms, (4) internet, and (5) linguistic features or characteristics. After the search was peer reviewed by a second health sciences librarian (CB), 5 databases were searched: ProQuest Central (ProQuest), which includes the arXiv repository; Scopus (Elsevier); IEEE Xplore (Institute of Electrical and Electronics Engineers); ACM Digital Library (Association for Computing Machinery); and Communication & Mass Media Complete (EBSCOhost). The keywords and search strategies are reported in Multimedia Appendix 1. Results were limited to citations published between January 2012 and December 2022. Within databases, results were limited to journal papers, conference proceedings, working papers, and book chapters.

Two reviewers (IF and DB) independently coded titles and abstracts in Covidence software (Veritas Health Innovation) [22] and resolved conflict in codes during research meetings. Papers were included if they focused on detecting misinformation and contained a "Methods" section describing an approach for algorithmically detecting misinformation (eg, reviews and viewpoints were excluded). Examples of the algorithms included supervised and semisupervised machine learning (eg, Bidirectional Encoder Representations from Transformers [BERT] classification) that was built on linguistic characteristics. Papers were excluded if they did not report specific linguistic characteristics, focused on misinformation in any language other than English, or used human coding but not algorithms. The detailed inclusion-exclusion criteria and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram are reported in Figure 2.

XSL•FO
RenderX

Figure 2. Flowchart of paper identification and extraction.



Identification of Linguistic Characteristics

Upon identifying eligible papers, 2 team members (IF and DB) reviewed the full text and extracted the linguistic characteristics. Around 11% (10/90) of papers underwent double-coding. After reviewers reached an agreement, we continued with single coding. The linguistic characteristics were extracted based on the following criteria: observability, applicability, and generalizability. The observability criterion was related to whether readers could easily observe the linguistic characteristics within the text; for example, positive emotions could be easily observed while morale or cognitive language styles may be difficult to distinguish. The applicability criterion distinguished linguistic characteristics that readers could easily apply while reading the text. For instance, common characteristics such as the number of words required substantial effort from readers to evaluate and, therefore, were deemed nonapplicable. In contrast, readers could easily use citations and hashtags in their post evaluations as the mere presence of these characteristics was determined to be helpful in identifying misinformation. The third criterion, generalizability, was chosen to ensure that linguistic characteristics were not related to a specific context but could be applied across various contexts. Thus, characteristics that were based on specific words such as "COVID-19," or "cure" were excluded.

Data Collection: Unproven Therapy

Overview

To test how extracted linguistic characteristics could distinguish social media posts from misinformation and factual information, we collected social media posts from X. Misinformation was operationalized here *as information that promoted cancer treatment that was known as ineffective or information that suggested cancer causes not supported by current scientific evidence* [23]. This definition focused our investigation on misinformation that could be harmful to patients with cancer or cancer survivors. Based on this operationalization, we searched existing resources that summarized unproven cancer therapy, such as "List of unproven cancer therapy" [24], a list of "Illegally sold cancer drugs" [25], and previous literature [23,26]. We extracted keywords and constructed 176 queries associated with unproven cancer treatments (Multimedia Appendix 2). Using these queries, we randomly selected up to 500 posts per query from social media. We used R software (R Foundation for Statistical Computing) to access the Academic X API. The data were manually evaluated to determine their relevance to the cancer context and unproven therapies. Queries

XSL•FO
RenderX

were edited to ensure relevance. Upon corrections, the data collection was implemented on a schedule every other week between July 2022 and August 2023. After data collection was completed, the duplicate posts were removed.

Data Labeling

To distinguish posts with misinformation from other discussions, 2 reviewers (IF and CR) double-coded a randomly chosen subset of 1064 posts, achieving an acceptable interrater agreement of 0.68 measured with Krippendorff a [27]. Since the agreement was rather on a lower bound, we followed the current recommendations [28] and resolved disagreements between coders during research meetings, reaching consensus case by case. The initial criterion for coding misinformation was developed deductively based on the definition of misinformation used in this study. A post was coded as containing misinformation if it promoted an unproven therapy as a cancer-directed treatment. For example, a post claiming that an alkaline diet can eliminate cancer would be classified as misinformation: "Cure for cancer is an alkaline diet and lots of alkaline water." As reviewers worked with the data, they developed additional criteria based on observed patterns. Specifically, posts were labeled as containing misinformation if they discussed unproven approaches to prevent cancer, for example, "Pygeum Bark is nature's defense against prostate cancer." Furthermore, if a post contained a combination of factual and false information it was labeled as "misinformation."

Posts that were labeled as non-misinformation fell into 4 distinct categories. First, posts mentioned complementary and alternative medicine but did not promote it as a cancer treatment, for example, "Acupuncture and acupressure seem to be helpful in reducing pain and anxiety in patients having surgery." Second, posts that used sarcasm and actively debunked misinformation related to cancer were in the non-misinformation category, for example, "If what you stated is true, then Gerson treatment for cancer is false." The third category included posts that discussed complementary and alternative therapies but not in the context of promotion of cancer treatment, for instance, "Grapes can help protect you from the sun! Who knew?" Finally, posts that presented information with ambiguity, lack of clarity, or insufficient context were categorized as non-misinformation, for instance, "As a pancreatic cancer patient providing myself with all the additional holistic care practices made all the difference." The author did not specify whether his symptoms were alleviated or cancer progression was slowed down because of holistic practices. Therefore, the post was coded as non-misinformation.

Once a subset of the database was labeled by 2 reviewers (IF and RC), we applied an algorithm to populate labels to the entire database. We worked with BERT [29], a machine learning model for natural language processing. The BERT model was chosen because it (1) worked well with short, informal text [30]; (2) was shown to be applicable to medical text extracted from X [31]; and (3) was successfully used in previous research to identify misinformation on X [32]. The BERT model was implemented with the programming language Python (Python Software Foundation). The manually prelabeled subset served as training data for the BERT model. Such semisupervised approaches are commonly used in similar classification tasks [33]. After training, BERT used its understanding of the language and context learned from the large corpus it was originally trained on and the specific examples from the manually prelabeled dataset. BERT predicted labels for each post in the rest of the data (unlabeled dataset), determining whether each was likely to contain misinformation or not based on the patterns and features it learned from the manually coded dataset.

After BERT algorithm assigned labels to the posts, a researcher (IF), blinded to the model's results, manually coded a random subset of the posts (n=960) using the same "misinformation" and "non-misinformation" labels, adhering to the same criteria that were used to prelabel the data. When compared with manual coding, the algorithm identified misinformation with an accuracy of 83%, with a higher 86% specificity, and a slightly lower sensitivity of 82%. Upon labeling, 2 datasets were created and used in the first experiment: the misinformation dataset included only posts with misinformation, and control BERT dataset 1 included only posts with non-misinformation (Figure 1).

Data Collection: Posts From Cancer Centers

Following the definition of misinformation as "information not supported by scientific evidence or expert consensus" [34] and the definition used for this research, we assumed that posts originating from cancer centers reflect scientific evidence and expert consensus. To collect posts with factual information, we retrieved X data posted by cancer centers. Cancer centers often shared internal announcements and organizational news on X. To make posts comparable between the dataset with misinformation and control datasets, we used the keywords "cancer," "treatment," "chemotherapy," "healing," and other words related to treating cancer or controlling cancer progress. With the help of R software, we sampled 300 posts per cancer center between June 2011 and November 2022. A researcher (IF) manually checked randomly chosen (n=100) posts. As expected, no misinformation was found in the posts originating from cancer centers. The dataset, therefore, was assumed to consist of non-misinformation posts from cancer centers and was designated as control dataset 2, which was used in the second experiment alongside the misinformation dataset.

Linguistic Characteristics Modeling

Upon data collection and labeling, we used algorithmic approaches to model linguistic characteristics. First, we used an automated text search using regular expressions in Python [35] to capture digital numbers, hashtags, and URLs in the text.

Second, we used the Linguistic Inquiry and Word Count (LIWC) software [36]. LIWC calculates the proportion of the words in the posts associated with distinct psychological dimensions [37]. In this study, LIWC identified when authors of posts used certain, absolute, or tentative language.

Third, we leveraged the natural language processing tool, Name Entity Recognition [38], which was trained on human-labeled datasets to extract names from unstructured text. Using Name Entity Recognition, we were able to identify which posts contained personal names, organizational names, or locations identified from text.

XSL•FO
RenderX

Fourth, we experimented with several models for sentiment analysis and identified the DistilBERT algorithm as an optimal approach for its accuracy in handling health-related X data [39]. DistilBERT is a black-box algorithm that is trained on a large corpus of data and is based on multiple deep stack layers. The DistilBERT algorithm identified positive, negative, and neutral tones present in the posts. To evaluate the algorithm's performance, we manually labeled 300 posts across the databases. On average, the DistilBERT algorithm achieved an 83% accuracy (82% for misinformation and 84% for the control database) in detecting the emotional tone within the posts.

Linguistic Characteristics Testing: Prediction of Misinformation Labels

Identified linguistic characteristics were used in an algorithm to test whether these could distinguish misinformation in posts. As shown in Figure 1, we conducted 2 experiments using *tidymodels* package in R software [40]. Using linguistic characteristics as predictors, we forecast the "misinformation" and "non-misinformation" labels in the datasets semimanually coded by researchers and BERT classification algorithm. Data were split 60:40 to enable evaluation of the predictive power of the model and reported the accuracy as a ratio of correctly classified posts to the total number of posts. We also reported area under the curve (AUC), which accounted for both false-positive and false-negative predictions. AUC value ranged from 0 to 1, where 0.5 indicated that the model performs no better than a random chance, and 1 was a perfect prediction. The model was built on the basis of Lasso ("Least Absolute Shrinkage and Selection Operator") regression, which allowed variable selection by shrinking the coefficients of less important predictors to zero [41]. Bootstrapping procedure was applied to optimize and stabilize the selection of variables [42]. Lasso was chosen to address multicollinearity and overfitting issues in the regression model. More importantly, Lasso regression helped identify a set of linguistic characteristics that effectively distinguished posts containing misinformation. To evaluate the significance of specific linguistic characteristics, we computed importance scores, with higher scores indicating greater relevance in distinguishing posts containing misinformation. Importance scores, a common measure in predictive modeling, indicates to what extent individual predictors contribute to the overall model performance. The assessment involves permutating the characteristic values through shuffling and measuring the subsequent decline in model performance, effectively revealing the critical factors influencing predictions. Finally, we conducted a permutation statistical test (with 1000 permutations) to determine whether models with linguistic characteristics significantly outperformed random chance.

Results

Structured Literature Review

A total of 5677 citations were initially identified across all databases. After removing 1598 duplicates, we screened 4070 unique citations in Covidence. Subsequently, 3605 were excluded during the title and abstract review phase, leaving 464 papers for full-text review. Ultimately, we extracted linguistic characteristics from 88 full-text papers. These papers featured algorithmic approaches for identifying misinformation through automated text analysis, spanning various contexts, including politics, social issues, and computer science. Exclusion reasons are detailed in Figure 2, and additional information about the included papers can be found in Multimedia Appendix 3.

Identified Linguistic Characteristics

The extracted linguistic characteristics and corresponding literature are detailed in Table 1. Representative examples that contain each linguistic characteristic were chosen by selecting posts from the misinformation dataset. We used results from linguistic characteristic modeling to identify such posts. The first category of characteristics pertains to the sentiment and emotional expression in the text and includes positive, negative emotions, and neutral sentiments (absence of either). Some papers delved into more nuanced emotions such as anger, fear, surprise, and others. We excluded these emotions due to the potential difficulty for readers to detect nuanced emotions reliably in the text.

The next category comprises linguistic characteristics that pertain to psychological concepts. It is worth noting that some psychological concepts consist of a combination of linguistic characteristics, such as social processes including references to family, friends, other people, and verbs indicating interactions. Although algorithms frequently use such combinations, we decided to exclude the following psychological concepts that consisted of combinations of linguistic characteristics such as cognitive, perceptual, social processes, and morality or deception. The rationale behind this exclusion is that users are unlikely able to observe and combine linguistic characteristics for evaluations of the posts. We also excluded characteristics mentioned in fewer than 4 studies, such as gratitude, insight, causation, and persuasion. Following our 3 criteria, we included negations, tentativeness, profanity (as a proxy of informality), and words associated with absolutes and certainty.

Other categories that met our inclusion criteria were linguistic characteristics such as names of individuals, locations, and organizations, as well as categories related to the presence of URLs, hashtags, personal pronouns, and numbers. Readers can identify these characteristics without additional efforts (observability criterion) and use them for evaluation of the text (applicability) because the presence of these characteristics in social media has historically been a distinguishable factor in detecting misinformation. Furthermore, these characteristics were not context-dependent and, therefore, satisfy the generalizability criterion.

XSL•FO
RenderX

Table 1. Linguistic characteristics and examples of misinformation.

Characteristics	Examples of linguistic characteristics and posts with misinformation[a]	Studies using characteristics for misinformation detection
Sentiment[b]		[43-93]
Negative emotions	• Chemo is costly and very painful. It seems to worsen illness and hasten life's end. • Sad this happened, to overcome cancer, consider utilizing cannabis oil in combination with vitamin B17. • Feeling frustrated that insurance doesn't cover certain treatments I believe in. Wish there were more options beyond the conventional cut, burn, and poison approach.	
Positive emotions	• Cure for cancer that works holistically, Vitamin B17, very good! • Please do some heavy doses of medical organic marijuana if possible let it eat that cancer. Wishing you healing and joy and comfort. • Wonderful treatment! Discover the incredible benefits of ProstateRelax, a natural herbal treatment for prostate cancer. ProstateRelax effectively treats and prevents the progression of prostate cancer.	
Neutral emotions	• Anyone with cancer. Check your body's pH level. Drink alkaline water, eat alkaline foods, and avoid acidic sugary treats and dairy. • Cancer cells thrive in low oxygen environments. B17, found in apricot seeds, can help. • Antineoplastons, a protein suppressed by cancer, could hold the key to a potential cure.	
Psycholinguistic		
Negation	• Unlock the potential of Acupuncture to modulate immunity and create an environment where cancer cannot thrive. Discover the holistic power of this ancient practice in bolstering your body's defenses against cancer. • I wonder why aren't we utilizing hyperbaric chambers for Cancer? Ask your doctor about the incredible potential of pure oxygen in rejuvenating and generating new cells to combat this disease. • Don't consume sugar (as cancer thrives on it), minimize or eliminate carb-rich foods like bread and pasta, and limit alcohol intake. Embrace the power of fasting to allow your body to heal itself.	[46,49,53,70,79,81,94-96]
Tentativeness	• 3 women with similar cancer, undergoing comparable treatments—2 passed away, but 1 is thriving Possible factor? She incorporated mistletoe & other non-pharma medicines into her regimen. • Concerns about [standard treatment] as a cancer solution persist, with claims of it being a harmful creation backed by influential medical forces. If it truly worked, wouldn't it have been banned long ago like Laetrile? • Listen or not: Vitamin B17, found in Apricot seeds and sold online as a "health supplement," has caught my attention as a potential cancer cure.	[49,51,59,61,62,66,81,94,96-100]
Absolute language or certainty	• I take sea buckthorn pills! They are an absolute lifesaver. • Vitamin B17 has definitely prevented my cancer from spreading. It's been a while, and there has been no growth. • During my time in a chemo clinic, alternative treatments were never allowed to be discussed or promoted. I left and started studying herbal medicine.	[43,51,59,61,94,97-101]
Profanity	• Create an alkaline environment that cancer can't thrive in! Incorporate herbs, vitamins, and minerals to support your healing journey. You are going to heal and beat that s*** • Go to a poor country and you get real tea with real ginger. Go to a rich country and you will get chemical b**** that will give you cancer • It damages healthy cells, no surefire cancer cure. It's like a c*** shoot for survival & recurrence. But I choose a different path: starving cancerous cells with therapeutic fasting & lifestyle shifts.	[48,57,62,63,66,69,81,89,96,98,102]

XSL•FO
RenderX

Characteristics	Examples of linguistic characteristics and posts with misinformation[a]	Studies using characteristics for misinformation detection
Named entities		[44,49,51,60,64,69,79,93,103-109]
Names	• I watched the documentary of Dr. B [name] on YouTube. He cured stage 4 cancer with no chemotherapy and no radiation.	
Location	• Fascinating, study from M [name of State]! Certain sound frequencies may aid the body in fighting cancer. Pair this with an alkaline diet - and the world is cured!	
Organization	• Must-watch documentary on YouTube! Unveiling a shocking cancer cure cover-up for over 40 years! B [name]: The Cancer Cure Cover-Up—Full documentary available now!	
URL	• Insights from Dr. N [name]! Learn how to transform the cancer terrain, boost immunity, and create an inhospitable environment for cancer using Acupuncture, Chinese herbal medicines, and food therapies. Check out the discussion here: [link provided].	[45,51,52,54,55,62,69,78,79 ,86-88,92,93,98,99,101,104,107-117]
Numeric data	• Cancer is nearly 100% curable but beware of certain hospital treatments. Explore alternative options for better outcomes.	[44,49,51,57,65,67,70,72,73,79,81,94,98,101,105]
Pronouns	• I love your positivity and your fight against cancer. Keep up the fight and adhere to Alkaline Diet for a healthier journey. • Your cancer can be cured by #fasting paired with no sugar alkaline diet. • A pro basketball player revealed how organic Wheatgrass healed his close friend from blood cancer. A testament to the power of natural remedies!	[61,66,68,72,78,79,93,97,99,103,106,108,112,118-121]
Hashtag	• #TualangHoney helps against skin Cancer with no side effects.	[43,44,47,52-55,59,64,66,77-79,82,87, 92,96,98,101,104,107,108,111,115,119,122,123]

[a]All posts were paraphrased to protect the author's anonymity.

[b]In sentiment analysis, emotions are identified by a "black box" model (DistilBERT). While we report here examples and highlight "negative/positive" words in the sentence, we must acknowledge that the algorithm may or may not use these words for detecting emotions.

Collected Data From X

We collected a total of 45,791 posts related to unproven cancer therapies. Among these, 13,046 posts were labeled as misinformation (forming the misinformation dataset), while 32,745 posts were categorized as non-misinformation (comprising control dataset 1). Furthermore, we gathered 6782 posts from the profiles of comprehensive cancer centers, which were used as control dataset 2, as shown in Figure 1. The content description of both the misinformation dataset and the control dataset 1 is shown in Table 2. To illustrate the dataset in this study, we categorized the X posts into 9 distinct categories. The examples of the posts with misinformation are shown in Table 1.

Table 2. Relevant prevalence of therapy categories within posts about unproven cancer therapy.

Categories of therapies	Total posts, n	Posts with misinformation, n (%)[a]	Examples of unproven cancer therapy
Diet based	5179	3069 (59)	Antioxidant, fasting, and alkaline diet
Alternative health system	7036	2250 (32)	Herbal therapy and ayurveda
Plant- and fungus-based	13,851	4386 (32)	Mushrooms
Synthetic substances	8471	2637 (31)	Antineoplastic Brudzinski and vitamin C
Spiritual and mental healing	2347	272 (12)	Meditation, praying, and tai chi
Electromagnetic and energy-based	2825	283 (10)	Polarity therapy and magnetic
Physical procedures	1144	49 (4)	Acupuncture
Other	4938	100 (2)	N/A[b]
Total	45,791	13,046 (28)	N/A

[a]Out of the total number of posts.

[b]N/A: not applicable.

XSL•FO
RenderX

Linguistic Characteristics Testing: Prediction of Misinformation Labels

As shown in Table 3, experiment 1 demonstrated that linguistic characteristics predicted misinformation with 60% accuracy. In experiment 2, they exhibited even stronger predictive power, achieving an accuracy of 77%. The importance scores for each linguistic characteristic are shown in Table 4.

Next, we selected linguistic characteristics with an impact score 0.05 and consistent predictive performance across experiments 1 and 2. These short-listed characteristics underwent further testing within the same experiments. In experiment 1, the short-listed characteristics achieved an accuracy rate of 50%, which did not significantly differ from random chance (*P*>.90). However, in experiment 2, these characteristics predicted misinformation with an accuracy rate of 73% and an AUC of 83. This performance was significantly better than random chance (McNemar $^{2}_{1}$=5.7 $\times 10^{7}$; *P*<.001). The importance scores for the short-listed characteristics are shown in Table 4. For a more detailed breakdown of the importance scores, we have summarized the percentage of posts containing these short-listed characteristics by dataset in Table 4 and the complete list in Multimedia Appendix 4.

Table 3. Lasso regression performance.

Name of the dataset	Total posts, n	Posts with misinformation, n	Accuracy, %
Experiment 1: misinformation dataset and control dataset 1	45,791	13,046	60
Experiment 2: misinformation dataset and control dataset 2	19,828	13,046	77

Table 4. Importance scores.

Linguistic characteristics	Experiment with control group 1		Experiment with control group 2		Experiment with short-listed characteristics (control group 2)	
	Predictors		Predictors		Predictors	
	Negative	Positive	Negative	Positive	Negative	Positive
Absolute language	—[a]	*0.11* [b]	—	*0.69*	—	*0.84*
Certainty	—	*0.21*	—	*1.13*	—	*1.02*
First-person pronoun	0.27	—	—	1.31	—	—
Hashtags	*0.56*	—	*1.55*	—	*1.6*	—
Location	*0.27*	—	*0.27*	—	*0.46*	—
Name	—	0.08	0.91	—	—	—
Negation	0.53	—	—	0.73	—	—
Negative emotions	0.24	—	0	—	—	—
Neutral emotions	0	—	—	0.07	—	—
Number	—	*0.17*	—	*0.29*	—	*0.28*
Organization	—	0.02	0.63	—	—	—
Positive emotions	—	0.31	0.46	—	—	—
Profanity	0.92	—	—	1.99	—	—
Second-person pronoun	0.02	—	0.45	—	—	—
Tentativeness	*0.08*	—	*0.16*	—	*0.08*	—
Third-person pronoun	0	—	0.23	—	—	—
URL	*0.3*	—	*2.28*	—	*2.47*	—

[a]Not applicable.

[b]Italicized values represent short-listed characteristics.

XSL•FO
RenderX

Table 5. The percentage of posts with short-listed linguistic characteristics.

Linguistic characteristics	Misinformation dataset (n=13,046), n (%)	Control dataset 1 (n=32,745), n (%)	Control dataset 2 (n=6782), n (%)
Positive predictors			
Certainty	1579 (12)	3044 (9)	208 (3)
Absolute	2741 (21)	7294 (22)[a]	630 (9)
Number	6358 (49)	14,360 (44)	2497 (37)
Negative predictors			
URL	6978 (53)	19,591 (60)	6560 (97)
Hashtags	2296 (18)	8512 (26)	4343 (64)
Location	1212 (9)	3373 (12)	975 (14)
Tentativeness	4154 (32)	11,171 (34)	1835 (27)[a]

[a]Valence of predictions is inferred from the model, which includes all characteristics simultaneously.

Discussion

Principal Findings

We have identified linguistic characteristics that can help people affected by cancer detect cancer misinformation on social media platforms such as X. Linguistic characteristics that were *likely* to be present in posts with misinformation were related to certain, absolute language, and numbers. Certain language included phrases that reflected a "degree of bravado" or "boasting of certainty." Examples of certain languages could be "I really believe," "it is definitely helpful," and similar others [36]. The absolute language referred to phrases that reflect black-and-white thinking and included words such as "none," "all," "never," and others [36]. The number category encompassed any information reported with digits such as percentages, count of any units, years, and priorities. Notably, all 3 linguistic characteristics could be united under the umbrella of definite, confident language. Linguistic characteristics that were *unlikely* to be present in posts with misinformation encompassed URLs, hashtags, and location mentions. Each of these attributes could be considered as a form of citation or reference. URLs offered direct links to the original source or further information, hashtags connected posts to broader relevant discussions, while locations mentioned in posts provided context and a sense of origin to the information shared. Our findings are consistent with some of the suggestions provided by previous guidelines for identifying misinformation. For instance, the Food and Drug Administration recommends being vigilant if patients read confident statements such as a drug definitely "cures cancer" or "guarantees results" [124]. Other guidelines encouraged users to search for references and original sources of health-related information [12-14].

While consistent with previous recommendations, our findings make a unique contribution. Previous work has based the guidelines on theoretical assumptions, while our study is one of the first to provide some empirical evidence based on a large dataset to support the recommendations for users. Another contribution is that we outlined ineffective linguistic characteristics for detecting cancer misinformation. Despite a substantial body of research showing that social media posts with sentiments predicted fake news, we did not find these relationships. A potential explanation could be the algorithm's limited efficiency in identifying emotions within cancer-related contexts. Furthermore, it is possible that authors express a limited range of emotions in cancer-related conversations, typically negative emotions toward cancer and both positive and negative emotions toward various treatments, including those that are unproven. These emotions may vary little across posts containing valid and nonvalid information, making emotions an unreliable factor for distinguishing misinformation.

Our work accumulates knowledge about misinformation detection from the literature covering a wide range of contexts—including political, social, and computer science—and translates this knowledge to the cancer context. The findings highlighted promising avenues for future research and could expedite the development of automated and augmented methods for identifying and verifying cancer-related misinformation on social media platforms. Finally, the robust labeled datasets developed by our research team are available to other researchers upon request to the corresponding author, thereby further supporting research on misinformation within the context of cancer and social media.

In practice, our work is at the forefront of customizing recommendations and contextualizing them for social network users. Our exploratory findings suggest a promising direction for studying linguistic characteristics that information users might apply when making quick judgments while scrolling through X feeds. Empowering users to stay vigilant in their initial evaluations could help reduce the spread of misinformation and the formation of erroneous beliefs. This is a crucial area for future research, which should explore how these findings apply in different cancer-related contexts and across various social networks.

Limitations

All the studies included in our analysis exclusively originate from peer-reviewed journals and conference proceedings; however, we must exercise caution when considering the potential for publication bias. Furthermore, in accordance with our selection criteria for linguistic characteristics, we included

XSL•FO
RenderX

only those papers that focused on text and excluded other forms of social media content, such as videos and images. We recommend that future research comprehensively explore social media, including multimedia content, as it could potentially provide additional insights for user-friendly recommendations.

In selecting linguistic characteristics, we prioritized observability, applicability, and generalizability. However, alternative criteria may be considered when users are open to a more thorough exploration of a post's validity. For example, future research should explore the use of metadata, link content analysis, and hashtag meanings. As misinformation evolves and its authors adjust to societal changes, the linguistic characteristics that identify misinformation may also shift. A longitudinal analysis is necessary to understand how linguistic characteristics perform in predicting misinformation over time.

Algorithms used in our analysis operate with a certain level of accuracy. Specifically, the accuracy of label identification in the dataset reached 83%, indicating that approximately 17% of posts were labeled incorrectly. This means that in experiment 1 some proportion of misinformation is included in the non-misinformation group and vice versa, making further exploration less accurate in experiment 1. This degree of uncertainty is common in algorithmic performance. Therefore, it is important to interpret our results in light of the inherent imperfections in algorithmic performance.

Furthermore, we encountered that the short-listed linguistic characteristics did not significantly outperform random chance in identifying misinformation in experiment 1. This outcome underscores a potential boundary condition of the effectiveness of the linguistic characteristics. Notably, experiment 1 encompassed more homogeneous data in contrast to experiment 2. Based on these findings, it becomes plausible to speculate that linguistic characteristics might provide limited help when a reader assesses posts within a closely knit community.

In experiment 2, the control dataset 2 consisted of posts shared by cancer centers and was compared with the misinformation dataset comprising random posts. To address this limitation, we collected posts from cancer centers that contain words related to cancer therapies. This step was taken to ensure a similar context of discussion as the posts with unproven therapy. Next, we exclude linguistic characteristics that are likely displayed differences between datasets due to the distinct nature of the information within control dataset 2. For example, linguistic traits such as "the use of profanity" or "first-person pronouns" were discarded. Furthermore, we decided to focus our analysis solely on the text within the posts and omitted other accompanying metainformation that users might observe, such as the user's name, location of the author, and posting time. This approach allowed us to assume that posts shared by cancer centers might be perceived more broadly, for instance, as posts shared by researchers, physicians, administrators, and patient advocates. Because of these measures, we anticipate that the linguistic characteristics identified in this research may help differentiate between health misinformation and factual posts on social media, irrespective of their sources. Despite our precautionary measures, we cannot fully guarantee that identified linguistics characteristics certainly distinguish between posts with misinformation and non-misinformation versus posts produced by the general public and posts by health experts from health care systems. However, there are factors that support the first conclusion more than the second. First, our findings are consistent with the previous theoretical and practical recommendations for identifying misinformation [12-14]. Second, the associated with misinformation linguistic characteristics, such as numbers and assertive language, are expected to be used by health experts. For instance, providers use numbers more confidently than the general public [125]. Professional guidelines for health providers encourage them to use numbers over verbal descriptions [126] as well as the use of assertive language in communication with patients [127,128]. Yet, our study associated these characteristics with misinformation shared by the general public on social media, which suggests that we might be finding more than just a mere distinction between the general public language and the health professional language. One study in and of itself is not yet a comprehensive body of evidence. Our findings will need to be validated and built upon via additional studies—including those that use posts from other types of entities and comparison groups.

Finally, our data were collected only on a single social network X. Many characteristics and customs of X are transferable to other social networks and our recommendations are likely to go beyond application on X, as demonstrated by the consistency of our recommendations with the recommendations of other researchers [12-14]. Given this limitation, our results need to be generalized cautiously, and further similar research is needed for different platforms (eg, Facebook, Pinterest, etc).

Conclusions

Our structured review synthesized knowledge from studies that used algorithmic approaches for text analysis to detect misinformation in social media. From this literature, we identified user-friendly linguistic characteristics that can assist individuals in distinguishing misinformation when they seek health-related information on social media. The linguistic characteristics, such as certainty, absolute language, and numbers, were positively associated with misinformation, while characteristics such as URLs, hashtags, and location mentions were negatively predictive of misinformation. Based on these findings, we suggested that users should be cautious of social media posts containing confident promises or specific numbers without proper references to the original information. According to our analysis, we expect that this approach will allow users to filter out two-thirds of posts with cancer-related misinformation. Yet, before drawing a definitive conclusion, further testing with different datasets is required.

XSL•FO
RenderX

Acknowledgments

ChatGPT 3.5 (OpenAI) [129] was used to assist with the professional editing of the manuscript. This study was supported by North Carolina Translational Research and Clinical Science Institute, Pilot Award Spring 2022.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Strategy for literature review.
[DOCX File , 29 KB - infodemiology_v5i1e62703_app1.docx]

Multimedia Appendix 2
List of unproven therapy.
[DOCX File , 25 KB - infodemiology_v5i1e62703_app2.docx]

Multimedia Appendix 3
Summary of the literature.
[DOCX File , 47 KB - infodemiology_v5i1e62703_app3.docx]

Multimedia Appendix 4
Summary of linguistic characteristics.
[XLSX File (Microsoft Excel File), 10 KB - infodemiology_v5i1e62703_app4.xlsx]

References

1. How much do you agree or disagree with the following statements? I use information from social media to make decisions about my health. HINTS. 2022. URL: https://hints.cancer.gov/view-questions/question-detail.aspx?PK_Cycle=14&qid=1849 [accessed 2023-09-20]
2. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
3. Johnson SB, Park HS, Gross CP, Yu JB. Use of alternative medicine for cancer and its impact on survival. J Natl Cancer Inst 2017;110:121-124.
4. Johnson SB, Park HS, Gross CP, Yu JB. Complementary medicine, refusal of conventional cancer therapy, and survival among patients with curable Cancers. JAMA Oncol 2018;4(10):1375-1381 [FREE Full text] [doi: 10.1001/jamaoncol.2018.2487] [Medline: 30027204]
5. Weil CR, Hutten RJ, Barney BM, Fagerlin A, Gaffney DK, Gill DM, et al. Shifting perceptions of alternative therapies in cancer patients during the COVID-19 pandemic: results from the complementary and alternative medicine exposure in oncology (CAMEO) study. J. Clin. Oncol 2022;40(16_suppl):e24130-e24130. [doi: 10.1200/jco.2022.40.16_suppl.e24130]
6. Johnson SB, Bylund CL. Identifying cancer treatment misinformation and strategies to mitigate its effects with improved radiation oncologist-patient communication. Pract Radiat Oncol 2023;13(4):282-285. [doi: 10.1016/j.prro.2023.01.007] [Medline: 36736620]
7. Johnson SB, Parsons M, Dorff T, Moran MS, Ward JH, Cohen SA, et al. Cancer misinformation and harmful information on facebook and other social media: a brief report. J Natl Cancer Inst 2022;114(7):1036-1039 [FREE Full text] [doi: 10.1093/jnci/djab141] [Medline: 34291289]
8. Alsyouf M, Stokes P, Hur D, Amasyali A, Ruckle H, Hu B. 'Fake News' in urology: evaluating the accuracy of articles shared on social media in genitourinary malignancies. BJU Int 2019;124(4):701-706. [doi: 10.1111/bju.14787] [Medline: 31044493]
9. Yussof I, Ab Muin NF, Mohd M, Hatah E, Mohd Tahir NA, Mohamed Shah N. Breast cancer prevention and treatment misinformation on Twitter: an analysis of two languages. Digit Health 2023;9:20552076231205742 [FREE Full text] [doi: 10.1177/20552076231205742] [Medline: 37808244]
10. Chen L, Wang X, Peng T. Nature and diffusion of gynecologic cancer–related misinformation on social media: analysis of tweets. J Med Internet Res 2018;20(10):e11515. [doi: 10.2196/11515] [Medline: 30327289]
11. Lazard AJ, Nicolla S, Vereen RN, Pendleton S, Charlot M, Tan H, et al. Exposure and reactions to cancer treatment misinformation and advice: survey study. JMIR Cancer 2023;9:e43749 [FREE Full text] [doi: 10.2196/43749] [Medline: 37505790]
12. Health misinformation. The US Surgeon General's Advisory. 2021. URL: https://www.hhs.gov/surgeongeneral/priorities/health-misinformation/index.html [accessed 2023-09-05]

XSL•FO
RenderX

13. Finding cancer information on the internet. American Cancer Society. 2020. URL: https://www.cancer.org/cancer/understanding-cancer/cancer-information-on-the-internet.html [accessed 2023-07-20]
14. How to find cancer resources you can trust. National Cancer Institute Internet. 2015. URL: https://www.cancer.gov/about-cancer/managing-care/using-trusted-resources [accessed 2022-06-28]
15. Bondielli A, Marcelloni F. A survey on fake news and rumour detection techniques. Inf Sci 2019;497:38-55. [doi: 10.1016/j.ins.2019.05.035]
16. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J. Lang. Soc. Psychol 2009;29(1):24-54. [doi: 10.1177/0261927x09351676]
17. Kapantai E, Christopoulou A, Berberidis C, Peristeras V. A systematic literature review on disinformation: toward a unified taxonomical framework. New Media Soc 2020;23(5):1301-1326. [doi: 10.1177/1461444820959296]
18. Zannettou S, Sirivianos M, Blackburn J, Kourtellis N. The web of false information. ACM J Data Inf Qual 2019;11(3):1-37. [doi: 10.1145/3309699]
19. Sendra A, Farré J, Vaagan RW. Seeking, sharing and co-creating: a systematic review of the relation between social support theory, social media use and chronic diseases. Soc Theory Health 2019;18(4):317-339. [doi: 10.1057/s41285-019-00106-z]
20. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
21. Molina MD, Sundar SS, Le T, Lee D. “Fake News” is not simply false information: a concept explication and taxonomy of online content. Am Behav Sci 2019;65(2):180-212. [doi: 10.1177/0002764219878224]
22. Babineau J. Product review: Covidence (systematic review software). J Can Health Libr Assoc 2014;35:68-71.
23. Ghenai A, Mejova Y. Fake cures. Proc ACM Hum Comput Interact 2018;2:1-20. [doi: 10.1145/3274327]
24. The Free Encyclopedia. List of unproven and disproven cancer treatments. Wikipedia [Internet]. Wikimedia Foundation, Inc. 2013. URL: https://en.wikipedia.org/wiki/List_of_unproven_and_disproven_cancer_treatments [accessed 2023-07-21]
25. Illegally sold cancer treatments. US Food and Drug Administration. 2017. URL: https://www.fda.gov/consumers/health-fraud-scams/illegally-sold-cancer-treatments [accessed 2023-09-01]
26. White RW, Hassan A. Content bias in online health search. ACM Trans Web 2014;8(4):1-33. [doi: 10.1145/2663355]
27. Marzi G, Balzano M, Marchiori D. K-Alpha Calculator-Krippendorff's Alpha Calculator: a user-friendly tool for computing Krippendorff's alpha inter-rater reliability coefficient. MethodsX 2024;12:102545 [FREE Full text] [doi: 10.1016/j.mex.2023.102545] [Medline: 39669968]
28. Oortwijn Y, Ossenkoppele T, Betti A. Interrater disagreement resolution: a systematic procedure to reach consensus in annotation tasks. 2021 Presented at: Proceedings of the Workshop on Human Evaluation of NLP Systems (HumEval); 2021 April 01; Dublin, Ireland p. 131-141.
29. Minaee S, Kalchbrenner N, Cambria E, Nikzad N, Chenaghlu M, Gao J. Deep learning--based text classification. ACM Comput Surv 2021;54(3):1-40. [doi: 10.1145/3439726]
30. Qasim R, Bangyal WH, Alqarni MA, Ali Almazroi A. A fine-tuned BERT-based transfer learning approach for text classification. J Healthc Eng 2022;2022:3498123 [FREE Full text] [doi: 10.1155/2022/3498123] [Medline: 35013691]
31. Roitero K, Bozzato C, Della MV, Mizzaro S, Serra G. Twitter goes to the doctor: detecting medical tweets using machine learning and BERT. 2020. URL: http://hdl.handle.net/11390/1189151 [accessed 2023-07-12]
32. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv Preprint posted online on October 11, 2018. [doi: 10.48550/ARXIV.1810.04805]
33. Elroy O, Yosipof A. Semi-supervised learning classifier for misinformation related to earthquakes prediction on social media. 2023 Presented at: Artificial Neural Networks and Machine Learning – ICANN 2023: 32nd International Conference on Artificial Neural Networks; September 26-29, 2023; Heraklion, Greece. [doi: 10.1007/978-3-031-44207-0_22]
34. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020;41:433-451 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
35. Chapman C, Stolee K. Exploring regular expression usage and context in Python. 2016 Presented at: Proceedings of the 25th International Symposium on Software Testing and Analysis; July 18, 2016; New York, NY. [doi: 10.1145/2931037.2931073]
36. Boyd RL, Ashokkumar A, Seraj S, Pennebaker JW. The development and psychometric properties of LIWC-22. 2022. URL: https://www.liwc.app [accessed 2023-07-19]
37. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 2009;29(1):24-54. [doi: 10.1177/0261927x09351676]
38. Goyal A, Gupta V, Kumar M. Recent named entity recognition and classification techniques: a systematic review. Comput. Sci. Rev 2018;29:21-43. [doi: 10.1016/j.cosrev.2018.06.001]
39. Ranganathan J, Tsahai T. Sentiment analysis of tweets using deep learning. In: Advanced Data Mining and Applications. Cham, Switzerland: Springer International Publishing AG; 2022:106-117.
40. Kuhn M, Wickham H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. 2020. URL: https://www.tidymodels.org [accessed 2023-07-24]
41. Bach F. Consistency of the group Lasso and multiple kernel learning. J Mach Learn Res 2008;9:1179-1225.

XSL•FO
RenderX

42. Bach FR. Bolasso: model consistent Lasso estimation through the bootstrap. 2008 Presented at: Proceedings of the 25th International Conference on Machine Learning; July 5, 2008; Helsinki, Finland p. 33-40. [doi: 10.1145/1390156.1390161]
43. Agarwal S, Samavedhi A. Profiling fake news: learning the semantics and characterisation of misinformation. In: Advanced Data Mining and Applications. ADMA 2022. Lecture Notes in Computer Science, vol 13087. Cham, Switzerland: Springer International Publishing; 2022:203-216.
44. Agarwal R, Gupta S, Chatterjee N. Profiling fake news spreaders on Twitter: a clickbait and linguistic feature based scheme. In: Natural Language Processing and Information Systems. New York City, NY: Springer International Publishing; 2022:345-357.
45. Alkhodair SA, Fung BCM, Ding SHH, Cheung WK, Huang S. Detecting high-engaging breaking news rumors in social media. ACM Trans Manage Inf Syst 2020;12(1):1-16. [doi: 10.1145/3416703]
46. Garg S, Kumar Sharma D. Linguistic features based framework for automatic fake news detection. Comput Ind Eng 2022;172(Part A):108432. [doi: 10.1016/j.cie.2022.108432]
47. Agrawal C, Pandey A, Goyal S. Fake news detection system based on modified bi-directional long short term memory. Multimed Tools Appl 2022;81(17):24199-24223. [doi: 10.1007/s11042-022-12772-9]
48. Ahmad I, Yousaf M, Yousaf S, Ahmad MO. Fake news detection using machine learning ensemble methods. Complexity 2020;2020:1-11 [FREE Full text] [doi: 10.1155/2020/8885861]
49. Alsmadi I, O'Brien MJ. Rating news claims: feature selection and evaluation. Math Biosci Eng 2019;17(3):1922-1939 [FREE Full text] [doi: 10.3934/mbe.2020101] [Medline: 32233515]
50. Aneja S, Aneja N, Kumaraguru P. Predictive linguistic cues for fake news: a societal artificial intelligence problem. arXiv Preprint posted online on November 26, 2022 [FREE Full text]
51. Ebrahimi Fard A, Mohammadi M, Chen Y, van de Walle B. Computational rumor detection without non-rumor: a one-class classification approach. IEEE Trans Comput Soc Syst 2019;6(5):830-846. [doi: 10.1109/tcss.2019.2931186]
52. Sharma AK, Gangeya RA, Kumar H, Chaurasia S, Srivastava DK. Ensemble rumor text classification model applied to different tweet features. Int J Fuzzy Log Intell Syst 2022;22(3):325-338. [doi: 10.5391/ijfis.2022.22.3.325]
53. Sharma D, Shrivastava P, Garg S. Utilizing word embedding and linguistic features for fake news detection. 2022 Presented at: 9th International Conference on Computing for Sustainable Global Development (INDIACom); March 23-25, 2022; New Delhi, India. [doi: 10.23919/indiacom54597.2022.9763294]
54. Shelke S, Attar V. Rumor detection in social network based on user, content and lexical features. Multimed Tools Appl 2022;81(12):17347-17368 [FREE Full text] [doi: 10.1007/s11042-022-12761-y] [Medline: 35282405]
55. Shelke S, Attar V. Role of various features in identification of rumors in the social network. 2021 Presented at: 12th International Conference on Computing Communication and Networking Technologies (ICCCNT); July 06-08, 2021; Kharagpur, India. [doi: 10.1109/icccnt51525.2021.9579856]
56. Sicilia R, Giudice S, Pei Y, Pechenizkiy M, Soda P. Health-related rumour detection on Twitter. 2017 Presented at: IEEE International Conference on Bioinformatics and Biomedicine (BIBM); November 13-16, 2017; Kansas City, MO. [doi: 10.1109/BIBM.2017.8217899]
57. Sun QJ. A machine learning analysis of the features in deceptive and credible news. arXiv Preprint posted online on October 5, 2019. [doi: 10.48550/arXiv.1910.02223]
58. Tuarob S, Tucker CS, Salathe M, Ram N. An ensemble heterogeneous classification methodology for discovering health-related knowledge in social media messages. J Biomed Inform 2014;49:255-268 [FREE Full text] [doi: 10.1016/j.jbi.2014.03.005] [Medline: 24642081]
59. van de Guchte L, Raaijmakers S, Meeuwissen E, Spenader J. Near real-time detection of misinformation on online social networks. In: Disinformation in Open Online Media. Cham, Switzerland: Springer International Publishing; 2020:240-260.
60. Vijeev A, Mahapatra A, Shyamkrishna A, Murthy S. A hybrid approach to rumour detection in microblogging platforms. 2018 Presented at: International Conference on Advances in Computing, Communications and Informatics (ICACCI); September 19-22, 2018; Bangalore, India. [doi: 10.1109/ICACCI.2018.8554371]
61. Volkova S, Jang JY. Misleading or falsification: inferring deceptive strategies and types in online news and social media. 2018 Presented at: Companion Proceedings of The Web Conference 2018; April 23, 2018; Lyon, France p. 575-583. [doi: 10.1145/3184558.3188728]
62. Xuan K, Xia R. Rumor stance classification via machine learning with text, user and propagation features. 2019 Presented at: International Conference on Data Mining Workshops (ICDMW); November 8-11, 2019; Beijing, China. [doi: 10.1109/icdmw.2019.00085]
63. Zhou X, Jain A, Phoha VV, Zafarani R. Fake news early detection: an interdisciplinary study. arXiv Preprint posted online on April 26, 2019. [Medline: 10.48550/ARXIV.1904.11679]
64. Wang L, Wang Y, de Melo G, Weikum G. Five shades of untruth: finer-grained classification of fake news. 2018 Presented at: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM); August 28-31, 2018; Barcelona, Spain. [doi: 10.1109/asonam.2018.8508256]
65. Gautam A, Jerripothula KR. SGG: spinbot, grammarly and glove based fake news detection. 2020 Presented at: IEEE Sixth International Conference on Multimedia Big Data (BigMM); September 24-26, 2020; New Delhi, India. [doi: 10.1109/bigmm50055.2020.00033]

XSL•FO
RenderX

66. Mahbub S, Pardede E, Kayes ASM. COVID-19 rumor detection using psycho-linguistic features. IEEE Access 2022;10:117530-117543. [doi: 10.1109/access.2022.3220369]
67. Niven T, Kao HY, Wang HY. Profiling spreaders of disinformation on Twitter: IKMLab and softbank submission. CLEF (Working Notes) 2020;2696:1-8 [FREE Full text]
68. Mondal T, Pramanik P, Bhattacharya I, Boral N, Ghosh S. Analysis and early detection of rumors in a post disaster scenario. Inf Syst Front 2018;20(5):961-979. [doi: 10.1007/s10796-018-9837-8]
69. Basharat S, Ahmad M. Inferring trust from message features using linear regression and support vector machines. In: Communications in Computer and Information Science. Singapore: Springer Singapore; 2018:577-598.
70. Butt S, Sharma S, Sharma R, Sidorov G, Gelbukh A. What goes on inside rumour and non-rumour tweets and their reactions: a psycholinguistic analyses. Comput Human Behav 2022;135:107345. [doi: 10.1016/j.chb.2022.107345]
71. Casillo M, Colace F, Conte D, de Santo M, Lombardi M, Mottola S. A multi-feature bayesian approach for fake news detection. In: Computational Data and Social Networks. Cham, Switzerland: Springer International Publishing; 2020:333-344.
72. Castillo JM, Fadera KDF, Ladao AAA, Go JG, Tamayo MB, Octaviano MV. Fake news detection on English news article?s title. 2021 Presented at: 1st International Conference in Information and Computing Research (iCORE); December 11-12, 2021; Manila, Philippines. [doi: 10.1109/icore54267.2021.00046]
73. Cerbin L, DeJesus J, Warnken J, Gokhale S. Unmasking the mask debate on social media. 2021 Presented at: IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC); July 12-16, 2021; Madrid, Spain. [doi: 10.1109/compsac51774.2021.00098]
74. Choudhary A, Arora A. Linguistic feature based learning model for fake news detection and classification. Expert Syst Appl 2021;169:114171. [doi: 10.1016/j.eswa.2020.114171]
75. Choudhry A, Khatri I, Jain M, Vishwakarma DK. An emotion-aware multitask approach to fake news and rumor detection using transfer learning. IEEE Trans Comput Soc Syst 2024;11(1):588-599. [doi: 10.1109/tcss.2022.3228312]
76. Felber T. Constraint 2021: machine learning models for COVID-19 fake news detection shared task. arXiv Preprint posted online on January 11, 2021 . [doi: 10.48550/arXiv.2101.03717]
77. Ghanem B, Ponzetto SP, Rosso P. FacTweet: profiling fake news Twitter accounts. In: Statistical Language and Speech Processing. Cham, Switzerland: Springer International Publishing; 2020:35-45.
78. Jogalekar NS, Attar V, Palshikar GK. Rumor detection on social networks: a sociological approach. 2020 Presented at: IEEE International Conference on Big Data (Big Data); December 10-13, 2020; Atlanta, GA. [doi: 10.1109/bigdata50022.2020.9378149]
79. Kumar A, Sangwan SR, Nayyar A. Rumour veracity detection on Twitter using particle swarm optimized shallow classifiers. Multimed Tools Appl 2019;78(4):24083-24101. [doi: 10.1007/s11042-019-7398-6]
80. Kumari R, Ashok N, Ghosal T, Ekbal A. A multitask learning approach for fake news detection: novelty, emotion, and sentiment lend a helping hand. 2021 Presented at: International Joint Conference on Neural Networks (IJCNN); July 18-22, 2021; Shenzhen, China. [doi: 10.1109/ijcnn52387.2021.9534218]
81. Kwon S, Cha M, Jung K. Rumor detection over varying time windows. PLoS One 2017;12(1):e0168344 [FREE Full text] [doi: 10.1371/journal.pone.0168344] [Medline: 28081135]
82. Li C, Chun SA, Geller J. Stemming the tide of fake news about the COVID-19 pandemic. Int FLAIRS Conf Proc 2022;35 [FREE Full text]
83. Rajabi Z, Shehu A, Purohit H. User behavior modelling for fake information mitigation on social web. In: Social, Cultural, and Behavioral Modeling. Cham, Switzerland: Springer International Publishing; 2019.
84. Sandrilla R, Devi MS. FNU-BiCNN: fake news and fake URL detection using Bi-CNN. Int J Adv Comput Sci Appl 2022;13(2) [FREE Full text]
85. Seddari N, Derhab A, Belaoued M, Halboob W, Al-Muhtadi J, Bouras A. A hybrid linguistic and knowledge-based analysis approach for fake news detection on social media. IEEE Access 2022;10:62097-62109. [doi: 10.1109/access.2022.3181184]
86. Singh JP, Kumar A, Rana NP, Dwivedi YK. Attention-based LSTM network for rumor veracity estimation of tweets. Inf Syst Front 2020;24(2):459-474. [doi: 10.1007/s10796-020-10040-5]
87. Singh JP, Rana NP, Dwivedi YK. Rumour veracity estimation with deep learning for Twitter. In: FIP Advances in Information and Communication Technology. Cham, Switzerland: Springer International Publishing; 2019.
88. Slimi H, Bounhas I, Slimani Y. URL-based tweet credibility evaluation. 2019 Presented at: IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA); November 3-7, 2019; Abu Dhabi, United Arab Emirates. [doi: 10.1109/aiccsa47632.2019.9035255]
89. Sotirakou C, Karampela A, Mourlas C. Evaluating the role of news content and social media interactions for fake news detection. In: Bright J, Giachanou A, Spaiser V, Spezzano F, George A, Pavliuc A, editors. Disinformation in Open Online Media. MISDOOM 2021. Lecture Notes in Computer Science, vol 12887. Cham, Switzerland: Springer International Publishing; 2021.
90. Verma PK, Agrawal P, Amorim I, Prodan R. WELFake: word embedding over linguistic features for fake news detection. IEEE Trans Comput Soc Syst 2021;8(4):881-893. [doi: 10.1109/tcss.2021.3068519]

XSL•FO
RenderX

91. Kula S, Choraś M, Kozik R, Ksieniewicz P, Woźniak M. Sentiment analysis for fake news detection by means of neural networks. In: Krzhizhanovskaya VV, Závodszky G, Lees MH, Dongarra JK, Sloot PMA, Brissos S, et al, editors. Lecture Notes in Computer Science. Cham, Switzerland: Springer International Publishing; 2020.
92. Vadavalli A, Al E. Feed forward neural network based effective feature extraction technique for better classification accuracy. TURCOMAT 2021;12(11):4576-4588.
93. Saikh T, Anand A, Ekbal A, Bhattacharyya P. A novel approach towards fake news detection: Deep learning augmented with textual entailment features. In: Natural Language Processing and Information Systems. Cham, Switzerland: Springer International Publishing; 2019:345-358.
94. Gravanis G, Vakali A, Diamantaras K, Karadais P. Behind the cues: a benchmarking study for fake news detection. Expert Syst Appl 2019;128:201-213. [doi: 10.1016/j.eswa.2019.03.036]
95. Hassan FM, Lee M. Political fake statement detection via multistage feature-assisted neural modeling. 2020 Presented at: IEEE International Conference on Intelligence and Security Informatics (ISI); November 9-10, 2020; Arlington, VA. [doi: 10.1109/isi49825.2020.9280531]
96. Patro J, Rathore PS. A sociolinguistic route to the characterizationdetection of the credibility of events on Twitter. 2020 Presented at: Proceedings of the 31st ACM Conference on Hypertext and Social Media; July 15, 2020; Virtual event, United States. [doi: 10.1145/3372923.3404795]
97. Kumar JM, Gopalani D, Kumar MY, Kumar R. Machine learning based fake news detection using linguistic features and word vector features. 2020 Presented at: 2020 IEEE 7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON); November 27-29, 2020; Prayagraj, India. [doi: 10.1109/upcon50219.2020.9376576]
98. Rastogi S, Bansal D. Disinformation detection on social media: an integrated approach. Multimed Tools Appl 2022;81(28):40675-40707 [FREE Full text] [doi: 10.1007/s11042-022-13129-y] [Medline: 35582207]
99. Reddy H, Raj N, Gala M, Basava A. Text-mining-based fake news detection using ensemble methods. Int J Autom Comput 2020;17(2):210-221. [doi: 10.1007/s11633-019-1216-5]
100. Volkova S, Shaffer K, Jang JY, Hodas N. Separating facts from fiction: linguistic models to classify suspicious trusted news posts on Twitter. 2017 Presented at: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers); December 24, 2024; Stroudsburg, PA. [doi: 10.18653/v1/p17-2102]
101. Lotfi S, Mirzarezaee M, Hosseinzadeh M, Seydi V. Rumor conversations detection in Twitter through extraction of structural features. Inf Technol Manag 2021;22(4):265-279. [doi: 10.1007/s10799-021-00335-7]
102. Khanam Z, Alwasel BN, Sirafi H, Rashid M. Fake news detection using machine learning approaches. IOP Conf Ser Mater Sci Eng 2021;1099(1):012040. [doi: 10.1088/1757-899x/1099/1/012040]
103. Braşoveanu AMP, Andonie R. Integrating machine learning techniques in semantic fake news detection. Neural Process Lett 2020;53(5):3055-3072. [doi: 10.1007/s11063-020-10365-x]
104. Hamidian S, Diab MT. Rumor detection and classification for Twitter data. arXiv Preprint posted online on November 25, 2019 [FREE Full text] [doi: 10.1090/mbk/121/79]
105. Iwendi C, Mohan S, Khan S, Ibeke E, Ahmadian A, Ciano T. Covid-19 fake news sentiment analysis. Comput Electr Eng 2022;101:107967 [FREE Full text] [doi: 10.1016/j.compeleceng.2022.107967] [Medline: 35474674]
106. Kasseropoulos D, Tjortjis C. An Approach Utilizing Linguistic Features for Fake News Detection. IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing; 2021.
107. Rezayi S, Soleymani S, Arabnia HR, Li S. Socially aware multimodal deep neural networks for fake news classification. 2021 Presented at: IEEE 4th International Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE; 2021; September 10, 2021; Tokyo, Japan. [doi: 10.1109/mipr51284.2021.00048]
108. Anggrainingsih R, Hassan GM, Datta A. BERT based classification system for detecting rumours on Twitter. arXiv Preprint posted online on September 7, 2021. [doi: 10.48550/ARXIV.2109.02975]
109. Tafannum F, Sharear Shopnil MN, Salsabil A, Ahmed N, Rabiul Alam MG, Tanzim Reza M. Demystifying black-box learning models of rumor detection from social media posts. 2021 Presented at: 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE; 2021; December 4, 2021; New York, NY. [doi: 10.1109/uemcon53757.2021.9666567]
110. Al-Rakhami MS, Al-Amri AM. Lies kill, facts save: detecting COVID-19 misinformation in Twitter. IEEE Access 2020;8:155961-155970. [doi: 10.1109/access.2020.3019600]
111. Kalyanam J, Velupillai S, Doan S, Conway M, Lanckriet G. Facts and fabrications about ebola: a Twitter based study. arXiv Preprint posted online on August 9, 2015. [doi: 10.48550/ARXIV.1508.02079]
112. Kwon S, Cha M, Jung K, Chen W, Wang Y. Prominent features of rumor propagation in online social media. 2013 Presented at: IEEE 13th International Conference on Data Mining; December 10, 2013; Dallas, TX. [doi: 10.1109/icdm.2013.61]
113. Mazzeo V, Rapisarda A, Giuffrida G. Detection of fake news on COVID-19 on web search engines. Front Phys 2021;9 [FREE Full text] [doi: 10.3389/fphy.2021.685730]
114. Puraivan E, Godoy E, Riquelme F, Salas R. Fake news detection on Twitter using a data mining framework based on explainable machine learning techniques. IET Conf Proc 2021;2021(1):157-162. [doi: 10.1049/icp.2021.1450]

XSL•FO
RenderX

115. Sahana VP, Pias AR, Shastri R, Mandloi S. Automatic detection of rumoured tweets and finding its origin. 2015 Presented at: International Conference on Computing and Network Communications (CoCoNet); December 16-19, 2015; Trivandrum, India. [doi: 10.1109/coconet.2015.7411251]
116. Sailunaz K, Kawash J, Alhajj R. Tweet and user validation with supervised feature ranking and rumor classification. Multimed Tools Appl 2022;81(22):31907-31927. [doi: 10.1007/s11042-022-12616-6]
117. Sharma D, Singhal S. Detection of fake news on social media using classification data mining techniques. Int J Eng Adv Technol 2019;9(1):3132-3138.
118. Afroz S, Brennan M, Greenstadt R. Detecting hoaxes, frauds, and deception in writing style online. 2012 Presented at: IEEE Symposium on Security and Privacy. IEEE; May 23, 2012; San Francisco, CA. [doi: 10.1109/sp.2012.34]
119. Khan S, Hakak S, Deepa N, Prabadevi B, Dev K, Trelova S. Detecting COVID-19-related fake news using feature extraction. Front Public Health 2021;9:788074 [FREE Full text] [doi: 10.3389/fpubh.2021.788074] [Medline: 35059379]
120. Levi O, Hosseini P, Diab M, Broniatowski D. Identifying nuances in fake news vs. satire: using semantic and linguistic cues. 2019 Presented at: Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda; November 01, 2019; Hong Kong, China. [doi: 10.18653/v1/d19-5004]
121. Argyris YA, Zhang N, Bashyal B, Tan PN. Using deep Learning to identify linguistic features that facilitate or inhibit the propagation of anti- and pro-vaccine content on social media. 2022 Presented at: 2022 IEEE International Conference on Digital Health (ICDH). IEEE; July 10-16, 2022; Barcelona, Spain p. 107-116 URL: https://europepmc.org/abstract/MED/37975063 [doi: 10.1109/icdh55609.2022.00025]
122. Maan M, Jain M, Trivedi S, Sharma R. Machine learning based rumor detection on Twitter data. In: Communications in Computer and Information Science. New York, NY: Springer International Publishing; 2022.
123. Raj C, Meel P. Is dynamic rumor detection on social media viable? An unsupervised perspective. arXiv Preprint posted online on November 23, 2021. [doi: 10.48550/arXiv.2111.11982]
124. Office of the Commissioner. 6 Tip-offs to rip-offs: don't fall for health fraud scams. US Food and Drug Administration. 2021. URL: https://www.fda.gov/consumers/consumer-updates/6-tip-offs-rip-offs-dont-fall-health-fraud-scams [accessed 2023-07-20]
125. Apter AJ, Paasche-Orlow MK, Remillard JT, Bennett IM, Ben-Joseph EP, Batista RM, et al. Numeracy and communication with patients: they are counting on us. J Gen Intern Med 2008;23(12):2117-2124 [FREE Full text] [doi: 10.1007/s11606-008-0803-x] [Medline: 18830764]
126. Büchter RB, Fechtelpeter D, Knelangen M, Ehrlich M, Waltering A. Words or numbers? Communicating risk of adverse effects in written consumer health information: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2014;14:76 [FREE Full text] [doi: 10.1186/1472-6947-14-76] [Medline: 25155972]
127. Omura M, Maguire J, Levett-Jones T, Stone TE. The effectiveness of assertiveness communication training programs for healthcare professionals and students: a systematic review. Int J Nurs Stud 2017;76:120-128. [doi: 10.1016/j.ijnurstu.2017.09.001] [Medline: 28964979]
128. Richard C, Lussier M, Millette B, Tanoubi I. Healthcare providers and patients: an essay on the importance of professional assertiveness in healthcare today. Med Educ Online 2023;28(1):2200586 [FREE Full text] [doi: 10.1080/10872981.2023.2200586] [Medline: 37040516]
129. ChatGPT 3.5. OpenAI. URL: https://chatgpt.com/g/g-F00faAwkE-open-a-i-gpt-3-5 [accessed 2025-01-16]

Abbreviations

API: application programming interface
AUC: area under the curve
BERT: Bidirectional Encoder Representations from Transformers
LIWC: Linguistic Inquiry and Word Count
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by T Mackey; submitted 31.05.24; peer-reviewed by D Boatman, A King, W Ahmed; comments to author 05.07.24; revised version received 22.08.24; accepted 23.11.24; published 12.02.25.

Please cite as:
Fridman I, Boyles D, Chheda R, Baldwin-SoRelle C, Smith AB, Elston Lafata J
Identifying Misinformation About Unproven Cancer Treatments on Social Media Using User-Friendly Linguistic Characteristics: Content Analysis
JMIR Infodemiology 2025;5:e62703
URL: https://infodemiology.jmir.org/2025/1/e62703
doi:10.2196/62703
PMID:

XSL•FO
RenderX

©Ilona Fridman, Dahlia Boyles, Ria Chheda, Carrie Baldwin-SoRelle, Angela B Smith, Jennifer Elston Lafata. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 12.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Model of Trust in Online COVID-19 Information and Advice: Cross-Sectional Questionnaire Study

Elizabeth Sillence[1], PhD; Dawn Branley-Bell[1*], PhD; Mark Moss[1*], PhD; Pam Briggs[1*], PhD

Department of Psychology, Northumbria University, Newcastle upon Tyne, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Elizabeth Sillence, PhD
Department of Psychology
Northumbria University
Ellison Terrace
Newcastle upon Tyne, NE1 8ST
United Kingdom
Phone: 44 1912437251
Email: elizabeth.sillence@northumbria.ac.uk

Abstract

Background: During the COVID-19 pandemic, many people sought information from websites and social media. Understanding the extent to which these sources were trusted is important in relation to health communication.

Objective: This study aims to identify the key factors influencing UK citizens' trust and intention to act on advice about COVID-19 found via digital resources and to test whether an existing model of trust in eHealth provided a good fit for COVID-19–related information seeking online. We also wished to identify any differences between the evaluation of general information and information relating specifically to COVID-19 vaccines.

Methods: In total, 525 people completed an online survey in January 2022 encompassing a general web trust questionnaire, measures of information corroboration, coping perceptions, and intention to act. Data were analyzed using principal component analysis and structural equation modeling. The evaluation responses of general information and COVID-19 vaccine information were also compared.

Results: The principal component analysis revealed 5 trust factors: (1) credibility and impartiality, (2) familiarity, (3) privacy, (4) usability, and (5) personal experiences. In the final structural equation modeling model, trust had a significant direct effect on intention to act (β=.65; P<.001). Of the trust factors, credibility and impartiality had a significant positive direct effect on trust (β=.82; P<.001). People searching for vaccination information felt less at risk, less anxious, and more optimistic after reading the information. We noted that most people sought information from "official" sources. Finally, in the context of COVID-19, "credibility and impartiality" remain a key predictor of trust in eHealth resources, but in comparison with previous models of trust in online health information, checking and corroborating information did not form a significant part of trust evaluations.

Conclusions: In times of uncertainty, when faced with a global emergent health concern, people place their trust in familiar websites and rely on the perceived credibility and impartiality of those digital sources above other trust factors.

(JMIR Infodemiology 2025;5:e59317) doi:10.2196/59317

KEYWORDS

eHealth; electronic health; digital intervention; trust; online information seeking; scientific credibility; digital resources; COVID-19; SARS-CoV-2; respiratory; infectious; pulmonary; pandemic; public health; health information; global health; surveys; social media

Introduction

Background

The COVID-19 pandemic understandably led to an increase in "official" sources of information and advice from politicians, public health officials, clinicians, and scientists. This public-facing information was communicated via the mainstream press, through live-streamed press briefings, and online. However, "unofficial" sources of information were also circulated, primarily via social media. For individuals, access to good quality information during the pandemic was critical,

XSL•FO
RenderX

not least because official messaging was constantly being updated in relation to recommended or mandated behaviors such as social distancing, mask-wearing, and self-isolation.

During this time, many people sought their information online [1] through websites, social media, and mobile apps. People looked for information on the signs and symptoms of the virus, measures to avoid catching and spreading the virus, self-care once infected, and vaccination information. In addition to health advice, people also sought related information on rules and guidance regarding self-isolating, masks, and social distancing.

Accurate and appropriate health communication is an important tool in tackling any pandemic and it can directly influence individuals' affective and behavioral responses to a crisis [2]. In relation to the COVID-19 pandemic, studies have shown that access to a larger and more diverse set of information sources led to increased worry [1,3] and greater confusion, in part because of the infodemic of misinformation and rumors that were promoted about the pandemic [4]. The UK Government's approach to tackling COVID-19 relied upon broad public trust, but issues with inconsistent and unclear messaging, as well as general political mistrust, were apparent [5]. In short, it sometimes became difficult for people to know who to trust in relation to taking appropriate actions to reduce the spread of COVID-19 and minimize personal risk.

Against this backdrop, the aim of this study was to understand more about the digital resources people in the United Kingdom used for COVID-19–related information and the extent to which they trusted these resources. Although we know that online health formed a key source of information for many people during the pandemic, we do not know how people evaluated these digital sources and what factors were important in trusting the information, the source, and ultimately deciding whether or not to act on the advice given. We also wished to test whether an existing model of trust in eHealth provided a good fit for COVID-19–related information seeking online. We begin by briefly reviewing the literature on trust and eHealth before introducing the COVID-19 context and outlining the study objectives.

Trust in Online Health Information

Over the last 20 years, research has consistently pointed to the importance of both the design and the content of websites in terms of establishing trustworthiness [6,7]. Commonly reported indicators of trust and credibility include site owners or sponsors; consensus among multiple sources; characteristics of writing and language; advertisements; content authorship; and interface design [8]. Related studies have looked at the quality of web-based health information and have highlighted navigability, aesthetics, and ease of understanding as important factors [9]. As digital resources for health have developed and diversified, we have seen a move away from government and medically driven sources towards more charity and patient-led sites [10] and the use of social media [11,12] meaning that shared patient experience has also become a critical factor in determining trust and appropriateness of online advice [13].

Despite concerns about the quality and reliability of some digital sources [14], they are often well-used and well-liked. Interestingly, they are not necessarily trusted and the advice they contain is not always acted upon. In part, this may relate to a dislike in the United Kingdom for commercial funding models underpinning health websites [10]. A recent model of trust in eHealth [15] found that credibility and impartiality are the key predictors of trust in eHealth websites and noted that websites containing patients' experiences can have a positive impact on trust but only if those sources have been checked against other sources first. The authors also noted that the need to corroborate digital information sources may be reduced in cases where there is strong familiarity with a well-used website.

COVID-19 Context

The COVID-19 pandemic led to a global surge in information seeking online in relation to the spread of the virus, best means of protection, access to health care, local rules and guidance, and, subsequently, information about COVID-19 vaccines, tracing apps and COVID-19 passports [16]. While official sources moved quickly to try and fill these information gaps, social media platforms provided a space for information and misinformation to circulate widely [17]. Conspiracy theories and rumors in relation to the virus and the vaccine were prevalent online as was poor-quality information [18-20]. The unique situation increased attention on governments as a source of information however historically government and official health sources have been subject to mistrust and their health messages resisted especially concerning vaccinations for example in the case of the Measles Mumps Rubella vaccination and the H1N1 (swine flu) vaccination program [21,22]. In these cases, trust in nonofficial information sources and the media is often higher.

United Kingdom Context

In response to the global pandemic, the UK prime minister announced a national lockdown on March 23rd, 2023 [23]. Daily press briefings followed, led by politicians and National Health Service (NHS) leaders providing coordinated information on COVID-19 legislation and guidance, health advice, and subsequently the vaccine rollout.

Survey data indicates there was a slight increase in political trust in the United Kingdom as the lockdown commenced [24] and most people supported the government enforcement of behavior in the early months [5] with positive views on government decision-making related to response transparency. Although people looked to government and health leaders for information and guidance these officials were not immune from criticism. Politicians and advisors often found themselves at the center of news stories that challenged perceptions of trust [24], and of privacy and security, for example in relation to the rollout of contact tracing apps [25] and COVID-19 passports. Low trust in scientists and medics was also associated with COVID-19 vaccine hesitancy [26].

The sudden onset of COVID-19 and its impact not just on UK citizens but worldwide highlighted the public's need for information. Understanding how individuals sought information from digital sources and whether they trusted this information is the focus of this study. Note that this distinct aim is different from many of the studies of information-seeking behavior during

XSL•FO
RenderX

the pandemic that were more focused on the motives that drive online interrogation. Typically, these searches adopted the Risk Information Seeking and Processing model [27] which sees risk information seeking as driven by factors such as information insufficiency, subjective norms, and relevant channel beliefs. Although the Risk Information Seeking and Processing model has been used effectively to model information-seeking behaviors in relation to COVID-19 [28,29] it says relatively little about the extent to which people decide whether to trust the information they are exposed to.

Other studies have examined overall levels of trust in traditional information sources concerning COVID-19 by comparing television, radio, and newspapers with websites [30] but to our knowledge, this is the first study that examines trust and the antecedents of trust in different digital resources in relation to COVID-19. Focusing on the antecedents of trust at this time alongside individuals' behavioral and attitudinal responses to the information they found is key for our future understanding of trusted health communication during health emergencies.

Rationale for This Study

The revised model of trust in eHealth [15] indicates a number of antecedents for trust in online health information and advice and for intention to act on that advice. This study builds upon that work by asking whether existing trust models are a good fit for COVID-19 information-seeking online. The uncertainty provided by the COVID-19 pandemic provides a unique opportunity to examine how people search for, evaluate, and make trust decisions about health information and advice.

The COVID-19 pandemic provides an opportunity to examine in more depth the type of health information seeking that has been taking place. As described previously, people's information needs vary including information on symptoms and symptom management, self-isolation, and vaccination. Vaccination in particular presents a unique opportunity to explore health information seeking within the context of heightened uncertainty and self-reported behavioral outcomes.

It may be that the global nature of the pandemic and people's desire for information exchange fueled social media sources of health information and increased visibility of patient experiences. On the other hand, information corroboration is effortful, and in times of heightened stress and uncertainty, it may not be appropriate or lead to better coping outcomes. Relying on a single source of information may be more straightforward but trust in government or health professionals may impact trust perceptions around such information sources.

Therefore, the study has three aims: (1) to examine whether an existing trust model is a good fit for COVID-19–related information seeking online, (2) to examine differences in affective responses to digital resources about COVID-19 vaccination versus general information about COVID-19, and (3) to examine whether searching had a self-reported impact on vaccination decisions or attitude toward COVID-19 passports.

Methods

Design

A cross-sectional survey was conducted in January 2022. At this time in the United Kingdom, the Omicron variant wave had just peaked, mask use was still advised but no longer compulsory in indoor settings, and self-isolation after a positive test result was still a legal requirement. We collected quantitative data from eHealth users regarding their use of health websites in relation to COVID-19. We used Prolific to recruit a representative UK sample.

Participants

A total of 600 people completed the survey. In total, 525 participants indicated they had looked for COVID-19 information online. Of these 85.3% (448/525) had looked for more general information and advice about COVID-19 while 14.7% (77/525) had looked for information specifically on the vaccine. Full details of participant demographics can be found in Table 1.

Participants were asked whether they had gone online to look for health advice and information about COVID-19. Those answering "yes" were asked to indicate whether they had been searching for general health advice about COVID-19 or whether they had been searching for health advice about COVID-19 vaccinations. Participants then completed a series of questions relating to the last time they searched for health advice about COVID-19 online. Specifically, they were asked to "think about any one digital source that you visited during that search" and to answer the remaining questions with respect to that source. They answered questions relating to the impact of health advice on their coping perceptions and intention to act on the advice, the degree to which they trusted the information and the digital source, their attitude toward COVID passports, for example, the NHS app that shows proof of vaccination and demographic information.

XSL•FO
RenderX

Table 1. Participant demographics (of those who reported looking for COVID-19 information, N=525). All participants were from the United Kingdom.

Characteristics	Values, n (%)
Age group (years)	
18-25	54 (10.3)
26-34	85 (16.2)
35-54	197 (37.5)
55-64	123 (23.4)
65 years or older	66 (12.6)
Sex	
Male	249 (47.4)
Female	273 (52)
Transgender	2 (0.4)
Other	1 (0.2)
Ethnicity	
Caucasian	430 (81.9)
Latino or Hispanic	3 (0.6)
Middle Eastern	5 (1)
African	11 (2.1)
Caribbean	10 (1.9)
South Asian	31 (5.9)
East Asian	11 (2.1)
Mixed	12 (2.3)
Other	7 (1.3)
Prefer not to say	5 (1)
Education level	
Less than secondary school	2 (0.4)
Secondary school	68 (13)
Further education (eg, college, A-level)	177 (33.7)
Bachelor's degree	194 (37)
Postgraduate degree (eg, MSc, PhD, MD)	82 (15.6)
Prefer not to say	2 (0.4)
Employment	
Full time	254 (48.4)
Part time	87 (16.4)
Retired	85 (16.2)
Unemployed	60 (11.4)
Student	29 (5.5)
Prefer not to say	10 (1.9)
Relationship status	
Single	143 (27.2)
Married or civil partnership or cohabiting	333 (63.4)
Divorced	30 (5.6)
Widowed	10 (1.9)
Prefer not to say	9 (1.7)

XSL•FO
RenderX

Measures

Unless stated otherwise, participants answered the following measures on a 5-point Likert scale (1=strongly disagree to 5=strongly agree).

General Web Trust Questionnaire

The general web trust questionnaire contained 36 items from the study by Sillence et al [15] alongside measures of coping, information corroboration, and affective responses also taken from Sillence et al [15]. Specifically, coping was measured by asking participants to respond to the following stem and variables "After I read the information about COVID-19 I felt…" (1) in control and (2) optimistic using a 5-point scale with the labels: 1=less, 2=slightly less, 3=no different, 4=slightly more, and 5=more (Cronbach α=.84.). Additional affective responses, worried, reassured, at risk, confused and anxious were measured using the same format.

Information corroboration with other sources of information was measured with the following 4 items: (1) "I checked other websites," (2) "I checked other sources," (3) I found the advice consistent across other websites or apps, and (4) I found the advice consistent across other sources (Cronbach α=.87).

Impact on vaccination decision was measured using a single item developed for this study: "To what extent did the information and advice you read online impact your decision regarding COVID vaccinations?" Responses were given on a 5-point scale from "1=It did not influence at all" to "5=It influenced to a very large degree."

Attitude toward COVID-19 passports was measured using a single item developed for this study, that is, "I think COVID passports are a good idea" (1=strongly disagree to 5=strongly agree).

Outcome Measures

Trust was measured following Sillence et al [15], using the mean response to the following 2 items: (1) "I trusted the site" and (2) "I felt I could trust the information on the site" (Cronbach α=.95). Intention to act was an outcome measure, assessed with 1 item "I intended to act upon the advice." This item was taken from Sillence et al [15].

Ethical Considerations

The study received full ethical approval from Northumbria University ethics committee (REF:33639). The survey was hosted on Qualtrics and all data was anonymized. The first page provided participants with information detailing the aim, length, data storage, contact details, and withdrawal process of the study. They were then asked to provide informed consent. Participants received £1.25 (€1.49; US $1.66) for taking part in the study and the average completion time was around 7 minutes.

Results

Overview

We first explored the general web trust questionnaire by performing principal component analysis (PCA). We then explored the relationship between the factor structure and outcomes by testing its fit to the sampled data using structural equation modeling (SEM).

Properties of the General Web Trust Questionnaire

The 36 items of the scale were entered into the PCA. All items loaded onto the extracted components but any items with factor loadings lower than 0.30 were suppressed (Table 2). The analysis indicated that 5 components possessed eigenvalues greater than 1 and together explained 68.7% of the variance in keeping with accepted conventions for successful PCA [31]. The Familiarity factor is the weakest of those extracted although it does meet the minimum threshold of comprising three items [32].

XSL•FO
RenderX

Table 2. Factor loadings for each item (factor loadings lower than .30 are suppressed).

Item	Rotation factor loadings				
	Personal experience (PEx)	Credibility and impartiality	Usability	Privacy	Familiarity
The language made it easy to understand	—[a]	—	.69	—	—
It helped me understand the issue better	—	—	.70	—	—
It was easy to use	—	—	.77	—	—
It told me most of what I needed to know	—	—	.59	—	—
The layout was consistent with other digital sources	—	—	.61	—	—
The advice appeared to be prepared by an expert	—	.69	—	—	—
The advice seemed to be offered in my best interests	—	.73	—	—	—
The advice came from a knowledgeable source	—	.73	—	—	—
The advice seemed credible	—	.80	—	—	—
It was owned by a well-known organization	—	—	—	—	.73
It featured familiar logos	—	—	—	—	.78
It had a professional design	—	—	—	—	.64
It had an attractive design	—	—	.47	—	—
It provided reassurances about my privacy	—	—	—	.66	—
It gave the option to post anonymously	—	—	—	.45	—
It gave reassurances about how they used your information	—	—	—	.78	—
It had a privacy policy	—	—	—	.82	—
It explained their use of cookies	—	—	—	.75	—
It contained accounts of other people's experiences	.87	—	—	—	—
There was a chance to share my experiences	.90	—	—	—	—
There were opportunities to interact with other people on the digital source	.87	—	—	—	—
I saw a wide range of experiences rather different to mine	.88	—	—	—	—
It offered powerful accounts of health experiences	.85	—	—	—	—
It felt like the advice was tailored to me personally	.62	—	—	—	—
I was offered the chance to see experiences from people just like me	.91	—	—	—	—
It contained contributions from likeminded people	.92	—	—	—	—
I was able to contribute to content on the digital source	.88	—	—	—	—
The personal accounts on the digital source were written by people similar to me	.91	—	—	—	—
I found personal accounts that reflected my own experience	.92	—	—	—	—
I found personal accounts that were relevant to my condition	.93	—	—	—	—
There were opportunities to gather information from the personal accounts on the digital source	.91	—	—	—	—
The personal accounts contained advice for readers	.91	—	—	—	—
The personal accounts provided social or emotional support	.89	—	—	—	—
The advice appeared to be impartial and independent	—	.78	—	—	—

XSL•FO
RenderX

Item	Rotation factor loadings				
	Personal experience (PEx)	Credibility and impartiality	Usability	Privacy	Familiarity
The advice seemed objective (ie, no hidden agenda)	—	.81	—	—	—
It was free from advertisements		.54	—	—	—
Eigenvalues	11.8	4.7	3.2	3.0	2.1
Variance explained (%)	32.7	13.1	8.9	8.2	5.8

[a]Not available.

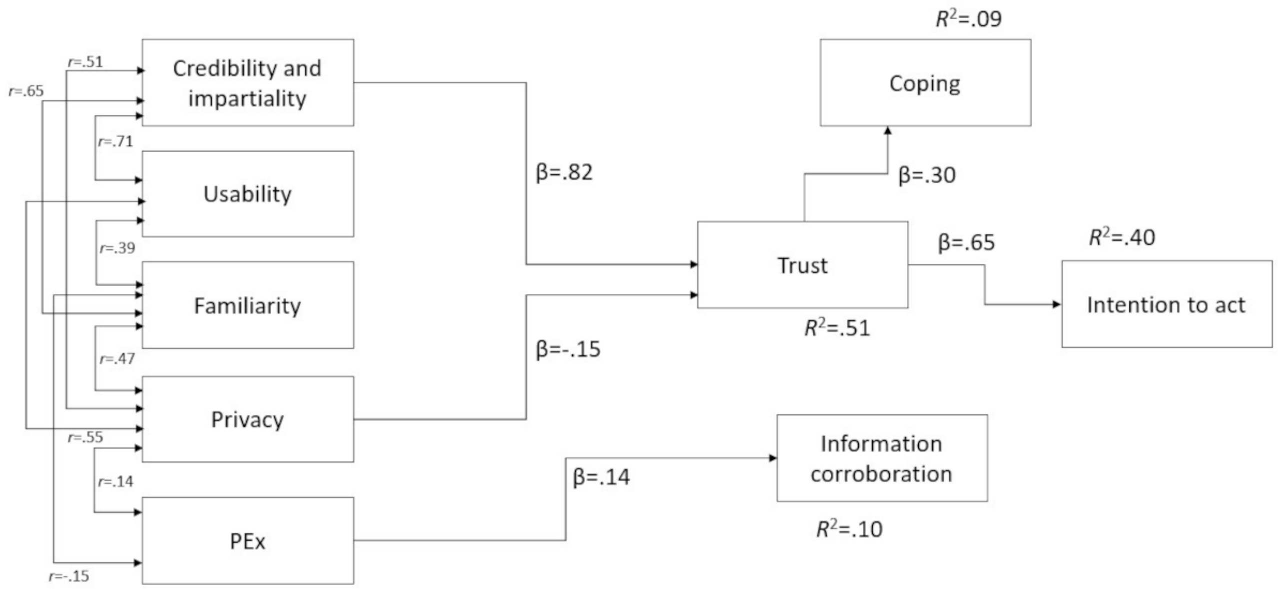
Exploring the Relationship Between the Trust Questionnaire and Self-Reported Behavioral Outcomes

The data were further analyzed using SEM performed in analysis of moment structures using the maximum likelihood estimation method on the item covariance matrix. The specified model was based on Sillence et al [15] and modified to incorporate the new 5-factor structure. The goodness of fit indices supports the specified model. The chi-square value indicated poor fit (χ^2_{773}=1265.5; P<.001). However, this test is considered too sensitive for samples over 200 and here the sample size is 448.

The Cmin/*df* value of 1.64 indicates a good fit. The goodness of fit and adjusted goodness of fit values of .89 and .86 respectively indicate adequate fit [33]. The comparative fit index value of .97 indicates good fit [34], as does the root mean square of approximation value of .04, 90% CI .034-.041 [35].

The model accounted for 64.7% of the variance in trust, 8.7% in coping, 9.7% in information corroboration, and 40.3% in intention to act. All beta path coefficients including those in Figure 1 and those that were not significant were inspected in evaluating the predictive power of the model and are presented for completeness in Table 3.

Figure 1. The trust model with significant standardized path coefficients.



XSL•FO
RenderX

Table 3. The unstandardized path weights and critical ratio (ie, *z* score) values for the main effects of the hypothesized full model.

Parameter	Unstandardized path coefficient	Critical ratio	*P* value
Credibility and impartiality			
Trust	.93	9.79	<.001
Information corroboration	.17	1.07	.29
Usability			
Trust	–.05	–.36	.72
Information corroboration	.39	1.56	.12
Familiarity			
Trust	–.04	–.64	.52
Information corroboration	.12	.98	.33
Privacy			
Trust	–.19	–2.43	.02
Information corroboration	.06	.41	.68
Personal experience			
Trust	–.001	–.03	.98
Information corroboration	.09	2.78	.01
Trust			
Coping	.27	4.89	<.001
Intention to act	.80	15.23	<.001
Coping–intention to act	–.04	–.67	.50
Information corroboration			
Trust	.001	.03	.98
Intention to act	–.02	–.61	.54

Only Credibility and Impartiality were found to possess a significant positive path to Trust. Privacy had a weaker yet significant negative path, meaning privacy assurances were associated with lower trust. Familiarity, usability, and personal experience (PEx) were not significantly predictive of Trust. Only Trust was found to significantly predict the intention to act on the advice. In addition, Trust significantly predicted Coping, suggesting that trustworthy websites heighten individuals' coping perceptions, making them feel more in control and optimistic. PEx significantly predicts Information Corroboration, suggesting that people are exploring a little further than the original digital source; however, this corroboration process does not appear to be affecting their level of trust or intention to act.

Comparison of Two Populations

Although the relatively small sample size for the vaccine information group meant that a comparable SEM model could not be reliably tested a series of independent samples *t* tests were used to compare the two groups on the key variables of interest (Tables 4 and 5).

Table 4. Mean (SD) values for key outcome variables.

Group	Trust	Intention to act	Corroboration	Impact on the decision regarding vaccination	Attitude toward COVID-19 passports
Searching for information on vaccinations (N=77)	4.22 (.91)	4.10 (1.05)	3.49 (1.24)	2.90 (1.21)	3.38 (1.51)
Searching for information on COVID-19 (N=448)	4.33 (.74)	4.13 (.89)	3.49 (1.06)	2.74 (1.39)	3.51 (1.36)

XSL•FO
RenderX

Table 5. Mean (SD) values for "after I read the information" variables.

Group	Worried	Reassured	At risk	Confused	Anxious	Optimistic	In control
Searching for information on vaccinations (N=77)	2.27 (1.11)	3.84 (.95)	2.40 (.98)	2.14 (1.13)	2.42 (1.20)	3.66 (1.11)	3.57 (1.13)
Searching for information on COVID-19 (N=448)	2.48 (.88)	3.68 (.77)	2.84 (.88)	2.15 (.98)	2.76 (.97)	3.27 (.81)	3.42 (.85)

Independent Sample *t* tests

There was no significant difference between groups for trust (t_{523}=–1.169; *P*=.24; Cohen *d*=–.14, 95% CI –.386 to .098), intention to act (t_{523}=–.187; *P*=.85; Cohen *d*=–.02, 95% CI –.265 to .219), corroboration (t_{523}=–.038; *P*=.97; Cohen *d*=–.01, 95% CI –.247 to .237), impact on decision regarding vaccination (t_{523}=.934; *P*=.35; Cohen *d*=.115, 95% CI –.127 to .357), or COVID-19 passports (t_{523}=–.773; *P*=.44; Cohen *d*=–.095, 95% CI –.337 to .146).

Those searching for information on vaccinations (mean 2.40) felt significantly less at risk than those searching for general information on COVID-19 (mean 2.84; t_{523}=3.988; *P*<.001; Cohen *d*=–49, 95% CI –.735 to –.2348) and felt significantly less anxious (mean 2.42) than those searching for general information on COVID-19 (mean 2.76; t_{523}=–2.758; *P*=.003; Cohen *d*=–.34, 95% CI –.583 to –.097). Those searching for information on vaccinations (mean=3.66) felt significantly more optimistic than those searching for general information on COVID-19 (mean=3.27; t_{523}=3.760; *P*<.001; Cohen *d*=.464, 95% CI .220-.707).

There was no significant difference for the variable "In Control" (t_{523}=1.335; *P*=.18; Cohen *d*=–.165, 95% CI –.077 to .407) or for "Confused" (t_{523}=–.054; *P*=.96; Cohen *d*=–.007, 95% CI –.248 to .235). Finally, the variables "Worried" and "Reassured" approached but did not reach statistical significance (t_{523}=–1.813; *P*=.07; Cohen *d*=–.224, 95% CI –.466 to .019 and t_{523}=1.712; *P*=.09; Cohen *d*=.211, 95% CI –.031 to .453, respectively).

Digital Sources of Information

Table 6 shows the digital sources used. The majority of participants used either the NHS health care sources or the governmental sources for both general information and vaccine-specific information.

Digital sources were categorized as: (1) Governmental sources: official UK government website (Gov.uk), World Health Organization, Office of National Statistics, and Centre for Disease Control. (2) NHS health care sources: any page hosted on the NHS website (nhs.uk). (3) Other health care sources: any non-NHS health care website. This included The Mayo Clinic, WedMD, patient.co.uk, and the Health Check podcast. (4) News websites: any of the mainstream news providers, the majority of those reported were the BBC. (5) Search engines: where participants did not go to one source but reported explicitly using search engines, such as Google, to intentionally search for COVID-19–related information, rather than, for example, visiting a particular source (perhaps a source perceived as authoritative or trusted), such as the NHS, government, or BBC websites, and browsing the content from there. (6) Scientific journal: any peer-reviewed journal publishing academic research. (7) Specific health condition websites: any website dedicated to a specified health condition rather than a general health website, including asthma.org and Crohn's & Colitis UK. (8) Social media and forums: any online forum or social networking platform defined as user-driven and facilitating sharing of content, dialogue creation, and communication by and between users (in keeping with Kapoor et al, 2018 [36]). (9) Other: all instances where resources were not explicitly specified or where participants reported visiting multiple sources. All other resources are named individually in Table 6.

Table 6. Digital sources used.

Source	General information (N=448), n (%)	Vaccine specific information (N=77), n (%)
National Health Service health care sources	262 (58.48)	39 (50.65)
Governmental sources	64 (14.30)	11 (14.29)
Multiple resources or unspecific	37 (8.30)	13 (16.88)
News websites	30 (6.70)	3 (3.90)
Other health care sources	6 (1.34)	1 (1.30)
Social media and forums	20 (4.46)	2 (2.60)
Search engines	19 (4.24)	7 (9.09)
Zoe COVID-19 study	6 (1.34)	0 (0)
Scientific journals	1 (0.22)	0 (0)
Specific health condition websites	2 (0.45)	0 (0)
Wikipedia	1 (0.22)	0 (0)
TripAdvisor	0 (0)	1 (1.30)

XSL•FO
RenderX

Discussion

Principal Results

Trust continues to significantly influence self-reported intention to act on health information. In terms of trust predictors, only credibility and impartiality have a significant, direct, and positive relationship with trust. Privacy has a significant negative relationship with trust. Other predictors (familiarity, usability, and PEx) may be indirect and mediated through other trust variables. The variable PEx had a significant direct effect on information corroboration and trust was found to significantly relate to coping perceptions. The findings suggest a number of important discussion points.

First, the Sillence et al [15] trust model provides a reasonable fit for COVID-19–related health information online. Trust continues to predict intention and the credibility and impartiality of the digital source remains the strongest predictor of trust in digital health sources. However, compared to the 2019 model, the picture here is of a simpler trust process in which the credibility and impartiality factor does the “heavy lifting” in relation to trust compared to the other variables. Another key difference is the lack of a relationship between corroboration and trust. In earlier models, health information seekers looked to verify the information they found online by cross-checking with other digital and nondigital sources. Here we see only a direct relationship between the credibility and impartiality of the website and trust. One reason for this, given the predominance of the NHS as the most popular site for information and advice, is that our health information seekers are simply taking the website at face value providing it appears sufficiently credible and impartial. However, it is interesting that in an American sample, information seekers relied heavily upon often unreliable social media sources for information and advice, yet still engaged in relatively low levels of fact-checking [37] and so we must consider the possibility that people are being bombarded with so much information in relation to the pandemic that they simply switch off.

The role of PEx within digital sources is interesting here. While PEx significantly predicts information corroboration there was no subsequent relationship with trust. In the 2019 model [15] it was suggested that patient experiences can positively influence trust but only if users first corroborate the information through other sources. In our study, we suggest that people are checking up on these patient stories and experiences simply out of interest rather than as a way of assessing the trustworthiness of the information. When faced with a high degree of uncertainty and with limited detailed information, assessments of risk may be emotion-based [38], and people may well seek out other people’s personal accounts of their COVID-19 experiences. Personal accounts are often engaging and are seen as more relatable than statistical information when it comes to decision-making [39]. While PEx is now embedded within a diverse range of digital resources, those more closely associated with personal content, for example, social media platforms or individual blogs, were generally underrepresented in the data we collected. Instead, we observed a reliance on official digital sources, in particular, the NHS website and government sources. In terms of pandemic or emergency, reliance on official sources may be more commonplace. Sillence et al [15] found that the majority of UK respondents cited the NHS website as their source of health information, and McNeill, Harris, and Briggs [40] noted that the main UK source to be retweeted during the H1N1 pandemic was NHS Choices. In this study, there was little reported use of social media, which is perhaps surprising and contrasts with other recent health pandemics in which social media use and misinformation have been prevalent [37,41,42] as well as in earlier studies examining the COVID-19 pandemic and the facilitation of conspiracy theories [43,44].

Despite generally high levels of mistrust in the government’s overall handling of the pandemic [5], UK citizens still sought information from government sites. Moreover, we see a reliance on health professionals and public health information. In a time of limited information, there may be fewer options available to information seekers and individuals may be satisfied with seeking official sources of information even if they contain basic knowledge as opposed to more detailed, specific information. This contrasts with earlier work on trust in digital health information in which personalization or tailoring is seen as important to trust. People with long-term experience of a particular health condition often become experts by experience and may seek more specific, tailored digital resources to support their health conditions. This involves making more fine-grained assessments of the personal relevance of the information before deciding to trust or act upon the advice it contains [10,45] and is especially true where the condition is rare or less well known [46]. In the case of COVID-19, a worldwide pandemic affecting all age groups, it might be that generic information applicable to all sufficed in this case. There was little sense that people were checking COVID-19 information in relation to their other, pre-existing health conditions and specific health websites may not have had that information readily available. In light of research that shows how health information overload may lead to increased anxiety [47], our participants’ reliance on relatively few, authoritative websites seems like a reasonable strategy. Too much, possibly conflicting, information about COVID-19 can leave an individual feeling overwhelmed and will ultimately lead to “information avoidance,” which is clearly a poor outcome in the face of a global pandemic.

Unlike Sillence et al’s [15] 2019 model, we note that privacy has a weak negative relationship with trust. The topic of privacy was raised repeatedly in relation to the discussion of contact tracing apps and COVID-19 passports and so while not directly related to the digital source being used it may be that being asked to think about the privacy features of sources stimulates a wider consideration of privacy and mistrust. Rather than privacy policies etc. being seen as an example of good practice, the very fact that these options were present on digital sources may have served as a reminder that data are being collected, processed, and often shared. Privacy nudges may well remind people of the need to be mindful of privacy but can also raise awareness of the data that is available for collection [48,49].

Second, trust significantly predicted coping suggesting that trustworthy websites heighten individuals’ coping perceptions, making them feel able to cope. Interestingly, Wang et al [1] did not find an association between the use of the internet as an

XSL•FO
RenderX

information source on COVID-19 and self-confidence in coping with COVID-19 but did not focus on trusted websites.

Looking at the affective variables in more detail for the two groups (general information seeking and vaccination information), we see that those searching for vaccination information felt more positive—specifically, they felt less at risk, less anxious, and more optimistic after reading the information. Wang et al [1] found that vaccination information sources have different effects on students' coping appraisal of COVID-19 with information from medical personnel leading to greater knowledge about the mechanism of vaccination and greater response efficacy of vaccination compared to information from coworkers or colleagues. In terms of coping, during the H1N1 pandemic, those people who adopted a more problem-focused coping strategy including seeking out information to help solve problems were more likely to indicate they would be vaccinated [22]. In our data, those individuals who have gone looking for information about vaccination feel better for having done so.

Zheng et al [50] noted that vaccine information seeking is related to vaccination intention and suggested that health information seeking can be viewed as a coping behavior when people do not have sufficient knowledge of a particular health topic. Although seeking vaccine-related information online was also positively related to perceived vaccine information overload [50], it may be that sticking with a single trusted source is preferable for improved coping. Finally, there were no differences in terms of trust, intention to act on information, or attitude toward COVID-19 passports between participants who were searching for general COVID-19 health information versus those who had searched for vaccination information. This is unsurprising given the similarity of digital sources used.

In summary, people searching for general COVID-19 information as well as those searching for COVID-19 vaccine-specific information sought out official sources of information online. In terms of uncertainty when faced with a global emergent health concern people place their trust in familiar websites and rely on the perceived credibility and impartiality of those digital sources.

Limitations

It is important to note that data was purposely not collected during a period of national lockdown in the United Kingdom. The vaccination program was already well underway and COVID-19 passports were very much still on the agenda. People may have sought information from alternative digital sources had data collection taken place during a period of lockdown. Focusing on the United Kingdom made sense given the local regulations and practices in place, but it would be interesting to make comparisons with other countries going forward. The reliance on the NHS website in the United Kingdom would be interesting to compare with countries where different funding models exist for example where health insurance schemes mean there is no single free at the point of service system. Vaccine hesitancy is relatively low in the U and has declined since the start of the vaccination rollout program from 10% to 3% in September 2021 [51]. Other countries, for example, France, have much higher levels of vaccine hesitancy [52], and comparisons here in relation to trust around digital health resources would warrant further investigation. Finally, it is interesting to note that although we have used a one-shot cross-sectional methodology, we mirror findings from Zhang et al [53], who examined trust over several waves earlier in the pandemic and noted a decrease in the use of social media over time and an increase in trust in government information.

Conclusion

In conclusion, in the context of COVID-19, "credibility and impartiality" remain a key predictor of trust in eHealth resources but in comparison with previous models of trust in online health information, checking and corroborating information did not form a significant part of trust evaluations. In times of uncertainty when faced with a global emergent health concern, people placed their trust in familiar websites and relied on the perceived credibility and impartiality of those digital sources.

Acknowledgments

This project contributes towards the EPSRC-funded Centre for Digital Citizens (grant number EP/T022582/1).

Data Availability

The datasets generated and analyzed during this study are available in the Open Science Framework repository [54].

Conflicts of Interest

None declared.

References

1. Wang PW, Lu WH, Ko NY, Chen YL, Li DJ, Chang YP, et al. COVID-19-related information sources and the relationship with confidence in people coping with COVID-19: facebook survey study in Taiwan. J Med Internet Res 2020;22(6):e20021 [FREE Full text] [doi: 10.2196/20021] [Medline: 32490839]
2. Finset A, Bosworth H, Butow P, Gulbrandsen P, Hulsman RL, Pieterse AH, et al. Effective health communication - a key factor in fighting the COVID-19 pandemic. Patient Educ Couns 2020;103(5):873-876 [FREE Full text] [doi: 10.1016/j.pec.2020.03.027] [Medline: 32336348]

XSL•FO
RenderX

3. Ho HY, Chen YL, Yen CF. Different impacts of COVID-19-related information sources on public worry: an online survey through social media. Internet Interv 2020;22:100350 [FREE Full text] [doi: 10.1016/j.invent.2020.100350] [Medline: 33072524]
4. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. J Public Health Policy 2020;41(4):410-420 [FREE Full text] [doi: 10.1057/s41271-020-00247-w] [Medline: 32826935]
5. Enria L, Waterlow N, Rogers NT, Brindle H, Lal S, Eggo RM, et al. Trust and transparency in times of crisis: results from an online survey during the first wave (April 2020) of the COVID-19 epidemic in the UK. PLoS One 2021;16(2):e0239247 [FREE Full text] [doi: 10.1371/journal.pone.0239247] [Medline: 33591985]
6. Kim Y. Trust in health information websites: a systematic literature review on the antecedents of trust. Health Informatics J 2016;22(2):355-369 [FREE Full text] [doi: 10.1177/1460458214559432] [Medline: 25518944]
7. Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. J Med Internet Res 2017;19(6):e218 [FREE Full text] [doi: 10.2196/jmir.7579] [Medline: 28630033]
8. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer evaluation of the quality of online health information: systematic literature review of relevant criteria and indicators. J Med Internet Res 2019;21(5):e12522 [FREE Full text] [doi: 10.2196/12522] [Medline: 31045507]
9. Zhang Y, Kim Y. Consumers' evaluation of web-based health information quality: meta-analysis. J Med Internet Res 2022;24(4):e36463. [doi: 10.2196/36463] [Medline: 35482390]
10. Sillence E, Hardy C, Briggs P. Why don't we trust health websites that help us help each other? An analysis of online peer-to-peer healthcare. 2013 Presented at: WebSci '13: Proceedings of the 5th Annual ACM Web Science Conference; 2013, May 2-4; Paris, France p. 396-404. [doi: 10.1145/2464464.2464488]
11. O'Leary K, Coulson N, Perez-Vallejos E, McAuley D. Towards understanding how individuals with inflammatory bowel disease use contemporary social media platforms for health-related discourse. Computers in Human Behavior 2020;112:106463. [doi: 10.1016/j.chb.2020.106463]
12. Sasse M, Ohrndorf S, Palmowski A, Wagner AD, Burmester GR, Pankow A, et al. Digital health information on autoinflammatory diseases: a YouTube quality analysis. Rheumatol Int 2023;43(1):163-171 [FREE Full text] [doi: 10.1007/s00296-022-05243-9] [Medline: 36374326]
13. Sillence E. Having faith in the online voice. Exploring contemporary issues of trust, language and advice in the context of e-health. LO 2017;87(8). [doi: 10.13092/lo.87.4175]
14. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. BMC Med Educ 2022;22(1):382 [FREE Full text] [doi: 10.1186/s12909-022-03446-z] [Medline: 35590410]
15. Sillence E, Blythe JM, Briggs P, Moss M. A revised model of trust in internet-based health information and advice: cross-sectional questionnaire study. J Med Internet Res 2019;21(11):e11125 [FREE Full text] [doi: 10.2196/11125] [Medline: 31710297]
16. Zhao X, Fan J, Basnyat I, Hu B. Online health information seeking using '#COVID-19 Patient Seeking Help' on Weibo in Wuhan, China: descriptive study. J Med Internet Res 2020;22(10):e22910 [FREE Full text] [doi: 10.2196/22910] [Medline: 33001838]
17. Van den Broucke S. Why health promotion matters to the COVID-19 pandemic, and vice versa. Health Promot Int 2020;35(2):181-186 [FREE Full text] [doi: 10.1093/heapro/daaa042] [Medline: 32297931]
18. Zeng J, Schäfer MS. Conceptualizing 'Dark Platforms'. Covid-19-related conspiracy theories on 8kun and Gab. Digital Journalism 2021;9(1):1-23. [doi: 10.1080/21670811.2021.1938165]
19. Islam MS, Kamal AM, Kabir A, Southern DL, Khan SH, Hasan SMM, et al. COVID-19 vaccine rumors and conspiracy theories: the need for cognitive inoculation against misinformation to improve vaccine adherence. PLoS One 2021;16(5):e0251605 [FREE Full text] [doi: 10.1371/journal.pone.0251605] [Medline: 33979412]
20. Chan C, Sounderajah V, Daniels E, Acharya A, Clarke J, Yalamanchili S, et al. The reliability and quality of YouTube videos as a source of public health information regarding COVID-19 vaccination: cross-sectional study. JMIR Public Health Surveill 2021;7(7):e29942 [FREE Full text] [doi: 10.2196/29942] [Medline: 34081599]
21. Brown KF, Kroll JS, Hudson MJ, Ramsay M, Green J, Long SJ, et al. Factors underlying parental decisions about combination childhood vaccinations including MMR: a systematic review. Vaccine 2010;28(26):4235-4248. [doi: 10.1016/j.vaccine.2010.04.052] [Medline: 20438879]
22. Taha SA, Matheson K, Anisman H. The 2009 H1N1 influenza pandemic: the role of threat, coping, and media trust on vaccination intentions in Canada. J Health Commun 2013;18(3):278-290. [doi: 10.1080/10810730.2012.727960] [Medline: 23301849]
23. Prime Minister's Statement on Coronavirus (COVID-19): 23 March 2020. GOV.UK. URL: https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020 [accessed 2025-10-15]
24. Davies B, Lalot F, Peitz L, Heering M, Cifterler HO, Babaian J, et al. Changes in political trust in Britain during the COVID-19 pandemic in 2020: integrated public opinion evidence and implications. Humanit Soc Sci Commun 2021;8(166):1-9. [doi: 10.1057/s41599-021-00850-6]

XSL•FO
RenderX

25. Samuel G, Roberts SL, Fiske A, Lucivero F, McLennan S, Phillips A, et al. COVID-19 contact tracing apps: UK public perceptions. Crit Public Health 2022;32(1):31-43 [FREE Full text] [doi: 10.1080/09581596.2021.1909707] [Medline: 35221546]
26. Allington D, McAndrew S, Moxham-Hall V, Duffy B. Coronavirus conspiracy suspicions, general vaccine attitudes, trust and coronavirus information source as predictors of vaccine hesitancy among UK residents during the COVID-19 pandemic. Psychol Med 2023;53(1):236-247 [FREE Full text] [doi: 10.1017/S0033291721001434] [Medline: 33843509]
27. Griffin RJ, Dunwoody S, Neuwirth K. Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. Environ Res 1999;80(2 Pt 2):S230-S245. [doi: 10.1006/enrs.1998.3940] [Medline: 10092438]
28. Zhou X, Roberto AJ. An application of the risk information seeking and processing model in understanding college students' COVID-19 vaccination information seeking and behavior. Sci Commun 2022;44(4):446-474 [FREE Full text] [doi: 10.1177/10755470221120415] [Medline: 38603181]
29. Ahn J, Kim HK, Kahlor LA, Atkinson L, Noh GY. The impact of emotion and government trust on individuals' risk information seeking and avoidance during the COVID-19 pandemic: a cross-country comparison. J Health Commun 2021;26(10):728-741. [doi: 10.1080/10810730.2021.1999348] [Medline: 34779340]
30. Ali SH, Foreman J, Tozan Y, Capasso A, Jones AM, DiClemente RJ. Trends and predictors of COVID-19 information sources and their relationship with knowledge and beliefs related to the pandemic: nationwide cross-sectional study. JMIR Public Health Surveill 2020;6(4):e21071 [FREE Full text] [doi: 10.2196/21071] [Medline: 32936775]
31. Hinkin TR. A brief tutorial on the development of measures for use in survey questionnaires. Organ Res Methods 1998;1(1):104-121. [doi: 10.1177/109442819800100106]
32. Yong AG, Pearce S. A beginner’s guide to factor analysis: focusing on exploratory factor analysis. TQMP 2013;9(2):79-94. [doi: 10.20982/tqmp.09.2.p079]
33. Marsh H, Grayson D. Latent variable models of multitrait-multimethod data. In: Structural Equation Modeling: Concepts, Issues, and Applications. Thousand Oaks: Sage; 1995:177-198.
34. Garson GD. Statnotes: Topics in Multivariate Analysis. 2009. URL: https://faculty.chass.ncsu.edu/garson/PA765/statnote.htm [accessed 2025-01-08]
35. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociological Methods & Research 1992;21(2):230-258. [doi: 10.1177/0049124192021002005]
36. Kapoor KK, Tamilmani K, Rana NP, Patil P, Dwivedi YK, Nerur S. Advances in social media research: past, present and future. Inf Syst Front 2017;20(3):531-558. [doi: 10.1007/s10796-017-9810-y]
37. Neely S, Eldredge C, Sanders R. Health information seeking behaviors on social media during the COVID-19 pandemic among American social networking site users: survey study. J Med Internet Res 2021;23(6):e29802 [FREE Full text] [doi: 10.2196/29802] [Medline: 34043526]
38. Leppin A, Aro AR. Risk perceptions related to SARS and avian influenza: theoretical foundations of current empirical research. Int J Behav Med 2009;16(1):7-29 [FREE Full text] [doi: 10.1007/s12529-008-9002-8] [Medline: 19214752]
39. Winterbottom A, Bekker HL, Conner M, Mooney A. Does narrative information bias individual's decision making? A systematic review. Soc Sci Med 2008;67(12):2079-2088. [doi: 10.1016/j.socscimed.2008.09.037] [Medline: 18951673]
40. McNeill A, Harris PR, Briggs P. Twitter influence on UK vaccination and antiviral uptake during the 2009 H1N1 pandemic. Front Public Health 2016;4:26 [FREE Full text] [doi: 10.3389/fpubh.2016.00026] [Medline: 26942174]
41. Dredze M, Broniatowski DA, Hilyard KM. Zika vaccine misconceptions: a social media analysis. Vaccine 2016;34(30):3441-3442 [FREE Full text] [doi: 10.1016/j.vaccine.2016.05.008] [Medline: 27216759]
42. Sharma M, Yadav K, Yadav N, Ferdinand KC. Zika virus pandemic-analysis of Facebook as a social media health information platform. Am J Infect Control 2017;45(3):301-302. [doi: 10.1016/j.ajic.2016.08.022] [Medline: 27776823]
43. Jennings W, Stoker G, Bunting H, Valgarðsson VO, Gaskell J, Devine D, et al. Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy. Vaccines (Basel) 2021;9(6):593 [FREE Full text] [doi: 10.3390/vaccines9060593] [Medline: 34204971]
44. Theocharis Y, Cardenal A, Jin S, Aalberg T, Hopmann DN, Strömbäck J, et al. Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. New Media & Society 2021;25(12):3412-3437. [doi: 10.1177/14614448211045666]
45. Entwistle VA, France EF, Wyke S, Jepson R, Hunt K, Ziebland S, et al. How information about other people's personal experiences can help with healthcare decision-making: a qualitative study. Patient Educ Couns 2011;85(3):e291-e298. [doi: 10.1016/j.pec.2011.05.014] [Medline: 21652162]
46. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007;68(2):173-178 [FREE Full text] [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]
47. Sultana T, Dhillon G, Oliveira T. The effect of fear and situational motivation on online information avoidance: the case of COVID-19. Int J Inf Manage 2023;69:102596 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2022.102596] [Medline: 36415624]

XSL•FO
RenderX

48. Zhang B, Xu H. Privacy nudges for mobile applications: Effects on the creepiness emotion and privacy attitudes. : ACM; 2016 Presented at: CSCW '16: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; 2016 27 February - 2 March; San Francisco California USA p. 1676-1690. [doi: 10.1145/2818048.2820073]
49. Wang Y, Leon P, Acquisti A, Cranor LF, Forget A, Sadeh N. A field trial of privacy nudges for Facebook. 2014 Presented at: CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2014 April 26; Toronto Ontario Canada p. 2367-2376. [doi: 10.1145/2556288.2557413]
50. Zheng H, Jiang S, Rosenthal S. Linking online vaccine information seeking to vaccination intention in the context of the COVID-19 pandemic. Science Communication 2022;44(3):320-346. [doi: 10.1177/10755470221101067]
51. Coronavirus and changing attitudes towards vaccination, England: 7 to 16 September 2021. Office for National Statistics. URL: https://tinyurl.com/bp8aczwh [accessed 2025-01-09]
52. Robinson E, Jones A, Lesser I, Daly M. International estimates of intended uptake and refusal of COVID-19 vaccines: a rapid systematic review and meta-analysis of large nationally representative samples. Vaccine 2021;39(15):2024-2034 [FREE Full text] [doi: 10.1016/j.vaccine.2021.02.005] [Medline: 33722411]
53. Zhang Y, Suhaimi N, Yongsatianchot N, Gaggiano JD, Kim M, Patel SA, et al. Shifting trust: examining how trust and distrust emerge, transform, and collapse in COVID-19 information seeking. : ACM; 2022 Presented at: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems; 2022 April 29; New Orleans LA USA p. 1-21. [doi: 10.1145/3491102.3501889]
54. Open software foundation. URL: https://osf.io/b4a9m/ [accessed 2025-01-24]

Abbreviations

NHS: National Health Service
PCA: principal component analysis
PEx: personal experience
SEM: structural equation modeling

Edited by R Cuomo; submitted 09.04.24; peer-reviewed by A Frik, J Rowley, S Pesälä; comments to author 06.09.24; revised version received 27.10.24; accepted 21.11.24; published 13.02.25.

Please cite as:
Sillence E, Branley-Bell D, Moss M, Briggs P
A Model of Trust in Online COVID-19 Information and Advice: Cross-Sectional Questionnaire Study
JMIR Infodemiology 2025;5:e59317
URL: https://infodemiology.jmir.org/2025/1/e59317
doi:10.2196/59317
PMID:

©Elizabeth Sillence, Dawn Branley-Bell, Mark Moss, Pam Briggs. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 13.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Transformer-Based Tool for Automated Fact-Checking of Online Health Information: Development Study

Azadeh Bayani[1,2], MSc; Alexandre Ayotte[1,2*], MSc; Jean Noel Nikiema[1,2,3*], PhD

[1]Laboratoire Transformation Numérique en Santé, LabTNS, Montreal, QC, Canada

[2]Centre de recherche en santé publique, Université de Montréal et CIUSSS du Centre-Sud-de-l'Île-de-Montréal, Montreal, QC, Canada

[3]Department of Management, Evaluation and Health Policy, School of Public Health, Université de Montréal, Montreal, QC, Canada

[*]these authors contributed equally

Corresponding Author:
Azadeh Bayani, MSc
Laboratoire Transformation Numérique en Santé, LabTNS
7101 Av. du Parc, Montréal,
Montreal, QC, H3N 1X9
Canada
Phone: 1 4389980241
Email: azadeh.bayani@umontreal.ca

Abstract

Background: Many people seek health-related information online. The significance of reliable information became particularly evident due to the potential dangers of misinformation. Therefore, discerning true and reliable information from false information has become increasingly challenging.

Objective: This study aimed to present a pilot study in which we introduced a novel approach to automate the fact-checking process, leveraging PubMed resources as a source of truth using natural language processing transformer models to enhance the process.

Methods: A total of 538 health-related web pages, covering 7 different disease subjects, were manually selected by Factually Health Company. The process included the following steps: (1) using transformer models of bidirectional encoder representations from transformers (BERT), BioBERT, and SciBERT, and traditional models of random forests and support vector machines, to classify the contents of web pages into 3 thematic categories (semiology, epidemiology, and management), (2) for each category in the web pages, a PubMed query was automatically produced using a combination of the "WellcomeBertMesh" and "KeyBERT" models, (3) top 20 related literatures were automatically extracted from PubMed, and finally, (4) the similarity checking techniques of cosine similarity and Jaccard distance were applied to compare the content of extracted literature and web pages.

Results: The BERT model for the categorization of web page contents had good performance, with F_1-scores and recall of 93% and 94% for semiology and epidemiology, respectively, and 96% for both the recall and F_1-score for management. For each of the 3 categories in a web page, 1 PubMed query was generated and with each query, the 20 most related, open access articles within the category of systematic reviews and meta-analyses were extracted. Less than 10% of the extracted literature was irrelevant; those were deleted. For each web page, an average of 23% of the sentences were found to be very similar to the literature. Moreover, during the evaluation, it was found that cosine similarity outperformed the Jaccard distance measure when comparing the similarity between sentences from web pages and academic papers vectorized by BERT. However, there was a significant issue with false positives in the retrieved sentences when compared with accurate similarities, as some sentences had a similarity score exceeding 80%, but they could not be considered similar sentences.

Conclusions: In this pilot study, we have proposed an approach to automate the fact-checking of health-related online information. Incorporating content from PubMed or other scientific article databases as trustworthy resources can automate the discovery of similarly credible information in the health domain.

(JMIR Infodemiology 2025;5:e56831) doi:10.2196/56831

KEYWORDS

fact-checking automation; transformers; infodemic; credible health information; machine learning; automated; online health information; misinformation; natural language processing; epidemiology; health domain

XSL•FO
RenderX

Introduction

With rapid progressions in the digital age, and the vast dissemination of textual information available online, the likelihood of coming across misinformation has surged [1,2]. Misinformation refers to information that is untrue, incorrect, or deceptive in nature [3]. It is prevalent across various domains, with social media being a particularly prominent source [4]. Indeed, many people seek health-related topics on modern platforms and websites available online [5]. Inaccurate health-related information, however, poses an even greater risk, as it can directly impact lives [6,7]. Health misinformation is considered "a health-related claim or information which is not correct due to a lack of scientific evidence or knowledge" [4,8]. The importance of trustworthy online health information became particularly clear during the COVID-19 pandemic, which triggered a new crisis known as the COVID-19 infodemic. An infodemic refers to the excessive spread of false or misleading information across both digital and physical spaces [9] causing confusion and detrimental outcomes, as it underscores the potential risks posed by inaccurate or deceptive information to individuals [3,10]. The infodemic often manifests across 4 key areas: scientific research, policy and health care practice, news outlets, and social media platforms [11]. As a result, distinguishing between true and reliable information and falsehoods has become increasingly challenging. The labor-intensive process of manually verifying information specifically in health-related fields demands expert oversight and consumes significant time [4,9,12]. Therefore, it is crucial to establish an automated fact-checking process to help users identify the accuracy of health-related information available online.

The fact-checking process involves evaluating the truthfulness of information and consists of 3 key tasks: claim detection, evidence retrieval, and claim verification [12]. The first 2 tasks can be considered as factual verification, while the third focuses on assessing the accuracy of claims, which involves distinguishing reliable information from falsehoods to establish their factual validity [13].

Several studies have explored automating the fact-checking process, primarily focusing on misinformation in the form of fake news on websites [4,14,15] or social media [2,7,16-18]. These studies have generated synthetic datasets as the gold standard to facilitate the automation of evidence-based fact-checking. Thus, they compiled datasets comprising information or claims along with their corresponding evidence from trusted sources. Models were then trained using these datasets to automate the fact-checking process [7,10,15,17-20]. To create a database of verified claims, they used methods such as modifying phrases from Wikipedia [20], manual selection of quotation sentences and handpicking of claims from health news sites [14,15,21], and automatic selection of verified claims that were manually done by experts of journalists from fact-checking websites [10]. For example, the FEVER dataset, generated by modifying sentences taken from Wikipedia, consisted of 185,400 claims [22]. PUBHEALTH is another dataset containing false, true, unproven, and a mixture of health-related claims. The dataset also had a column containing journalist-crafted, gold-standard explanations designed to substantiate the fact-check labels assigned to each claim [6,18]. While synthetic datasets provide valuable contributions to advancing automatic fact-checking efforts, they cannot fully address real-world challenges, particularly the need for real-time, dynamic information [23]. Therefore, there is a need that claims and their associated evidence to be automatically extracted [24]. A study [25] developed a Large Language Model called TrumorGPT, which addresses limitations in fact-checking by incorporating retrieval-augmented generation and using continually updated knowledge graphs. This approach uses few-shot learning, knowledge graph construction, and semantic reasoning, which enhances the model's ability to handle fact-checking tasks effectively. Another recent survey [12] explored automated techniques for predicting the veracity of claims, relying on natural language processing, knowledge representation, and databases. This study identified common challenges in fact-checking research and emphasized the importance of information retrieval and knowledge representation, particularly due to the rapid emergence of new claims.

Therefore, a key element of fact-checking involves identifying credible sources, and for health information, leveraging up-to-date scientific literature is essential as it is widely regarded as 1 of the most trustworthy references [26]. Indeed, numerous platforms and databases provide access to health-related and scientific literature, including Google Scholar, PubMed, ScienceDirect, and Web of Science, among others. These databases can be used as a reliable source for the automation of all the processes.

Numerous organizations have established guidelines to aid users in identifying trustworthy claims [27,28] where time-consuming manual recognition plays an important role in the process. In this pilot study, we proposed a novel automated evidence-based fact-checking approach that aims to identify and confirm accurate, truthful information using scientific literature and research databases as sources of truth. This exploratory evaluation highlights how using this approach may help users measure the extent of confidence in a web page and make informed decisions about accepting the health-related information of a website. Thus, the objective was to assess the truthfulness of health-related information through an evidence-based approach, without creating a synthetic database of claims-evidence but leveraging PubMed as a reliable source of fine-grained and up-to-date health-related information.

Methods

Approximately 1000 web pages were provided by Factually Health company on January 31, 2023. This company specializes in identifying reliable health-content websites [29]. The web pages were selected through random sampling within various disease categories to ensure a balanced dataset while minimizing the risk of overrepresentation of any single category. This approach accounted for variations in the number of available websites across disease categories. The web pages then underwent manual cleaning. Redundant pages were removed, and those unsuitable for research were excluded based on the

XSL•FO
RenderX

following criteria: pages primarily featuring video content, pages related to clinical studies, pages resembling anecdotes rather than factual health information, or pages that restricted data extraction by Python (Python Software Foundation) libraries.

After this process, a dataset comprising 538 web pages was finalized. These web pages represented a diverse range of diseases, including arthritis (81 pages), chronic obstructive pulmonary disease (79 pages), COVID-19 (66 pages), hypertension (66 pages), lung cancer (70 pages), prostate cancer (66 pages), and diabetes (110 pages).

The selection of diverse disease categories was intended to minimize potential bias in the analysis. However, our previous study demonstrated that the selected diseases did not significantly impact classification results [29]. Using the URLs of each web page, the content was extracted as text files using the "justext" library in Python, to remove additional links and extraneous content from websites, such as navigation links, headers, and footers.

The process included the following three steps: (1) Classification of web page content into 3 thematic categories, semiology, epidemiology, and management by evaluating various transformer models, including bidirectional encoder representations from transformers (BERT), SciBERT, and BioBERT, as well as traditional models such as random forest (RF) and support vector machine (SVM), (2) automating the creation of PubMed queries combining "WellcomeBertMesh" and "KeyBERT" models, (3) automatic extraction of top 20 related literatures from PubMed, and (4) applying similarity checking techniques of cosine similarity and Jaccard distance to compare the content of extracted literature and web pages vectorized using BERT tokenizer. As a reliable source of truth, PubMed was a suitable choice to find evidence for health-related claims. PubMed, an open-source platform dedicated to facilitating searches and retrieval of health-related literature, encompasses over 36 million papers [30].

Classification of Web Page Contents

One of the necessary stages before determining the veracity of a claim or information is to detect the sentences that need to be verified [31]. These claims are crucial to the content's main point but require verification through an annotation schema and developing a benchmark for automated claim detection [14,31]. To detect sentences that need to be verified, two major steps were taken: (1) the identification of 3 thematic categories of content and (2) the classification of web page content according to these categories.

The Content Categories

To compare web page content with materials from the scientific literature database, it was essential to categorize the content, ensuring that comparisons were made within the relevant subject. Three distinct thematic categories have been identified for analysis: epidemiology, semiology, and management. In the epidemiology category, we included all sentences related to the statistics of a disease, the population, the frequencies, the causes, the risk assessment of the disease, and all public health-related information about the disease (eg, as of 2014, the global prevalence rate of rheumatoid arthritis was about 0.24%). In the semiology category, we considered all sentences related to signs (eg, high blood pressure is another sign of the disease) and symptoms (eg, this disease has symptoms such as pain, discomfort, weakness, fatigue). Finally, for the management category, we considered all the sentences linked to therapeutic approach (eg, drug treatment and surgical intervention, prevention, and the element of paraclinical diagnosis of diseases (eg, a complete medical examination carried out by a doctor can better determine if a person has chronic obstructive pulmonary disease and the degree of severity of the disease).

Manual Annotation and Model Development

Two authors (AB and AA) independently annotated 200 web pages on a sentence-by-sentence basis considering the 3 categories of epidemiology, semiology, management, and neutral until reaching a roughly balanced amount of data across all classes [32]. We used the Cohen κ score to assess the agreement between the 2 reviewers AB and AA). Any discrepancies were resolved by the third author (JNN).

Neutral sentences were those that did not correspond to any of the defined thematic categories. Table 1 shows the distribution of sentences for each category. The portable serverless text annotation tool of MedTator-1.3-11 [33] was used for the annotation process. A total of 3 transformer models of BERT, SciBERT, and BioBERT were used to classify the sentences into the 4 mentioned categories. The BERT model has demonstrated superior performance in several text classification tasks [29,34,35]. SciBERT is an extension of BERT and is trained on a vast corpus of scientific literature spanning multiple domains [36] and BioBERT is pretrained using an extensive corpus comprising PubMed abstracts (PubMed) and full-text articles from PubMed Central [37]. We have also conducted a performance comparison between the transformer models and 2 traditional machine learning models: RF and SVM.

Table 1. The distribution of classes.

Category	Number of sentences
Neutral	3162
Semiology	851
Epidemiology	1171
Management	1066

XSL•FO
RenderX

The “BertTokenizer” library has been used to tokenize the incoming sentences, with the following parameters: We applied a maximum sequence length of 128 to standardize the size of each input sentence. To optimize the model's hyperparameters, we applied the Bayesian optimization approach using the ‘BayesianOptimization’ library in Python. The hyperparameter tuning spaces are detailed in Table 2.

Table 2. Hyper-parameter tuning search space.

Hyper-parameters	Range	Best trial
Learning rate	10^{-7}, 10^{-2}	3×10^{-5}
Weight decay	10^{-5}, 10^{-1}	10^{-3}
Number of epochs	(1:5)	3
Batch size	(8,16,32,64)	32

Automating PubMed Query Generation

Overview

Literature extraction involved identifying scientific articles within PubMed to support the process. To achieve this, the approach requires the formulation of a query by combining keywords and Medical Subject Headings (MeSH) terms, which can be extracted from web page content. This process included three steps: (1) Automating PubMed subquery creation from MeSH terms and creating a subquery using the “WellcomeBertMesh” model, (2) Automating PubMed subquery creation from keywords using KeyBert model and creating a subquery, and (3) Construction of the final query by combining the different subqueries.

Automating PubMed Subquery Creation Using MeSH Terms Extracted by Transformers

All the MeSH terms were extracted from the text using a pretrained model of “WellcomeBertMesh,” which takes its inspiration from “BertMesh,” which undergoes the pretraining using the entire text of biomedical publications and is built upon the foundation of the BioBert pretrained model [38]. Given that our evidence for the websites primarily comprised health-related articles from PubMed, we selected this model. Its architecture is rooted in the latest advancements in the biomedical field, prominently featuring Microsoft’s cutting-edge “PubMedBert” as its core framework [38].

To enhance the accuracy of the subquery, the identified MeSH terms were initially organized according to their MeSH categories to construct subsubqueries. The MeSH has a tree structure that is organized hierarchically, visually presenting descriptors in broader and narrower relationships. The top tier of the MeSH tree structure encompasses 19 comprehensive categories. While these terms are not included in MeSH data maintenance and distribution, they can be used to search PubMed by using the search term “category” [39]. Therefore, we have considered the MeSH terms under each head category together using the “OR” operator in this subsubquery. Then, we constructed the subquery using the “AND” operator between extracted MeSH terms in different categories. The pseudo-code for this step is presented in Figure 1.

XSL•FO
RenderX

Figure 1. MeSH (medical subject heading) subquery builder.

```
Input: A list of sentences belonging to a web page S = [s1, s2, …] for a specific category
Input: category to consider ∈ {Epidemiology, Semiology, Management}
Output: A PubMed query extracted from the web page
model ←Load the "WellcomeBertMesh" pre-trained model
/* iterating through sentences to compute their vector representation then extracting the MeSH terms corresponding
to each sentence: */
for i ← 1, n do:
        vi ← model_vector (si)
        meshi ← model (vi)
end for
/* identifying the head categories for each MeSH term extracted*/
for j ← 1, length(mesh) do:
        categoryj ← extract_mesh_head_category(meshj)
end for
/* creating subqueries based on the MeSH terms belonging to the same or different categories */
for i ← 1, n do:
        for k ← 1, K do:
            sub-subquery1, sub-subquery2 ← null
        /* put OR for mesh terms in the same category, put AND for different categories*/
                if meshi belong to same categoryk then
                    sub-subquery1 ← (meshi OR sub-subquery1)
                Else
                    sub-subquery2 ← (meshi AND sub-subquery2)
                end if
            MeSH-sub_query ← (sub-subquery1 AND sub-subquery2)
        end for
end for
```

Automating PubMed Subquery Creation Using Key Phrases Extracted by Transformers

The key phrases from web page contents have been extracted using the transformer model "KeyBERT" library, which is described in previous literature as having the best performance in extracting the key phrases [40], especially for long texts [41], which aligns with our need of extracting the key phrases of the scientific papers. The extracted keywords were combined with the "AND" operator to create a subquery.

Figure 2 shows the proposed pseudo-code to extract the keywords for the creation of the subquery.

XSL•FO
RenderX

Figure 2. Key phrase extractor and subquery builder.

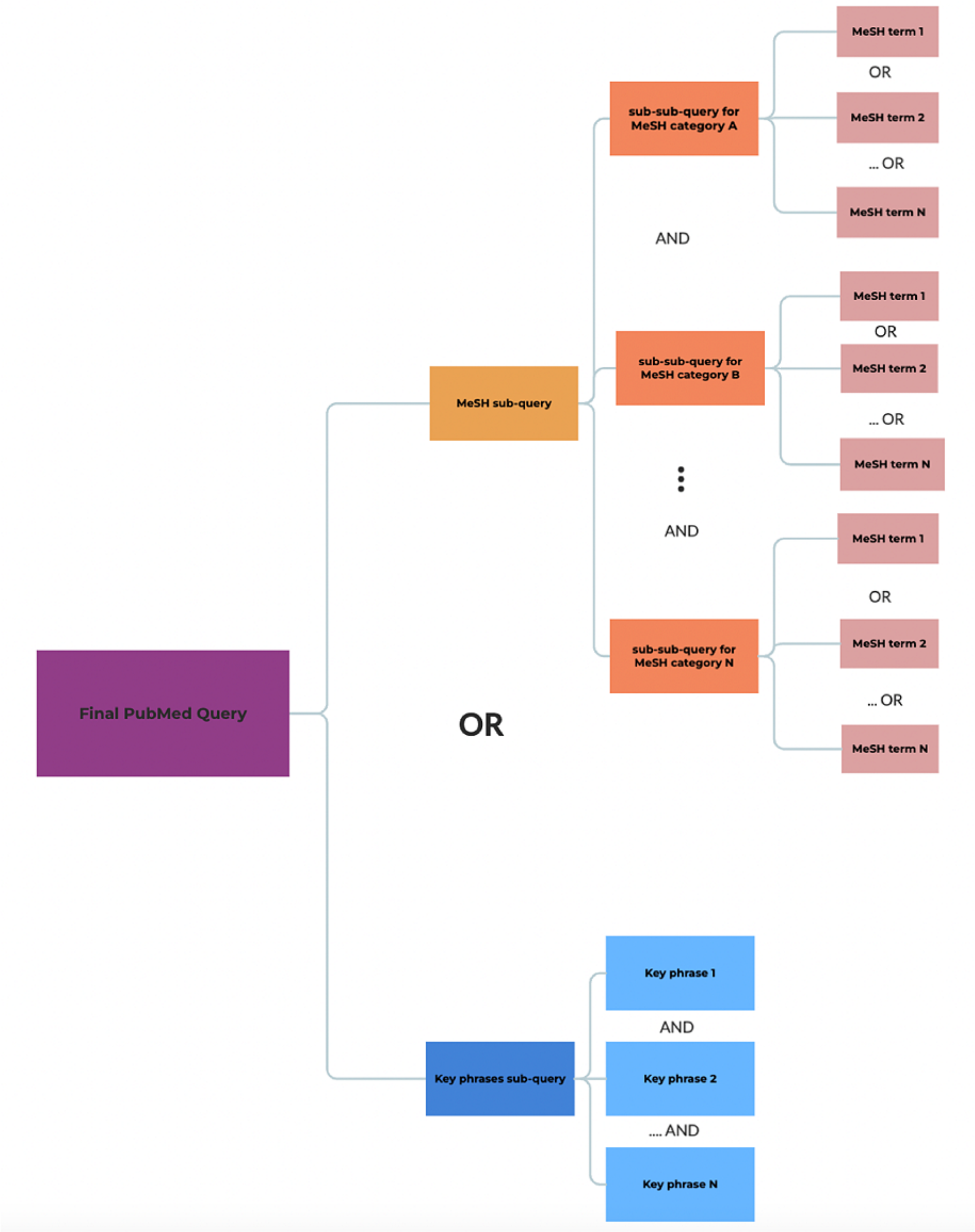
```
Input: A list of sentences belonging to a web page S = [s1, s2, …] for a specific category
Input: category to consider ∈ {Epidemiology, Semiology, Management}
Output: A PubMed query extracted from the web page
model ←Load the "KeyBERT" pre-trained model
/* computing a vector representation and extracting the key phrases corresponding to each sentence */
for i ← 1, n do
        vi ← model_vector (si)
        keyphrasei ← model (vi)
end for
/* creating key phrase subquery tailored to the specified categories */
keyphrase_query ← null
for i ← 1, length(keyphrases) do
        keyphrase_query ← (keyphrase_query AND keyphrasei)
end for
```

Construction of the Final Query

The subqueries extracted from the preceding processes were combined using the "OR" operator to construct the final query. Figure 3 presents a comprehensive overview of the process used to construct the final PubMed query, summarizing the structure and strategy behind its creation.

XSL•FO
RenderX

Figure 3. Detailed process diagram for the development of the comprehensive final PubMed query.



Automating Related Literature Extraction

The final query was used to retrieve a compilation of articles, from which the top open access 20 resulting papers were extracted. The "PMC_ids" of papers were extracted using the "Entrez" library of Python that provides integrated access to PubMed Medline [42]. To evaluate the quality of our query results, we conducted a comprehensive review of the obtained full-text papers. In our assessment of the extracted papers in PubMed, those subjected to filtering within the systematic reviews and meta-analysis category exhibited more related papers to the subject of the research, compared with papers that were not subject to such filtering. Consequently, we selected them to encompass a wider range of relevant articles.

Finally, the automatically extracted papers were manually checked to be pertinent considering the title of the papers, the irrelevant papers were removed and excluded from the final process.

Similarity Detection and Fact-Checking

For the process of computing the similarity measure between different sentences, for each disease, we randomly selected 5 web pages in our dataset. For each of the 3 predefined thematic categories in a web page, 1 PubMed query was generated and with each query, the 20 most related, open access articles within the type of systematic reviews and meta-analysis were extracted. The following steps were then carried out: (1) Categorizing the extracted related literature content based on the 3 thematic categories. This was necessary to analyze sentences (from websites and scientific articles) that are relevant to the same topics. (2) Comparing by thematic category, the content from scientific articles and web pages to identify similar sentences.

XSL•FO
RenderX

Finally, after conducting a manual evaluation of the identified similar sentences, we calculated the average number of categorized sentences for each randomly selected web page, as well as the average number of credible sentences detected. Credible sentences refer to those in the related literature that demonstrated similarity with the sentences from the web pages.

Categorizing the Extracted Literature

The more performant fine-tuned model on the web page contents was used to categorize literature contents into 3 thematic categories. This approach enabled us to facilitate a direct comparison between sentences sharing the same thematic context.

Comparing the Content From Literature and Web Pages to Identify Similar Sentences

For the sentence comparison, we used the BERT vectorizer to transform the texts into vectors. This allowed us to encode the semantic significance of sentences as numerical values, facilitating the application of different similarity detection algorithms [43].

Both scientific articles and web page sentences were transformed into vector representations, taking into account their respective thematic categories. Subsequently, each web page sentence was compared with scientific article sentences of the same category using the cosine similarity and Jaccard technique. A similarity threshold of 87% was chosen to determine sentence selection, ensuring that sentences with over 87% similarity were chosen.

Figure 4 shows the proposed pseudo-code for the similarity-checking part.

Figure 4. Paper similarity detection.

```
Input: A list of sentences belonging to a web page and papers S = [s1, s2, …], P = [p1, p2, …]
Input: category to check ∈ {Epidemiology, Semiology, Management}, similarity_threshold
Output: percentage of similarity between two contents
model ←Load the "bert_base_uncase" pre-trained model
/* computing vector representation of paper sentences */
for i ← 1, n do:
        vi , v'i ← model_vector (si, pi)

end for
/* computing the percentage of similarity between the contents of the web page and the papers*/
for k ← 1, length (vi) do
/* if the similarity between web page and the paper sentences be more than threshold*/
        if Cosine_similarity (vi, v'j∈{1…n}) > similarity_threshold then
          print (corresponding sentences of (vi, v'j))
        end if
end for
compute similarity percentage
```

For each disease, we randomly selected 5 web pages and extracted both their related papers and similar sentences. It was due to the inherent variability and specificity of medical information related to each disease. Diseases often exhibit unique characteristics, nuances, and clinical considerations. By prioritizing diseases, we aimed to provide a more granular and clinically relevant assessment of the similarity between the sentences. The outcomes, comprising sentences from the web pages and their corresponding similar sentences, underwent a manual verification by the authors to ensure semantic similarity between them. Subsequently, the proportion of semantically similar sentences between a web page and its related reference papers was calculated.

Ethical Considerations

This research relied solely on publicly accessible data and did not involve any human or animal participants, making it exempt from the need for ethical approval. The study strictly adheres to established data privacy norms to prevent any compromise of confidentiality or privacy. In addition, the project does not include any direct involvement or interactions with individuals, thereby minimizing potential ethical issues. The University of

XSL•FO
RenderX

Montreal's Research Committee has carefully examined our methodology and affirmed that this study falls outside the scope of Medical Research Involving.

Results

This section elaborates on the results of each part of the proposed pseudo-codes.

Classification of Web Page Contents

The annotation process for web page contents achieved a Cohen κ score of 87% among the 2 annotators (AA and AB), indicating high agreement between the annotators and ensuring the reliability of the data used for model evaluation.

The performance of transformer-based models (BERT, BioBERT, and SciBERT) was compared to traditional machine learning models (RF and SVM) for categorizing web page content into four categories. BERT emerged as the most effective model, consistently achieving superior precision, recall, and F_1-scores across all categories. Traditional models, in contrast, demonstrated lower performance, particularly in terms of F_1-scores, indicating limitations in balancing precision and recall effectively.

Table 3 illustrates the performance of the classification models used to classify the content of web pages. The performance matrix includes metrics such as precision, recall, and F_1-score.

Table 3. Performance evaluation of the BERT (Bidirectional Encoder Representations from Transformers) and machine learning models for web page content classification across considered categories.

	BERT[a]			BioBERT			SciBERT			RF[b]			SVM[c]		
Classes	Precision	Recall	F_1-score	Precision	Recall	F_1-score	Precision	Recall	F_1-score	Precision	Recall	F_1-score	Precision	Recall	F_1-score
Neutral	0.96	0.93	0.95	0.88	0.83	0.85	0.85	0.81	0.83	0.51	0.92	0.66	0.72	0.81	0.77
Semiology	0.91	0.94	0.93	0.81	0.81	0.81	0.77	0.79	0.78	0.96	0.05	0.09	0.71	0.59	0.64
Epidemiology	0.92	0.94	0.93	0.80	0.76	0.76	0.75	0.74	0.75	0.8	0.1	0.1	0.69	0.62	0.65
Management	0.95	0.96	0.96	0.83	0.89	0.89	0.83	0.87	0.85	0.59	0.58	0.59	0.74	0.73	0.74

[a]BERT: Bidirectional Encoder Representations from Transformers.

[b]RF: random forests.

[c]SVM: support vector machines.

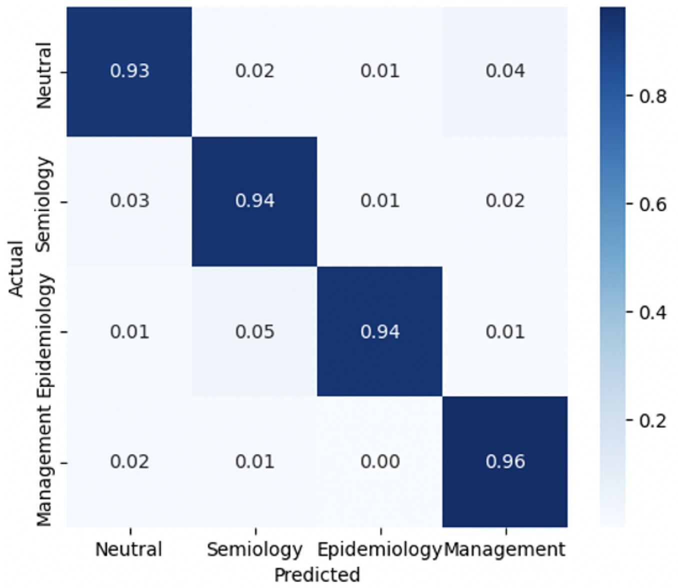
According to Table 3, among the transformer models, the BERT model had a promising performance with more than 93% recall for neutral sentences, 94% for semiology and epidemiology, and 96% for the management category. The model had an F_1-score of 95% for neutral sentences, 93% for semiology and epidemiology, and 96% for management. The model had 96% precision for neutral sentences, 91% for semiology, 92% for epidemiology, and 95% for management. Also, traditional models did not have high performance, the precision values for both RF and SVM were relatively low in some classes, indicating a high rate of false positives. Also, the F_1-scores for both RF and SVM were generally lower compared with the BERT model, indicating that they may not achieve a good balance between precision and recall. Therefore, the BERT model was selected for the classification of the web page contents.

The confusion matrix for the BERT model is shown in Figure 5, providing a detailed visualization of its classification performance across the different categories.

Figure 5 shows the confusion matrix for the BERT classifier, which correctly classified 0.93 of the neutral sentences, 0.94 for both the semiology and epidemiology sentences, and 0.96 for management sentences as true positives.

XSL•FO
RenderX

Figure 5. Bidirectional encoder representations from transformers model performance: confusion matrix for the classification of web page sentences into 3 thematic categories.



Automating PubMed Query Generation

To extract relevant literature for the web pages categorized thematically, a PubMed query was generated for each of the 7 diseases. Each query retrieved the 20 most related papers. The titles of the retrieved papers were manually evaluated, and less than 10% were deemed irrelevant, demonstrating the effectiveness of the generated queries. These irrelevant articles were excluded from further analysis.

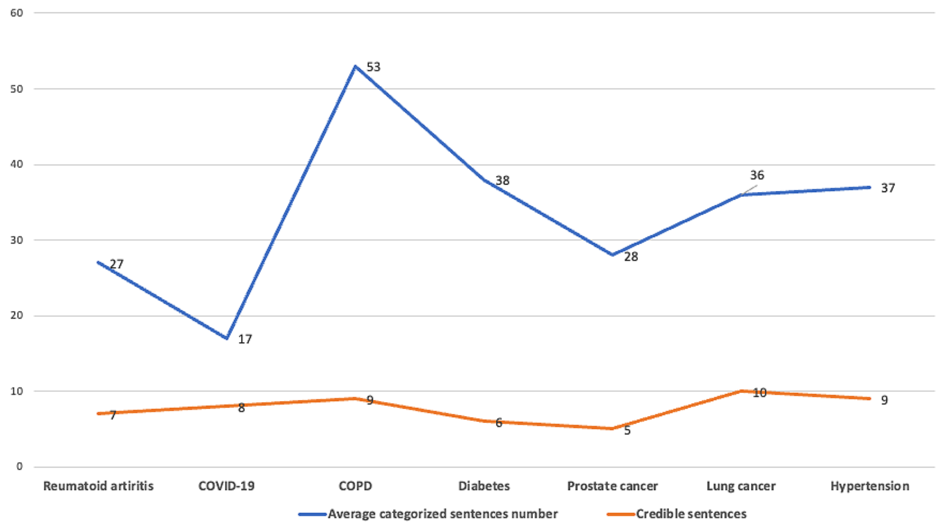
This result highlights the utility of using MeSH terms and key phrases in constructing PubMed queries, which efficiently yielded pertinent literature. The generated weblinks for accessing the papers followed the format: "https://pubmed.ncbi.nlm.nih.gov/PMID/," with PMIDs obtained directly from the PubMed queries.

Similarity Detection and Fact-Checking

Figure 6 illustrates the average percentage of credible information found in the 5 randomly selected web pages categorized by related diseases. Credible information is defined as sentences in the web pages that were successfully matched with corresponding sentences in PubMed articles.

On average, 23% of the sentences on each web page were identified as similar to statements in the scientific literature. While this demonstrates the potential of the system to detect credible content, a significant challenge arose with false positives. Some sentences achieved a similarity score exceeding 80% but were semantically dissimilar upon closer inspection.

Figure 6. The average number of credible sentences on web pages (red line) versus the average number of all sentences on each web page (blue line). COPD: chronic obstructive pulmonary disease.



XSL•FO
RenderX

For instance, the following sentences from an extracted paper and a web page had a similarity score of 88% yet conveyed different meanings:

1. “Previous studies have documented residual symptoms that continue 12 weeks after the onset of acute COVID-19, known as post-acute or long COVID-19.”

2. “The acute phase of COVID itself can last for up to 14 days.”

This highlights the need for more sophisticated approaches to accurately distinguish between syntactic similarity and genuine semantic alignment.

As an illustrative example, for the rheumatoid arthritis category, we randomly selected 5 web pages, each containing an average of 27 sentences distributed across 3 thematic categories: epidemiology, semiology, and management (represented by the blue line). Among these, an average of 7 sentences per web page were deemed credible and successfully matched to corresponding statements in the scientific literature (depicted by the red line).

Discussion

Principal Findings

In the present pilot study, our objective was to automate aspects of the fact-checking process for online health information. While previous research [21,26] has explored automation in various stages of fact-checking, such as evidence retrieval or claim identification, this pilot serves as an initial step toward achieving full automation in the fact-checking process. Our approach includes the automation of identifying verifiable sentences through a classification process. Notably, our study used a fine-tuned BERT model, which exhibited notable efficacy in categorizing health-related sentences. Although BioBERT and SciBERT models have been reported to outperform BERT in various downstream tasks [36,37], in our investigation, the BERT model demonstrated superior performance. This discrepancy could be attributed to BERT training on general-purpose texts, such as Wikipedia or Book Corpus [35], which align more closely with the content of websites targeted at general populations. In contrast, BioBERT and SciBERT are trained on more specialized texts, such as scientific publications [36,37].

Previous research [14,31,44] has shown that the identification of claim-worthy sentences or the recognition of key information needing verification from reliable sources is a fundamental first step in automating the fact-checking process akin to our approach. This process is commonly structured as a text classification task. The previous studies used human annotators [44] or crowdsourcing [31] to tag claim-worthy sentences and trained machine learning models to classify them. A previous study [14] focused on detecting claims within news and public information, assigning each sentence a likelihood score for containing significant factual claims. Also, automating the fact-checking process is far from straightforward, as it necessitates the utilization of artificial intelligence tools to struggle with the complexity of text and context [10]. Studies often considered the problem as a binary classification to split the contents into credible or non-credible information, however, the decision is more complex since there may be several ambiguities in the sentences. In addition, several parts of the process depend on human judgment, which needs further research in the area. Building on this groundwork, our study applied a BERT-based classification approach to detect health information requiring verification and automatically proposing a sentence for this process. Previous studies relied on reviewer selections to develop claim and evidence datasets, lacking attempts to automate claim identification with real-world resources [17,18,45].

In addition, rather than constructing a manual reference dataset as the evidence for verifiable sentences, we leveraged the PubMed database as our source of truth. We automated the detection of evidence for claims made on web pages in an unsupervised approach, streamlining the verification process. This aligns with previous studies [21,26] that used PubMed publications as evidence, using transformer models to generate queries and retrieve documents from PubMed. We demonstrated the effectiveness of using transformer models to extract MeSH terms and key phrases from web page content, enabling the efficient generation of PubMed queries. This approach facilitated the retrieval of related articles from scientific references without requiring supervision. According to a previous study [14], to verify the veracity of the claims, it is crucial to translate them into queries against the reference databases. However, other studies [6,20,22] created a knowledge database as the references to compare with the claims. Notably, Sarrouti et al [6] introduced a dataset comprising evidence-claim pairs, manually annotated as SUPPORT, REFUTE, and NEUTRAL. They used BERT-based models to create a realistic testing ground for evidence-based fact-checking systems.

To assess the alignment between claim sentences and extracted references, we measured their similarity, a practice supported by [46]. This study underscores the necessity for a model in claim verification to measure the semantic similarity between claims and verified factual knowledge or references. To compare the semantic similarity, we used a transformer-based representation that converted the textual content into vectorial representation, allowing us to capture the contextual nuances of each sentence consistent with previous approaches [19,43,47]. This approach is more efficient and produces semantically richer sentence representations than simply averaging the vectors of words that appear in each sentence, and facilitates the similarity detection for the algorithms [48]. We successfully identified factual evidence for 23% of the health-related information extracted from web pages, indicating the complexity inherent in health information. Further research is required to enhance contextual comparison between claims and verified references. Also, the cosine similarity outperformed the Jaccard distance measure for comparing the claims and evidence in this study, which is different from the previous study [4], as they reported that the Jaccard distance was better at the similarity selection measure. The reason may be due to differences in the nature of the datasets in the 2 studies.

XSL•FO
RenderX

Limitations

This study had several limitations. First, we faced a challenge in identifying sentences within the papers that closely matched the content of the web pages. Numerous methods have been devised to tackle this issue [19,43,46]; however, a comprehensive consideration of the complete meaning of sentences requires further investigation. In addition, 77% of the sentences did not have matching counterparts in the academic literature that we retrieved. Regarding this proportion, 2 possible assumptions can be made: either the sentences themselves were not valid or the algorithm was unable to locate their related counterparts. Another potential reason could be that the sentences, though addressing a common subject such as the same medical condition, exhibited variations in meaning or contextual interpretation. Consequently, it would be premature to assert that these unmatched sentences are inherently not credible, given the vast volume of published papers that renders comprehensive verification computationally infeasible. Expanding the number of selected papers for comparison could therefore increase the likelihood of identifying additional relevant sentences in the literature. Nonetheless, quantifying the proportion of credible sentences offers valuable insights to aid users in their trust assessment.

It is worth acknowledging that authors in the realm of health-related data often simplify and rephrase content to cater to their target audience, making it more challenging to identify credible references for their statements. Therefore, the researchers propose exploring other models such as text generation models as potential solutions to address this particular challenge including WordNet or sequence-to-sequence (Seq2Seq) models.

A second limitation was the sample size of the academic papers used in the comparison. Due to the extensive volume of health-related publications, the assessment was limited to a selection of 20 papers. Expanding this scope to include more papers per content type could enhance the discovery of factual evidence in PubMed publications. Thus, further investigation into paper retrieval approaches is recommended.

A third limitation was that, although the thematic categorization of web page content, such as epidemiology, semiology, and management, ensured that the generated PubMed queries were more precise and contextually relevant, the need for quality assessment of the extracted PubMed articles remains evident. While our method provides users with essential information to assess the accuracy of health information, the ultimate determination of its truthfulness may depend on individual judgment, expert evaluation, source credibility, scientific article quality (eg, journal quality, impact factor for the domain) and the contemporaneity of the information (eg, date of publication, retracted).

The retrieved articles may vary in quality, ranging from high-impact studies to potentially outdated or retracted articles that could influence the reliability of the fact-checking process and the conclusions drawn from matched content. Addressing these characteristics within an automated process remains a key challenge. In our previous research, the credibility of the sources was automatically assessed [29]. In this study, while we evaluate comparability with scientific articles, developing a credibility scoring strategy for these articles is also necessary. Combining an algorithm that evaluates website credibility and assigns a credibility score to scientific articles with 1 that determines truthfulness could significantly enhance the effectiveness of fact-checking. These models can change the structure of sentences and may improve the possibility of finding more similar sentences. Finally, while the process could not be automated entirely since each step needed human supervision for the results, the suggested techniques have the potential to substantially alleviate the human effort required to locate valid information.

Conclusions

Our approach aimed to empower users in the decision-making process regarding the truthfulness of information by providing relevant evidence and enabling informed judgments. As a pilot, this research serves as an initial step toward exploring the feasibility of automating fact-checking processes in health information. Specifically, the methods presented here could be applied to create tailored fact-checking workflows for specific disease areas, such as diabetes, arthritis, or cancer, which were among the categories included in this study. For instance, thematic categorization (eg, management and epidemiology) could improve the precision and relevance of fact-checking tools in health care contexts. Using state-of-the-art models such as transformers may improve the performance of the model since the BERT embedding captures the meaning of the sentences [49]. The investigation also revealed that incorporating PubMed publications as a trustworthy resource can enhance the discovery of similar credible information as evidence. Finally, while the process could not be entirely automated and required human supervision, the suggested techniques demonstrate significant potential for integration into fact-checking tools. This integration could reduce the effort required to validate health information, ultimately increasing accessibility and reliability for end-users. Future work should focus on expanding the dataset and testing the approach in real-world scenarios to further refine its applicability across various health domains.

Acknowledgments

The authors express their sincere gratitude to Factually Health Company for generously providing a factual dataset covering 7 diseases and to IVADO Labs for their support.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Conflicts of Interest

This research was funded by the Mitacs acceleration program in partnership with Factually Health Company, and the IVADO Funding for Collaborative Research in Data Science to Serve Sustainable Development.

References

1. Cuan-Baltazar JY, Muñoz-Perez MJ, Robledo-Vega C, Pérez-Zepeda MF, Soto-Vega E. Misinformation of COVID-19 on the internet: infodemiology study. JMIR Public Health Surveill 2020 Apr 09;6(2):e18444 [FREE Full text] [doi: 10.2196/18444] [Medline: 32250960]
2. Himelein-Wachowiak M, Giorgi S, Devoto A, Rahman M, Ungar L, Schwartz HA, et al. Bots and misinformation spread on social media: implications for COVID-19. J Med Internet Res 2021 May 20;23(5):e26933 [FREE Full text] [doi: 10.2196/26933] [Medline: 33882014]
3. Lee DKL, Ramazan O. Fact-checking of health information: the effect of media literacy, metacognition and health information exposure. J Health Commun 2021 Jul 03;26(7):491-500. [doi: 10.1080/10810730.2021.1955312] [Medline: 34292862]
4. Barve Y, Saini JR. Healthcare misinformation detection and fact-checking: a novel approach. IJACSA 2021;12(10). [doi: 10.14569/IJACSA.2021.0121032]
5. Ecker UKH, Lewandowsky S, Cook J, Schmid P, Fazio LK, Brashier N, et al. The psychological drivers of misinformation belief and its resistance to correction. Nat Rev Psychol 2022 Jan 12;1(1):13-29. [doi: 10.1038/s44159-021-00006-y]
6. Sarrouti M, Abacha AB, Mrabet Y, Demner-Fushman D. Evidence-based fact-checking of health-related claims. : Association for Computational Linguistics; 2021 Presented at: Findings of the Association for Computational Linguistics: EMNLP 2021; 2025 January 22; Punta Cana, Dominican Republic p. 3499-3512. [doi: 10.18653/v1/2021.findings-emnlp.297]
7. Sumayyia MD, Al-Madaney MM, Almousawi FH. Health information on social media. Perceptions, attitudes, and practices of patients and their companions. Saudi Med J 2019 Dec;40(12):1294-1298 [FREE Full text] [doi: 10.15537/smj.2019.12.24682] [Medline: 31828284]
8. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
9. Hang CN, Yu PD, Chen S, Tan CW, Chen G. MEGA: Machine Learning-Enhanced Graph Analytics for infodemic risk management. IEEE J Biomed Health Inform 2023 Dec;27(12):6100-6111. [doi: 10.1109/JBHI.2023.3314632] [Medline: 37713230]
10. Altuncu E, Nurse J, Bagriacik M, Kaleba S, Yuan H, Bonheme L, et al. aedFaCT: scientific fact-checking made easier via semi-automatic discovery of relevant expert opinions. arXiv:2305.07796 2023. [doi: 10.36190/2023.27]
11. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res 2020 Jun 29;22(6):e21820 [FREE Full text] [doi: 10.2196/21820] [Medline: 32589589]
12. Guo Z, Schlichtkrull M, Vlachos A. A survey on automated fact-checking. Transactions of the Association for Computational Linguistics 2022;10:178-206. [doi: 10.1162/tacl_a_00454]
13. Walter N, Cohen J, Holbert RL, Morag Y. Fact-checking: a meta-analysis of what works and for whom. Political Communication 2019 Oct 24;37(3):350-375. [doi: 10.1080/10584609.2019.1668894]
14. Hassan N, Arslan F, Li C, Tremayne M. Toward automated fact-checking: detecting check-worthy factual claims by ClaimBuster. : ACM; 2017 Presented at: Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining; 2017 August 13-17; Halifax NS Canada p. 1803-1812. [doi: 10.1145/3097983.3098131]
15. Samadi M, Mousavian M, Momtazi S. Deep contextualized text representation and learning for fake news detection. Information Processing & Management 2021 Nov;58(6):102723. [doi: 10.1016/j.ipm.2021.102723]
16. Zhang J, Featherstone JD, Calabrese C, Wojcieszak M. Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. Prev Med 2021 Apr;145:106408. [doi: 10.1016/j.ypmed.2020.106408] [Medline: 33388335]
17. Zhong B. Going beyond fact-checking to fight health misinformation: A multi-level analysis of the Twitter response to health news stories. International Journal of Information Management 2023 Jun;70:102626. [doi: 10.1016/j.ijinfomgt.2023.102626]
18. Akhtar M, Cocarascu O, Simperl E. PubHealthTab: a public health table-based dataset for evidence-based fact checking. : Association for Computational Linguistics; 2022 Presented at: Findings of the Association for Computational Linguistics: NAACL 2022; 2025 January 22; Seattle, United States p. 1-16. [doi: 10.18653/v1/2022.findings-naacl.1]
19. Martín A, Huertas-Tato J, Huertas-García Á, Villar-Rodríguez G, Camacho D. FacTeR-Check: semi-automated fact-checking through semantic similarity and natural language inference. Knowledge-Based Systems 2022. [doi: 10.1016/j.knosys.2022.109265]
20. Sathe A, Ather S, Le TM, Perry N, Park J. Automated fact-checking of claims from Wikipedia. : European Language Resources Association; 2020 Presented at: Proceedings of the Twelfth Language Resources and Evaluation Conference; 2025 January 22; Marseille, France p. 6874-6882.
21. Deka P, Jurek-Loughrey A, P D. Improved methods to aid unsupervised evidence-based fact checking for online heath news. JDI 2022;3(4):474-504. [doi: 10.26421/JDI3.4-5]

XSL•FO
RenderX

22. Thorne J, Vlachos A, Christodoulopoulos C, Mittal A. FEVER: a large-scale dataset for fact extraction and VERification. ACL Anthology 2018:1-9. [doi: 10.18653/v1/w18-5501]
23. Das A, Liu H, Kovatchev V, Lease M. The state of human-centered NLP technology for fact-checking. Information Processing & Management 2023 Mar;60(2):103219. [doi: 10.1016/j.ipm.2022.103219]
24. Nakov P, Corney D, Hasanain M, Alam F, Elsayed T, Barrón-Cedeño A, et al. Automated fact-checking for assisting human fact-checkers.. arXiv Preprint posted online on March 13, 2021. [doi: 10.48550/arXiv.2103.07769]
25. Hang CN, Yu PD, Tan CW. TrumorGPT: query optimization and semantic reasoning over networks for automated fact-checking. : IEEE; 2024 Presented at: 2024 58th Annual Conference on Information Sciences and Systems (CISS); 2024 March 13-15; Princeton, NJ, USA p. 1-6. [doi: 10.1109/ciss59072.2024.10480162]
26. Deka P, Jurek-Loughrey A, Deepak. Unsupervised keyword combination query generation from online health related content for evidence-based fact checking. : ACM; 2021 Presented at: iiWAS2021: The 23rd International Conference on Information Integration and Web Intelligence; 2021 November 29; Linz Austria p. 267-277. [doi: 10.1145/3487664.3487701]
27. Boyer C, Baujard V, Nater T, Scherrer J, Appel R. HONcode: a standard to improve the quality of medical/health information on the internet and HON's 5th survey on the use of internet for medical and health purposes. In: J Med Internet Res. 1999 Presented at: 6th Internet World Congress for Biomedical Sciences (INABIS 2000); 2000 December 7-16; Canada. [doi: 10.2196/jmir.1]
28. Padhy S, Rathore SS. An attention‐based deep learning model for credibility assessment of online health information. Computational Intelligence 2023 Jul 24;39(5):832-859. [doi: 10.1111/coin.12596]
29. Bayani A, Ayotte A, Nikiema JN. Automated credibility assessment of web-based health information considering health on the net foundation code of conduct (HONcode): model development and validation study. JMIR Form Res 2023;7:e52995 [FREE Full text] [doi: 10.2196/52995] [Medline: 38133919]
30. Giustini D, Barsky E. A look at Google Scholar, PubMed, and Scirus: comparisons and recommendations. J Can Health Libr Assoc 2005 Sep 01;26(3):85-89. [doi: 10.5596/c05-030]
31. Konstantinovskiy L, Price O, Babakar M, Zubiaga A. Toward automated factchecking. Digital Threats 2021 Apr 15;2(2):1-16. [doi: 10.1145/3412869]
32. Fernando KRM, Tsokos CP. Dynamically weighted balanced loss: class imbalanced learning and confidence calibration of deep neural networks. IEEE Trans Neural Netw Learn Syst 2022 Jul;33(7):2940-2951. [doi: 10.1109/TNNLS.2020.3047335] [Medline: 33444149]
33. He H, Fu S, Wang L, Liu S, Wen A, Liu H. MedTator: a serverless annotation tool for corpus development. Bioinformatics 2022 Mar 04;38(6):1776-1778 [FREE Full text] [doi: 10.1093/bioinformatics/btab880] [Medline: 34983060]
34. González-Carvajal S, Garrido-Merchán EC. Comparing BERT against traditional machine learning text classification.. arXiv Preprint published online May 26, 2020. [doi: 10.48550/arXiv.2005.13012]
35. Koroteev MV. BERT: a review of applications in natural language processing and understanding.. arXiv Preprint published online March 22, 2021. [doi: 10.48550/arXiv.2103.11943]
36. Beltagy I, Lo K, Cohan A. SciBERT: a pretrained language model for scientific text. ACL Anthology 2019:3615-3620. [doi: 10.18653/v1/d19-1371]
37. Lee J, Yoon W, Kim S, Kim D, Kim S, So C, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics 2020;36(4):1234-1240 [FREE Full text] [doi: 10.1093/bioinformatics/btz682] [Medline: 31501885]
38. You R, Liu Y, Mamitsuka H, Zhu S. BERTMeSH: deep contextual representation learning for large-scale high-performance MeSH indexing with full text. Bioinformatics 2021;37(5):684-692 [FREE Full text] [doi: 10.1093/bioinformatics/btaa837] [Medline: 32976559]
39. Dhammi IK, Kumar S. Medical subject headings (MeSH) terms. Indian J Orthop 2014 Sep;48(5):443-444 [FREE Full text] [doi: 10.4103/0019-5413.139827] [Medline: 25298548]
40. Priyanshu A, Vijay S. arXiv 2022. [doi: 10.48550/arXiv.2211.07499]
41. Giarelis N, Kanakaris N. A comparative assessment of state-of-the-art methods for multilingual unsupervised keyphrase extraction. In: Artificial Intelligence Applications and Innovations. Cham: Springer International Publishing; 2021:635-645.
42. Schuler GD, Epstein JA, Ohkawa H, Kans JA. Entrez: molecular biology database and retrieval system. Methods Enzymol 1996;266:141-162. [doi: 10.1016/s0076-6879(96)66012-1] [Medline: 8743683]
43. Qadrud-Din J, Rabiou A, Walker R, Soni R, Gajek M, Pack G, et al. Transformer based language models for similar text retrieval and ranking.. arXiv Preprint published online on May 10, 2020. [doi: 10.48550/arXiv.2005.04588]
44. Pathak A, Shaikh MA, Srihari R. Self-supervised claim identification for automated fact checking.. arXiv. :213-227 Preprint published online on February 3, 2021. [doi: 10.48550/arXiv.2102.02335]
45. Kotonya N, Toni F. Explainable automated fact-checking for public health claims. ACL Anthology 2020:7740-7754. [doi: 10.18653/v1/2020.emnlp-main.623]
46. Barve Y, Saini JR. Detecting and classifying online health misinformation with 'Content Similarity Measure (CSM)' algorithm: an automated fact-checking-based approach. J Supercomput 2023;79(8):9127-9156 [FREE Full text] [doi: 10.1007/s11227-022-05032-y] [Medline: 36644509]

XSL•FO
RenderX

47. Vijjali R, Potluri P, Kumar S, Teki S. Two stage transformer model for COVID-19 fake news detection and fact checking.. arXiv Preprint published online on November 26, 2020. [doi: 10.48550/arXiv.2011.13253]
48. Nabożny A, Balcerzak B, Wierzbicki A, Morzy M, Chlabicz M. Active annotation in evaluating the credibility of web-based medical information: guidelines for creating training data sets for machine learning. JMIR Med Inform 2021;9(11):e26065 [FREE Full text] [doi: 10.2196/26065] [Medline: 34842547]
49. Jawahar G, Sagot B, Seddah D. What does BERT learn about the structure of language? : HAL open scince; 2019 Presented at: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics; 2019 July 28 to August 2; Florence (Italy) p. 3651-3657. [doi: 10.18653/v1/p19-1356]

Abbreviations

BERT: bidirectional encoder representations from transformers
MeSH: medical subject heading
RF: random forest
SVM: support vector machines

Edited by T Mackey; submitted 27.01.24; peer-reviewed by P Deka, CN Hang, O Ismaila; comments to author 14.03.24; revised version received 08.05.24; accepted 24.12.24; published 21.02.25.

Please cite as:
Bayani A, Ayotte A, Nikiema JN
Transformer-Based Tool for Automated Fact-Checking of Online Health Information: Development Study
JMIR Infodemiology 2025;5:e56831
URL: https://infodemiology.jmir.org/2025/1/e56831
doi:10.2196/56831
PMID:39812653

©Azadeh Bayani, Alexandre Ayotte, Jean Noel Nikiema. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 21.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Beliefs in Misinformation About COVID-19 and the Russian Invasion of Ukraine Are Linked: Evidence From a Nationally Representative Survey Study

Dominika Grygarová[1,2], PhD; Marek Havlík[1], PhD; Petr Adámek[1,2], PhD; Jiří Horáček[1,2], PhD, Prof Dr Med; Veronika Juríčková[1,3], MA; Jaroslav Hlinka[1,4], PhD; Ladislav Kesner[1,5], PhD, Prof Dr

[1]Center for Advanced Studies of Brain and Consciousness, National Institute of Mental Health, Klecany, Czech Republic
[2]Third Faculty of Medicine, Charles University, Prague, Czech Republic
[3]First Faculty of Medicine, Charles University, Prague, Czech Republic
[4]Institute of Computer Science, Czech Academy of Sciences, Prague, Czech Republic
[5]Faculty of Arts, Masaryk University in Brno, Brno, Czech Republic

Corresponding Author:
Dominika Grygarová, PhD
Center for Advanced Studies of Brain and Consciousness
National Institute of Mental Health
Topolová 748
Klecany, 25067
Czech Republic
Phone: 420 283088264
Email: dominika.grygarova@nudz.cz

Abstract

Background: Detrimental effects of misinformation were observed during the COVID-19 pandemic. Presently, amid Russia's military aggression in Ukraine, another wave of misinformation is spreading on the web and impacting our daily lives, with many citizens and politicians embracing Russian propaganda narratives. Despite the lack of an objective connection between these 2 societal issues, anecdotal observations suggest that supporters of misinformation regarding COVID-19 (BM-C) have also adopted misinformation about the war in Ukraine (BM-U) while sharing similar media use patterns and political attitudes.

Objective: The aim of this study was to determine whether there is a link between respondents' endorsement of the 2 sets of misinformation narratives, and whether some of the selected factors (media use, political trust, vaccine hesitancy, and belief rigidity) are associated with both BM-C and BM-U.

Methods: We conducted a survey on a nationally representative sample of 1623 individuals in the Czech Republic. Spearman correlation analysis was performed to identify the relationship between BM-C and BM-U. In addition, multiple linear regression was used to determine associations between the examined factors and both sets of misinformation.

Results: We discovered that BM-C and BM-U were moderately correlated (Spearman ρ=0.57; P<.001). Furthermore, increased trust in Russia and decreased trust in the local government, public media, and Western allies of the Czech Republic predicted both BM-C and BM-U. Media use indicating frustration with and avoidance of public or mainstream media, consumption of alternative information sources, and participation in web-based discussions indicative of epistemic bubbles predicted beliefs in misinformation narratives. COVID-19 vaccine refusal predicted only BM-C but not BM-U. However, vaccine refusers were overrepresented in the BM-U supporters (64/161, 39.8%) and undecided (128/505, 25.3%) individuals. Both beliefs were associated with belief rigidity.

Conclusions: Our study provides empirical evidence that supporters of COVID-19 misinformation were susceptible to ideological misinformation aligning with Russian propaganda. Supporters of both sets of misinformation narratives were primarily linked by their shared trust or distrust in the same geopolitical actors and their distrust in the local government.

(JMIR Infodemiology 2025;5:e62913) doi:10.2196/62913

XSL•FO
RenderX

KEYWORDS
misinformation; COVID-19; war in Ukraine; political trust; digital media; belief rigidity; vaccine hesitancy; war; political; trust; belief; survey; questionnaire; national; false; association; correlation; correlation analysis; public opinion; media; news; health information; public health; COVID; misinformation; propaganda

Introduction

During the COVID-19 pandemic, many countries worldwide have experienced an increase and acceleration in the spread of conspiracies, hoaxes, misinformation, and intentionally disseminated disinformation [1,2]. A large body of scientific research has demonstrated the detrimental effects of the infodemic on vaccine hesitancy worldwide [3,4], hateful and divisive rhetoric [5], politicization of the issue [6], and radicalization [7].

Social epistemic structures known as echo chambers, which primarily emerge in web-based communities where members reinforce their shared views while actively discrediting other relevant voices [8], have been frequently identified as primary digital channels reinforcing beliefs in misinformation and fueling radicalization [9,10]. Similarly, in the Czech Republic, misinformation narratives have been monitored in web-based communities [11], as well as in chain emails, that have been massively forwarded [12,13]. The main COVID-19 misinformation narratives encompassed a wide range of claims, including the pandemic being a hoax, the assertion that the virus is not dangerous or was artificially developed, and the belief that vaccines are harmful, while PCR tests, face masks, and other preventive measures against COVID-19 pandemic are ineffective [14].

Apart from the spread of misinformation—false information disseminated without the intent to deceive—fueled by the uncertainty of pandemic developments and negative emotions on social media [15], it has been suggested that the issue of COVID-19 pandemic has also been “hijacked” and used by disinformation campaigns conducted for monetary [16] or political purposes [17]. Previous studies have indicated that worries about the harmful effects of vaccination and distrust in Western pharmaceutical companies and politicians have been exploited and reinforced by Russian disinformation campaigns, aiming to undermine public support for state authorities [18]. The Czech Security Information Service reported that pro-Russian activists, promoting antivaccination attitudes and pro-Russian narratives, used COVID-19 pandemic as a useful topic for spreading conspiracies and disinformation [13]. These activists operated largely in symbiosis with the anti–COVID-19 measures movement, particularly on Czech language fringe news websites [13] labeled “disinformation” or “antisystem” websites by media experts [12].

Another massive wave of infodemic began to spread after the Russian invasion of Ukraine in February 2022 [19]. The war has become a new global threat, dominating media coverage and social media attention. Consequently, the focus on COVID-19 pandemic has receded, along with COVID-19 misinformation in the web-based environment [20]. In the Czech Republic, misinformation, including pro-Russian narratives about the conflict in Ukraine and hostile targeting Ukrainian refugees, has spread on “antisystem” websites [21]. These narratives also proliferated via chain emails, which have steeply increased in number after the invasion [20], in social media communities [22], as well as in web-based discussions under web news articles, where increased troll and bot activity has been observed [20,21]. A direct comparison of fact-checking publications revealed that while hoaxes related to both COVID-19 pandemic and the Ukraine war were predominantly disseminated via social media, they differed in their preferred format. Fabricated content was more common in pandemic-related hoaxes, whereas out-of-context images were prevalent in disinformation surrounding the Russia-Ukraine war [23]. The flood of web-based disinformation during both COVID-19 pandemic and the Russian invasion of Ukraine galvanized fact-checking and verification efforts [24-26].

While previous research has shown that individuals who believed in COVID-19 conspiracy theories were more prone to believe in other unrelated, broader conspiracies [27-29], it remains an open question whether those who believe in misinformation about COVID-19 pandemic are also more susceptible to believe politically ideological misinformation. This question has become pressing since the onset of the Russian invasion of Ukraine and the massive spread of disinformation aligned with Russian propaganda. Such disinformation mixes elements of strategic narratives rooted in historical revisionism, imperial mythology, and war memories with factual lies and misinterpretations, aiming to manipulate public opinion and influence political decisions in European Union (EU) and North Atlantic Treaty Organization (NATO) member states [30]. Comparisons have been drawn between the disinformation narratives related to COVID-19 pandemic and those related to the Russia-Ukraine war [14,23]. Anecdotal observations suggest that individuals sharing rigid beliefs in misinformation narratives about COVID-19 pandemic (BM-C) may have also adopted beliefs in misinformation about the Russian invasion of Ukraine (BM-U), and that they tend to use specific digital media channels while avoiding public and mainstream media and share antisystem attitudes and political orientation toward Russia [21]. However, no empirical research has examined this social phenomenon population-wide. Therefore, to validate or refute these observations, we conducted a nationwide representative cross-sectional survey of the Czech Republic.

The first aim of this study was to determine whether there is an association between respondents’ endorsement of the 2 sets of misinformation narratives (BM-C and BM-U).

- Hypothesis 1: There is a correlation between BM-C and BM-U.

The second aim was to examine associations between beliefs in the 2 sets of misinformation (BM-C and BM-U) and factors anecdotally observed or suggested in both contexts. Media monitoring and official reports have indicated that both sets of misinformation have been spreading through specific digital

XSL•FO
RenderX

media channels, such as web-based discussions and web-based bubbles or echo chambers, political chain emails, and antisystem websites with political leanings toward Russia [13,21]. However, it remains unknown whether users of these channels are significantly more likely to believe the misinformation and to trust specific geopolitical powers on a nationwide scale. Therefore, we examined associations between (2a) political trust and the 2 sets of misinformation, as well as associations between (2b) media use factors and the 2 sets of misinformation.

- Hypothesis 2a: Distrust in the Czech government's decisions and public media, trust in Russia, and distrust in Russia's geopolitical opponents and Western allies of the Czech Republic (US, EU, and NATO) are shared factors that explain both BM-C and BM-U.
- Hypothesis 2b: The use of antisystem websites, emails, and social media as information sources, along with participation in web-based discussions and engagement in web-based bubbles, explains BM-C and BM-U.

The third aim of this study was to examine whether BM-C and BM-U are connected to COVID-19 vaccine refusal. Determining that this factor explains not only BM-C but also BM-U would indicate that this specific health-related behavior significantly reflects the politicization of the COVID-19 issue to such an extent that it increased susceptibility to ideological misinformation.

- Hypothesis 3: COVID-19 vaccine refusal explains both BM-C and BM-U.

In addition, we aimed to test whether beliefs in the 2 categories of misinformation are associated with belief rigidity. The underlying assumption is that individuals who endorse misinformation place greater emphasis on the importance of these beliefs, as they often provide complex collective narratives and transcend mere opinions on specific health, societal, or political issues. Rather, they may become a belief system infused with moral convictions, which tends to be fixed and rigid [31,32]. Belief rigidity has been connected to echo chambers [8,33], conspiracy thinking [34], and polarization [31,35,36].

- Hypothesis 4: Belief rigidity explains both BM-C and BM-U.

Methods

Procedure

The data were collected from April 25 to May 5, 2022, at the time when COVID-19 pandemic had subsided and 2 months after the start of the Russian invasion of Ukraine. The cross-sectional survey was completed by members of the Czech National Panel [37] as a part of a longitudinal study [38], using the standardized computer-assisted web interviewing method. Participation was voluntary, with financial compensation. The mean completion time of the survey was approximately 11 minutes, and participants were informed in advance about the length. The survey included sociodemographic data (gender, age, level of education, region of residence, and household income), as well as questions about beliefs in misinformation regarding COVID-19 pandemic and the Russian invasion of Ukraine, media use, political trust, belief rigidity, and whether and how many times they have been vaccinated against COVID-19. Only self-reported measures were used. To ensure the protection of personal information, all collected data were securely stored in an encrypted, password-protected institutional database hosted on National Institute of Mental Health servers. Only authorized personnel had access to the data. Any personal identifiers were anonymized during data processing to prevent unauthorized access or identification of participants.

Participants

Participants of the longitudinal study [38] were invited to participate in this study. We received responses from 1623 respondents (return rate: 55% of 2950 invited; 839/1623, 51.7% women) aged between 20 and 91 years (mean 55.04, SD 15.55). The proportions of participants' attained educational levels were as follows: 4.6% (76/1623) elementary school education, 29.1% (472/1623) certificate of apprenticeship, 36.2% (587/1623) high school education, and 30.2% (490/1623) university degree. The sample was constructed to be quota-representative of the adult population of the Czech Republic. To ensure repeated participation of various sociodemographic groups, it was necessary to adjust the current sample through poststratification weighting. This adjustment was based on current population distributions (using data from the Czech Statistical Office) for the following characteristics: gender, age, education, size of place of residence, region, crosscutting of age and education, crosscutting of age and gender, and employment status. The inclusion criteria were knowledge of the Czech language and being older than 18 years.

Measures

Beliefs in Misinformation Narratives

To measure BM-C and BM-U, we developed 2 questionnaires. The questionnaires were constructed based on the main misinformation related to COVID-19 published by the Center Against Hybrid Threats within the Ministry of the Interior of the Czech Republic [39]. The Ministry reported that such narratives had been spread in an attempt to exploit societal issues in accordance with the interests of foreign powers. We reduced the number of items from the original 15 to 6 based on results from our pilot study (N=423), excluding items according to item analysis, exploratory factor analysis (EFA), and the results of the Cronbach α coefficient. BM-C items are shown in Textbox 1. Similarly, the BM-U questionnaire was constructed, using the prevalent misinformation narratives related to the Russian invasion in Ukraine at the time of the study [40]. We selected 4 items from the original 8 based on pilot data according to the same procedure as in BM-C. BM-U items are shown in Textbox 1. Both questionnaires showed good internal consistency in both the pilot study (BM-C: Cronbach α=0.953; BM-U: Cronbach α=0.932) and in this study (BM-C: Cronbach α=0.846; BM-U: Cronbach α=0.891). Participants rated the items on a 5-point scale (1: "I do not agree at all"—5: "I completely agree").

XSL•FO
RenderX

Textbox 1. Items for beliefs in misinformation narratives (beliefs in misinformation narratives about COVID-19 pandemic [BM-C] and beliefs in misinformation about the Russian invasion of Ukraine [BM-U]).

Evaluate the extent to which you agree or disagree with the following statements.

BM-C

- "Western pharmacological vaccine companies are untrustworthy."
- "Vaccines are dangerous for the vaccinated."
- "The discrimination against Russian and Chinese vaccines is largely driven by political reasons."
- "The coronavirus was developed artificially, perhaps as a biological weapon."
- "The epidemic is fake, the situation has never been so serious."
- "Epidemic measures were ineffective and were counterproductive."

BM-U

- "The demilitarisation and de-Nazification of Ukraine is a legitimate objective for the Russian military operation in Ukraine."
- "The civilian casualties on the Ukrainian side are deliberately exaggerated by the European media."
- "Ukraine is developing banned biological weapons on its territory."
- "NATO and Western countries are exploiting Ukraine to serve their own interests."

COVID-19 Vaccination

Participants were asked whether and how many times they had been vaccinated against COVID-19 (0, 1, 2, or 3 times). It should be noted that at the time of the survey, the Ministry of Health of the Czech Republic recommended 3 doses of the vaccine.

Media Use

We used an adapted version of the media use questionnaire [41]. We omitted some items and included additional ones, while also rewording some items to better suit the research objectives of measuring media behavior and media effects that may be indicative of or contribute to the spread of misinformation. To compare responses to the 2 societal issues, we used identical wording for questions related to the COVID-19 pandemic (C), and the Russian invasion of Ukraine (U), with only a difference in the topic and time frame being questioned (eg, “How often did you search for news regarding COVID-19 at the height of the pandemic?” or “How often did you search for news on the Russian invasion of Ukraine last month?”). The mirrored items were placed in different locations within the questionnaire and never in sequence. The newly developed measures were tested in a pilot survey conducted via Facebook in April 2022 (N=423; response rate: 51.8% of 817 invited). Respondents were asked about their frequency of use of media channels categorized as public media, mainstream news websites [42], and those that have been previously connected to spreading misinformation: emails as a source of information (possibly indicating political chain emails), YouTube, social media, and “anti-system websites” that have been identified as such by various media experts [12,42]. However, at the time of our survey, in reaction to the Russian invasion of Ukraine and the uncertain development of the situation, most of the antisystem websites were evaluated as a threat to national security and were officially banned in the Czech Republic due to their open promotion of Russian disinformation narratives. Only 1 functioning, moderate news website, remained in our survey. Participants were also asked about their engagement in web-based discussions and web-based bubbles related to C/U. Furthermore, we decided to examine several other aspects of media use—searching and sharing the news (C/U), respondents’ interest in the 2 topics (C/U), and their frustration with public and mainstream media.

Political Trust

Perceptions of trust in the (1) Czech government and (2) public media were assessed in relation to both issues (C/U). Due to the high correlation of items 1 (C) and 2 (C) (r=0.783, n=1623; P<.001), as well as items 1 (U) and 2 (U) (r=0.849, n=1623; P<.001), we summed the items in 1 score for each topic: trust in the Czech government and public media regarding COVID-19 (*Trust in CZ-C*); trust in the Czech government and public media regarding Russian invasion of Ukraine (*Trust in CZ-U*). In addition, distrust in foreign geopolitical actors (Russia, United States, China, EU, and NATO) and belief rigidity was assessed. Detailed descriptions of the survey items and response scales for media use, political trust, and belief rigidity are shown in Multimedia Appendix 1.

Statistical Analysis

All data were analyzed using R software (R Core Team). The significance level was set at P≤.05. Poststratification weighting was applied using a quadratic programming algorithm based on current population distributions of the following characteristics: gender, age, education, region, residence size, job status, interaction between age and education, and interaction between age and gender. Descriptive statistics were used for demographic description. Shapiro-Wilk test did not confirm the normal distribution of BM-C and BM-U. EFA was conducted on both BM-C and BM-U items to uncover the latent structure based on interdependence between the items. The primary aim of the EFA was to clearly differentiate COVID-related and ideological items, ensuring that the correlation between BM-C and BM-U scales is not influenced by the ideological items possibly present in BM-C.

XSL•FO
RenderX

As the data were nonparametric, we used Spearman correlation to determine the relationship between BM-C and BM-U (Hypothesis 1). Multiple linear regression models were used to reveal the relationships between the examined factors according to Hypotheses 2-4 (COVID-19 vaccine refusal, media use, political trust, and belief rigidity) and beliefs in BM-C and BM-U. For the multiple linear regression models, we used normalization of nonparametric right-skewed data by square root. Two distinct models were constructed, 1 for BM-C and 1 for BM-U (dependent variables), with COVID-19 vaccine refusal, media use, political trust, and belief rigidity as independent variables. We also controlled for demographic characteristics (age, gender, education, and income). To compare the predictive power of the independent variables, we used a feature scaling approach. Specifically, we used normalization to standardize all continuous input variables to a uniform range of 1-5. This step guarantees comparability and stability in the regression analysis, establishing a standardized input space for the model and enabling the evaluation of the effect of each variable. However, categorical variables were maintained in their original scale to preserve their interpretability and intrinsic categorical distinctions.

Ethical Considerations

The procedure performed in this study was in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments. The study was approved by the ethics committee of the National Institute of Mental Health, Czech Republic (reference no. 181/21). The data were anonymized. Respondents were compensated by the Czech National Panel at a standard rate of 1 CZK (US $0.041) per minute for completing the questionnaire. The compensation was provided as credit, which could be transferred to a bank account, redeemed for a material reward, or donated to charity. In addition, 2 randomly selected participants had the chance to win a tablet. All participants provided informed consent. They were informed about the purpose of the study. Furthermore, they were informed that the data would be accessible only to authorized research staff and the principal investigator, whose name and contact information were provided for any follow-up questions or concerns. Participants were assured that their participation was voluntary.

Results

Exploratory Factor Analysis

EFA was conducted using parallel analysis to identify the underlying structure of the BM-C and BM-U items. Two factors were extracted, explaining 58.7% of the total variance, with factor 1 accounting for 32.4% of the variance and factor 2 accounting for 26.3%. The overall Kaiser-Meyer-Olkin measure of sampling adequacy was 0.91, indicating that the data were highly suitable for factor analysis. The Bartlett sphericity test (χ^2_{45}=9338; *P*<.001) further confirmed the appropriateness of conducting EFA. An Oblimin rotation was applied to enhance interpretability, allowing for correlations between factors. The first factor, labeled "Ideological," included all BM-U items and BM-C item 3 ("The discrimination against Russian and Chinese vaccines is largely driven by political reasons"). The second factor, labeled "COVID," comprised all remaining BM-C items (except item 3). Due to its significant loading on the ideological factor and theoretical considerations, BM-C item 3 was excluded from further analysis. Factor loadings are shown in Multimedia Appendix 2.

Correlation Between BM-C and BM-U and Descriptive Statistics for BM-C and BM-U

A moderate positive correlation was found between BM-C and BM-U (Spearman ρ=0.57; *P*<.001). For a more straightforward description of BM-C and BM-U, we considered 4 points ("I rather agree") and 5 points ("I completely agree") as an indication of belief in misinformation (*supporters*). Those who rated 3 points ("I neither agree nor disagree") were considered undecided whether they believe in misinformation or not (*undecided*). Those who rated 1 ("I completely disagree") or 2 ("I rather disagree") were considered *opponents* who do not endorse misinformation narratives. According to this grouping based on cumulative scores, the prevalence of BM-C supporters was 13.4% (217/1623), and the prevalence of BM-U supporters was 9.9% (161/1623). There were 50% (812/1623) of undecided respondents for BM-C and 31.1% (505/1623) for BM-U. The demographic description showed that supporters in BM-C were most represented in apprenticeship education degree (88/217, 41%), followed by high school degree (77/217, 36%) and university education level (42/217, 19%), with lowest numbers in elementary education level (10/217, 5%). BM-C opponents were most prevalent in the university education level (243/594, 40.9%). Supporters of BM-U were most prevalent in apprenticeship education level (60/505, 37%), followed by high school degree (54/161, 34%) and university degree (39/141, 24%). BM-U opponents were most prevalent in high school (357/957, 37.3%) and university education (342/957, 35.7%), followed by apprenticeship education (223/957, 23.3%). Overall, supporters and undecided both for BM-C and BM-C were less prevalent in the university education level and more in the apprenticeship education level compared with nonsupporters. Regarding household income, supporters and undecided (both for BM-C and BM-C) were represented less in the high-income group and more in the below poverty line income group compared with opponents. In terms of gender, noticeable differences were found in the undecided groups, particularly in BM-U, with female participants representing a higher proportion (304/505, 60.2%). Conversely, male participants were more prevalent among BM-U supporters (98/161, 61%). Differences in age compared with an average of the whole sample (mean 55.04, SD 15.56) were observed only in BM-U supporters, who were older (60.9 years), and BM-U opponents, who were younger (49.3 years). Vaccine refusers were minimally represented in BM-C opponents (33/594, 6%), more in BM-C undecided (146/812, 18%), and most in BM-C supporters (142/217, 65%). Moreover, 34.6% (75/217) of BM-C supporters were vaccinated despite their beliefs. Regarding BM-U, vaccine refusers were most represented in BM-U supporters (64/161, 40%), followed by BM-U undecided (128/505, 25.3%), with the lowest numbers in BM-U opponents (129/957, 13.5%; Figure 1). The descriptive statistics are shown in Table 1.

XSL•FO
RenderX

Figure 1. Distribution of unvaccinated (shown in red) and vaccinated (shown in cyan) against COVID-19 pandemic in relation to beliefs in misinformation regarding COVID-19 (BM-C) and the war in Ukraine (BM-U). The x-axis represents BM-C total score, and the y-axis represents BM-U total score.

Table 1. Sociodemographic characteristics of BM-C[a] and BM-U[b].

Sociodemographic variables	Opponents BM-C	Undecided BM-C	Supporters BM-C	Opponents BM-U	Undecided BM-U	Supporters BM-U
Male + female, n (%)	594 (36.6)	812 (50)	217 (13.4)	957 (59)	505 (31.1)	161 (9.9)
Female, n (%)	267 (44.9)	459 (56.5)	113 (52.1)	472 (49.3)	304 (60.2)	63 (39.1)
Male, n (%)	327 (55.1)	353 (43.5)	104 (47.9)	485 (50.7)	201 (39.8)	98 (60.9)
Age (years), mean (SD)	55.02 (16.62)	55.25 (15.11)	54.31 (14.16)	53.25 (15.19)	56.58 (14.89)	60.80 (13.28)
Elementary education, n (%)	17 (2.9)	47 (5.8)	10 (4.6)	35 (3.7)	31 (6.1)	8 (5)
Apprenticeship education, n (%)	119 (20)	265 (32.6)	88 (40.6)	223 (23.3)	189 (37.4)	60 (37.3)
High school education, n (%)	215 (36.2)	295 (36.3)	77 (35.5)	357 (37.3)	176 (34.9)	54 (33.5)
University education, n (%)	243 (40.9)	205 (25.2)	42 (19.4)	342 (35.7)	109 (21.6)	39 (24.2)
Income 1[c], n (%)	32 (5.4)	57 (7)	26 (12)	55 (5.7)	41 (8.1)	19 (11.8)
Income 2[d], n (%)	164 (27.6)	268 (33)	81 (37.3)	289 (30.2)	175 (34.7)	49 (30.4)
Income 3[e], n (%)	231 (38.9)	325 (40)	88 (40.6)	369 (38.6)	211 (41.8)	64 (39.8)
Income 4[f], n (%)	167 (28.1)	162 (20)	22 (10.1)	244 (25.5)	78 (15.4)	29 (18)
Vaccinated, n (%)	561 (94.4)	666 (82)	75 (34.6)	828 (86.5)	377 (74.7)	97 (60.2)
Unvaccinated, n (%)	33 (5.6)	146 (18)	142 (65.4)	129 (13.5)	128 (25.3)	64 (39.8)

[a]BM-C: beliefs in misinformation narratives about COVID-19 pandemic.
[b]BM-U: beliefs in misinformation about the Russian invasion of Ukraine.
[c]Below poverty line income (below 60% of the median).
[d]Low income (below the median).
[e]Upper middle income (up to 1.5 times the median).
[f]High income (above 1.5 times the median).

Factors Explaining BM-C

The multiple linear regression model explained 44.92% of the individual differences in BM-C ($F_{30,1592}$=45.1; adjusted R^2=0.45; P<.001). Descriptions of variables used in the BM-C model are shown in Multimedia Appendix 3. The results showed significant relationships between the 12 examined factors as the independent variables and BM-C total score as the dependent

XSL•FO
RenderX

variable (Table 2). Trust in the Czech government and public media, vaccination against COVID-19 pandemic, distrust in Russia, searching for news on COVID-19 pandemic, and participation in web-based discussions predicted lower levels of BM-C. Distrust in the United States, distrust in the EU, frustration from public and mainstream news, rigid beliefs, use of emails as a source of information, sharing COVID-19 news, and engagement in web-based bubbles predicted higher levels of BM-C. Regarding demographic factors, upper middle income (compared with high income), as well as elementary, apprenticeship, and high school education levels (compared with university education level) were associated with increased BM-C. The below poverty line income group (compared with high income) predicted lower levels of BM-C.

XSL•FO
RenderX

Table 2. The results of multiple linear regression models for BM-C[a] and BM-U[b,c].

Explaining variable	BM-C, coefficient (SE)	BM-C, *t* test (*df*)	BM-C, *P* value	BM-U, coefficient (SE)	BM-U, *t* test (*df*)	BM-U, *P* value
Intercept	2.11 (0.25)	8.48 (1592)	*<.001*	2.30 (0.28)	8.24 (1592)	*<.001*
COVID-19 vaccination	–0.17 (0.01)	–12.93 (1592)	*<.001*	–0.02 (0.01)	–1.55 (1592)	.12
Information from emails	0.04 (0.02)	2.05 (1592)	*.04*	0.04 (0.02)	2.13 (1592)	*.03*
YouTube	0.03 (0.02)	1.79 (1592)	*.07*	–0.02 (0.02)	–1.26 (1592)	.21
Antisystem websites	0.03 (0.02)	1.50 (1592)	.13	0.06 (0.02)	2.92 (1592)	*.004*
Public media	0.03 (0.02)	1.84 (1592)	.07	–0.02 (0.01)	–1.37 (1592)	.17
Mainstream websites	0.02 (0.02)	0.97 (1592)	.33	–0.04 (0.02)	–2.52 (1592)	*.01*
Exposure to social media	–0.03 (0.06)	–0.56 (1592)	.58	–0.07 (0.05)	–1.42 (1592)	.16
Social media information source	–0.001 (0.02)	–0.06 (1592)	.95	0.03 (0.02)	2.01 (1592)	*.045*
Discussions under news	0.02 (0.02)	1.06 (1592)	.29	0.05 (0.02)	2.68 (XX)	*.007*
Discussions on social media	0.32 (0.10)	3.18 (1592)	*.002*	–0.02 (0.15)	–0.12 (1592)	.90
Web-based bubbles	0.17 (0.06)	3.10 (1592)	*.002*	0.03 (0.07)	0.42 (1592)	.68
Search for news	–0.1 (0.03)	–3.81 (1592)	*<.001*	–0.06 (0.02)	–2.30 (1592)	*.02*
Sharing news	0.07 (0.03)	2.47 (1592)	*.01*	0.01 (0.03)	0.48 (1592)	.63
Interest in news	0.01 (0.02)	0.76 (1592)	.45	–0.03 (0.02)	–1.59 (1592)	.11
Frustration from media	0.12 (0.02)	6.27 (1592)	*<.001*	0.10 (0.02)	5.42 (1592)	*<.001*
Trust in Czech government	–0.22 (0.02)	–10.34 (1592)	*<.001*	–0.23 (0.02)	–11.81 (1592)	*<.001*
Distrust in Russia	–0.07 (0.02)	–2.94 (1592)	*.003*	–0.24 (0.02)	–10.26 (1592)	*<.001*
Distrust in United States	0.08 (0.03)	2.33 (1592)	*.02*	0.20 (0.03)	6.63 (1592)	*<.001*
Distrust in EU[d]	0.12 (0.04)	2.86 (1592)	*.004*	0.10 (0.04)	2.65 (1592)	*.008*
Distrust in China	0.02 (0.03)	0.82 (1592)	.41	–0.05 (0.02)	–0.22 (1592)	.82
Distrust in NATO[e]	–0.05 (0.04)	–1.20 (1592)	.23	0.07 (0.04)	1.81 (1592)	.07
Rigid beliefs	0.09 (0.02)	5.07 (1592)	*<.001*	0.08 (0.02)	4.99 (1592)	*<.001*
Income 1 (below poverty line)[f]	–0.18 (0.08)	–2.35 (1592)	*.02*	0.14 (0.07)	2.00 (1592)	*.046*
Income 2 (low)[f]	0.09(0.05)	1.75 (1592)	.08	0.07 (0.05)	1.42 (1592)	.16
Income 3 (upper middle)[f]	0.10 (0.05)	2.04 (1592)	*.04*	0.07 (0.05)	1.48 (1592)	.14
Elementary education[g]	0.31 (0.07)	4.37 (1592)	*<.001*	0.20 (0.07)	3.00 (1592)	*.003*
Apprenticeship education[g]	0.22 (0.05)	3.97 (1592)	*<.001*	0.06 (0.05)	1.14 (1592)	.25
High school education[g]	0.17 (0.05)	3.40 (1592)	*<.001*	0.02 (0.05)	0.32 (1592)	.75
Gender (female)[h]	–0.03 (0.04)	–0.89 (1592)	.37	–0.03 (0.04)	–0.81 (1592)	.42
Age (year)	0.001 (0.001)	1.01 (1592)	.31	0.07 (0.02)	3.23 (1592)	*<.001*

[a]BM-C: beliefs in misinformation narratives about COVID-19 pandemic.

[b]BM-U: beliefs in misinformation about the Russian invasion of Ukraine.

[c]Significant values are italicized.

[d]EU: European Union.

[e]NATO: North Atlantic Treaty Organization.

[f]Contrasted to high-income group.

[g]Contrasted to university degree.

[h]Contrasted to male.

XSL•FO
RenderX

Factors Explaining BM-U

The multiple regression model explained 62.21% of the variance in BM-U ($F_{30,1591}$=90.01; adjusted R^2=0.62; P<.001). Descriptions of variables used in the BM-U model are shown in Multimedia Appendix 3. We found significant relationships between the 12 examined factors as independent variables and BM-U total score as the dependent variable (Table 2). Trust in the Czech government and public media, distrust in Russia, consumption of mainstream news websites, and searching for news about the war in Ukraine predicted lower levels of BM-U. Conversely, distrust in the United States, distrust in the EU, frustration from public and mainstream news, consumption of "antisystem websites," use of emails as a source of information, use of social media as an information source, reading discussions under web news articles, and belief rigidity predicted higher levels of BM-U. Regarding demographic factors, below poverty line income (compared with high income), elementary education level (compared with university education level), and older age were associated with higher levels of BM-U.

Discussion

Principal Findings

Our study provides evidence of a connection between beliefs in COVID-19 misinformation (BM-C) and misinformation regarding the Russian invasion of Ukraine (BM-U) by identifying a correlation between these 2 sets of beliefs and several shared factors. Regarding political trust, higher trust in Russia and lower trust in local government, public media, and Western allies of the Czech Republic (the EU and the United States) were revealed as strong predictors of both BM-C and BM-U. In addition, frustration with public and mainstream media, using emails as a source of information—possibly indicating chain emails—and reduced frequency in searching for news related to COVID-19 pandemic or war in Ukraine, predicted both BM-C and BM-U. We also identified media use patterns commonly associated with the spread of misinformation, which predicted either BM-C or BM-U. These included participation in web-based bubbles, engagement in discussions under web news articles, use of antisystem websites, avoidance of mainstream media, use of social media as an information source, and sharing news. In addition, belief rigidity was a significant predictor for both BM-C and BM-U.

Correlation Between BM-C and BM-U

A moderate positive correlation discovered between BM-C and BM-U supports our hypothesis, indicating that a significant number of individuals believing in COVID-19 misinformation have also adopted ideological misinformation regarding the Russian invasion of Ukraine. This extends previous findings that beliefs in COVID-19 conspiracies correlate with beliefs in other, broader and unrelated conspiracies [27,28] to the politicized side of COVID-19 misinformation, which increased susceptibility to ideological misinformation aligning with Russian propaganda. Our finding provides further evidence for the so-called "conspiracy singularity" [43] suggesting the tendency of actors to spread and interconnect various conspiracy theories [44,45]. For instance, the same actors who spread COVID-19 conspiracies before the Russian invasion of Ukraine later disseminated anti-NATO and pro-Russian narratives in Finland [46] and Slovakia [47]. Our findings thus corroborate similar phenomena observed beyond the context of the Czech Republic and may provide further insights into the mechanisms by identifying underlying factors revealed in our analyses, which are discussed in the sections "Associations of Political Trust and Beliefs in Misinformation," "Associations of Media Use and Beliefs in Misinformation Narratives," "COVID-19 Vaccine Refusal," and "Belief Rigidity."

Associations of Political Trust and Beliefs in Misinformation

Our finding that lowered trust in governmental decisions and public media was associated with both increased BM-C and BM-U supported our hypothesis. Moreover, it was the strongest predictor explaining both BM-C and BM-U. It is in line with previous research linking distrust in public institutions to COVID-19 misinformation beliefs [48-51]. While most previous findings on associations between beliefs in COVID-19 misinformation and political attitudes report that conservatism is associated with increased susceptibility to misinformation [52-54], we did not inquire about partisanship but rather about trust in geopolitical powers. Our results showing increased trust in Russia in higher levels of both BM-C and BM-U indicate a leaning toward this geopolitical power in supporters of both sets of misinformation. In addition, we observed increased distrust toward the Czech Republic's geopolitical allies and Russia's main opponents—the United States and the EU—among individuals with higher levels of both BM-C and BM-U. While this ideological inclination is not surprising regarding BM-U, which openly promotes Russian propaganda, it is not as readily apparent in the case of BM-C. However, our result aligns with previous research that has suggested the role of Russian disinformation campaigns in supporting the antivaccination movement [18,55,56].

Our findings can thus be contextualized in light of the goals of Russia's hybrid war strategy, which aims to continually undermine the trustworthiness and legitimacy of foreign governments in the eyes of the target population by warping their beliefs, thoughts, decisions, and behavior over the long term [57]. The goal of this tactic is to gradually reconstruct the target population's prior beliefs in favor of Russia [58,59]. However, our study cannot establish a causal relationship in terms of direct influence of Russia's disinformation campaigns. The inclination toward Russia may also have deep historical roots, as the Czech Republic—former Czechoslovakia—was part of the Eastern Bloc under the direct influence of the Soviet Union for 4 decades. Increased trust in Russia may also represent an alternative to the current Western orientation of the Czech Republic as a member of the EU and NATO, reflecting a broader, socially driven epistemic mistrust that manifests in the rejection of authoritative information, as suggested by the socioepistemic model of belief in conspiracy theories [60].

Associations of Media Use and Beliefs in Misinformation Narratives

All of the identified media use factors linked to either BM-C or BM-U provided support for our hypothesis regarding media

XSL•FO
RenderX

use, formulated based on previous observations and theoretical or empirical associations with the dissemination of misinformation. However, it is noteworthy that not all of the examined factors demonstrated significant relationships with both BM-C and BM-U. The strongest media factor associated with higher levels of both beliefs was identified as frustration with the public and mainstream media. While previous research has established this factor as a predictor of higher anxiety and depression levels during the COVID-19 pandemic [41], our study extends its relevance to the context of misinformation susceptibility. This observation is complemented by another finding, which links less frequent searches for COVID-19 news with higher BM-C levels, and less frequent consumption of mainstream media and searches for the news about the war in Ukraine with BM-U. These findings align with previous research [1,61,62] and suggest that supporters of misinformation narratives engage in avoidance behavior, possibly due to their mistrust in information they perceive as misrepresented in public and mainstream media.

On the other hand, supporters of BM-C and BM-U showed higher engagement with other media channels. Specifically, there was an association between obtaining news information from emails—possibly indicating chain emails—and both BM-C and BM-U. In addition, reading discussions under web news articles and consuming information from antisystem websites was positively associated with BM-U. These findings corroborate observations regarding the role of such media channels in disseminating misinformation content and the susceptibility of their consumers to misinformation [13,20].

Next, the positive relationship between obtaining information from social media and increased BM-U, as well as the association between engagement in web-based bubbles and increased BM-C, indicates that the social media environment contributed to the spread of misinformation and their users' endorsement, as suggested by previous research [1,51,63-66]. While we acknowledge the limitations of the web-based survey method in assessing the phenomenon of web-based (epistemic) bubbles or echo chambers, it is plausible to assume that this phenomenon may have indeed been reflected in our results, as it aligns with prior findings [8-10,64].

Conversely, the negative relationship of engagement in discussions on social media and BM-C, as well as the lack of discernable associations between cumulative exposure to social media and BM-C/BM-U, underscores the reductive conclusions of associating social media platforms solely with the spread of misinformation. Indeed, social media offers users engagement in socializing and discussing a diverse array of content, as well as a broad spectrum of viewpoints on sociopolitical issues. Notably, in the context of nondemocratic regimes, digital media often serves as a primary source of obtaining reliable information. Research in nondemocratic regimes indicates that the use of digital media correlates with diminished adherence to misinformation, contrasting with users reliant solely on official information channels [67].

Our next finding of a positive association between sharing news and heightened levels of BM-C indicates that BM-C supporters demonstrated a propensity for active engagement with digital media. Speculatively, this could be due to heightened arousal triggered by specific content, frustration, or a sense of moral obligation to disseminate the alternative information on social media, perceived as accurate, compared with information reported by public and mainstream media, perceived as misleading or incomplete [68]. This inference is drawn from previous research indicating that the perceived accuracy of content significantly influences the likelihood of its sharing by users [69]. While our study did not directly explore the specific content shared by respondents, it is pertinent to note that previous studies have demonstrated that misinformation tends to be inherently more frequently shared than other types of news [69].

COVID-19 Vaccine Refusal

Our finding that vaccine refusal was a strong factor associated with BM-C supports our hypothesis and aligns with extensive prior research linking exposure to COVID-19 misinformation to COVID-19 vaccine hesitancy [48,62,70-73]. Our finding provides further evidence that COVID-19 vaccine refusal is a behavioral indicator of diverse attitudes that transcend medical concerns. However, it is important to note that 34.6% of BM-C supporters (75/217) reported being vaccinated, indicating a divergence from their beliefs. They may ultimately yield to social pressure and decide to get vaccinated, considering the practical difficulties posed by remaining unvaccinated in their daily lives during the pandemic.

Contrary to our hypothesis, COVID-19 vaccine refusal was not associated with BM-U, suggesting that this health-related behavior is a broader phenomenon that includes vaccine hesitancy due to health reasons, medical concerns, simple reluctance, and other factors. We conclude that vaccine refusal should not lead to the reductionist conclusion that COVID-19 vaccination was entirely politicized. However, we observed a higher prevalence of vaccine refusers in BM-U supporters (64/161, 39.8%), followed by BM-U undecided (128/505, 25.3%), with the lowest numbers in BM-U opponents (129/957, 13.5%). Special attention should be given to the BM-U undecided group, requiring longitudinal monitoring to assess whether they might become new adherents of BM-U.

Belief Rigidity

Our additional finding of a positive association between the rigidity of one's beliefs regarding sociopolitical issues with both BM-C and BM-U indicates that those who adhere to the alternative interpretations of both sociopolitical issues tend to harbor more fixed and rigid opinions than those who do not support such interpretations. Our finding is consistent with previous studies connecting belief rigidity to conspiratorial thinking [34] and beliefs in misinformation propagated through social media [74]. Rigid beliefs have been found to facilitate group cohesion, partisanship, polarization, and extremism [31,35,75]. It is thus plausible that beliefs such as BM-C or BM-U may serve as a group-shared alternative "truth" while being shared through the digital media environment as identified in our analysis. Furthermore, it is in line with our other finding (discussed in the "Associations of Media Use and Beliefs in Misinformation Narratives" section) indicating avoidance of public and mainstream information sources. This pattern is

XSL•FO
RenderX

consistent with previous research suggesting that belief rigidity is strengthened when individuals isolate themselves from contradictory information, thus reinforcing their confirmation bias [10].

Conclusions

Our findings support the hypothesis that individuals who endorsed COVID-19 misinformation were more susceptible to ideological misinformation, aligning with Russian propaganda. Supporters of both misinformation narratives shared common traits, including heightened distrust of local government, public media, the United States, and the EU, along with increased trust toward Russia. They also exhibited increased belief rigidity and demonstrated several common media use patterns, previously linked to the spread of misinformation. To gain a deeper understanding of these phenomena, longitudinal monitoring is essential. By tracking the development of BM-C, BM-U, and the examined factors over time, causal relationships can be uncovered.

Limitations

The primary shortcoming of this study was the constraint imposed by the short survey format. Due to time limitations, it was not feasible to use longer standardized questionnaires such as the Belief Rigidity Scale. Instead, we opted for a single statement specifically related to societal issues, such as politics, war, and pandemics, and we considered this finding as supplementary. On the other hand, we chose to investigate media use in more detail with practical implications in mind, aiming to identify specific media channels where misinformation is prevalent for targeted recommendations. However, some aspects of the media environment, such as web-based communities with an echo chamber effect and chain emails, were challenging to assess via survey. Consequently, our findings regarding these information sources should be interpreted with caution. In addition, while we acknowledge the availability of standardized COVID-19 conspiracy or misinformation scales, our objective was to study COVID-19 misinformation prevalent in the local context of the Czech Republic as identified by previous analytical sources.

Acknowledgments

The authors thank colleagues from PAQ Research, especially Daniel Prokop, for the opportunity to participate in their panel research and sharing sociological data, and Michaela Röschová for her valuable assistance with the technical aspects of the survey and data management. This work was supported by Czech Science Foundation (grant no. 20-13458S) and Ministry of Health of the Czech Republic (grant NU22-D-132).

Data Availability

The dataset generated during this study is available in the OSF data repository (osf.io/wtuqj).

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1

Items and response scales for variables on media use, political trust, and belief rigidity.

[DOCX File , 21 KB - infodemiology_v5i1e62913_app1.docx]

Multimedia Appendix 2

Factor loadings of BM-C and BM-U items resulting from exploratory factor analysis.

[DOCX File , 23 KB - infodemiology_v5i1e62913_app2.docx]

Multimedia Appendix 3

Descriptions of variables used in multiple linear regression models for BM-C and BM-U.

[DOCX File , 23 KB - infodemiology_v5i1e62913_app3.docx]

References

1. De Coninck D, Frissen T, Matthijs K, d'Haenens L, Lits G, Champagne-Poirier O, et al. Beliefs in conspiracy theories and misinformation about COVID-19: comparative perspectives on the role of anxiety, depression and exposure to and trust in information sources. Front Psychol 2021;12:646394 [FREE Full text] [doi: 10.3389/fpsyg.2021.646394] [Medline: 33935904]
2. Sherman SM, Smith LE, Sim J, Amlôt R, Cutts M, Dasch H, et al. COVID-19 vaccination intention in the UK: results from the COVID-19 vaccination acceptability study (CoVAccS), a nationally representative cross-sectional survey. Hum Vaccin Immunother 2021;17(6):1612-1621 [FREE Full text] [doi: 10.1080/21645515.2020.1846397] [Medline: 33242386]
3. Lee SK, Sun J, Jang S, Connelly S. Misinformation of COVID-19 vaccines and vaccine hesitancy. Sci Rep 2022;12(1):13681 [FREE Full text] [doi: 10.1038/s41598-022-17430-6] [Medline: 35953500]

XSL•FO
RenderX

4. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
5. Borges do Nascimento IJ, Beatriz Pizarro A, Almeida J, Azzopardi-Muscat N, André Gonçalves M, Björklund M, et al. Infodemics and health misinformation: a systematic review of reviews. Bull World Health Organ 2022;100(9):544-561. [doi: 10.2471/blt.21.287654]
6. Bolsen T, Palm R. Politicization and COVID-19 vaccine resistance in the U.S. Prog Mol Biol Transl Sci 2022;188(1):81-100 [FREE Full text] [doi: 10.1016/bs.pmbts.2021.10.002] [Medline: 35168748]
7. Levinsson A, Miconi D, Li Z, Frounfelker RL, Rousseau C. Conspiracy theories, psychological distress, and sympathy for violent radicalization in young adults during the COVID-19 pandemic: a cross-sectional study. Int J Environ Res Public Health 2021;18(15):7846 [FREE Full text] [doi: 10.3390/ijerph18157846] [Medline: 34360139]
8. Nguyen CT. Echo chambers and epistemic bubbles. Episteme 2018;17(2):141-161. [doi: 10.1017/epi.2018.32]
9. Gkinopoulos T, Truelsen Elbæk C, Mitkidis P. Morality in the echo chamber: the relationship between belief in COVID-19 conspiracy theories and public health support and the mediating role of moral identity and morality-as-cooperation across 67 countries. PLoS One 2022;17(9):e0273172 [FREE Full text] [doi: 10.1371/journal.pone.0273172] [Medline: 36070244]
10. Jiang B, Karami M, Cheng L, Black T, Liu H. Mechanisms and attributes of echo chambers in social media. arXiv Preprint posted online on 2021.06.09. [doi: 10.48550/arXiv.2106.05401]
11. Czech Elves (Cesti Elfove). Monitoring of social media—1/2 May 2021. 2021. URL: https://cesti-elfove.cz/monitor-socialnich-siti/ [accessed 2022-11-07]
12. Štětka V, Mazák J, Vochocová L. “Nobody tells us what to Write about”: the disinformation media ecosystem and its consumers in the Czech Republic. Javnost Public 2020;28(1):90-109. [doi: 10.1080/13183222.2020.1841381]
13. Czech Security Information Service. Annual report of Czech Security Information Service for 2021 (Vyrocni zprava 2021). 2022 Oct 17. URL: https://www.bis.cz/public/site/bis.cz/content/vyrocni-zpravy/2021-vz-cz-2.pdf [accessed 2022-11-24]
14. European Digital Media Observatory. From the pandemic to the war in Ukraine: One year and a half of countering disinformation. 2023. URL: https://edmo.eu/wp-content/uploads/2022/09/English.pdf [accessed 2024-10-20]
15. King KK, Wang B. Diffusion of real versus misinformation during a crisis event: a big data-driven approach. Int J Inform Manage 2023;71:102390. [doi: 10.1016/j.ijinfomgt.2021.102390]
16. Herasimenka A, Au Y, George A, Joynes-Burgess K, Knuutila A, Bright J, et al. The political economy of digital profiteering: communication resource mobilization by anti-vaccination actors. J Commun 2023;73(2):126-137 [FREE Full text] [doi: 10.1093/joc/jqac043] [Medline: 37016634]
17. Moy W, Kacper G. COVID-19 effects and Russian disinformation campaigns. Homeland Security Affairs. 2020. URL: https://www.hsaj.org/articles/16533 [accessed 2023-11-20]
18. Patel SS, Moncayo OE, Conroy KM, Jordan D, Erickson TB. The landscape of disinformation on health crisis communication during the COVID-19 pandemic in Ukraine: hybrid warfare tactics, fake media news and review of evidence. JCOM J Sci Commun 2020;19(5):A02 [FREE Full text] [doi: 10.22323/2.19050202] [Medline: 34504624]
19. Wenzel M, Stasiuk-Krajewska K, Macková V, Turková K. The penetration of Russian disinformation related to the war in Ukraine: evidence from Poland, the Czech Republic and Slovakia. Int Polit Sci Rev 2023;45(2):192-208. [doi: 10.1177/01925121231205259]
20. Czech Elves (Cesti Elfove). Fighters against foreign disinformation campaigns on the Czech internet. Monthly Report—February, March, April 2022. URL: https://cesti-elfove.cz/mesicni-report-unor-2022 [accessed 2022-11-08]
21. Ministry of the Interior of the Czech Republic. Manifestations of extremism and prejudicial hatred, situation report H2 2022 (Projevy extremismu a predsudecne nenavisti. Souhrnna situacni zprava 2. pololeti roku 2022). 2023 Jan 31. URL: https://www.mvcr.cz/clanek/pololetni-zprava-mv-o-extremismu-211168.aspx [accessed 2023-02-02]
22. Czech Elves. Disinformation on Facebook—February. 2022. URL: https://cesti-elfove.cz/monitoring-dezinfo-facebook-unor-2022/ [accessed 2022-11-19]
23. Sánchez del Vas R, Tuñón Navarro J. Disinformation on the COVID-19 pandemic and the Russia-Ukraine war: two sides of the same coin? Humanit Soc Sci Commun 2024;11(1):851. [doi: 10.1057/s41599-024-03355-0]
24. Magallón-Rosa R, Fernández-Castrillo C, Garriga M. Fact-checking in war: types of hoaxes and trends from a year of disinformation in the Russo-Ukrainian war. EPI 2023;32(5):1-14. [doi: 10.3145/epi.2023.sep.20]
25. Moreno-Gil V, Ramon-Vegas X, Rodríguez-Martínez R, Mauri-Ríos M. Explanatory journalism within European fact checking platforms: an ally against disinformation in the post-COVID-19 Era. Societies 2023;13(11):237. [doi: 10.3390/soc13110237]
26. Zecchinon P, Standaert O. The war in Ukraine through the prism of visual disinformation and the limits of specialized fact-checking. a case-study at. Digital J 2024:1-19. [doi: 10.1080/21670811.2024.2332609]
27. Georgiou N, Delfabbro P, Balzan R. COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. Pers Individ Dif 2020;166:110201 [FREE Full text] [doi: 10.1016/j.paid.2020.110201] [Medline: 32565592]
28. Pan W, Liu D, Fang J. An examination of factors contributing to the acceptance of online health misinformation. Front Psychol 2021;12:630268 [FREE Full text] [doi: 10.3389/fpsyg.2021.630268] [Medline: 33732192]

XSL•FO
RenderX

29. Myslivec J. It is all a conspiracy: conspiracy theories about the COVID-19 pandemic and the war in Ukraine. Cesk Psychol 2024;68(2):111-134. [doi: 10.51561/cspsych.68.2.111]
30. Bertamini G. Narrative warfare in the Russo-Ukrainian war: a comparison between Ukraine's and Russia's strategic communication. FINABEL The European Army Interoperability Centre. 2023. URL: https://finabel.org/narrative-warfare-in-the-russo-ukrainian-war-a-comparison-between-ukraines-and-russias-strategic-communication/ [accessed 2024-09-20]
31. Zmigrod L. The role of cognitive rigidity in political ideologies: theory, evidence, and future directions. Curr Opin Behav Sci 2020;34:34-39. [doi: 10.1016/j.cobeha.2019.10.016]
32. Leclercq S, Szaffarczyk S, Leptourgos P, Yger P, Fakhri A, Wathelet M, et al. Conspiracy beliefs and perceptual inference in times of political uncertainty. Sci Rep 2024;14(1):9001 [FREE Full text] [doi: 10.1038/s41598-024-59434-4] [Medline: 38637589]
33. Figà Talamanca G, Arfini S. Through the newsfeed glass: rethinking filter bubbles and echo chambers. Philos Technol 2022;35(1):20 [FREE Full text] [doi: 10.1007/s13347-021-00494-z] [Medline: 35308101]
34. Douglas KM, Sutton RM, Cichocka A. The psychology of conspiracy theories. Curr Dir Psychol Sci 2017;26(6):538-542 [FREE Full text] [doi: 10.1177/0963721417718261] [Medline: 29276345]
35. Cohen S. Cognitive rigidity, overgeneralization and fanaticism. In: Zeigler-Hill V, Shackelford T, editors. Encyclopedia of Personality and Individual Differences. Cham: Springer; 2017:1-7.
36. Zmigrod L, Rentfrow PJ, Robbins TW. The partisan mind: is extreme political partisanship related to cognitive inflexibility? J Exp Psychol Gen 2020;149(3):407-418. [doi: 10.1037/xge0000661] [Medline: 31380695]
37. Czech National Panel (Cesky narodni panel). URL: https://www.narodnipanel.cz/ [accessed 2022-10-08]
38. Prokop D, Markova L, Kudrnacova M, Cahlikova J, Bartos V, Bauer M, et al. Life during a pandemic (Zivot behem pandemie). 2020. URL: https://zivotbehempandemie.cz/projekt [accessed 2022-10-28]
39. Coronavirus: overview of the main disinformation messages (Koronavirus: přehled hlavních dezinformačních sdělení). Ministry of the Interior of the Czech Republic. URL: https://www.mvcr.cz/chh/clanek/koronavirus-prehled-hlavnich-dezinformacnich-sdeleni.aspx [accessed 2021-09-14]
40. Conspiratorial framing of Russian aggression against Ukraine in the Czech quasi-media scene (Konspirační rámování ruské agrese vůči Ukrajině na české kvazi-mediální scéně). Ministry of the Interior of the Czech Republic. URL: https://www.mvcr.cz/chh/clanek/konspiracni-ramovani-ruske-agrese-vuci-ukrajine-na-ceske-kvazi-medialni-scene [accessed 2022-06-05]
41. Grygarová D, Adámek P, Juríčková V, Horáček J, Bakštein E, Fajnerová I, et al. Impact of a long lockdown on mental health and the role of media use: web-based survey study. JMIR Ment Health 2022;9(6):e36050 [FREE Full text] [doi: 10.2196/36050] [Medline: 35605112]
42. Typology of Czech news websites (Typologie domacich zpravodajských webu). Czech Independent Journalism Foundation. 2018 Nov 13. URL: https://www.nfnz.cz/studie-a-analyzy/typologie-domacich-zpravodajskych-webu/ [accessed 2021-05-06]
43. Merlan A. The conspiracy singularity has arrived. 2020 Jul 17. URL: https://www.vice.com/en/article/the-conspiracy-singularity-has-arrived/ [accessed 2024-09-29]
44. Tuters M, Willaert T. Deep state phobia: narrative convergence in coronavirus conspiracism on Instagram. Convergence Int J Res New Media Technol 2022;28(4):1214-1238. [doi: 10.1177/13548565221118751]
45. Murru M. QAnon and its conspiracy milieu: the Italian case. In: Eslen-Ziya H, Giorgi A, editors. Populism and Science in Europe. Cham, Switzerland: Springer Nature; 2022:163-184.
46. Xia Y, Gronow A, Malkamäki A, Ylä-Anttila T, Keller B, Kivelä M. The Russian invasion of Ukraine selectively depolarized the Finnish NATO discussion on Twitter. EPJ Data Sci 2024;13(1):1. [doi: 10.1140/epjds/s13688-023-00441-2]
47. Lintner T, Diviák T, Nekardová B, Lehotský L, Vašečka M. Slovak MPs' response to the 2022 Russian invasion of Ukraine in light of conspiracy theories and the polarization of political discourse. Humanit Soc Sci Commun 2023;10(1):758. [doi: 10.1057/s41599-023-02276-8]
48. Franic J. What lies behind substantial differences in COVID-19 vaccination rates between EU member states? Front Public Health 2022;10:858265 [FREE Full text] [doi: 10.3389/fpubh.2022.858265] [Medline: 35757613]
49. Jennings W, Stoker G, Bunting H, Valgarðsson VO, Gaskell J, Devine D, et al. Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy. Vaccines (Basel) 2021;9(6):593 [FREE Full text] [doi: 10.3390/vaccines9060593] [Medline: 34204971]
50. Melki J, Tamim H, Hadid D, Makki M, El Amine J, Hitti E. Mitigating infodemics: the relationship between news exposure and trust and belief in COVID-19 fake news and social media spreading. PLoS One 2021;16(6):e0252830 [FREE Full text] [doi: 10.1371/journal.pone.0252830] [Medline: 34086813]
51. Sallam M, Dababseh D, Eid H, Al-Mahzoum K, Al-Haidar A, Taim D, et al. High rates of COVID-19 vaccine hesitancy and its association with conspiracy beliefs: a study in Jordan and Kuwait among other Arab Countries. Vaccines (Basel) 2021;9(1):42 [FREE Full text] [doi: 10.3390/vaccines9010042] [Medline: 33445581]
52. Gupta M, Dennehy D, Parra CM, Mäntymäki M, Dwivedi YK. Fake news believability: the effects of political beliefs and espoused cultural values. Inf Manag 2023;60(2):103745. [doi: 10.1016/j.im.2022.103745]

XSL•FO
RenderX

53. Calvillo DP, Ross BJ, Garcia RJB, Smelter TJ, Rutchick AM. Political ideology predicts perceptions of the threat of COVID-19 (and Susceptibility to Fake News About It). Soc Psychol Pers Sci 2020;11(8):1119-1128. [doi: 10.1177/1948550620940539]
54. Motta M. Republicans, not democrats, are more likely to endorse anti-vaccine misinformation. Am Polit Res 2021;49(5):428-438. [doi: 10.1177/1532673x211022639]
55. Eady G, Paskhalis T, Zilinsky J, Bonneau R, Nagler J, Tucker JA. Exposure to the Russian internet research agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. Nat Commun 2023;14(1):62 [FREE Full text] [doi: 10.1038/s41467-022-35576-9] [Medline: 36624094]
56. Hoyle A, Powell T, Cadet B, Kuijt J. Web of Lies: mapping the narratives, effects, and amplifiers of Russian COVID-19 disinformation. In: COVID-19 Disinformation: A Multi-National, Whole of Society Perspective (Advanced Sciences and Technologies for Security Applications). Cham: Springer; 2022:113-141.
57. Atkinson C. Hybrid warfare and societal resilience: implications for democratic governance. ISIJ 2018;39(1):63-76. [doi: 10.11610/isij.3906]
58. Galgoczy M, Phatak A, Vinson D, Mago V, Giabbanelli P. (Re)shaping online narratives: when bots promote the message of president trump during his first impeachment. PeerJ Comput Sci 2022;8:e947 [FREE Full text] [doi: 10.7717/peerj-cs.947] [Medline: 35494820]
59. Kasapoglu C. Russia's renewed military thinking: non-linear warfare and reflexive control. Research Division—NATO Defense College. 2015. URL: https://www.files.ethz.ch/isn/195099/rp_121.pdf [accessed 2022-11-12]
60. Pierre JM. Mistrust and misinformation: a two-component, socio-epistemic model of belief in conspiracy theories. J Soc Polit Psych 2020;8(2):617-641. [doi: 10.5964/jspp.v8i2.1362]
61. Boberg S, Quandt T, Schatto-Eckrodt T, Frischlich L. Pandemic populism: Facebook pages of alternative news media and the corona crisis—a computational content analysis. arXiv Preprint posted online on 2020.
62. Dhawan D, Bekalu M, Pinnamaneni R, McCloud R, Viswanath K. COVID-19 news and misinformation: do they matter for public health Prevention? J Health Commun 2021;26(11):799-808. [doi: 10.1080/10810730.2021.2010841] [Medline: 34865604]
63. Guelmami N, Ben Khalifa M, Chalghaf N, Kong JD, Amayra T, Wu J, et al. Development of the 12-item social media disinformation scale and its association with social media addiction and mental health related to COVID-19 in Tunisia: survey-based pilot case study. JMIR Form Res 2021;5(6):e27280 [FREE Full text] [doi: 10.2196/27280] [Medline: 34021742]
64. Chen E, Jiang J, Chang HH, Muric G, Ferrara E. Charting the information and misinformation landscape to characterize misinfodemics on social media: COVID-19 infodemiology study at a planetary scale. JMIR Infodemiology 2022;2(1):e32378 [FREE Full text] [doi: 10.2196/32378] [Medline: 35190798]
65. Pierri F, Piccardi C, Ceri S. Topology comparison of Twitter diffusion networks effectively reveals misleading information. Sci Rep 2020;10(1):1372 [FREE Full text] [doi: 10.1038/s41598-020-58166-5] [Medline: 31992754]
66. Wang X, Song Y. Viral misinformation and echo chambers: the diffusion of rumors about genetically modified organisms on social media. INTR 2020;30(5):1547-1564. [doi: 10.1108/intr-11-2019-0491]
67. Bryanov K, Vziatysheva V. Determinants of individuals' belief in fake news: a scoping review determinants of belief in fake news. PLoS One 2021;16(6):e0253717 [FREE Full text] [doi: 10.1371/journal.pone.0253717] [Medline: 34166478]
68. Balakrishnan V, Ng WZ, Soo MC, Han GJ, Lee CJ. Infodemic and fake news—a comprehensive overview of its global magnitude during the COVID-19 pandemic in 2021: a scoping review. Int J Disaster Risk Reduct 2022;78:103144 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.103144] [Medline: 35791376]
69. t'Serstevens F, Piccillo G, Grigoriev A. Fake news zealots: effect of perception of news on online sharing behavior. Front Psychol 2022;13:859534 [FREE Full text] [doi: 10.3389/fpsyg.2022.859534] [Medline: 35959002]
70. Oleksy T, Wnuk A, Gambin M, Łyś A. Dynamic relationships between different types of conspiracy theories about COVID-19 and protective behaviour: a four-wave panel study in Poland. Soc Sci Med 2021;280:114028 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114028] [Medline: 34023709]
71. Pavlović T, Azevedo F, De K, Riaño-Moreno JC, Maglić M, Gkinopoulos T, et al. Predicting attitudinal and behavioral responses to COVID-19 pandemic using machine learning. PNAS Nexus 2022;1(3):pgac093 [FREE Full text] [doi: 10.1093/pnasnexus/pgac093] [Medline: 35990802]
72. Thaker J, Subramanian A. Exposure to COVID-19 vaccine hesitancy is as impactful as vaccine misinformation in inducing a decline in vaccination intentions in New Zealand: results from pre-post between-groups randomized block Experiment. Front Commun 2021;6:721982. [doi: 10.3389/fcomm.2021.721982]
73. Verma G, Bhardwaj A, Aledavood T, De Choudhury M, Kumar S. Examining the impact of sharing COVID-19 misinformation online on mental health. Sci Rep 2022;12(1):8045 [FREE Full text] [doi: 10.1038/s41598-022-11488-y] [Medline: 35577820]
74. Bodrunova S, Nepiyuschikh D. Dynamics of distrust, aggression,conspiracy thinking in the anti-vaccination discourse on Russian Telegram. In: Social Computing and Social Media: Design, User Experience and Impact. HCII 2022. Cham: Springer; 2022:468-484.
75. Wu JS, Hauert C, Kremen C, Zhao J. A framework on polarization, cognitive inflexibility, and rigid cognitive specialization. Front Psychol 2022;13:776891 [FREE Full text] [doi: 10.3389/fpsyg.2022.776891] [Medline: 35401343]

XSL•FO
RenderX

Abbreviations

BM-C: beliefs in misinformation narratives about COVID-19 pandemic
BM-U: beliefs in misinformation about the Russian invasion of Ukraine
EFA: exploratory factor analysis
EU: European Union
NATO: North Atlantic Treaty Organization

Edited by T Purnat; submitted 04.06.24; peer-reviewed by M Klicperova-Baker, R Sánchez del Vas; comments to author 18.09.24; revised version received 23.10.24; accepted 22.01.25; published 10.03.25.

Please cite as:
Grygarová D, Havlík M, Adámek P, Horáček J, Juríčková V, Hlinka J, Kesner L
Beliefs in Misinformation About COVID-19 and the Russian Invasion of Ukraine Are Linked: Evidence From a Nationally Representative Survey Study
JMIR Infodemiology 2025;5:e62913
URL: https://infodemiology.jmir.org/2025/1/e62913
doi:10.2196/62913
PMID:

©Dominika Grygarová, Marek Havlík, Petr Adámek, Jiří Horáček, Veronika Juríčková, Jaroslav Hlinka, Ladislav Kesner. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 10.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Measurement, Characterization, and Mapping of COVID-19 Misinformation in Spain: Cross-Sectional Study

Javier Alvarez-Galvez[1,2], PhD; Carolina Lagares-Franco[3], PhD; Esther Ortega-Martin[1], MSc; Helena De Sola[2], PhD; Antonio Rojas-García[4], PhD; Paloma Sanz-Marcos[5], PhD; José Almenara-Barrios[6], PhD; Angelos P Kassianos[7], PhD; Ilaria Montagni[8], PhD; María Camacho-García[1], MSc; Maribel Serrano-Macías[1], PhD; Jesús Carretero-Bravo[9], PhD

[1]Computational Social Science DataLab (CS2 DataLab), University Institute for Sustainable Social Development (INDESS), University of Cadiz, Jerez de la Frontera (Cadiz), Spain
[2]Department of General Economy (Sociology), University of Cadiz, Cadiz, Spain
[3]Department of Statistic and Operative Research, University of Cadiz, Cadiz, Spain
[4]Department of Behavioural Science Methodology, University of Granada, Granada, Spain
[5]Department of Marketing and Communication, University of Cadiz, Jerez de la Frontera, Spain
[6]Department of Biomedicine, Biotechnology and Public Health, University of Cadiz, Cadiz, Spain
[7]Department of Nursing, School of Health Sciences, Cyprus University of Technology, Limassol, Cyprus
[8]Bordeaux Population Health Research Center, Université de Bordeaux - Inserm, Bordeaux, France
[9]Department of Quantitative Methods, Universidad Loyola Andalucía, Seville, Spain

Corresponding Author:
Javier Alvarez-Galvez, PhD
Computational Social Science DataLab (CS2 DataLab)
University Institute for Sustainable Social Development (INDESS)
University of Cadiz
Avda. de la Universidad, s/n, 11405
Jerez de la Frontera (Cadiz)
Spain
Phone: 34 956167217
Email: javier.alvarezgalvez@uca.es

Abstract

Background: The COVID-19 pandemic has been accompanied by an unprecedented infodemic characterized by the widespread dissemination of misinformation. Globally, misinformation about COVID-19 has led to polarized beliefs and behaviors, including vaccine hesitancy, rejection of governmental authorities' recommendations, and distrust in health institutions. Thus, understanding the prevalence and drivers of misinformation is critical for designing effective and contextualized public health strategies.

Objective: On the basis of a tailored survey on health misinformation, this study aims to assess the prevalence and distribution of COVID-19–related misinformation in Spain; identify population groups based on their beliefs; and explore the social, economic, ideological, and media use factors associated with susceptibility to misinformation.

Methods: A cross-sectional telephone survey was conducted with a nationally representative sample of 2200 individuals in Spain. The study developed the COVID-19 Misinformation Scale to measure beliefs in misinformation. Exploratory factor analysis identified key misinformation topics, and k-means clustering classified participants into 3 groups: convinced, hesitant, and skeptical. Multinomial logistic regression was used to explore associations between misinformation beliefs and demographic, social, and health-related variables.

Results: Three population groups were identified: convinced (1078/2200, 49%), hesitant (666/2200, 30.27%), and skeptical (456/2200, 20.73%). Conspiracy theories, doubts about vaccines, and stories about sudden death emerged as the most endorsed current misinformation topics. Higher susceptibility to misinformation was associated with the female sex, lower socioeconomic status, use of low-quality information sources, higher levels of media sharing, greater religiosity, distrust of institutions, and extreme and unstated political ideologies. Frequent sharing of health information on social networks was also associated with membership in the skeptical group, regardless of whether the information was verified. Interestingly, women were prone to COVID-19 skepticism, a finding that warranted further research to understand the gender-specific factors driving vulnerability

XSL•FO
RenderX

to health misinformation. In addition, a geographic distribution of hesitant and skeptical groups was observed that coincides with the so-called empty Spain, areas where political disaffection with the main political parties is greater.

Conclusions: This study highlights the important role of determinants of susceptibility to COVID-19 misinformation that go beyond purely socioeconomic and ideological factors. Although these factors are relevant in explaining the social reproduction of this phenomenon, some determinants are linked to the use of social media (ie, searching and sharing of alternative health information) and probably the political disaffection of citizens who have stopped believing in both the ideologically centrist mainstream parties and the institutions that represent them. Furthermore, by establishing the profile and geographic distribution of the convinced, hesitant, and skeptical groups, our results provide useful insights for public health interventions. Specific strategies should focus on restoring institutional trust, promoting reliable sources of information, and addressing structural drivers of health misinformation linked to gender inequalities.

(JMIR Infodemiology 2025;5:e69945) doi:10.2196/69945

KEYWORDS

misinformation; disinformation; conspiracy theories; health beliefs; polarization; vaccine hesitancy; COVID-19; artificial intelligence; AI

Introduction

Background

The COVID-19 pandemic has not only posed unprecedented challenges to global public health but also has been accompanied by a pervasive infodemic [1,2]. This term, defined by the World Health Organization (WHO) as "an over-abundance of information—some accurate and some not—that makes it hard for people to find trustworthy sources and reliable guidance when they need it" [3], underscores the critical role of information in shaping public responses to social and health emergencies [4]. While access to information is essential, the unprecedented volumes of data surrounding SARS-CoV-2 and COVID-19 has made it increasingly difficult for individuals and institutions to discern scientific evidence from misinformation or disinformation and anecdotal claims [5]. In fact, in this context of global uncertainty, we have been able to observe speculations and conspiracy theories [6] related to different topics, such as the supposed effectiveness of the RNA vaccines, the hidden intentions of global leaders and the pharmaceutical industry, false health treatments (eg, the hydroxychloroquine case), the real origin of the new virus, and even doubts about the existence of the virus [7].

During the recent pandemic, the dissemination of false information about scientific and health-related matters spread faster and more easily than the virus itself [8]. Furthermore, the blurring boundaries between evidence-based knowledge and false or misleading information linked to the new disease introduced substantial complexities to the population's ability to make informed decisions [5]. Specifically, the dissemination of new health guidelines and the heterogeneous political measures to contain the pandemic through traditional and social media caused an excess of noise around everything related to health (ie, vaccines, treatments, contention measures, etc), which ended up generating social polarization and social divisions that made it difficult to manage and control the health emergency situation [9]. In addition, information overload contributed to cognitive fatigue, which has also been shown to reduce people's ability to critically evaluate health sources and content [10]. In addition, the role of influential opinion and political leaders has further exacerbated the situation [11], propagating misconceptions and encouraging risky behaviors that may have undermined public health interventions [12].

The COVID-19 infodemic exemplifies how modern information ecosystems, particularly those based on social media communication, can amplify both evidence-based guidance and misinformation at an unprecedented scale [13,14]. These "misinformation ecosystems" have directly influenced social and health behaviors, contributing to vaccine hesitancy [15,16], noncompliance with health measures [17], and the proliferation of pseudoscientific remedies and theories [18], ultimately challenging the effectiveness of governmental and institutional countermeasures [19]. Moreover, as recent studies have shown, scientific content and sources also play a relevant role in the dissemination of misinformation, for example, through the design of poorly elaborated messages via social media platforms or the provision of weak evidence [20]. Therefore, when focusing on the concept of misinformation in health and, specifically, COVID-19, we are faced with a cross-cutting phenomenon that permeates all strata of our societies.

While it is becoming increasingly common to find studies addressing health misinformation and its impact on health systems, it still remains a difficult concept to define due to the fluctuating nature of the social media ecosystem and the diversity of health-related topics it covers, especially when analyzing the subject matter of COVID-19 [21]. To address this complexity, in this study, we adopt a broad definition of health misinformation, encompassing any health-related claim that relies on anecdotal evidence, is false, or is misleading due to the absence of scientific confirmation [22], regardless of whether this information has been issued intentionally or unintentionally [23]. Therefore, our definition incorporates 2 key forms: misinformation, which refers to false information shared without an intent to cause harm [24,25], and disinformation or malinformation, which involves false or partially true information deliberately crafted to deceive or harm specific individuals, social groups, institutions, or nations [26]. Our definition implicitly incorporates the flexibility of the definition by the National Academies of Sciences, Engineering, and Medicine report, which defines science-related misinformation as "information that asserts or implies claims that are inconsistent with the weight of accepted scientific evidence at

XSL•FO
RenderX

the time.” This conceptualization underscores the importance of distinguishing between misinformation that directly contradicts well-established evidence and misinformation that may reflect evolving or uncertain scientific consensus. In light of this, our study attempts to capture both types of claims in the COVID-19 context, acknowledging that the epistemic status of some items may be more contentious than others [27].

Although studies have been emerging since the beginning of the COVID-19 pandemic that have addressed misinformation sources, channels, and messages [28-31], in general, less attention has been paid to the exhaustive characterization of the profiles of people who embrace misinformation and particularly understanding which messages have penetrated the audiences. In the literature, cross-sectional studies can be found that have usually used opportunistic samples to understand specific attitudes and behaviors (eg, related to the new vaccines or preventive measures) [32-37], but, to our knowledge, there are no exhaustive and updated studies that allow us to determine the current prevalence of COVID-19 misinformation about the beliefs that may have emerged in recent years and that have progressively taken root in public opinion (eg, doubts about the need for additional vaccines, the origin of the virus, the existence of hidden governmental plans, and sudden deaths, among others). In addition, this highlights the need to provide a comprehensive sociodemographic and contextualized description of these social groups, which could ultimately help us identify the groups that we should target with information campaigns aimed at increasing literacy in COVID-19–related health issues.

In a recent review, political orientation has been found to be a key predictor of misinformation and beliefs in fake news [9]. However, most of the studies reviewed were based in the United States, reflecting a limited geographical focus. To date, this association has not been systematically explored in other contexts, underscoring the need for context-specific research to better understand how political orientation influences susceptibility to misinformation within different sociopolitical frameworks. Particularly in the Spanish case, a study has been carried out analyzing conspiracy beliefs and their association with ideological and religious values based on survey data from the Spanish Foundation for Science and Technology; however, the dataset is not specifically focused on the study of health misinformation related to COVID-19 but on the scientific aspects related to the pandemic [38]. Moreover, although there have been attempts to identify the diversity of misinformation and conspiracy groups through the application of latent profile techniques [39], there is still a lack of comprehensive characterizations of the explanatory factors (demographic, social, economic, political, and media use) of population profiles susceptible to misinformation. In other words, in addition to characterizing the sociodemographic groups of the different profiles that embrace health misinformation, there is a need for studies that provide detailed information on how these groups consume and share information as well as other health aspects that could explain their positioning (eg, general health status, chronic conditions, diagnoses, and COVID-19 vaccination).

Objectives

Given these knowledge gaps, the aim of this exploratory study was to assess the prevalence and distribution of COVID-19–related misinformation in Spain using data from a representative population sample obtained from the DCODES (Collective dynamics of health opinion contagion: the COVID-19 infodemic and its effects on decision making processes) project. Specifically, we aim to achieve the following specific objectives: (1) identify the issues that generate the greatest social division around the COVID-19 pandemic, (2) classify the main groups around these issues of misinformation about COVID-19, and (3) describe these social profiles by studying the association with a broad set of social determinants that could explain the positioning of these groups.

Methods

Design and Setting

This study used a cross-sectional design using telephone surveys conducted from January 2024 to March 2024, targeting individuals aged >18 years residing in Spain. Participants were selected from available databases, with consent obtained before their participation. The telephone survey method was chosen due to its numerous advantages, making it particularly suited to this study’s objectives. First, it allowed for rapid and extensive access to participants, including those in remote or rural areas. Second, it was cost-effective compared to face-to-face interviews, reducing logistical expenses. Third, the standardized administration of questionnaires minimized interviewer-related variability, ensuring greater consistency in data collection. Fourth, it was convenient for respondents, who could complete the survey from their homes, facilitating higher participation rates. Fifth, telephone surveys were particularly effective for addressing sensitive topics, as respondents often felt more secure and perceived greater confidentiality in such interactions. Collectively, these features made telephone surveys an optimal data collection method, ensuring both broad representativeness and logistical feasibility in this study.

Stratified sampling guaranteed representativeness from the different geographic areas of the country based on key demographic factors: age, sex, region, and population size of the areas of residence. Using a multistage sampling approach and drawing on nationally representative datasets, a sampling frame reflecting the demographic composition of Spain was composed. This rigorous methodology minimized selection bias and allowed accurate conclusions to be drawn. The final sample composed of 2200 individuals, with a confidence level of 95% and an estimated error of +2.1 or –2.1 percentage points.

This survey was part of the DCODES project, which aimed to conduct an in-depth analysis of the determinants of misinformation during the COVID-19 pandemic.

Measures

The questionnaire was developed by a team of 6 researchers with experience in survey methods and misinformation studies through 3 nominal group meetings in which the main contents and variables potentially associated with misinformation on COVID-19 were discussed and consensus was reached on which

XSL•FO
RenderX

were the most relevant in the general population. The questionnaire was divided into 4 fundamental blocks of indicators (Table 1).

Table 1. Thematic blocks and description of indicators.

Variables	Description of measures
COVID-19 misinformation indicators	The primary set of indicators included the COVID-19 Misinformation Scale, with 12 items related to erroneous, false, or misleading topics that were currently most widely disseminated about COVID-19. These items were extracted from the 3 main platforms for checking hoaxes and fake news in Spain (Newtral, Maldita.es, and EFE Verifica, belonging to the International Fact Checking Network). These indicators were presented on a Likert scale from 1 (totally disagree) to 5 (totally agree) points. The 12 items had high internal consistency (Cronbach α=0.87). The content of each statement is available in the Results section.
Health information use, media sharing, and digital health literacy	The survey incorporated questions about the information medium used the last time to obtain health information and the frequency of online media used to search for health topics. In addition, the survey asked about the frequency with which the individuals shared health topics through social networks and whether they checked this information before sharing it. As another means of measuring the individual's use and ability to use digital media to obtain health information, we also used the eHealth Literacy Scale [40], an 8-item instrument to measure the individual's digital health literacy, which has been previously validated in Spain [41].
COVID-19 diagnoses, protective measures, and health status	With the intention of evaluating the association between health and positioning around the different COVID-19 issues, respondents were asked whether they had taken the COVID-19 test, whether they had been vaccinated, and what was the degree of compliance with protective measures during the state of alarm in Spain. This last variable was obtained as the mean of compliance with confinement, social distancing, use of masks, hand washing, and diagnostic testing of contacts. In addition, they were also asked about their self-perceived health and whether they had any chronic disease.
Social and economic indicators	Demographic, social, and economic variables were also incorporated in the questionnaire. Specifically, sex, age, income, size of municipality, educational level, employment status, and nationality were collected. In addition, given the relationship shown with cultural and ideological aspects of the individual and the predisposition to misinformation, we also included political ideology (on a scale of 0-10, where 0 meant extreme left ideology and 10 meant extreme right ideology). Ideology was categorized so that a value of 0 to 1 represented far left wing, a value of 2 to 3 represented left wing, a value of 4 to 6 represented center, a value of 7 to 8 represented right wing, and a value of 9 to 10 represented far right wing. An extra category, nonresponse, for those who preferred not to reveal their ideology was also included. Individual religiosity was also assessed (where 0 meant not religious at all and 10 meant very religious) along with trust in institutions (the Spanish government, political parties, World Health Organization, health personnel, and the Spanish National Health System). Finally, as an analysis of the use of new technologies, we added a question about the knowledge and use of artificial intelligence models, which may pose a risk for increasing misinformation [42], and attitudes toward cryptocurrencies, a relationship that has been found to be associated with conspiracy beliefs [43].

Data Analyses

First, an exploration of the characteristics of the sample through descriptive statistics was carried out. After that, a dimensionality analysis was conducted on the COVID-19 Misinformation Scale, through an exploratory factor analysis (EFA). To determine the number of initial dimensions of the items, the criteria of the number of eigenvalues >1 and the conceptual meaning of the factors were considered. We used the weighted least squares mean and variance adjusted estimator as a factor extraction method [44] because it is the most suitable method for Likert scale items and the geomin rotation [45] as an oblique rotation method. This rotation method presents the advantage of balancing simplicity and realism when interpreting factor structures. Unlike orthogonal rotations, geomin allows correlations between factors and minimizes small factor loadings, making it particularly suitable for exploring complex latent structures where variables may load on multiple factors, thereby enhancing the interpretability and precision of the results.

Once the factors were derived from the EFA, we aimed to categorize individuals into groups based on their scores on the dimensions, using the k-means clustering technique. In an attempt to offer an exhaustive classification but at the same time a simple and operational one, participants were initially classified into three categories: (1) convinced-—those who believed or followed the majority opinion according to existing research evidence on COVID-19; (2) hesitant—individuals who expressed doubts, characterized by delays in accepting or rejecting COVID-19 information despite its availability; and (3) skeptical—those who tended to believe or follow minority opinion on COVID-19 issues that are erroneous, false, or misleading and commonly based on anecdotal evidence. This decision was made to simplify the results, as the silhouette method indicated that the 3-cluster solution provided a better classification than other groupings. Accordingly, these 3 categories were used to obtain the different individual profiles by testing their relationship with the variables associated with health information resources and use, COVID-19 diagnoses, health status, sociodemographic variables, ideology, and knowledge of technologies through a multinomial logistic regression model. The group of those individuals who were considered convinced was considered as the reference category, and independent variables with n categories were recategorized into n–1 dichotomous variables.

All the analyses were conducted using R (version 4.1.2; R Foundation for Statistical Computing) and RStudio (Posit PBC). To perform the EFA analyses, the *Lavaan* package [46] was

XSL•FO
RenderX

used, which incorporates the most up-to-date estimators for categorical variables.

Ethical Considerations

This study was approved by the Ethics Committee on Non-Biomedical Experimentation and Genetically Modified Organisms (CEENB-GMOs) of the University of Cadiz (005_2024). All data collected through the telephone survey were fully anonymized prior to analysis. No personally identifiable information was recorded or stored, ensuring complete confidentiality of participants and no financial or material compensation was provided to participants. Participation in the study was voluntary, and verbal informed consent was obtained at the beginning of the interview process.

Results

Table 2 shows the characteristics of the population of the representative sample of 2200 people that participated in this study. As can be seen, the sample was clearly balanced in accordance with the sex and age of the interviewees. In terms of the variables associated with COVID-19, a substantial percentage of vaccination (2088/2200, 94.91%) and a high follow-up of the measures during the pandemic (mean 4.47 out of 5, SD 0.69) were found. Almost 39.05% (859/2200) of the population had been diagnosed with COVID-19 at the time of the survey.

XSL•FO
RenderX

Table 2. Summary of the sample characteristics by sex (N=2200).

Characteristics	Overall	Male (n=1060)	Female (n=1140)
Age group (y), n (%)			
18-34	478 (21.73)	238 (22.45)	240 (21.05)
35-49	624 (28.36)	314 (29.62)	310 (27.19)
50-64	574 (26.09)	282 (26.6)	292 (25.61)
>64	524 (23.82)	226 (21.32)	298 (26.14)
Income level (€; US $1=€0.9267), n (%)			
<900	181 (8.23)	73 (6.89)	108 (9.47)
901-1200	230 (10.45)	85 (8.02)	145 (12.72)
1201-1800	334 (15.18)	150 (14.15)	184 (16.14)
1801-2400	332 (15.09)	161 (15.19)	171 (15)
2401-3000	327 (14.86)	198 (18.68)	129 (11.32)
3001-4500	292 (13.27)	161 (15.19)	131 (11.49)
>4500	177 (8.05)	104 (9.81)	73 (6.4)
N/A[a]	327 (14.86)	128 (12.08)	199 (17.46)
Size of the municipality (inhabitants), n (%)			
<10,000	467 (21.23)	228 (21.51)	239 (20.96)
10,001-50,000	563 (25.59)	263 (24.81)	300 (26.32)
50,001-100,000	227 (10.32)	116 (10.94)	111 (9.74)
100,001-400,000	470 (21.36)	203 (19.15)	267 (23.42)
>400,000	473 (21.5)	250 (23.58)	223 (19.56)
Education level, n (%)			
Basic or primary	253 (11.59)	112 (10.63)	141 (12.49)
Professional training or bachelor's degree	892 (40.86)	466 (44.21)	426 (37.73)
University degree	798 (36.56)	363 (34.44)	435 (38.53)
Postgraduate	240 (10.99)	113 (10.72)	127 (11.25)
Work status, n (%)			
Working	1264 (57.45)	664 (62.64)	600 (52.63)
Unemployed	536 (24.36)	258 (24.34)	278 (24.39)
Retired	203 (9.23)	76 (7.17)	127 (11.14)
Student	88 (4)	49 (4.62)	39 (3.42)
Unpaid work at home	73 (3.32)	1 (0.09)	72 (6.32)
N/A	36 (1.64)	12 (1.13)	24 (2.11)
Nationality, n (%)			
Spanish	2054 (93.36)	991 (93.49)	1063 (93.25)
Other (not Spanish)	146 (6.67)	69 (6.51)	77 (6.75)
Last source of health information , n (%)			
Internet	486 (22.09)	233 (21.98)	253 (22.19)
Family or friends	275 (12.5)	134 (12.64)	141 (12.37)
Books and newspapers	83 (3.77)	43 (4.06)	40 (3.51)
Health staff	1332 (60.55)	639 (60.28)	693 (60.79)
N/A	24 (1.09)	11 (1.04)	13 (1.14)
Online resources for health information, n (%)			

XSL•FO
RenderX

Characteristics	Overall	Male (n=1060)	Female (n=1140)
Social media	211 (9.59)	90 (8.49)	121 (10.61)
Online digital media	244 (11.09)	131 (12.36)	113 (9.91)
Wikipedia	343 (15.59)	148 (13.96)	195 (17.11)
Institutional webs	522 (23.73)	248 (23.4)	274 (24.04)
Google	375 (17.05)	170 (16.04)	205 (17.98)
Did not use online media	424 (19.27)	226 (21.32)	198 (17.37)
N/A	81 (3.68)	47 (4.43)	34 (2.98)
Sharing health information on social media, n (%)			
Never	1563 (71.05)	769 (72.55)	794 (69.65)
1 day a month	312 (14.18)	148 (13.96)	164 (14.39)
1 day a week	198 (9)	78 (7.36)	120 (10.53)
≥2 days a week	125 (5.68)	63 (5.94)	62 (5.44)
N/A	2 (0.09)	2 (0.19)	0 (0)
Evaluation of information shared, n (%)			
Yes	698 (34.23)	328 (33.74)	370 (34.68)
No	1502 (65.77)	732 (66.26)	770 (65.32)
Self-perceived health, n (%)			
Bad or fair	280 (12.73)	152 (14.34)	128 (11.23)
Good	1327 (60.32)	665 (62.74)	662 (58.07)
Very good	516 (23.45)	214 (20.19)	302 (26.49)
Excellent	77 (3.5)	29 (2.74)	48 (4.21)
Chronic conditions, n (%)			
Yes	1433 (65.14)	737 (69.53)	696 (61.05)
No	767 (34.86)	323 (30.47)	444 (38.95)
COVID-19 diagnoses, n (%)			
Yes	859 (39.05)	420 (39.62)	439 (38.51)
No	1341 (60.95)	640 (60.38)	701 (61.49)
COVID-19 vaccination, n (%)			
Yes	2088 (94.91)	1000 (94.34)	1088 (95.44)
No	112 (5.09)	60 (5.66)	52 (4.56)
AI[b] knowledge, n (%)			
I do not know anything about it	245 (11.14)	97 (9.15)	148 (12.98)
I have heard about it, but I do not know much about it	1085 (49.32)	477 (45)	608 (53.33)
I know AI, but I do not use it	410 (18.64)	217 (20.47)	193 (16.93)
I know AI and I use it	452 (20.55)	266 (25.09)	186 (16.32)
N/A	8 (0.36)	3 (0.28)	5 (0.44)
Cryptocurrency knowledge, n (%)			
I do not know anything about it	446 (20.27)	173 (16.32)	273 (23.95)
I have heard about it, but I do not know much about it	1159 (52.68)	504 (47.55)	655 (57.46)
I know cryptocurrencies, but I do not invest	457 (20.77)	286 (26.98)	171 (15)
I know cryptocurrencies and I invest in them	129 (5.86)	95 (8.96)	34 (2.98)
N/A	9 (0.41)	2 (0.19)	7 (0.61)

XSL•FO
RenderX

Characteristics	Overall	Male (n=1060)	Female (n=1140)
Political ideology, n (%)			
Far left wing	165 (7.5)	73 (6.89)	92 (8.07)
Left wing	418 (19)	195 (18.4)	223 (19.56)
Center	1,085 (49.32)	561 (52.92)	524 (45.96)
Right wing	211 (9.59)	103 (9.72)	108 (9.47)
Far right wing	111 (5.05)	57 (5.38)	54 (4.74)
No response	210 (9.55)	71 (6.7)	139 (12.19)
No trust in institutions, n (%)			
Spanish government	1077 (48.95)	533 (50.28)	544 (47.72)
Political parties	1613 (73.32)	794 (74.91)	819 (71.84)
World Health Organization	432 (19.64)	243 (22.92)	189 (16.58)
Health staff	68 (3.09)	34 (3.21)	34 (2.98)
Spanish National Health System	230 (10.45)	109 (10.28)	121 (10.61)
COVID-19 compliance measures (Gaussian), mean (95% CI)	4.47 (4.44-4.50)	4.33 (4.28-4.38)	5.00 (4.40-5.00)
Health online literacy (Gaussian), mean (95% CI)	3.29 (2.38-4.00)	3.25 (2.38-3.88)	3.38 (2.50-4.00)
Degree of religiosity (Gaussian), mean (95% CI)	3.85 (3.78-3.91)	3.58 (3.41-3.79)	4.09 (3.90-4.28)

[a]N/A: not available due to missing data.

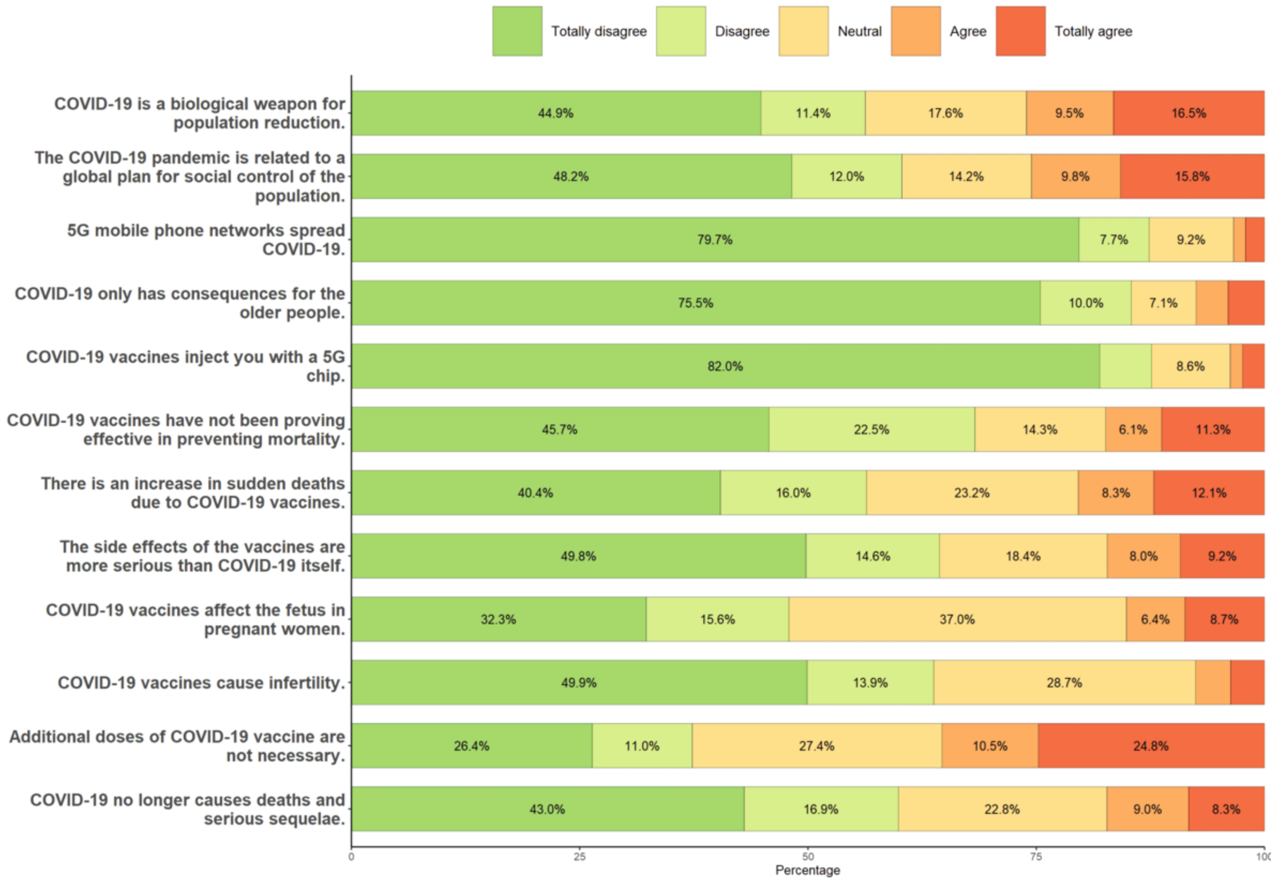
[b]AI: artificial intelligence.

Regarding the variables that could be associated with beliefs, attitudes, or knowledge that led to misinformation, there was a high level of distrust in political parties (1613/2200, 73.32%) and in the Spanish government (1077/2200, 48.95%). The general population had a medium level of digital health literacy (3.29 out of 5), with the value being somewhat higher in women. The main source of obtaining information on health was from health care professionals (1332/2200, 61.21%), while online information was obtained from institutional websites (522/2200, 24.63%), Google (375/2200, 17.7%), and Wikipedia (343/2200, 16.19%). Up to 34.23% (698/2200) of the participants stated that they did not check the health information they shared through online media.

Figure 1 shows the results of the COVID-19 Misinformation Scale. The items on the need for new doses of vaccines (777/2200, 35.32% thought there is no need for more) and the conspiracy surrounding COVID-19 (572/2200, 26% thought it was a biological weapon, and 559/2200, 25.41% thought it was related to a population control plan) stood out. They also negatively highlighted the sudden increase in deaths (862/2200, 39.18% strongly agreed, agreed or neutral). By contrast, the items concerning the beliefs about the relationship between COVID-19 and 5G mobile network were the ones that caused the least doubts in people, together with the statement that this disease only affected older people.

XSL•FO
RenderX

Figure 1. Summary of the COVID-19 misinformation statements.



To summarize the information and investigate the structure of these 12 items, a factor analysis was performed. The Kaiser-Meyer-Olkin statistic and Bartlett test showed that the data were suitable for factor analysis (Kaiser-Meyer-Olkin=0.91; P<.001). The results are presented in Table 3. As there were 3 eigenvalues >1 in the polychoric correlation matrix, factorial solutions from 2 to 5 dimensions were reviewed; it was found that the one that made the most conceptual sense was the one with 4 factors, with an explained variance of 64.8%. After performing the geomin rotation, all 12 items presented factor loadings >0.4 without cross loadings between factors, and only 1 item (ie, COVID-19 no longer caused sequelae) had a communality <0.3, although it remained in the model because it had a factor loading >0.4.

XSL•FO
RenderX

Table 3. Factorial structure of the COVID-19 Misinformation Scale.

COVID-19 Misinformation Scale items	Factor 1: global plan	Factor 2: health beliefs	Factor 3: vaccine hesitancy	Factor 4: fertility impact	R^2
COVID-19 is a biological weapon for population reduction.	*0.94* [a]	–0.011	0.001	0.009	0.883
The COVID-19 pandemic is related to a global plan for social control of the population.	*0.916*	0.026	0.058	–0.002	0.954
5G mobile phone networks spread COVID-19.	0.355	*0.654*	–0.097	0.001	0.703
COVID-19 only has consequences for the older people.	0.001	*0.849*	0.002	–0.235	0.515
COVID-19 vaccines inject you with a 5G chip.	0.182	*0.686*	0.037	0.024	0.726
COVID-19 no longer causes deaths and serious sequelae.	–0.002	*0.442*	0.263	–0.256	0.241
COVID-19 vaccines have not been shown to be effective in preventing mortality.	0.08	–0.018	*0.615*	–0.064	0.388
There is an increase in sudden deaths due to COVID-19 vaccines.	0.021	–0.029	*0.748*	0.086	0.653
The side effects of the vaccines are more serious than COVID-19 itself.	–0.11	0.031	*0.969*	0.015	0.842
Additional doses of COVID-19 vaccine are not necessary.	0.037	0.021	*0.553*	–0.011	0.347
COVID-19 vaccines affect the fetus in pregnant women.	0.062	–0.004	0.007	*0.843*	0.779
COVID-19 vaccines cause infertility.	–0.003	0.106	0.227	*0.61*	0.745
Eigenvalues	6.443	1.101	1.003	0.878	—[b]
Variance explained (%)	20.6	38.6	55	64.8	—

[a]Italic indicate the highest factor loading for each item.

[b]Not applicable.

Factor 1 grouped the 2 items associated with the COVID-19 conspiracy ideas of the existence of a global plan to reduce or control the population. The second factor grouped general health beliefs, such as the link between COVID-19 and 5G technologies, and the supposed impact of the new disease on older groups and mortality. The third factor grouped the claims associated with vaccines, while the fourth factor grouped the 2 items on the fertility consequences of the pandemic. A k-means cluster analysis was used to identify distinct groups of individuals based on their beliefs about the 4 factors extracted from the EFA. The aim of this analysis was to detect 3 distinct subgroups within the population, those who were informed of certain information about COVID-19 and those who had doubts and were skeptics (people who tended to agree with certain COVID-19 misinformation topics). The mean punctuation of the factors by groups can be seen in Table 4.

Table 4. Distribution of 3 clusters of individuals (k-means clustering; N=2200).

Social profiles	Factor 1: global plan, mean (SD)	Factor 2: health beliefs, mean (SD)	Factor 3: vaccine hesitancy, mean (SD)	Factor 4: fertility impact, mean (SD)	Individuals, n (%)
Convinced	1.13 (0.33)	1.30 (0.39)	1.71 (0.61)	1.61 (0.72)	1078 (49)
Hesitant	2.93 (0.88)	1.65 (0.57)	2.57 (0.75)	2.33 (0.85)	666 (30.27)
Skeptical	4.47 (0.70)	2.33 (0.91)	3.77 (0.83)	3.43 (0.96)	456 (20.73)

The largest group was composed of what we have called convinced individuals, accounting for almost half of the sample (1078/2200, 49%), while 30.27% (666/2200) of the individuals were classified as hesitant (ie, individuals reluctant to believe COVID-19–related information), and 20.73% (456/2200) of the individuals were in the skeptical group. Within the latter group, the score on the factor about conspiracies stood out, with a mean close to the maximum of 5 as well as a score that was also high on the statements about vaccines. The 2 factors, vaccines and conspiracies, were the main factors to classify the people who believed misinformation about COVID-19 (the factor with the biggest mean score in the hesitant and skeptical individuals).

The final step involved determining the variables associated with membership in these groups through a multinomial regression analysis, the results of which are presented in Table 5. With these variables, the regression model was able to detect

XSL•FO
RenderX

64.82% (1426/2200) of the people in their corresponding group. Among the socioeconomic variables, sex stood out, with a higher probability of women belonging to the skeptical group (OR 1.699; 95% CI (1.187-2.433); *P*=.004). In addition, both the education and income level variables acted in a similar way—the higher the level of education and income, the lower the probability of belonging to the skeptical group, adding also, in the case of educational level, a lower probability of belonging to the hesitant group. The size of the municipality also influenced the results with people in intermediate-sized municipalities (50,000-100,000 inhabitants) having a lower probability of belonging to the group with more skeptics (ie, the skeptical group).

XSL•FO
RenderX

Table 5. Multinomial logistic regression with group of misinformation as the dependent variable.

Convinced (reference)	Hesitant		Skeptical	
Category	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value
Sex				
Men	1	—[b]	1	—
Women	1.254 (0.953-1.65)	.12	1.699 (1.187-2.433)	.004[c]
Age group (y)				
18-34	1	—	1	—
35-49	1.014 (0.698-1.474)	.94	0.88 (0.546-1.416)	.60
50-64	0.661 (0.433-1.009)	.06	0.729 (0.43-1.237)	.24
>64	0.591 (0.317-1.101)	.098	0.578 (0.252-1.325)	.12
Income level (€; US $1=€0.9267)				
<900	1	—	1	—
901-1200	1.005 (0.561-1.801)	.99	0.628 (0.333-1.186)	.15
1201-1800	1.24 (0.723-2.129)	.43	0.415 (0.222-0.775)	.006[c]
1801-2400	0.936 (0.536-1.635)	.82	0.288 (0.149-0.56)	<.001[d]
2401-3000	0.793 (0.448-1.405)	.43	0.321 (0.164-0.629)	.001[c]
3001-4500	0.83 (0.461-1.495)	.54	0.237 (0.115-0.491)	<.001[d]
>4500	0.947 (0.494-1.814)	.87	0.317 (0.138-0.726)	.007[c]
Size of the municipality (inhabitants)				
<10,000	1	—	1	—
10,001-50,000	1.257 (0.866-1.824)	.23	1.227 (0.775-1.941)	.38
50,001-100,000	0.692 (0.426-1.124)	.14	0.456 (0.235-0.884)	.02[e]
100,001-400,000	0.993 (0.671-1.469)	.97	0.783 (0.473-1.295)	.34
>400,000	1.31 (0.876-1.958)	.19	0.955 (0.565-1.614)	.86
Education level				
Basic or primary	1	—	1	—
Professional training or bachelor's degree	0.589 (0.367-0.947)	.03[e]	0.481 (0.278-0.833)	.009[c]
University degree	0.453 (0.273-0.751)	.002[c]	0.272 (0.148-0.503)	<.001[d]
Postgraduate	0.405 (0.214-0.769)	.006[c]	0.122 (0.047-0.318)	<.001[d]
Work status				
Working	1	—	1	—
Unemployed	0.981 (0.577-1.667)	.94	0.626 (0.306-1.279)	.199
Retired	1.216 (0.751-1.971)	.43	0.791 (0.44-1.42)	.43
Student	1.235 (0.647-2.355)	.52	0.553 (0.213-1.436)	.22
Unpaid work at home	0.697 (0.329-1.478)	.35	0.413 (0.146-1.171)	.10
Nationality				
Spanish	1	—	1	—
Other (not Spanish)	1.437 (0.841-2.455)	.18	1.471 (0.774-2.795)	.24
Last source of health information				
Internet	1	—	1	—

XSL•FO
RenderX

Convinced (reference)	Hesitant		Skeptical	
Category	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value
Family or friends	0.589 (0.375-0.925)	.022[e]	0.659 (0.383-1.133)	.13
Books and newspapers	0.446 (0.211-0.943)	.04[e]	0.279 (0.097-0.8)	.02[e]
Health staff	0.705 (0.507-0.979)	.04[e]	0.452 (0.297-0.686)	<.001[d]
Online resources for health information				
Social media	1	—	1	—
Online digital media	0.884 (0.518-1.511)	.65	0.978 (0.496-1.927)	.95
Wikipedia	0.833 (0.503-1.381)	.48	1.223 (0.657-2.277)	.53
Institutional webs	0.641 (0.399-1.03)	.06	0.453 (0.24-0.856)	.02[e]
Google	0.753 (0.45-1.258)	.28	0.841 (0.443-1.597)	.60
Did not use online media	0.876 (0.498-1.542)	.65	1.258 (0.617-2.566)	.53
Sharing health information on social media				
Never	1	—	1	—
1 day a month	1.241 (0.872-1.767)	.23	1.313 (0.814-2.115)	.26
1 day a week	2.691 (1.745-4.149)	<.001[d]	2.498 (1.417-4.404)	.002[c]
≥2 days a week	1.475 (0.803-2.709)	.21	2.43 (1.195-4.942)	.01[e]
Contrast the information shared				
Yes	1	—	1	—
No	1.121 (0.846-1.485)	.43	0.836 (0.576, 1.215)	.35
Self-perceived health				
Bad or fair	1	—	1	—
Good	1.369 (0.9-2.084)	.14	0.929 (0.56-1.542)	.78
Very good	1.468 (0.903-2.388)	.12	1.145 (0.636-2.061)	.65
Excellent	1.427 (0.618-3.293)	.41	0.601 (0.194-1.86)	.38
Chronic conditions				
Yes	1	—	1	—
No	0.777 (0.581-1.04)	.09	1.169 (0.788-1.734)	.44
COVID-19 diagnoses				
Yes	1	—	1	—
No	1.095 (0.838-1.431)	.51	0.968 (0.681-1.375)	.86
COVID-19 vaccination				
Yes	1	—	1	—
No	2.489 (1.175-5.272)	.02[e]	7.649 (3.54-16.524)	<.001[d]
AI[f] knowledge				
I do not know anything about it	1	—	1	—
I have heard about it, but I do not know much about it	0.998 (0.598-1.667)	.99	0.62 (0.341-1.128)	.12
I know AI, but I do not use it	1.025 (0.562-1.869)	.94	0.77 (0.37-1.604)	.49
I know AI, and I use it	0.595 (0.324-1.091)	.09	0.492 (0.236-1.027)	.06
Cryptocurrency knowledge				
I do not know anything about it	1	—	1	—

XSL•FO
RenderX

Convinced (reference)	Hesitant		Skeptical	
Category	OR[a] (95% CI)	*P* value	OR (95% CI)	*P* value
I have heard about cryptocurrencies, but I do not know much about it	0.96 (0.658-1.402)	.83	0.731 (0.448-1.19)	.21
I know cryptocurrencies, but I do not invest	0.815 (0.509-1.306)	.40	0.688 (0.374-1.265)	.22
I know cryptocurrencies, and I invest in them	1.115 (0.594-2.094)	.74	1.035 (0.468-2.29)	.93
Political ideology				
Far left wing	1	—	1	—
Left wing	0.999 (0.574-1.74)	.99	1.578 (0.711-3.502)	.26
Center	1.323 (0.797-2.198)	.28	1.695 (0.818-3.511)	.16
Right wing	1.682 (0.905-3.126)	.10	1.175 (0.486-2.843)	.72
Far right wing	1.118 (0.505-2.473)	.78	3.138 (1.229-8.013)	.02[e]
No response	2.366 (1.14-4.911)	.02[e]	3.273 (1.289-8.309)	.01[e]
Trust in institutions				
Spanish government				
Yes	1	—	1	—
No	1.359 (0.983-1.88)	.06	1.198 (0.753-1.906)	.45
Political parties				
Yes	1	—	1	—
No	1.753 (1.298-2.369)	<.001[d]	2.271 (1.499-3.441)	<.001[d]
World Health Organization				
Yes	1	—	1	—
No	1.715 (1.176-2.501)	.005[c]	3.612 (2.335-5.588)	<.001[d]
Health staff				
Yes	1	—	1	—
No	1.249 (0.379-4.116)	.71	2.176 (0.658-7.201)	.20
Spanish National Health System				
Yes	1	—	1	—
No	1.898 (1.076-3.347)	.03[e]	4.441 (2.446-8.062)	<.001[d]
COVID-19 compliance measures (continuous)	0.776 (0.626-0.961)	.02[e]	0.673 (0.524-0.866)	.002[c]
Health online literacy (continuous)	0.986 (0.849-1.144)	.85	1.166 (0.965-1.409)	.11
Degree of religiosity (continuous)	1.145 (1.094-1.197)	<.001[d]	1.206 (1.138-1.279)	<.001[d]

[a]OR: odds ratio.

[b]Not available.

[c]Statistically significant at the level of <.01.

[d]Statistically significant at the level of <.001.

[e]Statistically significant at the level of <.05.

[f]AI: artificial intelligence.

Regarding the time of information and the search for health resources, it was observed that those who sought information through health personnel and written media (ie, books and newspapers) were less likely to belong to the skeptical group than those who searched on the internet. Similarly, among those who searched on the internet, those who searched on institutional websites were more likely to belong to the convinced group. Sharing on social media more frequently was associated with a greater probability of belonging to the skeptical group; however, there was no relationship with contrasting the

XSL•FO
RenderX

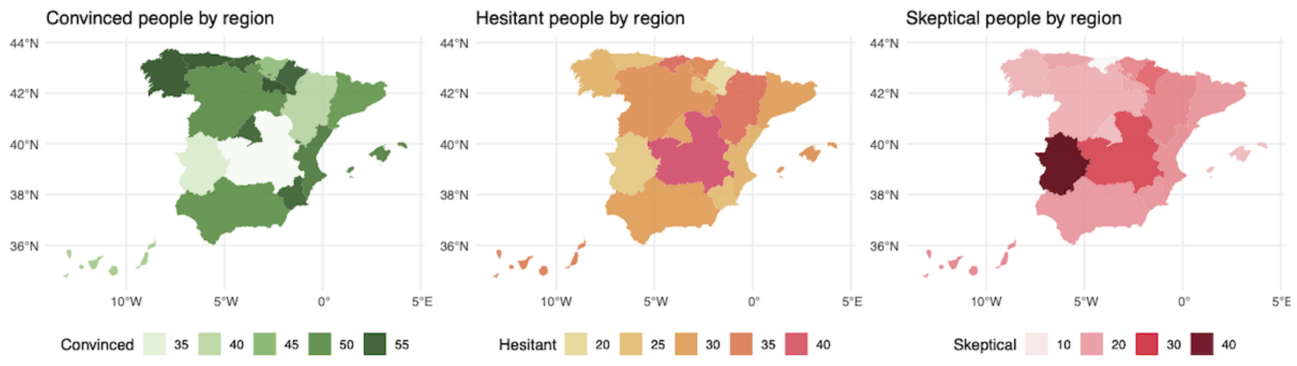
information before sharing it. The degree of digital health literacy of the individual, as measured by the eHealth Literacy Scale, also showed no relationship with the group to which they belonged.

With respect to certain individual thoughts, we found some relationships of interest. Political ideology showed how individuals on the far right and even those who did not want to declare their ideology (ie, possibly those showing political disaffection) were more likely to belong to the skeptical and hesitant groups with respect to individuals who held a moderate political position (whether they positioned themselves on the left or on the right on the ideological scale). In turn, distrust in certain institutions was shown to be a factor related to skepticism (specifically, distrust in political parties, the WHO, and the Spanish National Health System). The same relationship was shown for religiosity—the higher the degree of religiosity, the greater the probability of belonging to the skeptical group.

Reviewing the variables associated with the COVID-19 pandemic and health, not having been vaccinated for COVID-19 and lower compliance with protective measures during the pandemic were associated with individuals belonging to the skeptical group. By contrast, diagnosis of COVID-19, having chronic conditions, and self-perceived health did not present statistically significant relationships.

Finally, to obtain a contextualized representation of the different groups that we had characterized, we proceeded to study their regional distribution in the country as a whole. Figure 2 shows the maps for the 3 groups: convinced, hesitant, and skeptical. The spatial distribution of these groups on the map of Spain showed that the skeptical groups were found to a greater extent in the so-called empty Spain (particularly in regions such as Extremadura and Castilla-La Mancha), that is, those rural areas of the country characterized by a low population density due to the exodus of young people to the main urban areas. This is a trend that to some extent was also observed with the individuals in the hesitant group, who were also more prevalent among regions of lower population density.

Figure 2. Geographic distribution of convinced, hesitant, and skeptical groups.



Discussion

Principal Findings

This study sheds light on how misinformation related to COVID-19 spread through different social strata of the Spanish population. In line with previous research conducted in other countries, this study reinforces the assumption that social factors, such as political ideology, socioeconomic status, and trust in institutions, play a significant role in determining susceptibility to misinformation [47,48]. However, unlike other studies, our research provides a geographic identification and a comprehensive characterization of the social profiles that are at risk of using erroneous, false, or misleading information as well as a detailed description of the pandemic-related topics that, even today, continue to generate social polarization among the population. Specifically, we have detected that the social profiles most susceptible to misinformation in Spain—in addition to belonging to extreme right-wing ideologies, presenting a low level of socioeconomic status and having a greater distrust of institutions and public health systems—tend to a greater extent to be women without a defined political orientation (possibly due to political disaffection), who use lower-quality information sources (such as social media) and generally unofficial ones, even though they easily share the health information they find through these unofficial media. Likewise, they are groups that reject the COVID-19 vaccines and have shown a lower follow-up of the protective measures decreed by the health authorities during the pandemic.

Among the most relevant topics on misinformation related to COVID-19, it can be emphasized that the themes of the ineffectiveness of additional vaccine doses, pandemic conspiracy theories, and claims of a sudden rise in mortality rates were the topics that elicited the most doubt or agreement within the Spanish population. More than 20% of respondents believed that COVID-19 vaccines lead to a sudden rise in mortality, a perception that likely undermines public motivation to receive further doses. Although Spain is a country with high vaccination rates, these data show that it is necessary to improve the quality of the information provided to the population about the beneficial effects of vaccination and, in particular, existing evidence related to COVID-19 vaccination data. By contrast, the items on the relationship between COVID-19 and 5G mobile

XSL•FO
RenderX

phone network are the ones that cause the least doubts in people, as in other studies that have shown that the narrative linking the onset of the pandemic to 5G antennas is the content that people were least likely to believe [30]. In any case, it is necessary to take into account that a small percentage of the population believes this type of hoax.

Cluster analysis allowed us to obtain 3 groups of individuals based on their beliefs toward COVID-19 content. Being a study with a representative sample of the Spanish population, it is relevant to mention that 20.73% (456/2200) of the respondents have been classified in the group of skeptical individuals or the cluster that tends to believe erroneous, false, or misleading information (ie, commonly based on anecdotal evidence more than on real facts and existing scientific evidence). In addition, 30.27% (666/2200) of the participants could be classified as hesitant individuals who express doubts about the current evidence regarding COVID-19. This percentage is notably high for a general population study and reflects a significant level of uncertainty around COVID-19 among the Spanish population. According to this finding, a work conducted by the Spanish Foundation for Science and Technology in 2022 on scientific misinformation in Spain found that more than a quarter of the population living there received false or misleading information about science on a weekly basis [49]. These data, combined with our findings, underscore the magnitude of the problem of misinformation and the risk of having a high percentage of hesitant population (ie, those who are more susceptible to behavioral changes based on the influence of external information sources independent of their quality). Furthermore, it becomes clear that misinformation not only affects all social strata but also has a clear impact on public health through health behaviors (eg, the rejection of vaccines and protective measures against COVID-19).

The factors related to COVID-19 that best differentiated the groups were those associated with conspiracy theories and COVID-19 vaccination. This finding aligns with previous research [50], demonstrating a strong link between belief in misinformation and a conspiratorial worldview, where individuals suspect that both government and institutions are the root of societal problems. Besides, the strength of content related to antivaccine movements stands out in our study, which is positioned as one of the most relevant misinformation topics in other studies [23], and it is one of the most relevant topics among COVID-19 skeptics who embrace misinformation and conspiracy theories. From these findings, it is evident that the hoaxes associated with vaccines are one of the most difficult narratives for institutions to alleviate [35].

In addition to the analyzed data on the prevalence of misinformation in Spain and its social determinants, our study has identified an interesting geographic distribution of misinformed groups that curiously coincides with the rural areas belonging to the so-called empty Spain, regions that, despite being led by both right-wing and left-wing political parties (eg, Castilla-La Mancha and Extremadura, respectively), have the common characteristic of having a low density of population compared with other country areas. Thus, the higher prevalence of the skeptical group in these areas could be due to the high degree of rurality that could imply, on average, both a lower socioeconomic status of the population (ie, lower education, income, and occupational status) and a certain degree of political disillusionment with the country's main political parties in view of the scarcity of policies to reverse the demographic situation of these areas. As discussed by Southwell et al [51], limited community infrastructure and reduced social cohesion in geographically isolated areas can constrain the diffusion of health information and facilitate the persistence of misinformation, particularly in rural areas such as the empty Spain. This factor may partially explain the observed geographic distribution.

Our findings indicate that individuals with lower socioeconomic status were more likely to be linked to the skeptical group. This aligns with international studies showing that this factor is a clear predictor of health misinformation [52]. However, women were also found to be prone to skepticism, a finding that warrants further investigation to understand gender-specific factors driving COVID-19 misinformation vulnerability, as other studies on gendered scientific misinformation in Spain show that women differentiate misinformative content better than men [49], contrary to the data from our study. It is likely that the role played by misinformation narratives about COVID-19 vaccines and problems with fertility or births is a factor influencing these gender differences [53].

Interestingly, this study highlights the association between institutional trust and misinformed beliefs and conspiracy theories about COVID-19. In the Spanish context, the strong distrust in political parties, the government, and institutions further exacerbated skepticism [38]. While political polarization has been shown to influence misinformation in other countries, our study found that mistrust in health-related institutions, such as the Spanish National Health System and the WHO, was a key determinant of scientifically unsupported opinions linked to skeptical groups in Spain [54]. Therefore, as highlighted in previous studies, this finding underscores the critical role of misinformation in shaping institutional and governmental trust [55], reinforcing the necessity for health institutions to re-establish credibility through transparent communication strategies and consistent messaging. Moreover, as Agley and Xiao [30] point out, it is necessary to keep in mind that believing in false information or conspiracy theories does not necessarily imply that a person cannot simultaneously believe in the official scientific versions. Consequently, simply repeating accepted scientific explanations will not necessarily stop people from believing false information (eg, people who have trust in vaccines but reject COVID-19 vaccines) or even, in practice, combining scientific and anecdotal evidence to make health decisions. Thus, new strategies are needed to strengthen scientific literacy and confidence in science through research transparency and outreach by those individuals working in the scientific field, along with specific knowledge about new technologies that, like the new RNA vaccines, raise more doubts among the population in our countries.

Contrary to some findings that suggest a strong protective role for digital health literacy against misinformation, this study did not find a significant relationship between digital literacy (as measured by the eHealth Literacy Scale) and susceptibility to health misinformation [56]. Instead, the quality of the

XSL•FO
RenderX

information sources emerged as a more decisive factor. Individuals who relied on health professionals, institutional websites, and official media (eg, newspapers) were less likely to be misinformed and subsequently skeptical about COVID-19–related topics. These findings suggest that public health interventions should promote reliable and well-identifiable information sources over simply enhancing digital literacy skills. In addition, consistent with global studies, our analysis also found that individuals in the skeptical group were more likely to share health-related information on social media frequently [57]. This behavior can be linked to the psychological need for validation within communities, which also drives the spread of misinformation and the formation of echo chambers in which possibly many of these beliefs are fed back [58]. Thus, public health campaigns should not only aim at promoting critical thinking but also address the emotional and psychosocial aspects of information and misinformation sharing.

A distinctive aspect of this study was the nuanced role of political ideology in misinformation beliefs. Thus, although like other international studies we identified the strong association between ideologically extreme positions and greater susceptibility to misinformation [16], our results suggest that those who did not disclose their political stance were more likely to belong to the skeptical or hesitant groups, a finding that could be linked to the political disaffection that currently exists in Spain among voters on both sides of the political spectrum. This finding, combined with data on the geographic distribution of the hesitant and skeptical groups around the empty Spain, underscores the importance of considering both political disaffection and extremism as key drivers of health misinformation and social polarization in health behaviors in the Spanish context [59].

Finally, in interpreting our findings, it is essential to recognize that not all misinformation operates with equal epistemic weight or public impact. As emphasized in the report by National Academies [27], some claims, such as those concerning vaccine safety, are clearly at odds with established scientific evidence, whereas others may stem from misunderstandings of evolving science or reflect contested interpretations. Therefore, this complexity in apprehending the inherently fluid nature of misinformation highlights the need for nuanced classification and targeted interventions, particularly in public health communication strategies that aim to distinguish deliberate disinformation from more benign forms of confusion or outdated beliefs.

Limitations

Despite these interesting findings, our work is not without limitations. Given the observational nature of this study, it is not possible to draw causal inferences from the results. Similarly, the cross-sectional design of this study does not allow us to determine the extent to which the responses remain influenced by the social changes brought about by the measures implemented during the pandemic. In addition, the clustering techniques used are not free from potential errors in classifying individuals into groups. Nevertheless, in this study, we have opted to combine the best model fit with criteria of simplicity, interpretability, parsimony, and applicability to other international contexts. In any case, it is also important to highlight the strengths of our work. First, we have identified misinformation topics related to COVID-19 that continue to generate doubts and polarization among the Spanish population. Second, to the best of our knowledge, our study is the first to measure the prevalence of convinced, hesitant, and skeptical groups around COVID-19–related topics in Spain using a nationally representative survey. Third, we have provided a comprehensive characterization of the different social profiles that are susceptible to misinformation in Spain and described their geographic distribution. In summary, our study offers crucial insights into the scope of measures that should be adopted and, particularly, the social determinants that could be targeted to combat misinformation.

Conclusions

This study provides a comprehensive overview of COVID-19–related misinformation in Spain, offering valuable insights into the social, economic, and ideological factors that influence susceptibility to false or misleading information. By identifying and profiling the convinced, hesitant, and skeptical groups, we have demonstrated the significant polarization surrounding health-related issues, such as vaccine hesitancy about the new vaccines and conspiracy theories. These findings emphasize the importance of targeted interventions to improve public understanding of complex health information to combat the global threat of misinformation. Our findings highlight the critical role of institutional trust in shaping public attitudes and health behaviors, underscoring the need for health and governmental institutions to rebuild credibility through transparent communication and consistent messaging. In addition, the influence of socioeconomic status, political ideology, and information sources on misinformation susceptibility suggests that addressing these structural determinants is essential for the design of effective public health strategies.

From a broader perspective, this study contributes to the global research stream on the impact of misinformation on health behaviors, particularly in the context of future pandemics or health crises. Future research should explore longitudinal approaches to examine how misinformation beliefs evolve over time and assess the effectiveness of interventions designed to reduce the spread of false information. Furthermore, exploring the role of digital literacy, alongside strategies to promote reliable, evidence-based information sources, is crucial to addressing the challenges of emerging information and misinformation ecosystems, especially as we await the full impact of new artificial intelligence technologies. Our findings highlight the necessity of a multifaceted approach that integrates education, health policy development, media platforms, and community engagement to effectively counter the persistent menace of health misinformation.

XSL•FO
RenderX

Acknowledgments

We would like to acknowledge the support of the University Research Institute for Sustainable Social Development and the University of Cádiz. The publication is part of project DCODES (PID2020-118589RB-I00), granted by the Spanish Ministry of Science and Innovation and financed by MCIN/AEI/10.13039/501100011033. The present study has also been supported by the project NETDYNAMIC (CNS2022-135907), funded by MCIN/AEI/10.13039/501100011033 and by the European Union "Next Generation EU"/PRTR.

Conflicts of Interest

None declared.

References

1. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
2. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat Hum Behav 2020 Dec;4(12):1285-1293 [FREE Full text] [doi: 10.1038/s41562-020-00994-6] [Medline: 33122812]
3. Novel coronavirus (2019-nCoV): situation report-13. World Health Organization. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200202-sitrep-13-ncov-v3.pdf [accessed 2025-05-29]
4. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb 29;395(10225):676 [FREE Full text] [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
5. Briand SC, Cinelli M, Nguyen T, Lewis R, Prybylski D, Valensise CM, et al. Infodemics: a new challenge for public health. Cell 2021 Dec 09;184(25):6010-6014 [FREE Full text] [doi: 10.1016/j.cell.2021.10.031] [Medline: 34890548]
6. Alvarez-Galvez J, Suarez-Lledo V, Rojas-Garcia A. Determinants of infodemics during disease outbreaks: a systematic review. Front Public Health 2021;9:603603 [FREE Full text] [doi: 10.3389/fpubh.2021.603603] [Medline: 33855006]
7. Suarez-Lledo V, Ortega-Martin E, Carretero-Bravo J, Ramos-Fiol B, Alvarez-Galvez J. Unraveling the use of disinformation hashtags by social bots during the COVID-19 pandemic: social networks analysis. JMIR Infodemiology 2025 Jan 09;5:e50021 [FREE Full text] [doi: 10.2196/50021] [Medline: 39786891]
8. Naughton J. Fake news about COVID-19 can be as dangerous as the virus. The Guardian. 2020 Mar 14. URL: https://www.theguardian.com/commentisfree/2020/mar/14/fake-news-about-covid-19-can-be-as-dangerous-as-the-virus [accessed 2025-05-29]
9. Van Bavel JJ, Gadarian SK, Knowles E, Ruggeri K. Political polarization and health. Nat Med 2024 Nov;30(11):3085-3093 [FREE Full text] [doi: 10.1038/s41591-024-03307-w] [Medline: 39455731]
10. Ahmed S, Rasul ME. Examining the association between social media fatigue, cognitive ability, narcissism and misinformation sharing: cross-national evidence from eight countries. Sci Rep 2023 Sep 18;13(1):15416 [FREE Full text] [doi: 10.1038/s41598-023-42614-z] [Medline: 37723265]
11. Gramacho W, Turgeon M, Kennedy J, Stabile M, Mundim PS. Political preferences, knowledge, and misinformation about COVID-19: the case of Brazil. Front Polit Sci 2021 May 14;3:646430. [doi: 10.3389/fpos.2021.646430]
12. Hatcher W. A failure of political communication not a failure of bureaucracy: the danger of presidential misinformation during the COVID-19 pandemic. Am Rev Public Adm 2020 Aug 11;50(6-7):614-620 [FREE Full text] [doi: 10.1177/0275074020941734]
13. Rocha YM, de Moura GA, Desidério GA, de Oliveira CH, Lourenço FD, de Figueiredo Nicolete LD. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review. Z Gesundh Wiss 2021 Oct 09;31(7):1-10 [FREE Full text] [doi: 10.1007/s10389-021-01658-z] [Medline: 34660175]
14. Aïmeur E, Amri S, Brassard G. Fake news, disinformation and misinformation in social media: a review. Soc Netw Anal Min 2023 Feb 09;13(1):30 [FREE Full text] [doi: 10.1007/s13278-023-01028-5] [Medline: 36789378]
15. Dib F, Mayaud P, Chauvin P, Launay O. Online mis/disinformation and vaccine hesitancy in the era of COVID-19: why we need an eHealth literacy revolution. Hum Vaccin Immunother 2022 Dec 31;18(1):1-3 [FREE Full text] [doi: 10.1080/21645515.2021.1874218] [Medline: 33625960]
16. Lee SJ, Lee C, Hwang H. The role of deliberative cognitive styles in preventing belief in politicized COVID-19 misinformation. Health Commun 2023 Dec;38(13):2904-2914. [doi: 10.1080/10410236.2022.2125119] [Medline: 36134653]
17. Nivette A, Ribeaud D, Murray A, Steinhoff A, Bechtiger L, Hepp U, et al. Non-compliance with COVID-19-related public health measures among young adults in Switzerland: insights from a longitudinal cohort study. Soc Sci Med 2021 Jan;268:113370 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113370] [Medline: 32980677]
18. Zhang D, Fang B, Yang L, Cai Y. Disinformation of text mining online about tobacco and the COVID-19 discussed on Sina Weibo. Tob Induc Dis 2021;19:83 [FREE Full text] [doi: 10.18332/tid/142776] [Medline: 34720798]
19. Pomeranz JL, Schwid AR. Governmental actions to address COVID-19 misinformation. J Public Health Policy 2021 Jun;42(2):201-210 [FREE Full text] [doi: 10.1057/s41271-020-00270-x] [Medline: 33510401]

XSL•FO
RenderX

20. Southwell BG, Brennen JS, Paquin R, Boudewyns V, Zeng J. Defining and measuring scientific misinformation. Ann Am Acad Pol Soc Sci 2022 May 05;700(1):98-111 [FREE Full text] [doi: 10.1177/00027162221084709]
21. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
22. Chou WS, Oh A, Klein WM. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
23. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
24. Fetzer JH. Disinformation: the use of false information. Minds Mach 2004 May;14(2):231-240 [FREE Full text] [doi: 10.1023/b:mind.0000021683.28604.5b]
25. Wu K, Yang S, Zhu KQ. False rumors detection on Sina Weibo by propagation structures. In: Proceedings of the 31st International Conference on Data Engineering. 2015 Presented at: ICDE '15; April 13-17, 2015; Seoul, South Korea p. 651-662 URL: https://ieeexplore.ieee.org/document/7113322
26. Hernon P. Disinformation and misinformation through the internet: findings of an exploratory study. Gov Inf Q 1995;12(2):133-139 [FREE Full text] [doi: 10.1016/0740-624X(95)90052-7]
27. Understanding and addressing misinformation about science. National Academies of Sciences, Engineering, and Medicine. 2025. URL: https://doi.org/10.17226/27894 [accessed 2025-05-29]
28. León B, Martínez-Costa MP, Salaverría R, López-Goñi I. Health and science-related disinformation on COVID-19: a content analysis of hoaxes identified by fact-checkers in Spain. PLoS One 2022 Apr 13;17(4):e0265995 [FREE Full text] [doi: 10.1371/journal.pone.0265995] [Medline: 35417493]
29. Helgeson SA, Mudgalkar RM, Jacobs KA, Lee AS, Sanghavi D, Moreno Franco P, National COVID Cohort Collaborative (N3C). Association between X/Twitter and prescribing behavior during the COVID-19 pandemic: retrospective ecological study. JMIR Infodemiology 2024 Nov 18;4:e56675 [FREE Full text] [doi: 10.2196/56675] [Medline: 39556417]
30. Agley J, Xiao Y. Misinformation about COVID-19: evidence for differential latent profiles and a strong association with trust in science. BMC Public Health 2021 Jan 07;21(1):89 [FREE Full text] [doi: 10.1186/s12889-020-10103-x] [Medline: 33413219]
31. Farhoudinia B, Ozturkcan S, Kasap N. Emotions unveiled: detecting COVID-19 fake news on social media. Humanit Soc Sci Commun 2024 May 18;11(1):640 [FREE Full text] [doi: 10.1057/S41599-024-03083-5]
32. Druckman JN, Ognyanova K, Baum MA, Lazer D, Perlis RH, Volpe JD, et al. The role of race, religion, and partisanship in misperceptions about COVID-19. Group Process Intergr Relat 2021 May 31;24(4):638-657 [FREE Full text] [doi: 10.1177/1368430220985912]
33. Lanier HD, Diaz MI, Saleh SN, Lehmann CU, Medford RJ. Analyzing COVID-19 disinformation on Twitter using the hashtags #scamdemic and #plandemic: retrospective study. PLoS One 2022 Jun 22;17(6):e0268409 [FREE Full text] [doi: 10.1371/journal.pone.0268409] [Medline: 35731785]
34. Bernacer J, García-Manglano J, Camina E, Güell F. Polarization of beliefs as a consequence of the COVID-19 pandemic: the case of Spain. PLoS One 2021 Jul 13;16(7):e0254511 [FREE Full text] [doi: 10.1371/journal.pone.0254511] [Medline: 34255781]
35. Fang Y. Why do people believe in vaccine misinformation? The roles of perceived familiarity and evidence type. Health Commun 2024 Nov 21;39(13):3480-3492. [doi: 10.1080/10410236.2024.2328455] [Medline: 38514925]
36. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman AL, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020 Oct 14;7(10):201199 [FREE Full text] [doi: 10.1098/rsos.201199] [Medline: 33204475]
37. Alvarez-Galvez J, Anastasiou A, Lamnisos D, Constantinou M, Nicolaou C, Papacostas S, et al. The impact of government actions and risk perception on the promotion of self-protective behaviors during the COVID-19 pandemic. PLoS One 2023;18(4):e0284433 [FREE Full text] [doi: 10.1371/journal.pone.0284433] [Medline: 37068083]
38. Nefes TS, Präg P, Romero-Reche A, Pereira-Puga M. Believing in conspiracy theories in Spain during the COVID-19 pandemic: drivers and public health implications. Soc Sci Med 2023 Nov;336:116263 [FREE Full text] [doi: 10.1016/j.socscimed.2023.116263] [Medline: 37797543]
39. Stubenvoll M. Investigating the heterogeneity of misperceptions: a latent profile analysis of COVID-19 beliefs and their consequences for information-seeking. Sci Commun 2022 Dec;44(6):759-786 [FREE Full text] [doi: 10.1177/10755470221142304] [Medline: 38603223]
40. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
41. Paramio Pérez G, Almagro BJ, Hernando Gómez Á, Aguaded Gómez JI. Validación de la escala eHealth Literacy (eHEALS) en población universitaria española. Rev Esp Salud Publica 2015;89(3):329-338 [FREE Full text] [doi: 10.4321/S1135-57272015000300010] [Medline: 26388346]
42. Barman D, Guo Z, Conlan O. The dark side of language models: exploring the potential of LLMs in multimedia disinformation generation and dissemination. Mach Learn Appl 2024 Jun;16:100545. [doi: 10.1016/j.mlwa.2024.100545]

XSL•FO
RenderX

43. Littrell S, Klofstad C, Uscinski JE. The political, psychological, and social correlates of cryptocurrency ownership. PLoS One 2024;19(7):e0305178 [FREE Full text] [doi: 10.1371/journal.pone.0305178] [Medline: 38959184]
44. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. Behav Res Methods 2016 Sep;48(3):936-949. [doi: 10.3758/s13428-015-0619-7] [Medline: 26174714]
45. Yates A. Multivariate Exploratory Data Analysis: A Perspective on Exploratory Factor Analysis. New York, NY: State University of New York Press; Mar 1989.
46. Rosseel Y. An R package for structural equation modeling. J Stat Softw 2012;48(2):1-36 [FREE Full text] [doi: 10.18637/jss.v048.i02]
47. Scherer LD, McPhetres J, Pennycook G, Kempe A, Allen LA, Knoepke CE, et al. Who is susceptible to online health misinformation? A test of four psychosocial hypotheses. Health Psychol 2021 Apr;40(4):274-284 [FREE Full text] [doi: 10.1037/hea0000978] [Medline: 33646806]
48. Wang T, Yu W. Alternative sources use and misinformation exposure and susceptibility: the curvilinear moderation effects of socioeconomic status. Telemat Inform 2022 May;70:101819 [FREE Full text] [doi: 10.1016/j.tele.2022.101819]
49. Desinformación científica en España. Fundación Española para la Ciencia y la Tecnología. URL: https://www.fecyt.es/publicaciones/desinformacion-cientifica-en-espana [accessed 2025-05-29]
50. Hanley HW, Kumar D, Durumeric Z. A golden age: conspiracy theories' relationship with misinformation outlets, news media, and the wider internet. Proc ACM Hum Comput Interact 2023 Oct 04;7(CSCW2):1-33 [FREE Full text] [doi: 10.1145/3610043]
51. Southwell BG, Slater JS, Rothman AJ, Friedenberg LM, Allison TR, Nelson CL. The availability of community ties predicts likelihood of peer referral for mammography: geographic constraints on viral marketing. Soc Sci Med 2010 Nov;71(9):1627-1635. [doi: 10.1016/j.socscimed.2010.08.009] [Medline: 20864236]
52. Gaitán-Rossi P, Mendez-Rosenzweig M, García-Alberto E, Vilar-Compte M. Barriers to COVID-19 vaccination among older adults in Mexico City. Int J Equity Health 2022 Jun 19;21(1):85 [FREE Full text] [doi: 10.1186/s12939-022-01685-6] [Medline: 35717236]
53. Moschese V, De Angelis L, Capogna MV, Graziani S, Baglivo F, Pietropolli A, Italian Society of Pediatric Allergology and Immunology (SIAIP) Vaccine Committee. Vaccine hesitancy and knowledge regarding maternal immunization among reproductive age women in central Italy: a cross sectional study. Front Glob Womens Health 2023 Sep 14;4:1237064 [FREE Full text] [doi: 10.3389/fgwh.2023.1237064] [Medline: 37779637]
54. Pickup M, Stecuła D, Van der Linden C. Who shares conspiracy theories and other misinformation about COVID-19 online: survey evidence from five countries. J Quant Descr Digit Media 2022 Oct 28;2:1-38 [FREE Full text] [doi: 10.51685/jqd.2022.024]
55. Ognyanova K, Lazer D, Robertson RE, Wilson C. Misinformation in action: fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. Harv Kennedy Sch Misinformation Rev 2020 Jun 2:66 [FREE Full text] [doi: 10.37016/MR-2020-024]
56. Anthonysamy L, Sivakumar P. A new digital literacy framework to mitigate misinformation in social media infodemic. Glob Knowl Mem Commun 2022 Nov 14;73(6/7):809-827 [FREE Full text] [doi: 10.1108/gkmc-06-2022-0142]
57. Vraga EK, Tully M. News literacy, social media behaviors, and skepticism toward information on social media. Inf Commun Soc 2019 Jul 18;24(2):150-166 [FREE Full text] [doi: 10.1080/1369118X.2019.1637445]
58. Paletz SB, Auxier BE, Golonka EM. Motivation to share. In: Paletz SB, Auxier BE, Golonka EM, editors. A Multidisciplinary Framework of Information Propagation Online. Cham, Switzerland: Springer; 2019:37-45.
59. Angerer S, Glätzle-Rützler D, Lergetporer P, Rittmannsberger T. Beliefs about social norms and gender-based polarization of COVID-19 vaccination readiness. Eur Econ Rev 2024 Apr;163:104640 [FREE Full text] [doi: 10.1016/j.euroecorev.2023.104640]

Abbreviations

EFA: exploratory factor analysis
WHO: World Health Organization

XSL•FO
RenderX

Edited by R Cuomo; submitted 11.12.24; peer-reviewed by B Southwell; comments to author 22.04.25; revised version received 29.04.25; accepted 29.04.25; published 16.06.25.

Please cite as:
Alvarez-Galvez J, Lagares-Franco C, Ortega-Martin E, De Sola H, Rojas-García A, Sanz-Marcos P, Almenara-Barrios J, Kassianos AP, Montagni I, Camacho-García M, Serrano-Macías M, Carretero-Bravo J
Measurement, Characterization, and Mapping of COVID-19 Misinformation in Spain: Cross-Sectional Study
JMIR Infodemiology 2025;5:e69945
URL: https://infodemiology.jmir.org/2025/1/e69945
doi:10.2196/69945
PMID:

©Javier Alvarez-Galvez, Carolina Lagares-Franco, Esther Ortega-Martin, Helena De Sola, Antonio Rojas-García, Paloma Sanz-Marcos, José Almenara-Barrios, Angelos P Kassianos, Ilaria Montagni, María Camacho-García, Maribel Serrano-Macías, Jesús Carretero-Bravo. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 16.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Modularity of Online Social Networks and COVID-19 Misinformation Spreading in Russia: Combining Social Network Analysis and National Representative Survey

Boris Pavlenko[1], MS

International Center for the Study of Institutions and Development, HSE University, Moscow, Russian Federation

Corresponding Author:
Boris Pavlenko, MS
International Center for the Study of Institutions and Development
HSE University
20 Myasnitskaya ulitsa
Moscow, 101000
Russian Federation
Phone: 7 9253740978
Email: bvpavlenko@hse.ru

Abstract

Background: The outbreak of SARS-CoV-2 in 2019 was accompanied by a rise in the popularity of conspiracy theories. These theories often undermined vaccination efforts. There is evidence that the spread of misinformation about COVID-19 is associated with online social media use. Online social media enables network effects that influence the dissemination of information. It is important to distinguish between the effects of using social media and the network effects that occur within the platform.

Objective: This study aims to investigate the association between the modularity of online social networks and the spread of, as well as attitudes toward, information and misinformation about COVID-19.

Methods: This study used data from the social network structure of the online social media platform Vkontakte (VK) to construct an adjusted modularity index (fragmentation index) for 166 Russian towns. VK is a widely used Russian social media platform. The study combined town-level network indices with data from the poll "Research on COVID-19 in Russia's Regions" (RoCIRR), which included responses from 23,000 individuals. The study measured respondents' knowledge of both fake and true statements about COVID-19, as well as their attitudes toward these statements.

Results: A positive association was observed between town-level fragmentation and individuals' knowledge of fake statements, and a negative association with knowledge of true statements. There is a strong negative association between fragmentation and the average attitude toward true statements ($P<.001$), while the association with attitudes toward fake statements is positive but statistically insignificant ($P=.55$). Additionally, a strong association was found between network fragmentation and ideological differences in attitudes toward true versus fake statements.

Conclusions: While social media use plays an important role in the diffusion of health-related information, the structure of social networks can amplify these effects. Social network modularity plays a key role in the spread of information, with differing impacts on true and fake statements. These differences in information dissemination contribute to variations in attitudes toward true and fake statements about COVID-19. Ultimately, fragmentation was associated with individual-level polarization on medical topics. Future research should further explore the interaction between social media use and underlying network effects.

(JMIR Infodemiology 2025;5:e58302) doi:10.2196/58302

KEYWORDS

COVID-19; Russia; infodemic; modularity; online social networks; VK; misinformation; beliefs; information spreading; social network analysis

XSL•FO
RenderX

Introduction

Background

The COVID-19 pandemic challenged many societies. Governments had to reduce mortality risks while minimizing economic shocks. One of the most efficient ways to mitigate such impacts is vaccination. There is considerable variation in people's behavior regarding vaccination [1]. Evidence suggests that one of the factors associated with increased vaccine hesitancy was the use of social networking and belief in COVID-19 fake news and conspiracy theories [2-4].

Pivetti et al [5] proposed a mechanism that links social media usage, COVID-19 fake news, and vaccination hesitancy. First, online social media users are more likely to encounter misinformation about COVID-19 compared with traditional media users [6]. Second, exposure to fake news increases the likelihood that some users will believe in conspiracy theories [3]. Belief in conspiracy theories, in turn, raises the perceived risk of vaccination while lowering the perceived risk of the disease [7]. Finally, these altered risk perceptions reduce the likelihood of vaccination and negatively impact other preventive behaviors [8,9].

However, a causal relationship between online social media use and belief in conspiracy theories has not yet been established. Moreover, there is evidence that the unavailability of social media positively influences the popularity of searches for COVID-19 fake news [10]. Finally, social media can also promote proscience or provaccine messages that reduce vaccination hesitancy [11,12]. In other words, what matters is the content observed on social media, not social media use itself.

It is important to study the Russian case of COVID-19 pandemic attitudes for several reasons. Russia exhibited relatively high levels of vaccine hesitancy [1], which may have contributed to one of the highest excess mortality rates in the world [13,14]. Additionally, the availability of data and variation at both the individual and town levels offer valuable tools for examining the issue.

Theories concerning the consumption of information on online social media generate mixed hypotheses. On the one hand, prior beliefs may be reinforced by filter bubbles created by social networking site algorithms that tailor users' feeds to match their interests [15]. Echo chambers further enable individuals to interact primarily with others who share similar views [16]. On the other hand, individuals may also encounter differing perspectives online [17], which can increase the diversity of their views.

The spread of misinformation about COVID-19 is supported by echo chamber theory and evidence that homogeneous groups are more likely to disseminate fake news [18-20]. However, when information escapes an echo chamber, only a minority of outsiders accept it [18]. According to the opinion dynamics model [21], individuals are likely to change their beliefs only when the ideological gap with the person they interact with is small. Similarly, there are homogeneous groups that disseminate scientific information as well [16]. Previous review studies have shown that it is primarily human users, rather than bot accounts, who spread fake news to most users [22], underscoring the importance of studying the networks of ordinary users.

This paper provides evidence on the indirect effects of online social media on misinformation diffusion and the formation of personal beliefs. The study findings indicate that it is not only the platform that matters but also the structure of citizens' online interpersonal networks. In this study, data on the characteristics of online social media networks on Vkontakte (VK) were combined with polling data from users in the corresponding towns. It was shown that town-level fragmentation of these networks is associated with the spread of misinformation about COVID-19 and the share of fake statements encountered by respondents. The results remain robust after controlling for individual characteristics, including the use of online and traditional media, fear of COVID-19, and household experience with the disease. This study contributes to the growing literature on the spread of health misinformation on social media [23].

The presented approach differs from studies that rely solely on online social media data. Numerous studies have examined the spread of specific conspiracy theories within online environments [23]. However, a key question remains: how does the social environment—measured through a network index—affect an individual's likelihood of encountering true or false statements about COVID-19? Finally, how do networks shape opinions about these statements, and what role does online social media play in this process?

Related Work

This section is divided into a brief literature review covering studies on information diffusion in social networks, network modularity, and its outcomes. These areas are explored to understand how information spreads within networks and how modularity influences social outcomes. Following this, the estimation of fragmentation will be discussed and the formulation of hypotheses will be presented based on the literature review and methodological approach.

Information Spread in the Network

Information diffusion refers to the process of spreading information among agents or communities within a network [24]. The efficiency of this diffusion is influenced by the network's structure and the characteristics of its nodes and links. The concept of the "strength of ties" introduces the role of interpersonal trust between individuals in a network [25]. Strong ties exist between nodes that share many neighbors, while weak ties connect nodes with fewer common neighbors. In social interactions, strong ties typically form between individuals who know each other well. While strong ties are more effective for persuading people within the network, weak ties can facilitate broader and more efficient information spread [26]. Homophily—the tendency of similar nodes to group together based on shared traits—also enhances the speed of information diffusion within these clusters [27]. At the structural level, the concept of communities is central; communities are groups of nodes clustered based on their connectivity [28]. Modularity, which measures the strength of the division of a network into communities, plays a significant role in shaping how information diffuses [29-31]. The foundational theoretical framework for

XSL•FO
RenderX

studying information diffusion is the Epidemic Spread Model [32].

All of these theories are interdependent; for example, higher homophily is expected among nodes within the same community. In other words, individuals with more ties to each other are likely to share common characteristics. In this study, no personal information about the nodes in the network was used, which limits the range of analytical tools available. Instead, the modularity index was used to measure potential polarization within a town [29].

Network's Modularity and Its Outcomes

Network modularity influences the speed of information flow within a network [30]. In animal networks, modularity has been shown to affect the efficiency of information transmission in a nonlinear manner [31]. Similarly, in the context of COVID-19, the structure of the network influences disease spread even when the total number of links remains constant [33]. Strategies that incorporate knowledge of community structure significantly improve the efficiency of containment efforts at the town level [33-35].

Previous studies have examined the fragmentation of online social networks and the spread of misinformation within them. For a review of the spread of fake news and belief change, see [36]. These studies have shown that individuals within the same community are more likely to influence each other's opinions than outsiders. A concept particularly relevant to this study is that of epistemic echo chambers [37], which arise when individuals rely primarily on their social networks for information, and when those networks are fragmented—thereby limiting exposure to diverse viewpoints.

Several studies have examined the relationship between network modularity and economic or social outcomes. In this context, modularity is used as a tool to capture social capital. For instance, the modularity of networks at the municipal level has been shown to predict economic development [38], while at the town level, it has been linked to levels of corruption [39]. More broadly, network modularity serves as one indicator for predicting societal polarization at the network level [29].

Most studies have focused on how misinformation about COVID-19 spreads online, how social media use influences vaccination intentions [3], or knowledge of specific conspiracy theories [18]. For a review of studies on this topic, see [3,40].

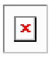
What distinguishes this study from previous work is its focus on the indirect effects of network modularity on the town-level spread of information and misinformation. Specifically, the author posits that the likelihood of epistemic echo chambers emerging in a town is correlated with the fragmentation of its network. The study estimates the association between the fragmentation of VK's social media network at the town level and the likelihood of individuals encountering fake news and related opinions, while controlling for individual characteristics. The study findings suggest that online social media polarization and fragmentation can have broader consequences, extending beyond individual opinion formation to affect information dynamics at the community level.

Methods

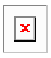
Fragmentation Index as Network's Modularity Measure

As shown previously, network structure plays a key role in the transmission of information. At the macro level, the focus shifts to communities, where understanding how members are interconnected becomes essential. To capture this, a modularity index was used as a measure of network interconnectedness. In this study, an improved version of the modularity index was applied to better account for differences in network size.

The *fragmentation index* indicates the extent to which nodes tend to cluster together and remain separate from others. The Louvain algorithm [41] divides a network into relatively segregated groups (communities). It should be noted that the problem of community detection is NP-hard; therefore, the algorithm provides a heuristic solution. The fragmentation index estimates the likelihood that edges occur within small groups rather than between them.



Equation 1 represents the modularity index of the network that captures the ratio of ties with and between community members, where L is the number of edges in the network, L_k is the number of edges adjacent to members of community k, L_k^w is the number of edges within community k, and $Q(S)$ is the modularity of the network.



Equation 2 shows the adjustment for modularity calculated on the alternative network if all ties were within community members, where $Q_{max}(S)$ is the theoretical maximum if all edges were within the communities.

Equation 3 represents the fragmentation index:

$F_s=[Q(S)]/[Q_{max}(S)]$ (3)

In other words, the fragmentation index is a modularity measure adjusted for network size, allowing for meaningful comparison of network structures across different sizes. This methodology follows previous research in the field [39,42].

The modularity index has been shown to influence the spread of information [30,31,43]. Moreover, it has been established that a network's modular structure, as measured by the modularity index, affects the spread of diseases [33]. Both processes are modeled using the same basic theoretical framework [32].

Hypothesis Formulation

Higher fragmentation indicates that there are more connections within communities than between them. In such networks, information is more likely to circulate within a community rather than spread across communities. If we assume that community members share certain characteristics that make them more similar to one another than to outsiders, they are more likely to agree with the information they encounter and

XSL•FO
RenderX

to disseminate it further [18]. This, in turn, increases the likelihood of repeated exposure to the same information.

Building on the theory of the strength of weak ties [25], the diversity-bandwidth trade-off—comparing the properties of weak ties with strong ties—can also be applied at the community level [26]. Existing literature shows that both conspiracy theories and scientific news tend to spread within closed, homogeneous groups [16,44]. In more fragmented networks, such groups are more likely to exist, as higher homogeneity increases the probability that individuals belong to the same community. While both weak and strong ties can facilitate the spread of new information, the effectiveness depends on the type of information and the surrounding environment [26]. In the context of online news dissemination, the prevalence of strong ties does not necessarily offer an advantage—unless trusted users (those with higher bandwidth) are more likely to believe in and share accurate information.

- Hypothesis 1(a): Fragmentation is negatively associated with the relative number of true statements about COVID-19 encountered.

Lower fragmentation implies the presence of more weak ties, which can facilitate the spread of accurate information in the context of news dissemination. In less fragmented networks, individuals who believe in fake news are more likely to be connected to others with differing views, increasing the chances of exposure to corrective information.

In other words, while less fragmented networks tend to exhibit a greater diversity of views, the likelihood of encountering fake information online is lower. This is because individuals who believe in fake news are less likely to share it in such environments, as the audience response to these posts tends to be weaker. This behavior can be explained by the concept of strategic self-presentation [45,46]. Users are more likely to share content that portrays them favorably and garners positive feedback from others. In diverse networks, posts expressing extreme views are likely to be received less favorably due to greater ideological distance between the poster and the audience. As a result, more neutral content is amplified, as it attracts broader approval. While this mechanism is particularly pronounced on online platforms due to algorithmic reinforcement [36], it also holds relevance in offline social interactions.

- Hypothesis 1(b): Fragmentation is positively associated with the relative number of fake statements about COVID-19 encountered.

In towns with more fragmented networks, individuals are more likely to be divided into numerous closed communities. This structure fosters epistemic echo chambers, where exposure to diverse viewpoints is limited [37]. Within such communities, the spread of fake news and conspiracy theories is more likely to occur, increasing the chances that individuals connected to or adjacent to these communities will encounter misinformation.

- Hypothesis 2(a): Fragmentation is negatively associated with belief in true statements about COVID-19.

Groups within fragmented networks are more likely to share similar views and exert strong influence on one another's opinions [36]. That is, individuals within the same community tend to hold similar beliefs. However, different communities within the same network may hold opposing views, contributing to overall polarization. Within individual communities, lower diversity of opinion, fewer weak ties, and reduced efficiency of information flow may limit exposure to accurate information. As a result, individuals in more fragmented networks may be less informed and less likely to believe true statements.

- Hypothesis 2(b): Fragmentation is positively associated with belief in fake statements about COVID-19.

In more fragmented networks, the spread of fake statements is more likely, as the existence of closed communities where misinformation is accepted and reinforced becomes more probable. Individuals within or adjacent to these communities are more likely to encounter and believe such statements, which are then further propagated. However, the overall effect may be limited. When fake news circulates beyond these echo chambers, it often encounters resistance or negative reactions. Such exposure can prompt critical evaluation among those outside the community, leading to skepticism and potential revisions in belief [18].

Causality

The methodology used in this study does not permit causal inference between network fragmentation, the likelihood of encountering information, and an individual's belief in that information. However, reviews of prior research offer insights into potential causal mechanisms linking information spread and opinion change [36]. Previous studies indicate that fragmented networks are prone to increasing polarization [47], and the dissemination of polarized content can further intensify opinion divergence [48]. Additionally, exposure to opposing viewpoints may reinforce existing beliefs—a phenomenon known as belief entrenchment [49]—whereas exposure to more neutral information can reduce polarization and shift opinions [21].

Summing up, in fragmented networks, the emergence of closed groups that believe in fake news is more likely due to the formation of epistemic echo chambers. Strong believers within these groups may share misinformation beyond their immediate community. However, such exposure is unlikely to change the views of others—particularly if those individuals already hold conflicting beliefs.

Data

The dataset combines the "Research on COVID-19 in Russia's Regions" (RoCIRR) database on COVID-19 in Russia with social network data collected from the most popular online social network, VK. The RoCIRR dataset includes responses from over 23,000 individuals. Data were collected between November 4 and December 1, 2020, through an online survey of respondents from 61 Russian regions, designed to be representative at the regional level. Detailed information about the poll is provided in Multimedia Appendix 1. Respondent recruitment was conducted by Online Market Intelligence, an online polling company operating in Russia and analogous to

XSL•FO
RenderX

Amazon's Mechanical Turk (MTurk) in the United States. Online Market Intelligence maintains a panel of the adult population in cities with over 100,000 residents. In addition, a subsample of respondents was drawn from smaller towns and rural areas for the RoCIRR database. For more details on the RoCIRR dataset, see [50].

From this dataset, towns with more than 10 respondents were selected, resulting in 168 towns. For these towns, VK users were identified using VK API instruments [51] to construct social networks. Additional data sources included official municipal statistics from Rosstat [52] and the main regional socioeconomic statistics from Rosstat [53], which were used to create town-level controls for wages and population. Municipal statistics were missing for 2 towns due to changes in municipality structure; detailed information about this is provided in Multimedia Appendix 2. All responses marked as "difficult to say" were treated as missing observations and excluded from the survey. The final dataset included 16,587 respondents from 166 towns, with an average of 119.2 respondents per town and 88,700 nodes per network. The dataset covered 60 of the 61 regions included in the original sample.

The final dataset includes both smaller towns with populations under 100,000 and major cities such as Moscow and Saint Petersburg. Most towns in the sample (95/166) have populations below 250,000, though large cities (with populations over 500,000) are also represented. The sample comprises both regional capitals and ordinary towns. A detailed distribution of town populations in the sample is provided in Multimedia Appendix 3.

About VK

VK is a Russian online social network where users can post public messages, add other users to their "friend" lists, and send private messages. As of November 2023, VK was the fifth most popular website in Russia [54] and the most popular online social network, with over 80 million monthly active users.

Data from VK were collected using the VK API between February 7 and February 25, 2023. It is assumed that networks based on VK data remain relatively stable over time. To support this assumption, 2 samples of the fragmentation index from 2023 and 2024 were used and statistical differences were assessed. It should be noted that community detection algorithms are heuristic, and therefore different runs may result in different node groupings [55]. As a result, observed differences in fragmentation could arise solely from the algorithm's allocation process. Descriptive statistics for the datasets are provided in Multimedia Appendix 4. A paired *t* test (1-tailed) shows no significant differences between the fragmentation indexes from 2023 and 2024 (P=.07). Moreover, it has been shown that the use of retrospective data in online social network analysis is possible, though its effectiveness may be limited by account bans [56].

This section explains the algorithm used to select accounts from VK. For each town, the algorithm attempts to identify accounts based on the following criteria: age group (3 age groups ranging from 18 to 65 years), gender, and number of friends (at least 100 from the same town and no more than 500 in total). In total, accounts are selected from 6 groups—3 age groups, each split by gender. The algorithm was run 3 times, resulting in the selection of up to 18 accounts per town. It should be noted that in larger towns, it is easier to find accounts that meet these criteria due to the larger pool of available accounts. Thus, this method allows for the construction of networks with sizes roughly proportional to the actual population sizes of the towns. The criteria were chosen to simplify calculations, filter out bot accounts at an early stage, and prioritize accounts belonging to individuals likely residing in the towns they list. The threshold of 100 friends is consistent with the mode number of friends observed on platforms such as Facebook and Twitter [57]. Statistical characteristics are presented in Table 1.

Table 1. Characteristics of networks (n=166).

Statistic	Mean (SD)	Range
Fragmentation index	0.449 (0.093)	0.293-0.799
Number of nodes	88,749.800 (81,418.540)	375-541,123
Number of edges	688,859.500 (487,270.800)	2783-2,463,411
Density	0.001 (0.003)	0.00001-0.040
Clustering	0.225 (0.075)	0.030-0.442

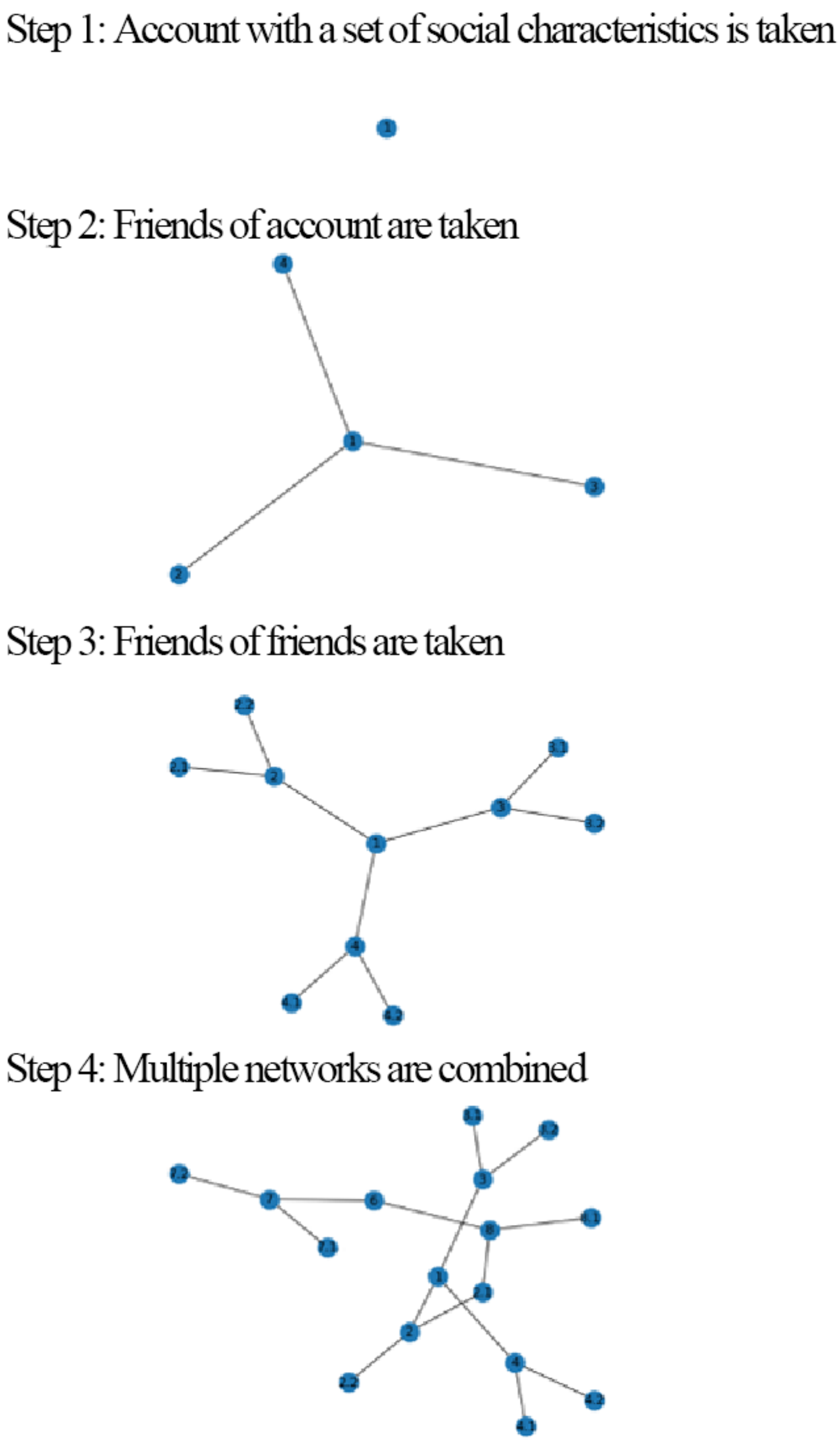
Networks for each account are sampled using breadth-first search to a depth of 2. In other words, for each selected account, their friends and the friends of those friends are included—but acquaintances of friends of friends are not sampled. A network representation of the data collection procedure is shown in Figure 1. While breadth-first search introduces a sampling bias toward high-degree nodes, this bias decreases as the proportion of sampled nodes increases [58]. Furthermore, by using multiple accounts to construct each network, randomness is introduced into the sampling process, which helps reduce potential bias [59].

To estimate node and edge coverage, the author proposes calculating proportions relative to the population as the lower limit of coverage, and proportions relative to VK users as the upper limit. The lower limit reflects the representation of a hypothetical social network encompassing all citizens, while the upper limit captures the representation of the actual VK user network. Based on this, the lower bound indicates that 88,750 of the 435,663 (20.37%) 2019 town populations were screened, suggesting that a significant portion of the population is represented in the network. The upper bound of node coverage—adjusted using the share of respondents who reported

XSL•FO
RenderX

using VK—is estimated to be slightly over 30% (88,750/272,766, 32.54%). The lower limit of edge coverage is nearly 3% (688,859/21,783,150, 3.16%), representing the proportion of edges in the sampled graph relative to all social connections of citizens, with the Dunbar number used as the average number of edges [57]. The upper limit of edge coverage reflects the proportion of edges in the sampled network relative to the expected number of edges among VK users from the same town.

Figure 1. Network data collection procedure.



The upper limit of edge coverage was estimated as the product of the proportion of VK users in the total population and the average share of friends from the same town. These calculations yield an upper estimate of edge coverage of nearly 13% (688,859/5,416,402, 12.72%). Details on the estimation of the expected degree are provided in Multimedia Appendix 5. Previous studies have shown that sampling 15% of a graph is sufficient [60]. While the study's sample is smaller, it is

XSL•FO
RenderX

important to note that the effects of sampling bias on community structure are more limited than on other network characteristics. This is because high-degree nodes—which are more likely to appear in a sample than in the full network—are particularly relevant for determining network structure [61]. Therefore, reducing bias by increasing the share of low-degree nodes in a sample may actually harm the preservation of community structure [61]. Furthermore, the relationship between changes in the fragmentation index and the number of accounts sampled was analyzed, as detailed in Multimedia Appendix 5. As expected, an overlap among the combined networks was observed, resulting in diminishing returns in both edge coverage and node coverage as the network size increases. The most significant spikes in edge coverage occur when the first 3 networks are combined. On average, only a modest increase in edge coverage was observed after combining 10 networks. These estimations suggest that some bias may be introduced by the chosen sampling technique, although the study aims to demonstrate that this bias is minimal. Multimedia Appendix 5 presents edge coverage for each account selected in the sampling process. In a subsample of 49 towns, the average edge coverage reaches nearly 16% (543,687/3,406,978, 15.96%) after selecting 14 accounts and increases to over 20% (726,626/3,389,593, 21.44%) with 18 accounts. Note that the number of iterations to create each of the 49 networks varies by city. Thus, as we examine changes in the average edge coverage introduced by each iteration of sampling, the total number of towns in the sample may vary. However, even with all 18 accounts, some towns still show edge coverage below 10% (637,297/6,600,559, 9.66%). Findings from this study indicate that 15 of the 22 (68%) towns in the subsample achieve edge coverage above 24% (779,875/3,136,689, 24.86%) using this method. This suggests that some bias may result from the network sampling approach. To address this, a robustness check was conducted, restricting the analysis to towns with average edge coverage above 15% (739,677/4,917,172, 15.04%). The results are presented in Multimedia Appendix 6.

Note that it is possible for a person to have multiple accounts or to incorrectly state their hometown, and the poll sample may overrepresent social media users. Data on the number of network nodes for each town and corresponding population census figures are presented in Multimedia Appendix 3.

Main Variables

The main variables of interest are the proportion of known fake and true statements about COVID-19, and the sum of attitudes toward fake and true statements. Additionally, a measure of polarization was constructed.

The control variables are social media usage (VK) and network size. Individual-level controls include gender (dummy), education level, household income, age, fear of COVID-19, a dummy variable indicating household experience with COVID-19, and institutional trust measured as trust in the president. The analysis also controls for television use, as these variables have been used in previous studies on attitudes toward COVID-19 beliefs [6,50]. At the town level, the analysis controls for the natural logarithm of average wage (2019) and the natural logarithm of population (2019). Descriptive statistics for all variables are provided in Table 2.

This study combines individual-level survey data from towns with the network characteristics of those towns. It is assumed that individuals from the same town are influenced by the characteristics of that town's network, as constructed from VK data. A weaker assumption is that differences in VK users' network structures reflect underlying mechanisms of network formation, such as social capital [62]. The estimates are therefore conservative—the study assumes that all respondents are equally affected by the level of fragmentation in their town. In the descriptive statistics (Table 2), town-level characteristics are weighted by the number of respondents in each town; such variables are marked accordingly.

XSL•FO
RenderX

Table 2. Descriptive statistics for used variables (n=16,587).

Statistic	Mean (SD)	Range
Fragmentation index (town level)	0.487 (0.103)	0.293 to 0.799
Gender dummy (1 is female)	0.620 (0.485)	0 to 1
Education level	4.247 (1.055)	1 to 5
Household income group	3.403 (0.820)	1 to 5
Age (years)	36.569 (9.664)	18 to 80
Network size (natural logarithm of the number of nodes; town level)	11.731 (0.790)	5.927 to 13.201
Number of nodes in the network (town level)	155,199.500 (93,322.270)	375 to 541,123
VK use dummy	0.617 (0.486)	0 to 1
Natural logarithm of real wages 2019 (town level)	10.743 (0.147)	10.191 to 11.262
Natural logarithm of population 2019 (town level)	13.310 (0.819)	10.497 to 16.350
Household COVID-19 experience dummy	0.124 (0.330)	0 to 1
Fear of COVID-19	2.936 (0.891)	1 to 4
Trust in president	3.020 (1.466)	1 to 5
Television as a source of news	0.566 (0.496)	0 to 1
Share of encountered fake statements about COVID-19	0.458 (0.286)	0.000 to 1.000
Share of encountered true statements about COVID-19	0.534 (0.257)	0.000 to 1.000
Attitude to true statements about COVID-19	2.143 (2.058)	–8 to 8
Attitude to fake statements about COVID-19	–2.223 (2.792)	–10 to 10
Share of true statements agreed	0.415 (0.257)	0.000 to 1.000
Share of fake statements agreed	0.105 (0.161)	0.000 to 1.000
Difference in attitude between fake and true statements	–4.366 (3.634)	–18 to 12
Misinformation error	0.184 (0.220)	0.000 to 1.750

This section of the paper provides a detailed overview of the primary questions included in the questionnaire used in this study. Table 3 presents the exact translated wording of these key questions. In the first question, which concerns the encounter with true and false statements, respondents are given a list of statements and asked to indicate which ones they have come across, with the option to select "none of the above." In the following question, respondents are asked to indicate their attitudes toward the plausibility of the statements they had previously selected.

Table 3. Translated wordings of the main questions.

Question number	Question wording
Question 1: knowledge of statements	Select statements that you have previously come across
Question 2: attitude to statements (only for questions that the respondent chooses in the list in previous questions)	Rate your attitude to the following statement (Likert scale from 1 to 4, and difficult to answer): 1=The statement is absolutely NOT reliable; 4=The statement is absolutely reliable.

A measure of the spread of COVID-19–related information and misinformation was created. To do this, the number of encountered statements about COVID-19 was summed and the proportion relative to the total number of statements included in the questionnaire was calculated. These calculations are done separately for true and fake statements. As a result, the proportion of encountered true statements and the proportion of encountered fake statements were obtained. Calculating proportions allows us to consider the full set of statements rather than analyzing each one individually. The relationship between the network's fragmentation at the town level and the individual-level encounter with information was interpreted as an indicator of the average level of information spread within a city.

To examine whether information exposure translates into differences in beliefs, attitudes toward the statements were calculated. The relevant survey question measures the extent to which respondents believe a given statement is true. Attitudes were quantified by transforming responses on the Likert scale (originally ranging from 1 to 4) to a scale from –2 to +2, but only for the statements the respondent recognized. If a respondent had never encountered a statement or found it difficult to answer, the corresponding value was set to 0. Then, the transformed scores for all true and fake statements were

XSL•FO
RenderX

summed separately to form the true attitude and fake attitude variables. These transformations were designed to create an opinion variable that neutralizes agreement and disagreement across statements, while retaining observations even when a respondent skipped or was unsure about a particular item. Descriptive statistics for attitudes toward each statement are provided in Table 4. Statements 1-5 are classified as fake, and statements 6-9 as true.

Table 4. Attitude toward COVID-19 statements (n=16,587).

Statistic	Mean (SD)	Range	Never came across, n
1. COVID-19 does not exist	–1.219 (1.158)	–2 to 2	570
2. COVID-19 was developed by the United States	–0.112 (0.807)	–2 to 2	10,998
3. COVID-19 was developed by China	–0.229 (0.910)	–2 to 2	8775
4. 5G towers affect COVID-19 immunity	–0.448 (0.906)	–2 to 2	11,083
5. COVID-19 development was financed by Bill Gates	–0.214 (0.728)	–2 to 2	12,911
6. Russia has already developed its first vaccine	0.860 (1.096)	–2 to 2	3651
7. The Prime Minister of Russia got sick with COVID-19	0.570 (0.963)	–2 to 2	8497
8. Russian Government manipulates official COVID-19 statistics	0.674 (1.024)	–2 to 2	6161
9. Trump takes hydroxychloroquine to prevent getting COVID-19	0.040 (0.393)	–2 to 2	15,220

The share of statements a respondent agrees with was also calculated. This is done by summing binary variables coded as 1 if the respondent agrees (either strongly or weakly) and 0 if they do not. The total is then divided by the number of true or fake statements, respectively. This variable does not differentiate between strong and weak agreement, as that distinction is already captured by the attitude variable. Instead, it is used to examine whether fragmentation is associated with a higher overall share of agreed-upon statements, regardless of the strength of the respondent's stance.

Finally, to estimate the overall levels of misinformation about COVID-19 associated with fragmentation, 2 additional variables were created. First, a measure of the difference between opinions on true and fake statements was constructed, which captures individual-level polarization in beliefs about COVID-19. Second, a total misinformation error score, defined as the sum of 2 components, was calculated: the share of fake statements a respondent agrees with and the share of true statements they disagree with.

VK use was measured using a binary variable indicating whether a respondent reported using the VK social media platform. Network size was calculated as the natural logarithm of the number of nodes in the combined network for each town. This variable serves as an alternative measure of VK's relative popularity, derived from network data rather than self-reported survey responses or census statistics.

The final model that estimates the association with COVID outcomes is as follows:

$$\text{COVID\_outcomes}_i = \alpha + \beta_1\text{Fragmentation}_j + \beta_2\text{VK}_i + \text{Controls}_i + \text{Controls}_j + \varepsilon_i$$

where i is the respondent from town j; COVID_outcomes_i is the share of known statements, attitude measures, differences in attitude toward true and fake statements, and misinformation error; Fragmentation_j is a town-level fragmentation measure; VK_i is a dummy for VK online social media usage; Controls_i is a set of individual controls; and Controls_j is a set of town-level controls.

The model assumes that the estimated value of the fragmentation index is uniform across all respondents within a given town. The relationship between fragmentation and individual responses is treated as linear, allowing us to estimate the average effect of fragmentation at the town level. While this study does not differentiate the effects of network fragmentation between VK users and nonusers, it does estimate the direct effects of VK usage separately.

Ethical Considerations

Only public information was collected from social media for this study. All relevant information about social networks is presented in an aggregated form and users cannot be reidentified. The survey used in the study was approved by the Columbia Institutional Review Board (protocol number IRB-AAAT4453).

Results

Overview

The results are presented in 2 sections: (1) the relationship between fragmentation and the spread of information, and (2) the relationship between fragmentation and attitudes toward statements. The first section demonstrates that network fragmentation is associated with differential patterns in the dissemination of true versus false statements. The second section highlights how fragmentation contributes to a widening gap between belief in misinformation and agreement with accurate statements about COVID-19, suggesting that fragmented online social networks play a role in shaping polarized beliefs.

The observed effect is primarily driven by a strong negative association between network fragmentation and attitudes toward true statements about COVID-19, while the coefficient for attitudes toward fake statements is statistically insignificant (P=.55).

XSL•FO
RenderX

Part 1: Fragmentation and Information Spreading

Results in Table 5 indicate that the fragmentation index is positively associated with the number of encountered fake statements and negatively associated with the number of encountered true statements. Both findings support the initial hypotheses regarding the relationship between tie strength and information diffusion. Although the estimated coefficients are relatively small, they apply uniformly to all individuals within a given town. An increase in fragmentation from the lowest to the highest level observed in the sample would result in a 0.028 increase in the proportion of known fake statements—equivalent to approximately 64% (0.028/0.044) of the estimated effect of VK social media use. For true statements, the effect is even more pronounced, with a coefficient of –0.034, which exceeds the estimated effect of social media usage. Notably, the coefficients for VK social media use are both significant (*P*<.001) and positive, suggesting that online platforms facilitate the dissemination of information, regardless of its veracity.

Table 5. Information spreading.[a,b,c]

Information spread	Dependent variables	
	Share of encountered fake statements about COVID-19	Share of encountered true statements about COVID-19
Fragmentation index	0.057 (0.028); .04	–0.068 (0.025); .006
Size of the network	–0.016 (0.004); <.001	–0.004 (0.003); .18
Vkontakte use dummy	0.044 (0.005); <.001	0.030 (0.004); <.001
Socioeconomic town-level controls[d]	+	+
Individual-level controls[e]	+	+
Observations, n	16,587	16,587
R^2	0.018	0.037

[a]Data for the first 3 rows are presented as estimated β coefficients of linear regression (SE); *P* value.

[b]Exact *P* values are reported.

[c]Robust SEs by town are given in brackets.

[d]Socioeconomic town-level controls include the natural log of wages 2019 and the natural log of population 2019.

[e]Individual-level controls include gender dummy, education level, household income, age, fear of COVID-19, a dummy variable for household experience with COVID-19, institutional trust measured as a trust to the president, and use of television.

Part 2: Fragmentation and Attitude Toward Statements

Columns 1 and 2 of Table 6 show that fragmentation is associated with a lower average attitude toward true statements, while no significant relationship is observed for fake statements (*P*=.55). These serve as the baseline results for the association between network fragmentation and belief in true and fake COVID-19 statements. However, a more nuanced analysis is needed to understand the specific ways in which fragmentation influences belief. In particular, it is important to differentiate between respondents who moderately agree with several statements and those who strongly agree with only one. To capture this distinction, this study uses the share of statements with which a respondent agrees as an additional measure in the estimation.

If we look into columns 3 and 4, we observe that, for true statements, fragmentation leads to a lower share of people agreeing with true statements, but shows an insignificant relationship for fake statements (*P*=.12). Also, we observe that social media use is significantly negatively associated with attitude toward fake statements (*P*<.001), but at the same time, it is positively associated with the share of fake statements the respondent believes are true.

The difference in results between true and fake statements can be explained by the idea of ideological distance. When fake news spreads beyond its bubble of believers, it is eventually observed by those who hold different prior beliefs or have been exposed to different information. This leads to double-checking by people outside the initial bubble, which results in a stronger antifake stance. In other words, in more fragmented towns, fake news is spread more widely, but it does not translate into changes in beliefs. Similarly, VK users have a significantly lower average attitude toward fake news but agree with a higher share of fake statements. The result for VK use corresponds with the findings of Bursztyn et al [63], who show that VK use leads to greater variation in beliefs. While they attribute this to network echo-chamber effects, the relationship may be more complex, as no significant relationship was found between network fragmentation and the popularity of extreme opinions about COVID-19.

When looking at the combined results for differences in opinions about fake and true statements, polarization introduced by fragmentation was observed. Column 5 shows the difference between respondents' views regarding fake and true statements. A change in fragmentation from the lowest level (Kostroma) in the sample to the highest (Astrakhan) is associated with a 0.5-point increase in the difference between attitudes. This corresponds to a nearly 14% (0.5/3.64) increase relative to the SD of the variable. At the same time, VK use decreases polarization, although the coefficient is smaller.

XSL•FO
RenderX

Finally, column 6 shows that fragmentation is positively associated with misinformation error—that is, agreeing with untrue statements and disagreeing with true statements. This indicates that, while the overall difference in attitude toward fake statements is indistinguishable from 0 (column 2), fragmentation leads to errors in respondents' ability to distinguish between true and fake statements. VK use is also positively associated with misinformation errors, suggesting that people on online social media are less likely to distinguish fake news from real.

Summing up the results, the findings show that fragmentation is associated with the spread of misinformation and leads to disbelief in true statements, greater polarization of opinions, and increased misinformation errors.

Table 6. Relationship between the fragmentation of networks and attitudes toward statements.[a,b,c]

Statement relationship	Dependent variables					
	Average attitude to true statements about COVID-19	Average attitude to fake statements about COVID-19	Share of true statements agree	Share of fake statements agree	Difference in attitude between fake and true statements	Misinformation error
Fragmentation index	–0.848 (0.194); <.001	0.150 (0.274); .55	–0.102 (0.024); <.001	0.025 (0.016); .12	0.998 (0.355); .005	0.052 (0.022); .02
Size of the network	0.030 (0.027); .27	0.041 (0.038); .28	–0.0001 (0.003); .97	–0.003 (0.002); .13	0.011 (0.050); .82	–0.005 (0.003); .01
Vkontakte use dummy	0.055 (0.034); .10	–0.161 (0.046); <.001	0.022 (0.004); .001	0.006 (0.003); .04	–0.216 (0.060); <.001	0.014 (0.004); <.001
Socioeconomic town-level controls[d]	+	+	+	+	+	+
Individual-level controls[e]	+	+	+	+	+	+
Observations, n	16,587	16,587	16,587	16,587	16,587	16,587
R^2	0.077	0.058	0.059	0.025	0.069	0.027

[a]Data for the first 3 rows are presented as estimated β coefficients of linear regression (SE); *P* value.

[b]Exact *P* values are reported.

[c]Robust SEs by town are given in brackets.

[d]Socioeconomic town-level controls include the natural log of wages 2019 and the natural log of population 2019.

[e]Individual-level controls include gender dummy, education level, household income, age, fear of COVID-19, a dummy variable for household experience with COVID-19, institutional trust measured as a trust to the president, and use of television.

Discussion

Principal Findings

This study adds to the body of literature on the COVID-19 infodemic [40,64,65] and health misinformation more broadly [23]. The study's findings show how the characteristics of the network in which an individual is embedded can lead to both polarization and the spread of misinformation. Moreover, network formation itself may be influenced by cultural and social traits of society, such as social capital [62]. This suggests that, more generally, some societies are structurally more prone to misinformation, and that online social media may amplify these effects by reinforcing polarization.

The study shows that social media users are more likely to be aware of both true and fake news, but the association is stronger for fake news. At the same time, social network fragmentation is positively associated with knowledge of fake news and negatively associated with knowledge of true statements. The combined impact of social media use and network fragmentation reveals a difference in the likelihood of encountering true versus fake news about COVID-19. For example, a respondent from the town with the highest fragmentation who uses VK is aware of 7.3% (0.0895-0.0167) more fake statements but 0.4% (–0.0243 to 0.0199) fewer true statements compared with a respondent who does not use social media and lives in the town with the lowest fragmentation. This result provides empirical evidence of differences in the spread of true and fake statements beyond online social media [28].

The study goes beyond medical research and contributes to the literature on the relationship between online social media and ideological polarization. Bursztyn et al [63] proposed a theoretical mechanism explaining how the distribution of preferences shifts with the penetration of social media. The study results align with the model's predictions—echo chamber effects resulting from fragmentation increase ideological distance in opinions about true and fake statements. At the same time, social media use reduces this distance. Thus, the study offers a potential mechanism through which social media penetration leads to polarization. In fragmented networks, echo chambers are more prominent, which amplifies differences in opinions. As a result, the effects of online social network usage may be more constrained when accounting for the structure of such networks.

XSL•FO
RenderX

The study findings highlight the distinction between the effects of online social media use and those of social network structure. Previously, the effects of social media use and social media penetration were often interpreted as consequences of social media bubbles [63]. These are distinguishable correlations that can move in different directions. However, understanding the interrelation between social media use (or penetration) and network structure was beyond the scope of this study and should be explored in future research.

The results presented provide an important bridge between studies focusing on the spread of misinformation within online social media and those examining the propagation of fake news in society more broadly. It was also shown that the fragmentation of online social networks influences the average opinion of individuals within the same town—a result that was overlooked in previous research on this topic [36,65].

However, the findings do not control for active engagement with COVID-19 news. While the study analysis controls for household experience with COVID-19, institutional trust, and fear of COVID-19—which may partially account for active news consumption and abstention—there is evidence of personal network effects on news abstention related to COVID-19 [66]. Part of the results may thus reflect network effects on abstention; that is, in certain networks, individuals may be less willing to discuss specific topics. The study hypothesized that such situations are more likely in smaller networks, where each tie holds relatively greater value. However, the study findings show that network size, if anything, is negatively associated with the popularity of fake statements. This suggests that people in smaller towns are not necessarily more likely to abstain from discussing COVID-19.

Limitations

This study had several limitations, which the author attempted to address and clarify in terms of their potential impact on the results. First, the network data were collected in 2023, whereas the survey was conducted in 2020. This time gap introduces a measurement error bias, which likely reduces the statistical significance of the results [67]. Additionally, the method used for sampling networks has limitations, as discussed in the "About VK" section. The study assumes that connections between accounts persist over time and that there has been no substantial drift in the fragmentation of towns—an assumption necessary for interpreting the observed correlations. Moreover, VK social networks are not representative of the entire population and do not capture all social interactions among Russian citizens. To estimate the potential bias introduced by using VK networks as a proxy for actual social interaction networks, demographic data from VK accounts and survey responses were compared. This information is provided in Multimedia Appendix 7. A similar age and gender structure was observed between the 2 groups, but there was a substantial difference in the proportion of individuals with higher education. The survey data likely overrepresent individuals with higher education compared with the general population. Additionally, users with higher education may not consistently disclose their educational background on VK, as this information is not required to use the platform. Importantly, this research does not claim a causal relationship between fragmentation and the spread of conspiracy theories.

To partially test the assumption of network structure stability, additional data in June 2023 and January 2024 (a 7- and 11-month gap, respectively) were collected. The results of the paired *t* test are presented in Multimedia Appendix 4. No significant differences in town-level fragmentation were found between 2023 and 2024. This suggests that network structure is either persistent over time or shaped by stable societal traits. While network scientists have increasingly focused on the emotional aspects of social ties, the formation of social ties and network structures remains understudied [68]. Social scientists often link network structure to generalized trust, a key component of social capital [62], and generalized trust is known to be stable over long periods [69]. Therefore, even if the 2023 fragmentation measure does not serve as a perfect proxy for actual fragmentation in 2020, it likely still captures elements of social capital that are persistent over time.

A robustness check was conducted using a subsample of towns with higher edge coverage. Previous research has indicated that an edge coverage of 15% yields a network sample that effectively preserves its structural characteristics [60]. This study analyzed a sample of 149 towns, excluding those with the lowest edge coverage, as they are more prone to bias. This adjustment increased the average edge coverage from 12.72% (688,859/5,416,402) to 15.04% (739,677/4,917,172). The results of these robustness checks are presented in Multimedia Appendix 6. The main outcomes discussed in this study remain robust even after excluding towns that may produce biased estimates of network variables.

Conclusions

The results presented are particularly relevant in the context of the increasing spread of fake news, misinformation, and disinformation. They suggest that it is not only the sources of information that matter but also the structure of interpersonal networks. More broadly, these findings highlight the critical role of social networks in the dissemination of information.

This study demonstrates the indirect effects of online social media structures on the spread of both information and misinformation about COVID-19, as well as on changes in public attitudes. The results underscore the relative importance of the social networks individuals belong to and how information flows through these networks.

For future research, it is important to develop a clearer understanding of the mechanisms linking social media use, personal networks, and town-level modularity to fully capture the multilevel effects of social media on the spread of misinformation and attitudes toward it. Additionally, future studies should consider using more advanced methods of network sampling.

XSL•FO
RenderX

Acknowledgments

The article was prepared within the framework of the HSE University Basic Research Program. The author acknowledges the valuable comments provided by Anton Kazun (HSE University), Andrey Tkachenko (Nazarbayev University), Valery Kalyagin (HSE University), and participants of the conferences “New Advances in the Political Economy of Development in Eurasia,” “Contemporary Trends and Prospects in the Political Economy of Development,” and “Data Analytics, Networks and Approximation.”

Conflicts of Interest

None declared.

Multimedia Appendix 1
Parameters of the survey.
[DOCX File , 20 KB - infodemiology_v5i1e58302_app1.docx]

Multimedia Appendix 2
Information about municipal statistics data for the construction of control variables.
[DOCX File , 13 KB - infodemiology_v5i1e58302_app2.docx]

Multimedia Appendix 3
Information about towns in the sample.
[DOCX File , 43 KB - infodemiology_v5i1e58302_app3.docx]

Multimedia Appendix 4
Information about fragmentation index from different periods.
[DOCX File , 15 KB - infodemiology_v5i1e58302_app4.docx]

Multimedia Appendix 5
Additional information about network sampling.
[DOCX File , 1383 KB - infodemiology_v5i1e58302_app5.docx]

Multimedia Appendix 6
Robustness checks. Using the dataset with over 15% average edge coverage.
[DOCX File , 17 KB - infodemiology_v5i1e58302_app6.docx]

Multimedia Appendix 7
Comparison of demographic information between accounts of Vkontakte (VK) users and respondents in the poll.
[DOCX File , 14 KB - infodemiology_v5i1e58302_app7.docx]

References

1. Sallam M. COVID-19 vaccine hesitancy worldwide: a concise systematic review of vaccine acceptance rates. Vaccines (Basel) 2021 Feb 16;9(2):160 [FREE Full text] [doi: 10.3390/vaccines9020160] [Medline: 33669441]
2. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. Soc Sci Med 2020 Oct;263:113356 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113356] [Medline: 32967786]
3. Cascini F, Pantovic A, Al-Ajlouni YA, Failla G, Puleo V, Melnyk A, et al. Social media and attitudes towards a COVID-19 vaccination: a systematic review of the literature. EClinicalMedicine 2022 Jun;48:101454. [doi: 10.1016/j.eclinm.2022.101454] [Medline: 35611343]
4. Theocharis Y, Cardenal A, Jin S. Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. New Media Soc. Published online October 9 :2021. [doi: 10.1177/14614448211045666]
5. Pivetti M, Di BS, Paleari FG, Hakoköngäs E. Conspiracy beliefs and attitudes toward COVID-19 vaccinations: A conceptual replication study in Finland. J Pac Rim Psychol. 2021 :15. [doi: 10.1177/18344909211039893]
6. Chipidza W, Akbaripourdibazar E, Gwanzura T, Gatto NM. Topic Analysis of Traditional and Social Media News Coverage of the Early COVID-19 Pandemic and Implications for Public Health Communication. Disaster Med Public Health Prep 2022 Oct;16(5):1881-1888 [FREE Full text] [doi: 10.1017/dmp.2021.65] [Medline: 33653437]
7. Scrima F, Miceli S, Caci B, Cardaci M. The relationship between fear of COVID-19 and intention to get vaccinated. The serial mediation roles of existential anxiety and conspiracy beliefs. Personal Individ Differ. 2022 :184. [doi: 10.1016/j.paid.2021.111188]

XSL•FO
RenderX

8. Jennings W, Stoker G, Bunting H, Valgarðsson VO, Gaskell J, Devine D, et al. Lack of Trust, Conspiracy Beliefs, and Social Media Use Predict COVID-19 Vaccine Hesitancy. Vaccines (Basel) 2021 Jun 03;9(6) [FREE Full text] [doi: 10.3390/vaccines9060593] [Medline: 34204971]
9. Bierwiaczonek K, Gundersen A, Kunst J. The role of conspiracy beliefs for COVID-19 health responses: A meta-analysis. Current Opinion in Psychology 2022 Aug;46:101346 [FREE Full text] [doi: 10.1016/j.copsyc.2022.101346] [Medline: 35486966]
10. Motta M, Hwang J, Stecula D. What Goes Down Must Come Up? Pandemic- Related Misinformation Search Behavior During an Unplanned Facebook Outage. Health Communication 2024 Sep;39(10):2041-2052. [doi: 10.1080/10410236.2023.2254583] [Medline: 37674255]
11. Zhang KC, Fang Y, Cao H, Chen H, Hu T, Chen Y, et al. Behavioral Intention to Receive a COVID-19 Vaccination Among Chinese Factory Workers: Cross-sectional Online Survey. J Med Internet Res 2021 Mar 09;23(3):e24673 [FREE Full text] [doi: 10.2196/24673] [Medline: 33646966]
12. Daley MF, Narwaney KJ, Shoup JA, Wagner NM, Glanz JM. Addressing Parents' Vaccine Concerns: A Randomized Trial of a Social Media Intervention. Am J Prev Med 2018 Jul;55(1):44-54 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.010] [Medline: 29773490]
13. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. Elife 2021 Jun 30;10:10 [FREE Full text] [doi: 10.7554/eLife.69336] [Medline: 34190045]
14. Wang H, Paulson KR, Pease SA. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020?21. The Lancet. 2022 ;399(10334):a. [doi: 10.1016/S0140-6736(21)02796-3]
15. Pariser E. The Filter Bubble: What the Internet Is Hiding from You. London: Penguin Press; 2011.
16. Del VM, Bessi A, Zollo F. The spreading of misinformation online. Presented at: Proc Natl Acad Sci. ;113(3); 2016 p. 554-559. [doi: 10.1073/pnas.1517441113]
17. Flaxman S, Goel S, Rao JM. Filter Bubbles, Echo Chambers, and Online News Consumption. PUBOPQ 2016 Mar 22;80(S1):298-320. [doi: 10.1093/poq/nfw006]
18. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G Conspiracy Theory: Social Network Analysis of Twitter Data. J Med Internet Res 2020 May 6;22(5):e19458. [doi: 10.2196/19458]
19. Yang KC, Pierri F, Hui PM. The COVID-19 Infodemic: Twitter versus Facebook. Big Data Soc. 2021 ;8(1):205395172110138. [doi: 10.1177/20539517211013861]
20. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
21. Deffuant G, Neau D, Amblard F, Weisbuch G. Mixing beliefs among interacting agents. Adv Complex Syst. (01n04) 2000;03:87-98. [doi: 10.1142/S0219525900000078]
22. Vziatysheva V. How fake news spreads online? Int J Media Inf Lit. 2020 ;5(2).
23. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of Health Misinformation on Social Media: Systematic Review. J Med Internet Res 2021 Jan 20;23(1):e17187. [doi: 10.2196/17187] [Medline: 33470931]
24. Li M, Wang X, Gao K, Zhang S. A Survey on Information Diffusion in Online Social Networks: Models and Methods. Information 2017;8(4):118. [doi: 10.3390/info8040118]
25. Granovetter MS. The Strength of Weak Ties. American Journal of Sociology 1973 May;78(6):1360-1380. [doi: 10.1086/225469]
26. Aral S. The Future of Weak Ties. American Journal of Sociology 2016 May;121(6):1931-1939. [doi: 10.1086/686293]
27. De CM, Sundaram H, John A, Seligmann D, Kelliher A. "Birds of a Feather": Does User Homophily Impact Information Diffusion in Social Media? ArXiv. 2010 Jun. URL: http://arxiv.org/abs/1006.1702 [accessed 2023-09-21]
28. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 09;359(6380):1146-1151. [doi: 10.1126/science.aap9559]
29. Interian RG. Marzo R, Mendoza I, Ribeiro CC. Network polarization, filter bubbles, and echo chambers: an annotated review of measures and reduction methods. Int Trans Oper Res. 2023 ;30(6):a. [doi: 10.1111/itor.13224]
30. Peng H, Nematzadeh A, Romero DM, Ferrara E. Network modularity controls the speed of information diffusion. Phys Rev E 2020 Nov;102(5-1):052316. [doi: 10.1103/PhysRevE.102.052316] [Medline: 33327110]
31. Romano V, Shen M, Pansanel J, MacIntosh AJJ, Sueur C. Social transmission in networks: global efficiency peaks with intermediate levels of modularity. Behav Ecol Sociobiol 2018;72(9):154. [doi: 10.1007/s00265-018-2564-9]
32. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. Phys Rev Lett 2001 Apr 02;86(14):3200-3203. [doi: 10.1103/PhysRevLett.86.3200] [Medline: 11290142]
33. Pérez-Ortiz M, Manescu P, Caccioli F, Fernández-Reyes D, Nachev P, Shawe-Taylor J. Network topological determinants of pathogen spread. Sci Rep 2022 May 11;12(1):7692 [FREE Full text] [doi: 10.1038/s41598-022-11786-5] [Medline: 35545647]
34. Block P, Hoffman M, Raabe IJ, Dowd JB, Rahal C, Kashyap R, et al. Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. Nat Hum Behav 2020 Jun;4(6):588-596. [doi: 10.1038/s41562-020-0898-6] [Medline: 32499576]

XSL•FO
RenderX

35. Pires MA, Oestereich AL, Crokidakis N, Duarte Queirós SM. Antivax movement and epidemic spreading in the era of social networks: Nonmonotonic effects, bistability, and network segregation. Phys Rev E 2021 Sep;104(3-1):034302. [doi: 10.1103/PhysRevE.104.034302] [Medline: 34654182]
36. Altoe F, Moreira C, Pinto H, Jorge J. Online Fake News Opinion Spread and Belief Change: A Systematic Review. Human Behavior and Emerging Technologies 2024 Apr 30;2024:1-20. [doi: 10.1155/2024/1069670]
37. Donkers T, Ziegler J. The dual echo chamber: Modeling social media polarization for interventional recommending. In; 2021 Presented at: Proceedings of the 15th ACM conference on recommender systems; 13 September 2021; Amsterdam p. 12-22. [doi: 10.1145/3460231.3474261]
38. Norbutas L, Corten R. Network structure and economic prosperity in municipalities: A large-scale test of social capital theory using social media data. Soc Netw 2018;52:A.
39. Wachs J, Yasseri T, Lengyel B, Kertész J. Social capital predicts corruption risk in towns. R Soc Open Sci 2019;6(4):182103. [Medline: 31183137]
40. Gisondi MA, Barber R, Faust JS, Raja A, Strehlow MC, Westafer LM, et al. A Deadly Infodemic: Social Media and the Power of COVID-19 Misinformation. J Med Internet Res 2022 Feb 1;24(2):e35552. [doi: 10.2196/35552]
41. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech Theory Exp 2008;2008(10).
42. Sah P, Leu ST, Cross PC, Hudson PJ, Bansal S. Unraveling the disease consequences and mechanisms of modular structure in animal social networks. Presented at: Proc Natl Acad Sci. ;114(16); 2017 p. 4165-4170. [doi: 10.1073/pnas.1613616114]
43. Xiong X, Ma J, Wang M, Zhou G, Xu K. Information diffusion model in modular microblogging networks. World Wide Web 2015;18:1051-1069.
44. Bessi A, Coletto M, Davidescu GA, Scala A, Caldarelli G, Quattrociocchi W. Science vs Conspiracy: Collective Narratives in the Age of Misinformation. PLoS ONE 2015 Feb 23;10(2):e0118093. [doi: 10.1371/journal.pone.0118093] [Medline: 25706981]
45. Baumeister R, Leary M. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. In: Interpers Development. London: Ashgate Publishing; 2007.
46. Bareket-Bojmel L, Moran S, Shahar G. Strategic self-presentation on Facebook: Personal motives and audience response to online behavior. Comput Hum Behav 2016;55:788-795. [doi: 10.1016/j.chb.2015.10.033]
47. Biondi E, Boldrini C, Passarella A, Conti M. Dynamics of opinion polarization. IEEE Trans Syst Man Cybern Syst. 2023 ;53(9):5381-5392.
48. Tu S, Neumann S. A Viral Marketing-Based Model For Opinion Dynamics in Online Social Networks. : Proceedings of the ACM Web Conference 2022. ACM; 2022 Presented at: WWW '22: The ACM Web Conference 2022; April 25 - 29, 2022; Lyon p. 1570-1578. [doi: 10.1145/3485447.3512203]
49. Bail CA, Argyle LP, Brown TW, Bumpus JP, Chen H, Hunzaker MBF, et al. Exposure to opposing views on social media can increase political polarization. Proceedings of the National Academy of Sciences 2018 Sep 11;115(37):9216-9221 [FREE Full text] [doi: 10.1073/pnas.1804840115] [Medline: 30154168]
50. Borisova E, Frye T, Schoors KJ, Zabolotskiy V. Fear, Trust and Demand for Regulationvidence from the COVID-19 Pandemic in Russia. Published online :2022.
51. Published online 2024. VK developer API instruments overview. URL: https://dev.vk.com/en/guide [accessed 2024-03-11]
52. Vedenkov M, Komin M, Tsygankov M. Database of municipal indicators: combined and processed data for 2006 - 2020 // Rosstat. Published online January 27, 2022. URL: http://data.rcsi.science/data-catalog/datasets/115/ [accessed 2024-02-25]
53. Gosteva I, Kazancev K. Large cities of Russia: combined data on main socio-economic indicators for 1985-2019 // Rosstat. Published online 2021. URL: http://data.rcsi.science/data-catalog/datasets/187/ [accessed 2024-02-25]
54. Cross Web. Mediascope Website Audience Measurement. URL: https://mediascope.net/data/ [accessed 2024-09-04]
55. Hairol ASH, Abas ZA, Yunos NM. Comparison between Louvain and Leiden Algorithm for Network Structure: A Review. J Phys Conf Ser. 2021 Presented at: p. A. [doi: 10.1088/1742-6596/2129/1/012028]
56. Elmas T. The Impact of Data Persistence Bias on Social Media Studies. In: The Impact of Data Persistence Bias on Social Media Studies. In: Proceedings of the 15th ACM Web Science Conference 2023. ACM; 2023 Presented at: WebSci '23: 15th ACM Web Science Conference 2023; 30 April 2023- 1 May 2023; Austin p. 196-207. [doi: 10.1145/3578503.3583630]
57. Arnaboldi V, Conti M, Passarella A, Dunbar RI. Online Social Networks and information diffusion: The role of ego networks. Online Social Networks and Media 2017 Jun;1:44-55. [doi: 10.1016/j.osnem.2017.04.001] [Medline: 40541500]
58. Kurant M, Markopoulou A, Thiran P. Towards Unbiased BFS Sampling. IEEE J. Select. Areas Commun 2011 Oct;29(9):1799-1809. [doi: 10.1109/JSAC.2011.111005] [Medline: 40358844]
59. Peng S, Roth AR, Perry BL. Random sampling of alters from networks: A promising direction in egocentric network research. Soc Netw. 2023 :72. [doi: 10.1016/j.socnet.2022.09.004]
60. Jure L, Leskovec J, Christos F, Faloutsos C. Sampling from large graphs. In: Knowl Discov Data Min. Published online August. 2006 Presented at: KDD06: The 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 20 - 23, 2006; Philadelphia p. 20-636. [doi: 10.1145/1150402.1150479]
61. Jiao B. Sampling unknown large networks restricted by low sampling rates. Sci Rep 2024 Jun 10;14(1):13340 [FREE Full text] [doi: 10.1038/s41598-024-64018-3] [Medline: 38858487]

XSL•FO
RenderX

62. Putnam R. Bowling Alone: The Collapse and Revival of American Community. New York: Simon and Schuster; 2000.
63. Bursztyn L, Egorov G, Enikolopov R, Petrova M. Social Media and Xenophobia: Evidence from Russia. National Bureau of Economic Research. 2019. URL: https://www.nber.org/papers/w26567 [accessed 2024-02-03]
64. Pian W, Chi J, Ma F. The causes, impacts and countermeasures of COVID-19 "Infodemic": A systematic review using narrative synthesis. Inf Process Manag 2021 Nov;58(6):102713 [FREE Full text] [doi: 10.1016/j.ipm.2021.102713] [Medline: 34720340]
65. Balakrishnan V, Ng WZ, Soo MC, Han GJ, Lee CJ. Infodemic and fake news ? A comprehensive overview of its global magnitude during the COVID-19 pandemic in 2021: A scoping review. Int J Disaster Risk Reduct. 2022 :78. [doi: 10.1016/j.ijdrr.2022.103144]
66. Qu Y, Saffer AJ, Austin L. What Drives People Away from COVID-19 Information?: Uncovering the Influences of Personal Networks on Information Avoidance. Health Commun 2023 Feb;38(2):216-227. [doi: 10.1080/10410236.2021.1944457] [Medline: 34187260]
67. Pischke S. Lecture notes on measurement error. Lond Sch Econ Lond. Published online. 2007. URL: https://econ.lse.ac.uk/staff/spischke/ec524/Merr_new.pdf [accessed 2024-12-31]
68. Brass DJ. New Developments in Social Network Analysis. Annu Rev Organ Psychol Organ Behav. 2022 ;9(1):225-246. [doi: 10.1146/annurev-orgpsych-012420-090628]
69. Wellander B, Sanandaji T. Tracing the Historic Roots of Generalized Trust. Essays in Economic and Business History 2020:38-141.

Abbreviations

MTurk: Mechanical Turk
RoCIRR: Research on COVID-19 in Russia’s Regions
VK: Vkontakte

Edited by R Sivakumaran, M Haupt; submitted 12.03.24; peer-reviewed by D Wang, KC Yang; comments to author 19.09.24; revised version received 31.12.24; accepted 02.06.25; published 26.06.25.

Please cite as:
Pavlenko B
Modularity of Online Social Networks and COVID-19 Misinformation Spreading in Russia: Combining Social Network Analysis and National Representative Survey
JMIR Infodemiology 2025;5:e58302
URL: https://infodemiology.jmir.org/2025/1/e58302
doi:10.2196/58302
PMID:

©Boris Pavlenko. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 26.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Breast Cancer Vlogs on YouTube: Descriptive and Content Analyses

Nina Morena[1], MA; Elly Dimya Htite[2], HBSc, MSc; Yitzchok Ahisar[3], BSc, MDCM; Victoria Hayman[4], BSc, MDCM; Carrie A Rentschler[1], PhD; Ari N Meguerditchian[4,5,6], MS, MD

[1]Department of Art History and Communication Studies, McGill University, McCall McBain Arts Building, Room 155, 853 Sherbrooke Street West, Montreal, QC, Canada

[2]Faculty of Medicine, McGill University, Montreal, QC, Canada

[3]Department of Surgery, General Surgery, University of British Columbia, Vancouver, BC, Canada

[4]Department of Surgery, Faculty of Medicine, McGill University, Montreal, QC, Canada

[5]St Mary's Research Centre, Montreal, QC, Canada

[6]Gerald Bronfman Department of Oncology, Faculty of Medicine, McGill University, Montreal, QC, Canada

Corresponding Author:
Nina Morena, MA
Department of Art History and Communication Studies, McGill University, McCall McBain Arts Building, Room 155, 853 Sherbrooke Street West, Montreal, QC, Canada

Abstract

Background: Many women with breast cancer document their experiences in YouTube vlogs, which may serve as peer-to-peer and community support.

Objective: This study aimed to determine (1) the forms of content about breast cancer that tend to be discussed in vlogs, (2) the reasons why women choose to vlog their breast cancer experiences, and (3) the potential for breast cancer vlogs to serve as an alternative or complement to peer-to-peer support as well as a site of digital community overall.

Methods: YouTube was searched in incognito mode in November 2023 using the search terms "breast cancer vlog." A maximum of 10 videos/creator were included based on viewership and date created. Video characteristics collected included title; length; number of views, likes, comments; and playlist inclusion. Videos were assessed for sponsorship; presence of explanation and discussion on breast cancer; type of content; and themes. Creator characteristics included age, location, and engagement approaches. Descriptive and content analyses were performed to analyze video content and potential areas where peer-to-peer support may be provided.

Results: Ninety vlogs by 13 creators were included, all from personal accounts. The mean (SD) video length, number of views, and number of comments were 21.4 (9.1) minutes, 266,780 (534,465), and 1485 (3422), respectively. Of the 90 videos, 35 (39%) included hashtags, and 11 (12%) included paid sponsorships. The most common filming location was the home (87/90; 97%), followed by the hospital (28/90; 31%) and car (19/90; 21%). Home vlogs were most often set in the living room (43/90; 44%), bedroom (32/90; 33%), or kitchen (20/90; 21%). Thirty-four of 60 videos (57%) included treatment visuals and physical findings. Creators addressed motivation for vlogging in 44/90 videos (49%); the two most common reasons were wanting to build a community and helping others. In 42/90 videos (47%), creators explicitly expressed emotion. Most common themes were treatment (77/90; 86%), mental health (73/90; 81%), adverse effects (65/90; 72%), appearance (57/90; 63%), and family relationships (33/90; 37%). Patient-directed advice was offered in 52/90 videos (58%), mostly on treatment-related issues. In 51/90 videos (57%), creators provided explicit treatment definitions. Chemotherapy was discussed in 63/90 videos (70%); surgery in 52/90 (58%), primarily mastectomy; radiation in 27/90 (30%); and general adverse effects in 64/90 (71%). Twenty-two of 90 videos (24%) were about a new diagnosis. When mentioned (40/90; 44%), the most common creator location was the United States. When mentioned (27/90; 30%), the most common age was 20 - 29 years.

Conclusions: The dedication to building community support by vlog creators, and the personal nature of their storytelling, may make vlogs a potential resource for peer-to-peer support.

(JMIR Infodemiology 2025;5:e66812) doi:10.2196/66812

KEYWORDS

breast cancer vlog; YouTube; social media; experience; video; content analysis; breast; cancer; women; oncology; descriptive analysis

XSL•FO
RenderX

Introduction

Vlogs, or "video blogs," are personally and individually created experiential videos based on wide-ranging topics, usually posted to YouTube. A popular subset of vlogs are created by women with breast cancer, in which they document their breast cancer experiences in a publicly accessible digital format. As videos that tend to invite viewers into the lives of their creators, vlogs typically receive high levels of engagement and draw audiences who will continue watching in order to keep up with what creators are doing next. Online tools such as these vlogs are important parts of patients' experience with processing and managing chronic illness [1]. They can help patients find support, community, and information in spaces they may not have access to in real life [2]. Overall, social support is a prominent theme in literature about breast cancer communication on social media [2].

YouTube is growing as a source of "peer-to-peer health information sharing and support" [3]. The importance of peer support across cancer is widely acknowledged [4]. Breast cancer peer support programs have been shown to be effective in enhancing patients' quality of life [5], particularly regarding the alleviation of depression and anxiety [6]. Moreover, peer programs are similarly shown to be successful in providing emotional and psychosocial support [7]. Patient support groups can fulfill patient needs and improve quality of life [6,8,9]. Peer education was also found to reduce levels of psychological pain [10]. Research demonstrates that women with breast cancer need and expect online support programs, despite platform-related challenges such as inconsistent content moderation and that "well-organized and tailored" peer support is important to enhance their quality of life [11].

While vlogs about breast cancer may have the potential to serve as peer-to-peer support and provide community, their content, quality, and role in the breast cancer experience are understudied. The objectives of this study were to (1) determine what forms of content about breast cancer tend to be discussed in vlogs, (2) inquire into the reasons why women choose to vlog their breast cancer experiences, and (3) consider the potential for breast cancer vlogs to serve as an alternative or complement to peer-to-peer support, as well as a site of digital community overall.

Methods

Study Design

YouTube was searched in incognito mode in November 2023, using the search terms "breast cancer vlog." Breast cancer vlog creators were identified, and their video characteristics were collected. If the creator produced less than 10 videos, all were included. If they produced more than 10, a maximum of 10 were included, based on viewership and most recent date created. Creators were identified based on whether they produced English-language breast cancer-related videos in the last 5 years. Prior to evaluating the videos in this dataset, reviewers were trained to collect vlog data using a standardized data collection tool. A sample was collected, and data collection quality, as well as procedures for any disagreement in evaluations, were assessed before initiating the formal data collection.

Reviewers collected creator information (ie, age, location, profession, and cancer stage). They collected the video titles; length; date; number of views, likes, and comments; and hashtags; and noted whether the videos were part of a playlist. Reviewers then assessed videos based on consumer details (ie, sponsorships, product recommendations, and endorsements), and the creators' explanations and discussions of their breast cancer. Engagement levels were considered based on vlog creators' discussion with their audience, whether they commented on their experience with their audience, or asked the audience to share insights of their own. Types of content (diagnosis, surgery, or other event) were also collected.

Video themes were extracted both deductively and inductively. The themes collected deductively are described in Table 1. Descriptive and content analyses were performed to assess and analyze video content and potential areas where peer-to-peer support may be provided. Reviewers were encouraged to add subthemes inductively in cases where the content exceeded or was more specific than the deductive themes [12]. An additional set of questions was applied to videos that contained advice, in order to characterize the nature, source, and potential validity of the advice provided (Table 2).

XSL•FO
RenderX

Table . Deductive themes.

Theme	Definition
Appearance	Refers to one’s experiences with their appearance, including how it may have changed over the course of treatment
Mental health	Refers to questions of mental health, including stress, anxiety, depression, or other related issues
Fear of recurrence	Refers to fears or anxieties related to the future possibility of a cancer recurrence
Gender identity	Refers to one’s gender identity, particularly in the context of breast surgery or reconstruction choices
Sexuality	Refers to experiences of sexual intimacy, including how these may have changed over the course of treatment
Fertility	Refers to one’s experiences with fertility or infertility, including fertility treatment
Motherhood	Refers to the specific relationship between mother and child, how to disclose to one’s children, as well as wanting to be a mother
Spousal relationship	Refers to the patient’s relationship with their spouse, including stress on the spouse who takes on a caregiving role
Family relationship	Refers to family-based experiences, such as how one may disclose their diagnosis to their family and how a diagnosis shifts the family dynamic
Path to diagnosis	Refers to the story or experience of being diagnosed with breast cancer, such as discovering a breast lump
Treatment	Refers to a wide array of topics, ranging from treatment choices to the experience of receiving treatment
Adverse effects	Refers to specific experiences with adverse effects, whether due to surgery, chemotherapy, or other forms of treatment

Table . Advice assessment.

Characteristic	Definition
Vlog creator verbalizes a reference or source for advice.	Refers to the vlog creator verbally explaining where or from whom they learned about the advice provided.
Vlog creator confirms having tried the advice personally.	Refers to the vlog creator describing their own personal experience by following the advice provided.
Proposed advice recommends adding something (addition) or not doing something (omission).	Refers to whether the vlog creator suggests making some sort of addition to their care or cancer management or ceasing or removing an aspect of their care or cancer management.
Proposed advice involves a product to be applied or consumed.	Refers to whether the proposed advice involves applying or consuming a product.
Proposed advice suggests a modification or reduction in the treatment plan.	Refers to whether the proposed advice suggests some sort of modification to or reduction in one’s cancer care management or treatment plan.
Proposed advice is potentially beneficial, neutral, or potentially harmful.	Refers to whether the advice provided is of potential benefit to a patient, neutral, or of potential harm to a patient.

Ethical Considerations

Institutional ethics review was not required for the completion of the study, as all the data including patient- and disease-specific information and opinions or experiences were volunteered into the public domain by the creators. No patient- or disease-specific information was collected. No identifiable information is published.

Results

A total of 90 vlogs by 13 vlog creators were included in the study, all of which originated from personal YouTube accounts. The mean (SD) video length was 21.4 (9.1) minutes. The mean (SD) number of views was 266,780 (534,465). The mean (SD) number of comments was 1485 (3422). Hashtags—words or phrases preceded by the “#” symbol that serve to categorize social media content—were included in 35 videos (39%), the majority of which were breast cancer-related. Paid sponsorships were present in 11 videos (12%). Creators promoted their own

XSL•FO
RenderX

channel in a large majority of videos (80/90; 89%); for instance, by encouraging channel subscriptions. Most creators (88/90; 98%) included a title that effectively summarized the video topic, meaning that the title described the events discussed in the video. Where mentioned, the range of 20 - 29 years was the most common age group (14/27; 52%), followed by 30 - 39 years (10/27; 37%). Where mentioned, the most common creator location was the United States of America (25/40; 62%).

Visuals were present in 60/90 videos (67%); of these 60 videos, 34 (57%) included images or videos of vlog creators undergoing treatment (such as receiving chemotherapy or radiation, or undergoing magnetic resonance imaging) as well as physical features of treatment, including port scars, surgical drains, and breast contour after expander placement. A portion of videos (36/90; 40%) included inserted recorded clips, for instance, playing a recording of a phone call with their pathologist discussing results or footage of entering the radiation machine.

Videos were mainly filmed at home (87/90; 97%), at the hospital (28/90; 31%), or in the car (19/90; 21%). It is possible that vlogs were filmed in more than one setting. When filmed at home, vlogs were most often set in the creator's living room (43/90; 44%), bedroom (32/90; 33%), or kitchen or office (both 20/90; 21%).

In half of the vlogs (45/90, 50%), the creator commented on how their audience makes them feel, and in 44/90 (49%), the creator explained why they decided to make vlogs about their breast cancer experience, the most frequent reasons being: (1) enjoying filming vlogs, (2) wanting to build a community, (3) having a predominantly female viewership, (4) wanting others with cancer to feel less alone, (5) sharing information on surgery, and (6) providing details about signs of their recurrence. For example, one vlog creator described the process of filming and posting vlogs about her metastatic breast cancer as therapeutic "because it feels like she sat and talked to someone about everything on her mind"; this creator also referred to the viewers who have expressed that her vlogs helped them through their own experiences as making her feel like her vlogs have a purpose. It is important to her that her audience—specifically, others facing a similar diagnosis—knows that they are "not alone in this at all." In 42/90 videos (47%), creators expressed emotion in an explicit way; for instance, one vlog creator filmed her last chemotherapy treatment and was emotional while ringing the celebratory bell and thanking her nurses.

Advice was offered in 52/90 (58%) videos, with the most common topics being cold capping, hair regrowth, clothing, nutrition, mental health habits, chemotherapy preparation, saline soaking for radiation burns, wig use, and cancer prevention. References or sources were rarely cited, with advice usually originating from the video creator themselves. In most cases, the creator confirmed the advice was based on their personal experience. Advice overwhelmingly involved making some sort of addition to one's care or cancer management rather than an omission. None of the proposed advice was considered of potential harm to patients. Additional details on videos containing advice are outlined in Table 3.

Table . Advice provided in vlogs.

Characteristic	Frequency (%)	Example
Vlog creator verbalizes a reference or source for advice.	• Yes: 3 (5%) • No: 49 (95%)	Friend, website, or book
Vlog creator confirms having tried the advice personally.	• Yes: 42 (81%) • No: 10 (19%)	N/A[a]
Proposed advice recommends adding something (addition) or not doing something (omission).	• Yes: 51 (98%) • No: 1 (2%)	Cold capping, purchasing products, stretching, attending therapy sessions, using sleep aids, maintaining a positive outlook, or praying
Proposed advice involves a product to be applied or consumed.	• Yes: 15 (20%) • No: 37 (80%)	Vaseline, compression socks, ginger supplement, cold capping equipment, hair or beauty products, blankets, or pajamas
Proposed advice suggests a modification or reduction in the treatment plan.	• Yes: 14 (27%) • No: 38 (73%)	Modifying chemotherapy administration method by using a port-a-cath instead of a peripheral venous infusion
Proposed advice is potentially beneficial, neutral, or potentially harmful.	• Harmful: 0% • Neutral: 24 (46%) • Potentially beneficial: 28 (54%)	Reduced risk of deep vein thrombosis, reduced hair loss, and reduced diarrhea

[a]N/A: not applicable.

Where the cancer stage was mentioned, stage IV was most common (13/90; 14%). Of the 90 videos, 22 (24%) were about a new diagnosis. Chemotherapy was the predominant treatment form discussed in the majority of videos (63/90, 70%); surgery in 58% (52/90), primarily mastectomy (20/52, 38%); and radiation in 30% (27/90). The general adverse effects were discussed in 71% of the videos (64/90). In over half of the videos (50/90; 57%), creators provided a structured definition to some aspect of their treatment.

The most common themes were treatment (77/90; 86%), mental health (73/90; 81%), adverse effects (65/90; 72%), appearance (57/90; 63%), and family relationships (33/90; 37%). Subthemes included young age, finances, the importance of online

XSL•FO
RenderX

community support, social life, fear of surgery, egg retrieval, and confidence and redefining beauty standards.

Discussion

Principal Findings and Comparison With Previous Works

This study demonstrates that vlogs by women with breast cancer receive significant levels of engagement and represent an important site of online community for women. However, the nature of their content, their authorship, and the potential to be integrated into care plans are underexplored. Our study shows that most commonly discussed themes in breast cancer vlogs include treatment, mental health, adverse effects, and appearance, but that a wide range of subthemes are also present. Moreover, they are often filmed in a personal home setting. These patient-created videos, which included less paid sponsorship than what researchers have identified in previous analyses [12], included detailed and overt expressions about why creating vlogs about their breast cancer is a valuable experience. The production behind vlogs, combined with the settings in which they are filmed, contribute to the feeling of connection between the creator and the audience. That vlogs are often filmed at home may be important to building a sense of community. The seemingly close and casual environment of filming in one's own living room or bedroom creates a sense of proximity to the vlog creator that is more aligned with the support group model that might not necessarily be present in a professional setting.

There is a rich community-centered aspect to breast cancer vlogs, which may position them as complementary forms of peer-to-peer social support and as unique methods for effective coping. Tailoring peer support to the moments when patients are most in need is crucial [13]. In their study on factors of engagement and patient-reported outcomes in a stage IV breast cancer Facebook group, Kashian and Jacobson [14] conclude, "Optimal social support plays a critical role facilitating engagement in online breast cancer support groups. It is not enough that members exchange social support in online support groups, rather members must exchange the type of support that facilitates effective coping." Power and Hegarty [15] also found that support programs need to be tailored to the needs of women with breast cancer and have identified "the need to allow more informal sharing to occur in facilitated peer support programs." Vlogs are easily findable and watchable. While different peer support models and approaches will have varying outcomes on different patients, web-based support without training and/or moderation should be used with caution [16]. YouTube is heavily engaged with for public health reasons. However, there is a need for higher-quality content [17]. Research evaluating YouTube videos about radiotherapy in breast cancer concludes that while videos were inconsistent in following best practice guidelines, YouTube still has potential toward disseminating health information [18]. Research on YouTube videos about radiotherapy in lung cancer draw similar conclusions [19].

It is possible that informal sharing also happens in online spaces, where patients might feel they have more control over how much they can share and can do so in the format, style, narrative, and medium of their choosing. Ziegler and colleagues [20] have shown a moderate positive association between peer support and psychological improvement in cancer patients. People who post and share about their experiences in online communities are not the only ones who benefit from the peer support in these spaces; research shows that "lurkers," or those who watch or consume content without sharing their own, can also benefit from the advice and insights shared [21]. Many messages in online support groups are requests for information and opinion from those in the same situation and are designed to reach like-minded people [22]. These networks are informal, grow organically, and are accessed when needed.

The use of visuals and additional recorded clips in vlogs, such as the inserted phone recordings or hospital footage, may contribute to a sense of closeness between the creator and their audience. Such instances of additional video editing are indications of vlog creators' platform expertise and demonstrate the production efforts behind vlog creation [23]. However, in the context of documenting one's breast cancer experiences, additional video editing may actually serve to provide even more detail for one's audience: playing a phone recording or showing hospital footage (of, for instance, receiving chemotherapy infusions) allows the audience to see and hear parts of the creator's breast cancer experience that they were not present for, thereby creating a further sense of closeness with the creator. Inviting one's audience into private hospital or clinical encounters also holds implications for standardized practices of medical confidentiality. For these creators, there is something important about sharing what tends to be considered private. The perceived closeness afforded by hospital footage and recorded phone calls, for example, may also serve to demystify these experiences for audience members who might be about to begin their own breast cancer treatment. In this context, typical conceptions of what is considered public or private are blurred, and viewers may begin to develop emotional bonds with vloggers they have never met: a phenomenon not unique to vlogs about breast cancer, but which occurs across different types of social media content and platforms [24]. Scholars have discussed the notion of the microcelebrity, particularly the ways in which celebrity and fame are connected to different media [25]. Given this context, it is plausible that breast cancer vloggers may achieve certain levels of microcelebrity status on YouTube, which may result in a feeling of parasocial closeness—defined as "nonreciprocal socio-emotional connections with media figures such as celebrities or influencers"—on the part of the audience, and which can continue in cases where the video creator dies [26].

Key features of breast cancer vlogs, which include explanations of the diagnosis experience, are associated with receiving empathic support from audiences [27]. The emotionally intense context afforded by the medium of video is demonstrated to lead to community-building and social support among vloggers and their audiences [28]. In their study of breast cancer narratives on social media, Ma and colleagues [29] found, "Stories that were longer, less emotionally intensive, told from the cancer survivor's perspective, with gender identity–related information, describing the act of providing social support, explicitly requesting engagement and/or donation, and using

XSL•FO
RenderX

more vivid forms of visuals such as linked images tended to be more engaging." However, while Ma and colleagues [29] found that emotionally intensive stories may be less engaging, social media research suggests, "Crying and anxiety blogs can function as a means of demonstrating vloggers' 'authenticity,' and thus fostering valuable intimacy between vloggers and their audience" [30]. The sense of intimacy shared by vloggers—especially when demonstrating explicit emotion in their vlogs—is valuable in this context because it serves to develop a bond between the creator and the audience, who may decide to begin regularly following the creator's breast cancer trajectory. In our study, creators explicitly expressed emotion in almost half of the videos. Emotion is an important part of the breast cancer stories that vlog creators tell, and it resonates with their audiences: several videos in this dataset, particularly those relating to metastatic disease and end of life, were challenging for reviewers to watch and remained distinctly memorable after data collection was completed. The vulnerability and emotion present in these vlogs—and thus, their relatability—may provide a form of connection between creator and audience that health care professionals cannot provide due to the objectivity they have to maintain in delivering care.

Vlogging about breast cancer may also hold therapeutic value [31]. Creators in this dataset have discussed what they perceive to be the benefits of vlogging, among which is the idea that vlogging is a way to speak about everything on their minds, not unlike sitting and chatting with someone. When creators hear from their audience that watching their vlogs helped them to feel less alone, or helped them in their own journeys, they express that this audience response helps them to feel like their vlogs have a purpose. Vlogs are a source of support and can help patients cope with isolation not only in dealing with cancer, but also with other chronic illnesses such as fibromyalgia, diabetes, and HIV [32-34]. The value of vlogging about breast cancer can extend beyond patient communities to health care settings; one reviewer, a surgical resident, commented that watching the vlogs in this dataset was important to their training experience. For this reviewer, seeing how patients understand the course of their treatment as explained to them by their doctor/surgeon, what they focus on, what they fear, and how they take the information they have received and share it in a digestible way with their viewers was very useful, and would contribute to how they engage with cancer patients going forward. In this way, watching vlogs may also hold value in medical training settings, as examples of voluntarily provided patient experience.

Questions of how to handle the transfer of information in vlogs were considered throughout this study's design. In some cases, reviewers were able to anecdotally or casually identify some vloggers who provided high-quality explanations of their cancer and its treatment, to the point that the reviewer would suggest it to their patients if needed. That being said, it is challenging even for health care professionals to evaluate the quality of information as it is presented intertwined with personal experiences (often indistinguishable) in experience-based videos. Validated tools for evaluating YouTube video quality now exist; however, Gabarron and colleagues [35] demonstrated, in their 2013 review, that as recently as 10 years ago, guidelines for such evaluations were "unclear and not standardized." While standardized instruments such as the DISCERN tool and the Patient Education Materials Assessment Tool can be used to evaluate different forms of media, their applicability is limited in regards to patient-created, experiential content. Vlogs are an experience-based media format grounded in personal storytelling, where it is often difficult to distinguish information from advice or lived experience. As such, the goals behind both creating vlogs and watching them may be more closely related primarily to building and seeking forms of community and connection, rather than imparting information or learning about medical facts. Nevertheless, our assessment revealed that the majority of advice offered in vlogs consisted of patient-centered concerns and personal preferences that were unlikely to affect cancer treatment trajectories. The proposed advice was not found to be of potential harm to patients (even if questionably beneficial) and was largely experiential.

Limitations

The limitations of this study include restrictions based on language and challenges related to collecting demographic data. This study was limited to English-speaking videos, which may influence views or perceptions of the breast cancer experience. In addition, because vloggers may or may not disclose specific information such as their age, nationality, location, profession, or cancer stage, the consistent collection of these data points was not possible. Future research on breast cancer vlogging practices in specific linguistic, geographic, or ethno-cultural communities represents an opportunity to understand the circulation of health information within communities that may not have easy access to mainstream health services. While some of the peer support literature cited in this paper predates the development and widespread use of social media, the authors contend that this earlier foundational literature functions as a precursor to contemporary understandings of the benefits of social media communities, their effects on patients' outlooks, as well as their reasons for participating in them.

Given that the purpose of vlogging is not to educate one's peers but to share experiences and build an online community, the potential for using currently existing standardized tools to assess information quality is limited. In light of the growth of online peer-support communities and the lack of methodology regarding the quality and accuracy of information that is woven into patient experiences reported in the vlog format, there is a need for the development of a methodology that specifically validates the quality of information transferred in these settings. The next phase of our research will address breast cancer vlogs, and other cancer-related media content, created by and from the perspectives of health care professionals where, in contrast, health care professionals have a primary goal of transferring information to support patients. Future research may evaluate the differences in engagement metrics (such as views, likes, or comments) between breast cancer video content created by patients for patients and content created by health care professionals.

Conclusion

This study demonstrated that the awareness of and dedication to building community that vlog creators show in this context,

XSL•FO
RenderX

as well as the personal nature of their storytelling, their advice and suggestions, and their discussions of wide-ranging yet specific topics all position vlogs by women with breast cancer as a potential resource for peer-to-peer support in breast cancer. The experiences of both creating and watching breast cancer vlogs hold significant potential benefits for peer-to-peer support in breast cancer care. This study aligns with Kashian and Jacobson's [14] conclusions, which suggest that given the association between optimal social support and community engagement, "Hopefully practitioners will use this information to encourage patients to join quality online support groups for positive experiences." While there are risks associated with consuming online content, the potential benefits of community and support offered by breast cancer vlogs should not be overlooked. Future research will consider patient perspectives and further address how the specific themes discussed in vlogs may be used to improve the cancer care experience for breast cancer and other cancers.

Acknowledgments

The authors thank the Shades of Pink fund of the Cedars Cancer Foundation for their financial support toward the completion of this research.

Conflicts of Interest

None declared.

References

1. Işikman NG. Coping with hard times through video narratives. International Journal of Communication Research 2022;12(3):205-210 [FREE Full text]
2. Shetty N, Yang Y, Asan O. A scoping review of communication on social media - breast cancer. Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care 2021 Jun;10(1):28-32. [doi: 10.1177/2327857921101003]
3. Kressler J. Women's stories of breast cancer: sharing information through YouTube video blogs [Dissertation]. : The University of Western Ontario (Canada); 2014.
4. Kowitt SD, Ellis KR, Carlisle V, et al. Peer support opportunities across the cancer care continuum: a systematic scoping review of recent peer-reviewed literature. Support Care Cancer 2019 Jan;27(1):97-108. [doi: 10.1007/s00520-018-4479-4] [Medline: 30293093]
5. Ashbury FD, Cameron C, Mercer SL, Fitch M, Nielsen E. One-on-one peer support and quality of life for breast cancer patients. Patient Educ Couns 1998 Oct;35(2):89-100. [doi: 10.1016/s0738-3991(98)00035-4] [Medline: 10026552]
6. Zhang S, Li J, Hu X. Peer support interventions on quality of life, depression, anxiety, and self-efficacy among patients with cancer: A systematic review and meta-analysis. Patient Educ Couns 2022 Nov;105(11):3213-3224. [doi: 10.1016/j.pec.2022.07.008] [Medline: 35858869]
7. Mirrielees JA, Breckheimer KR, White TA, et al. Breast cancer survivor advocacy at a university hospital: development of a peer support program with evaluation by patients, advocates, and clinicians. J Cancer Educ 2017 Mar;32(1):97-104. [doi: 10.1007/s13187-015-0932-y] [Medline: 26477478]
8. Pandey V, Saha SK, Mukherjee SK, Jha A. Development of self-constructed tool to measure factors influencing patient empowerment through online patient support groups. Academy of Marketing Studies Journal 2024;28(1):1-12 [FREE Full text]
9. Taleghani F, Babazadeh S, Mosavi S, Tavazohi H. The effects of peer support group on promoting quality of life in patients with breast cancer. Iran J Nurs Midwifery Res 2012 Feb;17(2 Suppl 1):S125-S130. [Medline: 23833594]
10. Jiang L, Liu Z, Liao Y, Wang J, Li L. Influence of peer support education on psychological adaptation of breast cancer patients. Chin J Pract Nurs 2021;36:1367-1372.
11. Ahn J, Lee KE. Experiences of peer support activities and the need for a metaverse-based program in young women with breast cancer: a qualitative study. Asia Pac J Oncol Nurs 2023 Jul;10(7):100253. [doi: 10.1016/j.apjon.2023.100253] [Medline: 37448532]
12. Morena N, Ahisar Y, Wang X, Nguyen D, Rentschler CA, Meguerditchian AN. Content quality of YouTube videos about metastatic breast cancer in young women: systematic assessment. JMIR Cancer 2023 Nov 14;9:e45101. [doi: 10.2196/45101] [Medline: 37737837]
13. Dunn J, Steginga SK, Occhipinti S, Wilson K. Evaluation of a peer support program for women with breast cancer—lessons for practitioners. J Community Appl Soc Psychol 1999 Jan;9(1):13-22. [doi: 10.1002/(SICI)1099-1298(199901/02)9:1<13::AID-CASP488>3.0.CO;2-F]
14. Kashian N, Jacobson S. Factors of engagement and patient-reported outcomes in a stage IV breast cancer Facebook group. Health Commun 2020 Jan;35(1):75-82. [doi: 10.1080/10410236.2018.1536962] [Medline: 30351185]

XSL•FO
RenderX

15. Power S, Hegarty J. Facilitated peer support in breast cancer: a pre- and post-program evaluation of women's expectations and experiences of a facilitated peer support program. Cancer Nurs 2010;33(2):E9-E16. [doi: 10.1097/NCC.0b013e3181ba9296] [Medline: 20142745]
16. Hu J, Wang X, Guo S, et al. Peer support interventions for breast cancer patients: a systematic review. Breast Cancer Res Treat 2019 Apr;174(2):325-341. [doi: 10.1007/s10549-018-5033-2] [Medline: 30600413]
17. Arias RK. Descriptive analysis of the most viewed youtube videos related to breast cancer survivors [Dissertation]. : EdD Teachers College, Columbia University; 2023.
18. Chai BS, Ingledew PA. Characteristics assessment of online YouTube videos on radiotherapy for breast cancer. Clin Breast Cancer 2023 Jun;23(4):e230-e238. [doi: 10.1016/j.clbc.2023.02.011] [Medline: 36922301]
19. Li JZH, Giuliani M, Ingledew PA. Characteristics assessment of online YouTube videos on radiotherapy for lung cancer. Cureus 2021 Oct;13(10):e19150. [doi: 10.7759/cureus.19150] [Medline: 34868784]
20. Ziegler E, Hill J, Lieske B, Klein J, dem OV Knesebeck, Kofahl C. Empowerment in cancer patients: Does peer support make a difference? A systematic review. Psychooncology 2022 May;31(5):683-704. [doi: 10.1002/pon.5869] [Medline: 34981594]
21. Setoyama Y, Yamazaki Y, Namayama K. Benefits of peer support in online Japanese breast cancer communities: differences between lurkers and posters. J Med Internet Res 2011 Dec 29;13(4):e122. [doi: 10.2196/jmir.1696] [Medline: 22204869]
22. Sillence E. Giving and receiving peer advice in an online breast cancer support group. Cyberpsychol Behav Soc Netw 2013 Jun;16(6):480-485. [doi: 10.1089/cyber.2013.1512] [Medline: 23659724]
23. Abidin C. Mapping internet celebrity on TikTok: exploring attention economies and visibility labours. Cultural Science Journal 2020 Jan 19;12(1):77-103. [doi: 10.5334/csci.140]
24. Marwick AE. Status Update: Celebrity, Publicity, and Branding in the Social Media Age: Yale University Press; 2013.
25. Marwick AE. You may know me from YouTube: (micro-)celebrity in social media. In: Marshall PD, editor. A Companion to Celebrity: Wiley-Blackwell; 2015:333-350. [doi: 10.1002/9781118475089]
26. Hoffner CA, Bond BJ. Parasocial relationships, social media, & well-being. Curr Opin Psychol 2022 Jun;45:101306. [doi: 10.1016/j.copsyc.2022.101306] [Medline: 35219157]
27. Hale BJ, Gonzales AL, Richardson M. Vlogging cancer: predictors of social support in YouTube cancer vlogs. Cyberpsychol Behav Soc Netw 2018 Sep;21(9):575-581. [doi: 10.1089/cyber.2018.0176] [Medline: 30132690]
28. Huh J, Liu LS, Neogi T, Inkpen K, Pratt W. Health vlogs as social support for chronic illness management. ACM Trans Comput Hum Interact 2014 Aug;21(4):23. [doi: 10.1145/2630067] [Medline: 26146474]
29. Ma Z, Ma R, Zhao X, Wang X. Stories that engage the audience: an investigation of popular breast cancer narratives on social media. Telematics and Informatics 2023 Nov;85(1):102048. [doi: 10.1016/j.tele.2023.102048]
30. Dobson AS, Robards B, Carah N. Digital intimate publics and social media: towards theorising public lives on private platforms. In: In Digital Intimate Publics and Social Media: Springer; 2018:3-27. [doi: 10.1007/978-3-319-97607-5_1]
31. Pennebaker JW, Seagal JD. Forming a story: the health benefits of narrative. J Clin Psychol 1999 Oct;55(10):1243-1254. [doi: 10.1002/(SICI)1097-4679(199910)55:10<1243::AID-JCLP6>3.0.CO;2-N] [Medline: 11045774]
32. Liu LS, Huh J, Neogi T, Inkpen K, Pratt W. Health vlogger-viewer interaction in chronic illness management. Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 27 to May 2, 2013; Paris, France p. 49-58 URL: https://dl.acm.org/doi/proceedings/10.1145/2470654 [accessed 2025-03-26] [doi: 10.1145/2470654.2470663] [Medline: 24634895]
33. Huh J, Liu LS, Neogi T, Inkpen K, Pratt W. Health vlogs as social support for chronic illness management. ACM Trans Comput Hum Interact 2014 Aug;21(4):1-31. [doi: 10.1145/2630067] [Medline: 26146474]
34. Isika N, Mendoza A, Bosua R. "Because people can see me": appropriation of YouTube for chronic illness management [Abstract]. Presented at: Proceedings of the 27th European Conference on Information Systems (ECIS); Jun 8-14, 2019; Stockholm & Uppsala, Sweden URL: https://aisel.aisnet.org/ecis2019_rp/13/ [accessed 2025-03-26]
35. Gabarron E, Fernandez-Luque L, Armayones M, Lau AY. Identifying measures used for assessing quality of YouTube videos with patient health information: a review of current literature. Interact J Med Res 2013 Feb 28;2(1):e6. [doi: 10.2196/ijmr.2465] [Medline: 23612432]

Edited by T Mackey; submitted 23.09.24; peer-reviewed by B Gomaa, M O'Connor; revised version received 14.02.25; accepted 22.02.25; published 31.03.25.

Please cite as:
Morena N, Htite ED, Ahisar Y, Hayman V, Rentschler CA, Meguerditchian AN
Breast Cancer Vlogs on YouTube: Descriptive and Content Analyses
JMIR Infodemiology 2025;5:e66812
URL: https://infodemiology.jmir.org/2025/1/e66812
doi:10.2196/66812

XSL•FO
RenderX

© Nina Morena, Elly Dimya Htite, Yitzchok Ahisar, Victoria Hayman, Carrie A Rentschler, Ari N Meguerditchian. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 31.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Tracking Public Interest in Rare Diseases and Eosinophilic Disorders in Germany: Web Search Analysis

Michael Hindelang[1,2,3], BSc, MSc; Sebastian Sitaru[1], Dr med; Alexander Zink[1], MD, MBA, MPH, PhD

[1]Department of Dermatology and Allergy, TUM School of Medicine and Health, Technical University of Munich, Ismaningerstraße 22, Munich, Germany

[2]Pettenkofer School of Public Health, Munich, Germany

[3]Institute for Medical Information Processing, Biometry, and Epidemiology (IBE), Ludwig-Maximilians-Universität München, Munich, Germany

Corresponding Author:
Michael Hindelang, BSc, MSc
Department of Dermatology and Allergy, TUM School of Medicine and Health, Technical University of Munich, Ismaningerstraße 22, Munich, Germany

Abstract

Background: Eosinophilia and hypereosinophilic syndrome (HES) are rare disorders grouped under the term hypereosinophilic disorders. They are diagnosed based on an increased number of eosinophils. They can also cause serious symptoms, including skin, lung, and gastrointestinal problems. These disorders are very rarely recognized due to their rarity and misdiagnosis.

Objective: This study analyzes public interest in hypereosinophilic disorders using data on internet search volume in Germany between 2020 and 2023. Objectives include identifying frequently searched terms, evaluating temporal trends, analyzing seasonal patterns, evaluating geographic differences in search behavior, and identifying unmet information needs and frequently searched risk factors.

Methods: A retrospective analysis using Google Ads Keyword Planner gathered monthly search volume data for 12 German terms related to hypereosinophilic disorders. These terms were selected based on their medical relevance and common usage identified from medical literature. Data were analyzed descriptively, with trends, seasonal variations, and geographical distributions examined. Chi-square tests and correlation analysis assessed statistical significance.

Results: A total of 178 keywords were identified, resulting in a search volume of 1,745,540 queries. The top keyword was "eosophile," a misspelling, followed by "eosinophilia" and "HES." The main categories included "Eosinophilia," "Eosinophils," and "Churg-Strauss syndrome." Temporal analysis showed seasonal growth in search volumes, peaking in January 2023, with higher interest during winter. Geographical analysis showed regional variations.

Conclusions: This research shows a growing public interest in eosinophilic diseases, reflected by a steadily increasing search volume over time. This is particularly evident in searches for basic definitions and diagnostic criteria, such as "eosinophils" or "symptoms of eosinophilic diseases." This increase in search volume, which peaked in January 2023, indicates an increased interest in accurate and readily available information for rare conditions.

(JMIR Infodemiology 2025;5:e69040) doi:10.2196/69040

KEYWORDS

hypereosinophilia; eosinophilia; public health informatics; web search analysis; rare diseases

Introduction

Background

Hypereosinophilic diseases are defined as the presence of persistently elevated eosinophil counts that can cause tissue damage and inflammation of various organs, including the skin, lungs, or gastrointestinal tract [1,2]. These diseases cause symptoms such as severe itching, breathing problems, and gastrointestinal complaints [1,3-5]. Some severe forms of these diseases are characterized by hypereosinophilic syndrome (HES), in which there is tissue infiltration of eosinophils above accepted thresholds for eosinophilia, which can lead to abnormal damage to organ systems [6-9].

Hypereosinophilic disorders can have a significant impact on the individual, but they are often unrecognized and misdiagnosed, mainly due to their rarity and the heterogeneity of clinical presentation [10-13]. Diagnosis is typically based on the exclusion of other causes of eosinophilia (eg, allergic diseases or parasitic infections) and histopathological and immunohistochemical studies, permitting classification in the following subtypes: myeloproliferative (M-HES), lymphocytic (L-HES), idiopathic (I-HES), and chronic eosinophilic leukemia not otherwise specified (CEL-NOS). [1,10]. Treatment is based

XSL•FO
RenderX

on subtype; corticosteroids plus systemic immunosuppression may be used, as well as cytotoxic agents (hydroxyurea and methotrexate), tyrosine kinase inhibitors (imatinib for M-HES), or targeted biologics (mepolizumab) [4]. Prognosis for HES varies; mortality rates can be as low as 7% to 10% for some subtypes, while mortality rates for CEL-NOS can be as high as 33% over a period of 19 - 90 months [4].

It would be helpful to understand how the public navigates information about hypereosinophilic disorders to identify how this process could be altered to lead to successful diagnosis and treatment [4,14]. Digital technologies have developed to the extent that there is widespread access to online health information [15]. Patients are also more likely to turn to the internet to research their health-related questions, starting with a search for their specific symptoms and continuing with possible underlying causes and treatment plans [16-21]. Therefore, this shift from patients becoming passive recipients of medical advice to active participants in health care has seen a steep increase.

The patient journey—the patient experience from symptom awareness to treatment itself—has become more fluid and individualized [15]. This process may involve many phases, such as seeking signs of symptoms, possible diagnoses, or treatment options. These phases are dependent on individual, emotional, and contextual factors and lead to a series of nonlinear pathways through the health care system. These pathways are particularly convoluted in hypereosinophilic disorders [22] due to changing patient presentations, providers, diagnostic challenges, and lack of public and clinical awareness.

The Google Ads Keyword Planner is a valuable tool that allows users to analyze public interest and engagement with health topics using search volume data [23,24]. This approach allows researchers to monitor real-time data on public interest, identify trends in information tracking, and analyze spatial and temporal variations in interest [25-27]. In contrast, Google Trends provides a broader view of relative interest over time, offering insights into how frequently terms are searched relative to all searches made on Google. While Google Trends normalizes search volume data and provides general trends, it lacks the detailed search volume metrics that Google Ads Keyword Planner offers. Unlike social media, where shared information may be curated and filtered [28], search queries provide a more direct and unfiltered view of individuals' health concerns and information needs. This perspective is particularly relevant for rare and underrecognized conditions such as hypereosinophilic disorders, where traditional sources of public health data may be limited.

Objective

Given the rarity and diagnostic challenges of hypereosinophilic disorders, understanding how the public seeks information about these conditions can provide valuable insights into awareness gaps and unmet informational needs. Therefore, the primary aim of this study is to analyze public interest and information-seeking behavior related to eosinophilic disorders in Germany, using Google search data from 2020 to 2023. By examining trends, seasonal patterns, and geographical variations in search volumes, this study seeks to identify key areas of concern and opportunities to enhance public education on these underrecognized conditions.

Methods

Study Design and Data Collection

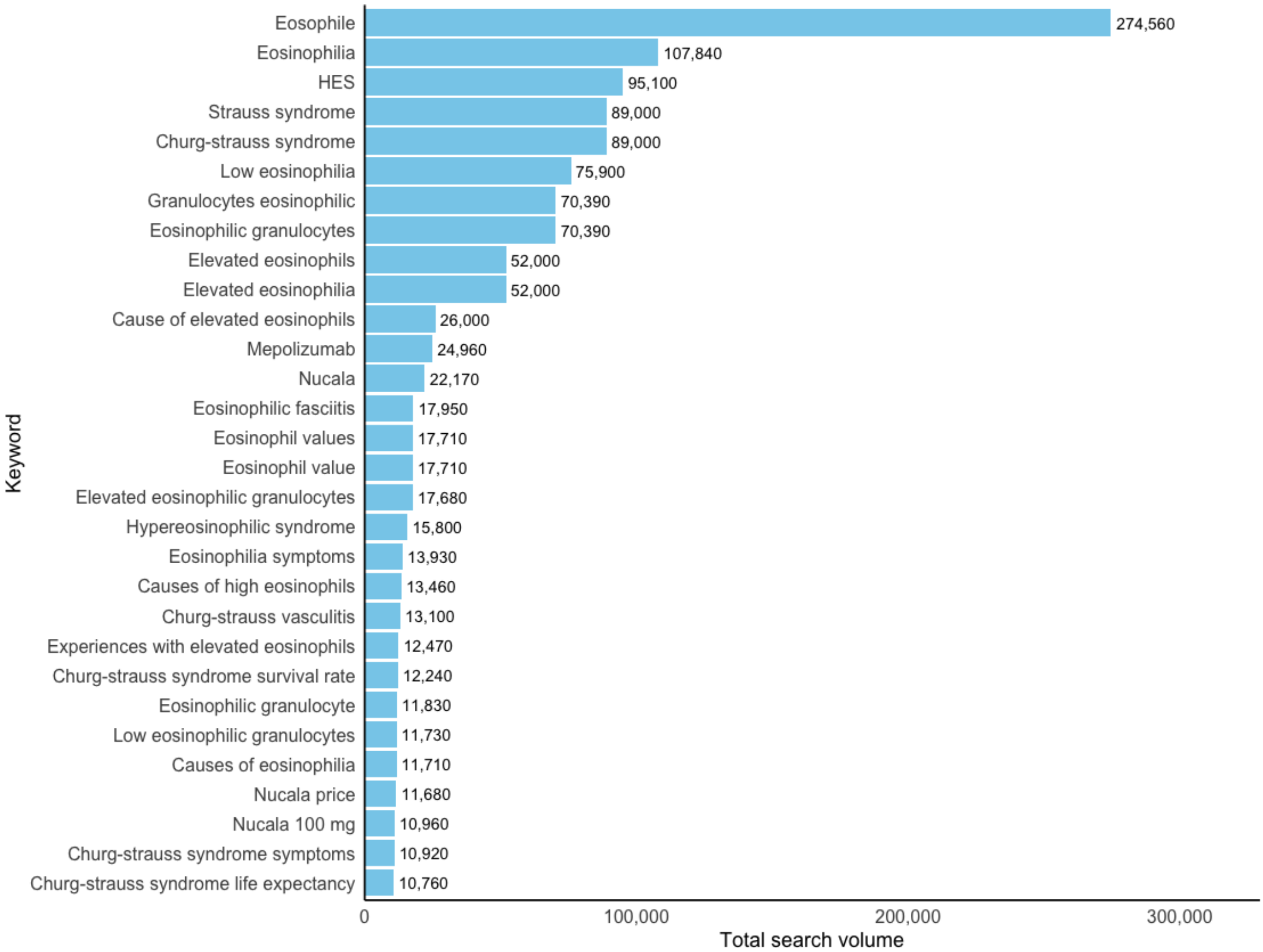
In this retrospective analysis, the Google Ads Keyword Planner was used to gather monthly search volume data. Although initially designed for marketing campaigns, this tool effectively provides monthly web search volume data (ie, monthly number of web searches) for research purposes [27,29-31]. To determine the search volume in a specific area, relevant search terms are entered into the planner. The language and geographical settings can then be configured, and the most relevant keywords and phrases for the topic entered.

Search Terms and Keyword Identification

For this study, 12 German search terms related to hypereosinophilia and associated conditions were entered (Figure 1). The goal was to obtain related keywords and phrases and their monthly search volume in Germany between January 2020 and December 2023. The search terms were "Hypereosinophilia," "Hypereosinophilic Syndrome," "Hes," "Eosinophilia," "Blood Eosinophilia," "Reactive Eosinophilia," "Tissue Eosinophilia," "Eosinophilic Granulocytes," "Eosinophilic Syndrome," "FIP1L1," "Mepolizumab," and "Nucala."

XSL•FO
RenderX

Figure 1. Total search volume for the top 30 keywords (2020 - 2023). HES: hypereosinophilic syndrome.



The 178 keywords were reviewed for relevance to hypereosinophilia and grouped inductively into categories based on their association with the disease, clinical subtypes, treatments, diagnostic markers, and related conditions: "Eosinophilia," "Eosinophils," "Churg Strauss Syndrome," "HES," "Eosinophilic Granulocytes," "Nucala," "Mepolizumab," "Eosinophilic Fasciitis," "Eosinopenia," "EGPA (Eosinophilic Granulomatosis with Polyangiitis)," "Eosinophilic Granulomatosis," "Blood Eosinophilia," and "FIP1L1."

Categories for recurring topics were further subdivided into subcategories (eg, diagnostic information). For each keyword, only one subcategory was assigned. The data were analyzed descriptively.

Geographical Scope

The search volume in all of Germany was examined. Search data for all 16 German federal states and cities were analyzed. Cites were selected based on their population and geographic location in order to obtain a representative overview of all of Germany. For a more in-depth view, cities that are of particular interest due to their unique demographics or health-related infrastructure were also included.

The predefined list of cities included in the study comprised cities that are well distributed across the country and include both large and small cities: Bad Bramstedt, Berlin, Bremen, Cologne, Dortmund, Dresden, Erfurt, Frankfurt, Freiburg, Giessen, Hamburg, Hanover, Heidelberg, Jena, Kassel, Kiel, Kirchheim Teck, Leipzig, Magdeburg, Mainz, Mannheim, Munich, Nuremberg, Regensburg, Rostock, and Stuttgart.

Statistical Methods

To summarize and visualize the search volume data, we used descriptive statistics. Means and measures of dispersion (SD and IQR) were calculated for the monthly search volumes in different categories and subcategories. Frequencies and percentages were used to describe the distribution of searches among the identified categories. We applied a time series decomposition to monthly search volume data since January 2020, using the seasonal and trend decomposition using Loess to extract seasonal, trend, and remainder components. This allowed us to quantify seasonal patterns with CIs and to measure the variability in the data.

Temporal and Geographical Analysis

Trends during the study reporting period (January 2020 to December 2023) were assessed using a time series analysis. We applied seasonal decompositions to locate and quantify seasonal fluctuations in search volume (seasonal decomposition of time series). Search queries were processed for each federal state and city, and the search queries per 100,000 inhabitants were calculated to analyze the geographical distribution of the search

XSL•FO
RenderX

volume. This allowed us to identify areas with increased or decreased search activity related to eosinophilic disorders.

Statistical Testing

Chi-square tests were conducted to assess the significance of differences in search volume between categories and regions. This test was useful for determining whether the distribution of searches across categories was statistically different from what would be expected by chance.

Correlation Analysis

Prior to statistical analysis, we tested the normality of the search volume data using the Shapiro-Wilk test. The results indicated a normal distribution (P=.89). A chi-square test was conducted to examine the associations between search volumes and different regions. This analytical strategy helped to explore potential demographic or health-related factors associated with higher search activity. In this map, each tile represents a search term category, and red areas highlight regions with high relative search volume, resulting in an easy-to-understand heat map visualization that makes it simple to see how interest is distributed across regions in the categories. The intensity of the displayed color reflects the number of searches per 100,000 people on average, with light colors indicating minimal search interest and dark colors indicating higher search interest.

Identification of Significant Rises

We calculated z scores for each time point to detect significant increases in search volume. The z score was computed by subtracting the mean search volume and dividing by the SD. Points with z scores greater than 2 were marked as significant rises, indicating substantial deviations from the average trend.

Validation of Methodology

To ensure the validity of the methodology, the search volumes from the Google Ads Keyword Planner were compared with data from Brandwatch (Giles Palmer) [32]. Brandwatch analyzes mentions on various platforms such as Facebook Public, forums, internet-based news, and X (formerly Twitter). This comparison was used to verify whether the identified peaks and trends in search volume can also be found on other platforms and whether similar patterns exist across multiple sources. This ensured that the observed trends are consistent and not distorted by the commercial orientation of Google Ads.

Software and Tools

All statistical analyses were performed using R (version 4.1.2; R Core Team). Spatial data analysis used *rnaturalearth* packages, and visualizations were created with *ggplot2* [33-40].

Ethical Considerations

As the study was based on publicly accessible Google search terms, there was no requirement for institutional review board approval, and informed consent was not applicable.

Results

Overview of Search Volume

Overall, 178 keywords related to hypereosinophilia were identified, resulting in a search volume of 1,745,540 queries from January 2020 to December 2023.

The analysis of search volumes for keywords related to eosinophilic disorders revealed several key insights. Interestingly, the top keyword was “eosophile” with 274,560 searches, despite being a misspelling and lacking medical meaning (Figure 1). This indicates a potential gap in public understanding or a common typographical error. Following this, the correct medical term “eosinophilia” had 107,840 searches, and “HES” had 95,100 searches. Other top keywords with the highest search volumes included “Churg-Strauss syndrome” and “low eosinophilia” (Table 1).

XSL•FO
RenderX

Table . Top 30 unique keywords for each main category and subcategory.

Main category	Subcategory	Frequency of unique keywords, n
Eosinophilia	Diagnosis	18
Eosinophil granulocytes	Diagnosis	15
Eosinophils	Diagnosis	13
Eosinophilia	Causes/associated diseases	11
HES[a]	General information	11
Nucala	General information	11
Mepolizumab	General information	10
Eosinophilia	Symptoms	7
Churg-Strauss syndrome	Treatment	5
Churg-Strauss syndrome	General information	5
Eosinophilic fasciitis	General information	5
Eosinophilia	General information	5
Churg-Strauss syndrome	Causes/associated diseases	4
Eosinophilia	In animals	4
Churg-Strauss syndrome	Diagnosis	3
Churg-Strauss syndrome	Symptoms	3
Eosinopenia	General information	3
Nucala	Dosage	3
Nucala	Costs	3
Blood eosinophilia	General information	2
Churg-Strauss syndrome	Localization	2
Churg-Strauss syndrome	Survival	2
EGPA[b]	Treatment	2
Eosinopenia	Diagnosis	2
Eosinophil granulocytes	Causes/associated diseases	2
Eosinophilia	Treatment	2
Eosinophils	Causes/associated diseases	2
HES	Treatment	2
HES	Symptoms	2

[a]HES: hypereosinophilic syndrome.

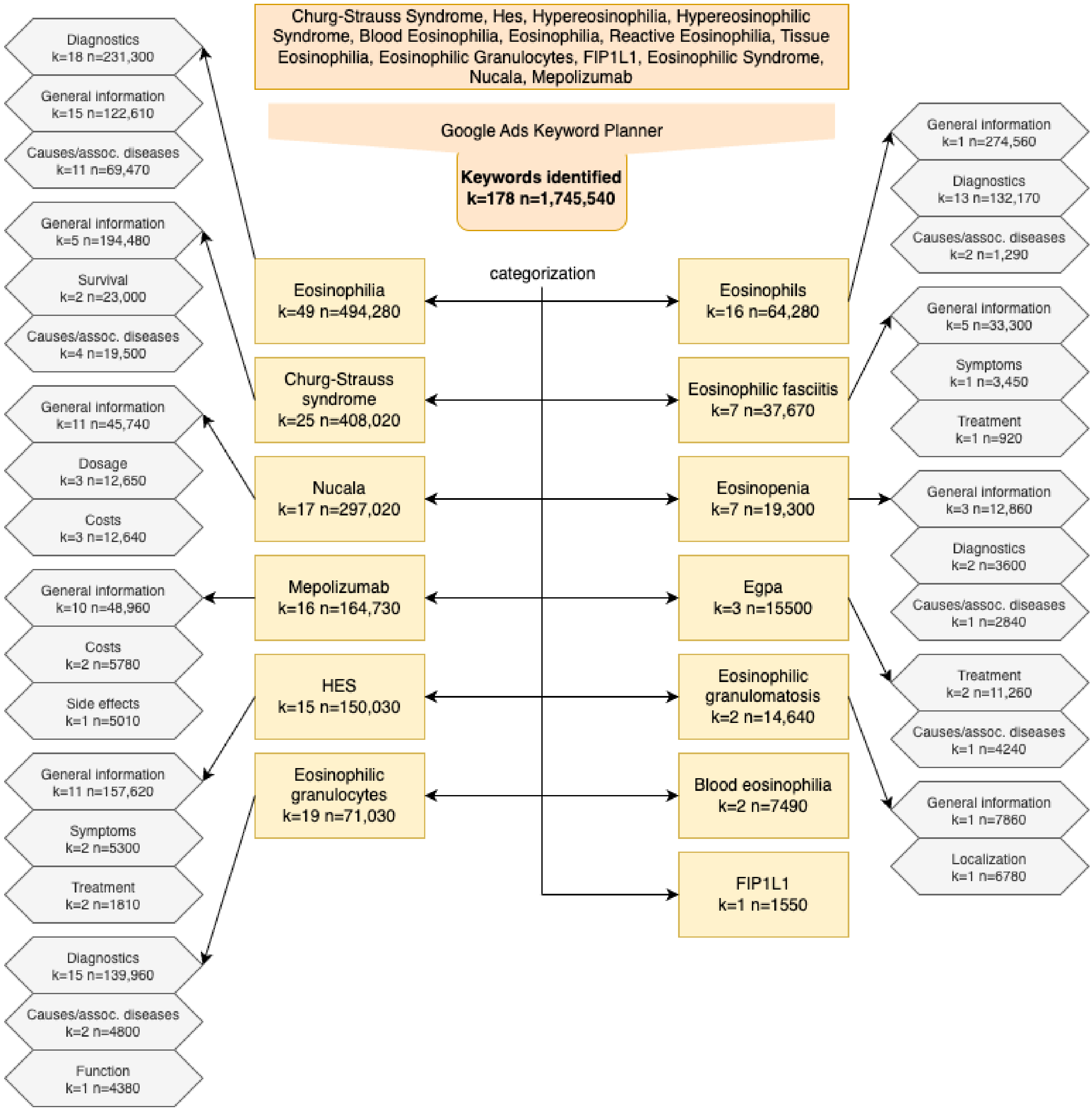
[b]EGPA: eosinophilic granulomatosis with polyangiitis.

Categorization of Search Terms

These keywords were assigned to the following categories (Figure 2). In terms of main categories, "Eosinophilia" topped the list with a total of 494,280 searches, accounting for 28.32% of the total search volume. This was followed by "Eosinophils" with 408,020 searches (23.38%) and "Churg-Strauss syndrome" with 297,020 searches (17.02%). These categories highlight the primary areas of interest among the public.

XSL•FO
RenderX

Figure 2. Flowchart of data generation and content categorization. This flowchart illustrates the process used to gather and categorize the keywords related to eosinophilic disorders. First, relevant search terms were identified based on their medical significance and prevalence in the literature. Next, data were collected from Google Ads Keyword Planner, which provided monthly search volume information for each term. The identified keywords were then grouped into broad categories such as "Eosinophilia," "HES," "Churg-Strauss Syndrome," and others based on their relevance. Within these categories, further subcategories were created, focusing on specific topics such as diagnosis, symptoms, treatment, and related diseases. For clarity, only the top three subcategories are listed when more than three subcategories exist. HES–related web searches in Germany from 2020 to 2023 were analysed. If no subcategory was created, the group was too small or only general information was searched for. assoc.: associated; EGPA: eosinophilic granulomatosis with polyangiitis; FIP1L1: gene involved in hypereosinophilia; HES: hypereosinophilic syndrome; k: number of keywords; n: number of searches.



Within these main categories, the top subcategories provided further insight into specific areas of interest (Figure 2). For "Eosinophilia," the most searched subcategory was "diagnosis" with 231,300 searches (13.25% of the total search volume), followed by "general information" with 122,610 searches (7.02%), and "causes/associated diseases" with 69,470 searches (3.98%). For "Eosinophils," the leading subcategories were "general information" (274,560 searches, 15.73%), "diagnosis" (132,170 searches, 7.57%), and "causes/associated diseases" (1290 searches, 0.07%). In the "Churg-Strauss syndrome" category, "general information" led with 194,480 searches (11.14%), followed by "survival" with 23,000 searches (1.32%) and "causes/associated diseases" with 19,500 searches (1.12%).

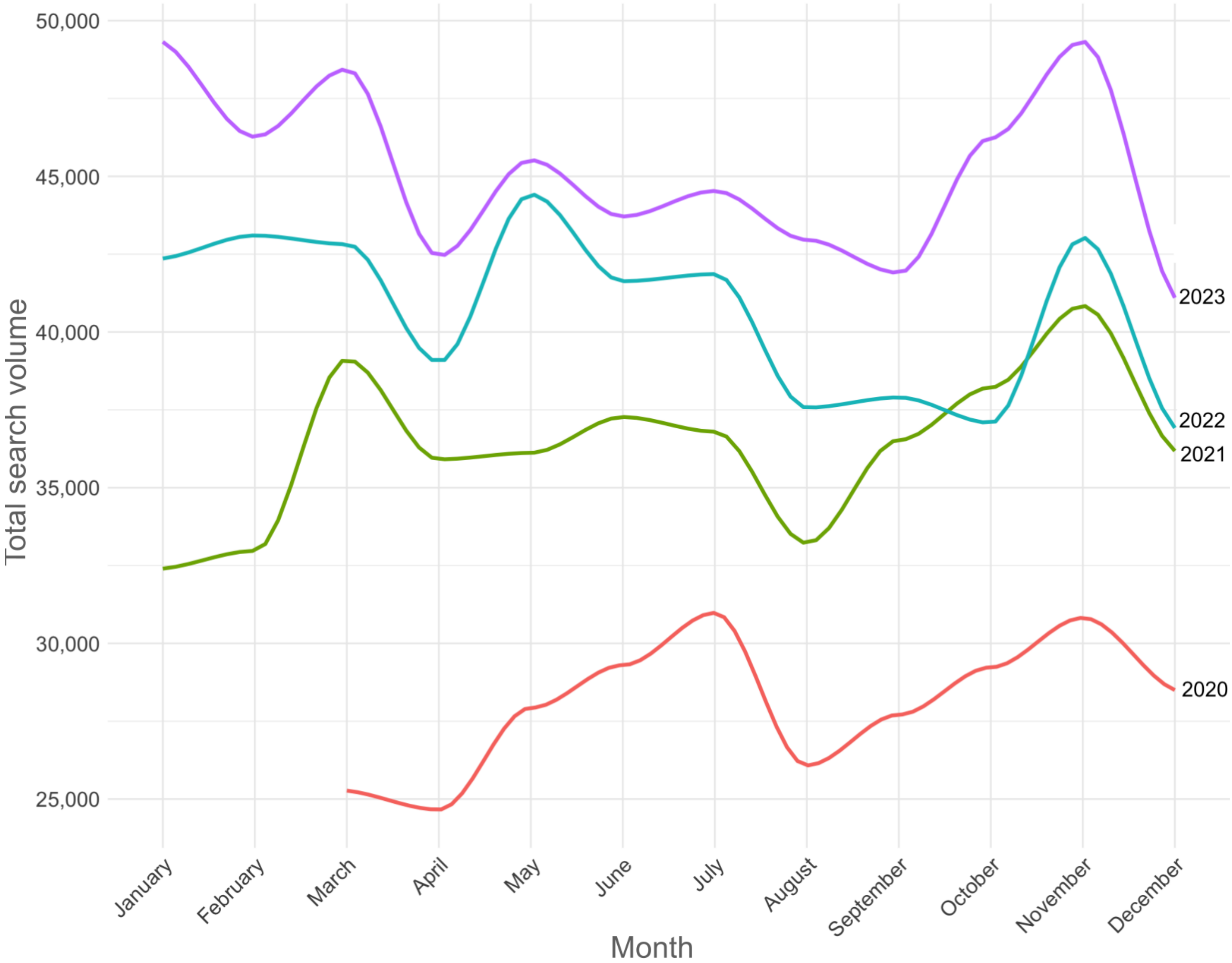
Temporal Trends in Search Volume

From March 2020 to December 2023, eosinophilic disorders exhibited significant growth, peaking in January 2023 at 49,320 queries (Figure 3). Seasonal patterns reveal higher interest during winter months and slight declines in summer, with a consistent yearly increase in overall search volumes. This trend

XSL•FO
RenderX

highlights rising public awareness and interest in eosinophilic disorders over time (see also Figure S3 in Multimedia Appendix 1) .

Figure 3. Total search volume by month over 4 years.



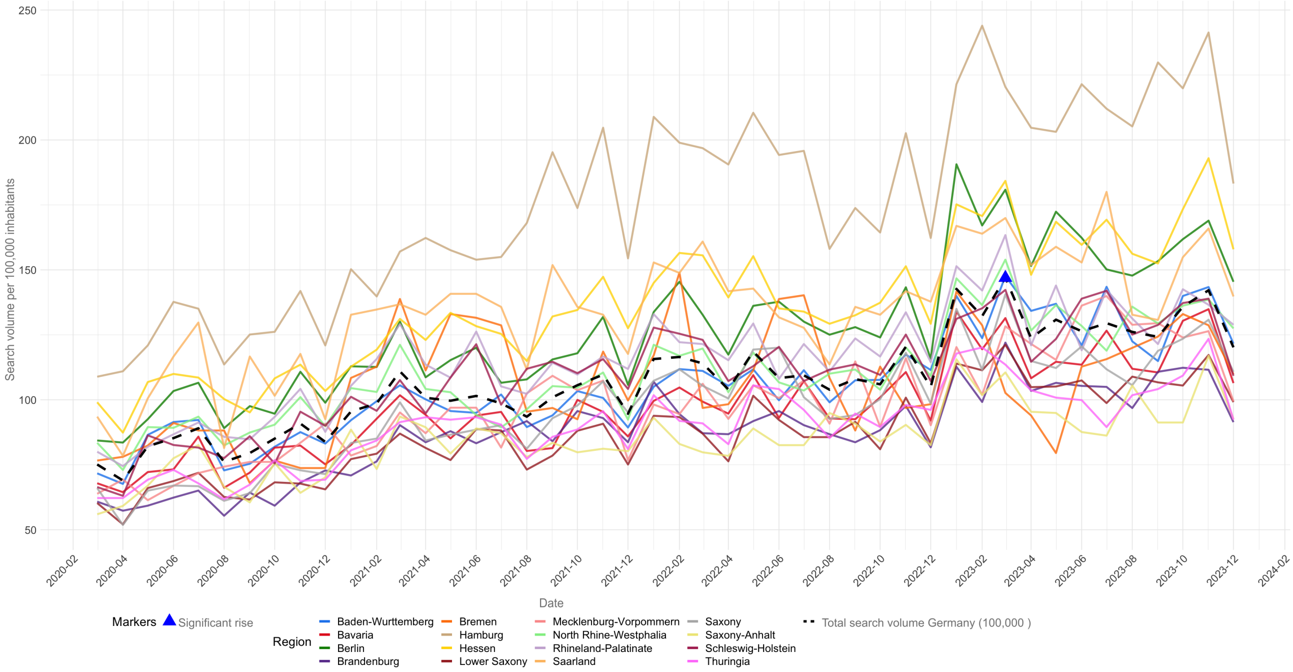
The decomposition of the search volume time series revealed key patterns (Figures S5 and S6 in Multimedia Appendix 1), with Figure S5 in Multimedia Appendix 1 outlining the seasonal and trend decomposition using Loess breakdown into seasonal, trend, and remainder components, and Figure S6 in Multimedia Appendix 1 confirming the seasonal estimates' reliability through diagnostic plots and CIs. The seasonal component showed periodic fluctuations, with notable peaks and troughs, such as a peak in March 2020 (2889) and a low in December 2020 (–4165). This indicates regular cyclical variations in search volume. The trend component exhibited a consistent upward trajectory, increasing from 24,925 in March 2020 to 46,654 by December 2023, suggesting sustained growth in search interest. The remainder component displayed random fluctuations, reflecting irregular variations not explained by the seasonal or trend components. CIs for the seasonal component showed variability, particularly in March 2020, with intervals ranging from –1687 to 7467. Over time, these intervals became narrower, indicating more precise seasonal estimates.

Geographic Distribution

From March 2020 to December 2023, the total search volume per 100,000 inhabitants in Germany showed notable variations (Figure 4 and Figure S1 in Multimedia Appendix 1). Nationally, the peak search volume was 147 per 100,000 inhabitants in March 2023, and the lowest was 69 in April 2020. At the state level, Hamburg recorded the highest peak of 244 searches per 100,000 inhabitants in February 2023. In contrast, Saxony-Anhalt's peak was 117 in November 2023. Seasonal patterns were observed, with increased search volumes across most regions at the start and end of the year. The lowest volumes generally occurred in April 2020 across various states, reflecting consistent national trends.

XSL•FO
RenderX

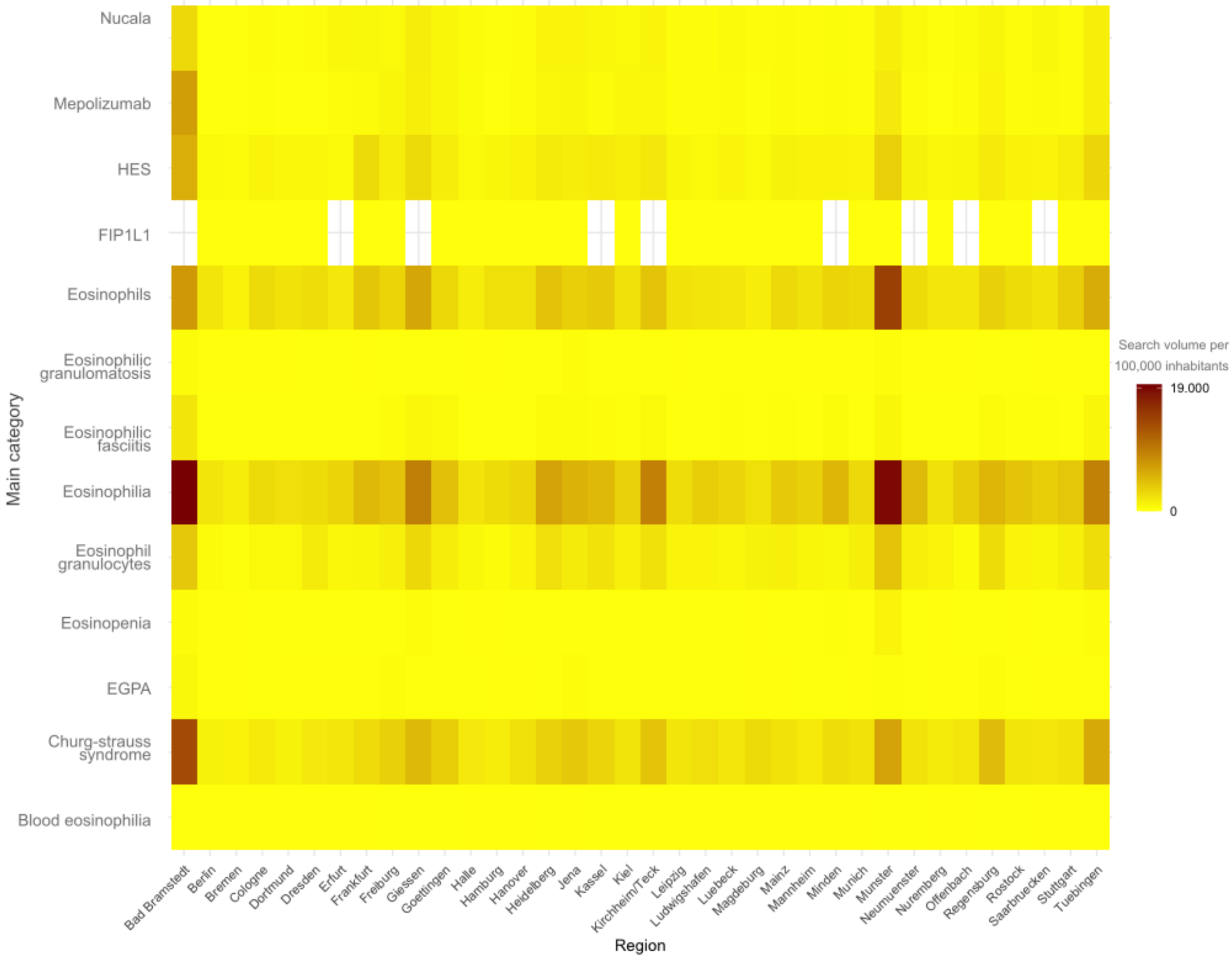
Figure 4. Total search volume over time per 100,000 inhabitants for Germany and by state.



The heatmap (Figure 5 and Figure S4 in Multimedia Appendix 1)) illustrates the search volumes per 100,000 inhabitants across the different German cities. “Eosinophilia” exhibits the highest search volume in Bad Bramstedt with 19,523, while Bremen shows the lowest with 1513. Similarly, for “Churg-Strauss syndrome,” Bad Bramstedt has the highest value at 13,562, compared with Bremen’s 1043. For medications such as “Mepolizumab” and “Nucala,” the highest values are also in Bad Bramstedt (7004 and 2832, respectively), with Bremen having the lowest (207 and 218, respectively). Other cities such as Munster and Tuebingen also display high search volumes, particularly for “eosinophils” (14,529 in Munster) and “Churg-Strauss syndrome” (6154 in Tuebingen).

XSL•FO
RenderX

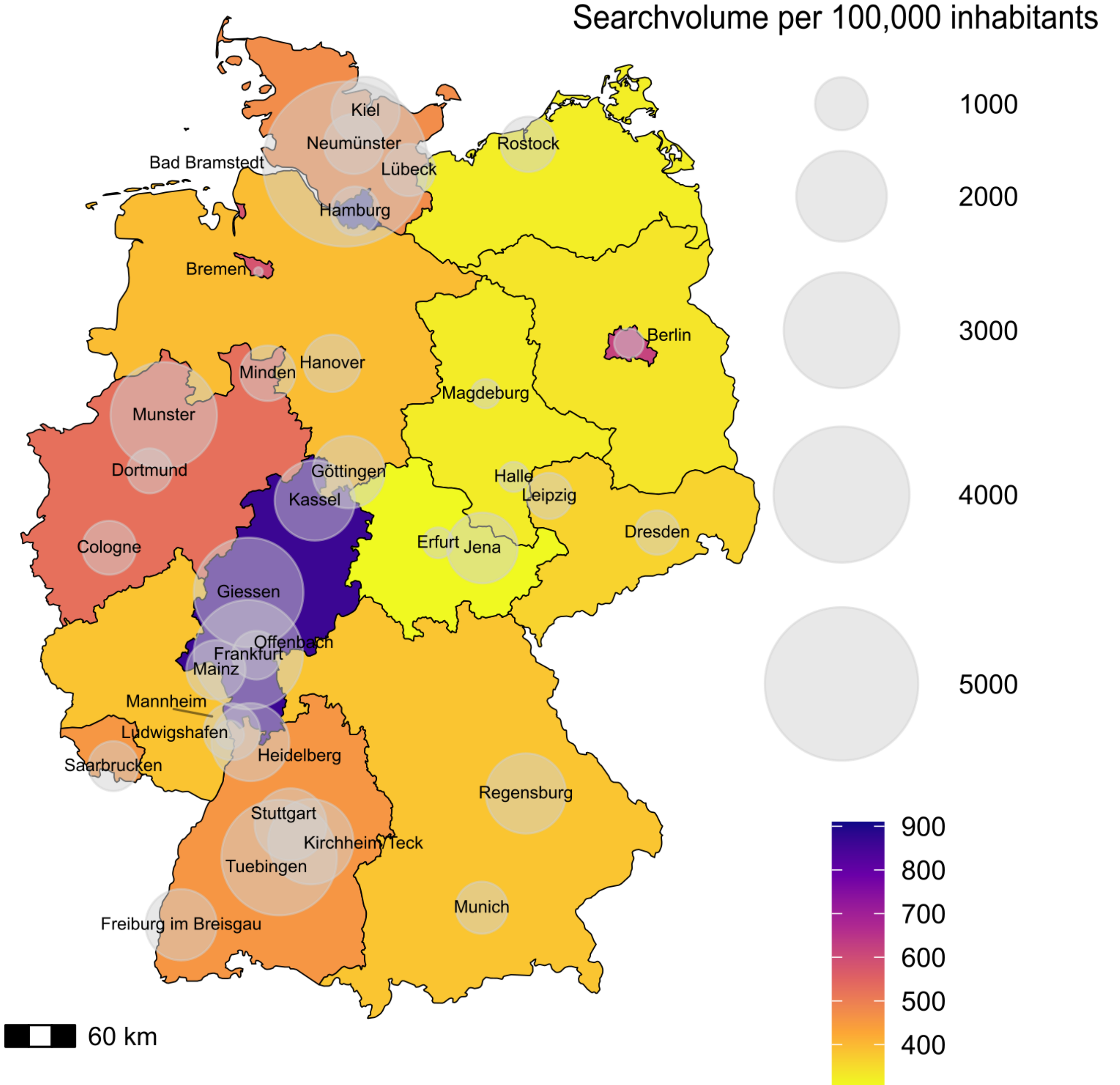
Figure 5. Search volume per 100,000 inhabitants for key categories across regions in Germany from 2020 to 2023. EGPA: eosinophilic granulomatosis with polyangiitis; FIP1L1: gene involved in hypereosinophilia; HES: hypereosinophilic syndrome.



The analysis of the geographical distribution of the total 4-year search volume for the category "HES" per 100,000 inhabitants only shows remarkable differences across different regions and cities (Figure 6). Hamburg has the highest search volume at 854, followed by Hessen with 805. Thuringia, Saxony-Anhalt, and Mecklenburg-Vorpommern have the lowest volumes at 304, 318, and 327 per 100,000 inhabitants, respectively. Among cities, Bad Bramstedt has the highest search volume at 5738 per 100,000 inhabitants. Tübingen, Giessen, and Frankfurt follow with 2989, 2736, and 2719 respectively. Münster reports 2600 searches, Kirchheim Teck reports 1825, Kassel reports 1666, and Regensburg reports 1665. Heidelberg, Göttingen, Stuttgart, and Freiburg im Breisgau show volumes of 1598, 1441, 1433, and 1413, respectively. The lowest search volumes are in Erfurt, Halle, Berlin, and Magdeburg with 690, 683, 680, and 676, respectively, and Bremen has 587 searches per 100,000 inhabitants.

XSL•FO
RenderX

Figure 6. Search volume per 100,000 inhabitants for the “HES” category by region. HES: hypereosinophilic syndrome.



Validation of Search Volume Trends With External Data Sources

To validate the findings from Google Ads Keyword Planner, we compared them with data from a platform [32] that analyzes mentions across social media, forums, internet-based news, and X (formerly Twitter) (Figure S2 in Multimedia Appendix 1). The data revealed notable peaks in January 2023, particularly in internet-based news, and a peak in October 2022 for forum mentions. These trends aligned with the spikes observed in the Google Ads search volume data, supporting the hypothesis that the increases in search activity are reflected in broader public discussions, indicating genuine rises in interest about hypereosinophilic disorders.”

Correlation Analysis

The chi-square test result was highly significant (P<.001), indicating a strong association between the categories of search terms and the regions after normalization per 100,000 inhabitants. This suggests that the frequency and type of searches for HES-related information are significantly influenced by regional factors (Table S1 in Multimedia Appendix 1).

Identification of Significant Rises

Significant rises in search volume are marked with blue triangles in Figure 4. These were detected using z scores, indicating periods where the search volume increased substantially beyond typical variations. Notably, in March 2023, there was a significant rise in search volume, reaching 147 per 100,000 inhabitants, well above the average with a z score of 2.08. This points to a notable event or increased interest during this period, potentially indicating a surge in public concern or awareness about the topics searched.

XSL•FO
RenderX

Discussion

Principal Findings

From March 2020 to December 2023, a total of 1,745,540 search queries on eosinophilic diseases were documented, suggesting a high level of public interest in eosinophilic diseases in Germany. The increasing number of searches conducted, particularly during the winter months, may indicate a growing public interest in rare diseases, possibly due to external influences (eg, seasonal worsening of symptoms and increased media coverage) or displacement effects on search engine suggestion lists. It is striking how often people enter search terms incorrectly, such as "eosophile," which accounted for 274,560 search queries, suggesting gaps in the public's understanding of medical terminology. The frequent misspelling of terms indicates a barrier to finding correct medical information. These spelling mistakes show that there are gaps in the public's knowledge of medical terms, which also points to the need for digital literacy campaigns. Raising awareness and educating the public about the correct terminology can help improve the search for information and support self-education, especially with regard to rare diseases.

Regional differences in search volume suggest that cities like Hamburg, Bad Bramstedt, and Freiburg show higher engagement, likely due to the presence of specialized health centers and greater access to medical information. In comparison, Saxony-Anhalt and Bremen show a lower search volume because there are fewer specialists in these regions and consequently less awareness of the disease. These rates may vary depending on disease prevalence and regional health campaigns, as indicated by the peak in January 2023. Future studies should investigate how access to health care, disease prevalence, and regional campaigns may impact public interest in rare diseases. Interestingly, the most common subcategory for searches related to the broader topic of eosinophilia was diagnosis, suggesting a desire among the public to learn more about the identification and symptoms of eosinophilic disorders. This suggests that those searching for such terms may not yet have been diagnosed and are looking for possible causes of their symptoms. The spike to 49,320 searches in January 2023 may be the result of public health campaigns or media coverage and illustrates how external events influence public health information behavior.

Insights into user preferences highlight some important issues related to eosinophilic diseases, particularly in the area of diagnosis. Conversely, public health efforts should focus on simplified and accessible materials for patients in the early stages, containing credible knowledge about symptoms and evaluations to support efficient and correct diagnosis. Searching and geographically examining keywords can enable targeted public health campaigns in unconscious areas, which can shorten the diagnostic delay and lead to earlier detection and better outcomes. Seasonal peaks in search volume, particularly during the winter months, may reflect worsening of respiratory symptoms due to the colder weather [41]. Due to the exacerbation of eosinophilic diseases, including HES, during this period, people feel the need for more information [41]. The peaks can also be explained by increased public health campaigns or media attention, as respiratory problems tend to receive more attention during the flu season [42].

Regions with lower engagement, such as Saxony-Anhalt and Bremen, may benefit from targeted outreach strategies to increase awareness and education about eosinophilic disorders. Health campaigns in these areas could focus on improving access to information, addressing gaps in public knowledge, and raising awareness about available specialized care centers [43-46]. Additionally, digital platforms, including social media and search engines, can be leveraged to disseminate accurate and timely information, directly reaching individuals in these regions who may not have access to traditional sources of health education. Such approaches could help reduce diagnostic delays by guiding individuals to seek appropriate care sooner and more effectively [43].

To validate the findings, data from social media and internet-based news were examined. The trends, especially the peaks in January 2023 and October 2022, align with search volume spikes. However, social media data may reflect specific user groups rather than the broader population, and news data can be influenced by external factors such as public health campaigns. While these sources provide useful insights, they should be interpreted cautiously as they may not fully capture the overall public interest in eosinophilic disorders.

Comparison With Literature

The findings of our study are consistent with previous research on rare diseases, which highlights the significant challenges faced by patients in obtaining a timely and accurate diagnosis. Hypereosinophilic disorders present with complex clinical features that can lead to diagnostic delays and misdiagnoses, similar to other rare conditions. The 2013 report on the impact of rare diseases shows that patients with rare diseases often have to undergo long diagnostic pathways. In the United States, it takes an average of 7.6 years, and in the United Kingdom, 5.6 years before a correct diagnosis is made [44]. During this period, patients typically consult up to 8 physicians and receive 2-3 misdiagnoses, with 82% of social media comments on HES reporting diagnostic delays, highlighting the complexity of diagnosing rare eosinophil-driven disorders [45].

In our study, the diagnosis-related subcategory was the most frequently searched, reflecting the public's struggle to find accurate information and the challenges associated with diagnosing hypereosinophilic disorders [46,47]. This pattern of information-seeking behavior suggests a significant unmet need for awareness and educational resources, both for the public and health care professionals [48].

This finding aligns with previous research on atopic dermatitis and pollen allergies, which also indicated that online search data could reveal significant gaps in public knowledge and help identify areas requiring targeted health education [26,29,49-51].

Strengths and Limitations

A key strength of this study is its use of Google Ads Keyword Planner, providing real-time search data to assess public interest in eosinophilic disorders. This approach demonstrates the

XSL•FO
RenderX

potential of infodemiology for rare disease research, offering insights into awareness, informational needs, and regional variations, which traditional data sources often lack. Compared with studies such as Pauer et al [52], which used search queries for rare disease epidemiology, our study focuses specifically on eosinophilic disorders, adding nuance to public engagement analysis [53,54]. Additionally, unlike Pauer et al [52] and Tozzi et al [55], which examined information quality and user demographics, we provide real-time trends in public interest, enhancing the understanding of disease awareness.

Several limitations should be considered when interpreting these findings. Google Ads Keyword Planner, while effective for data collection, is primarily a marketing tool and may introduce biases or limitations in data accuracy for research purposes [56,57]. The high search volume for misspelled terms such as “eosophile” may have skewed the results and points to gaps in public education that were not fully explored in this study. Future research could investigate the reasons behind common misspellings and their impact on information retrieval and understanding. Furthermore, the generalizability of the results may be limited by demographic and access biases. Younger people are generally more tech-savvy and therefore more likely than older cohorts to search for health information on Google, skewing the search data toward their demographic group. This study did not account for differences in internet use by age, socioeconomic status, and region, and the results may not generalize to different segments of the population [58]. The final limitation concerns the temporal relevance of the data, which extend to December 2023 but may not account for newer trends or developments, such as newer treatment options or educational campaigns, that could influence interest and search behavior outside the time frame of this study.

Conclusion

The findings of this study provide valuable insights into public interest and information-seeking behaviors related to hypereosinophilic disorders in Germany. The data suggest that there is a growing awareness and concern about these rare conditions, as evidenced by the increasing search volumes over time. The pronounced regional differences highlight the need for localized health education and resource allocation to address potential disparities in awareness and access to information.

Our results emphasize the urgent need for faster, more accurate diagnostic pathways and enhanced awareness among health care professionals to improve the management and outcomes for patients with rare diseases such as HES and EGPA. The study underscores the potential of using Google search trends as a tool for public health surveillance, particularly for rare and underrecognized conditions such as hypereosinophilic disorders. Future research should aim to integrate demographic data and explore the impact of public health campaigns and health care access on search behaviors. Additionally, efforts should be made to improve public understanding of these disorders through accurate and accessible information, potentially leveraging the very platforms where information-seeking is occurring. Addressing these knowledge gaps could lead to better patient outcomes through earlier diagnosis and more informed health decision-making.

Acknowledgments

Parts of this study were funded by an unrestricted research grant from GSK GmbH and Co KG.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MH was responsible for conceptualization, methodology, investigation, formal analysis, and writing the original draft. SS contributed to conceptualization, methodology, and writing, as well as review and editing. AZ also contributed to conceptualization, methodology, and writing, as well as review and editing.

Conflicts of Interest

MH once received a personal honorarium from GSK, AZ received personal honoraria from GSK, and SS has no conflict of interest.

Multimedia Appendix 1
A correlation matrix of search volume data for various keywords related to eosinophilic disorders across 27 cities in Germany. This matrix highlights the relationships between different terms such as "eosinophilia," "Churg-Strauss syndrome," "HES," and related treatments and conditions. The data reflect regional variations with notable differences in search volume across cities. These correlations provide insights into public interest in specific aspects of eosinophilic disorders and help identify patterns in search behaviors linked to geographic location.
[DOCX File, 2057 KB - infodemiology_v5i1e69040_app1.docx]

References

XSL•FO
RenderX

1. Shomali W, Gotlib J. World Health Organization-defined eosinophilic disorders: 2022 update on diagnosis, risk stratification, and management. Am J Hematol 2022 Jan 1;97(1):129-148. [doi: 10.1002/ajh.26352] [Medline: 34533850]
2. Valent P, Klion AD, Horny HP, et al. Contemporary consensus proposal on criteria and classification of eosinophilic disorders and related syndromes. J Allergy Clin Immunol 2012 Sep;130(3):607-612. [doi: 10.1016/j.jaci.2012.02.019] [Medline: 22460074]
3. Akuthota P, Weller PF. Spectrum of eosinophilic end-organ manifestations. Immunol Allergy Clin North Am 2015 Aug;35(3):403-411. [doi: 10.1016/j.iac.2015.04.002] [Medline: 26209892]
4. Requena G, van den Bosch J, Akuthota P, et al. Clinical profile and treatment in hypereosinophilic syndrome variants: a pragmatic review. J Allergy Clin Immunol Pract 2022 Aug;10(8):2125-2134. [doi: 10.1016/j.jaip.2022.03.034] [Medline: 35470096]
5. Curtis C, Ogbogu P. Hypereosinophilic syndrome. Clin Rev Allergy Immunol 2016 Apr;50(2):240-251. [doi: 10.1007/s12016-015-8506-7] [Medline: 26475367]
6. Ogbogu PU, Rosing DR, Horne MK. Cardiovascular manifestations of hypereosinophilic syndromes. Immunol Allergy Clin North Am 2007 Aug;27(3):457-475. [doi: 10.1016/j.iac.2007.07.001] [Medline: 17868859]
7. Schwaab J, Lübke J, Reiter A, et al. Idiopathic hypereosinophilic syndrome—diagnosis and treatment. Allergo J Int 2022 Oct;31(7):251-256. [doi: 10.1007/s40629-022-00221-w]
8. Noh HR, Magpantay GG. Hypereosinophilic syndrome. Allergy Asthma Proc 2017 Jan 1;38(1):78-81. [doi: 10.2500/aap.2017.38.3995] [Medline: 28052805]
9. Bonanno D, Fedele R, Minciullo P, Quattrocchi P, Ferlazzo B. Idiopathic hypereosinophilic syndrome associated with rheumatoid arthritis. a case report [Article in Italian]. Reumatismo 2003;55(3):181-183 [FREE Full text] [doi: 10.4081/reumatismo.2003.181] [Medline: 14513118]
10. Klion AD. How I treat hypereosinophilic syndromes. Blood 2015 Aug 27;126(9):1069-1077. [doi: 10.1182/blood-2014-11-551614] [Medline: 25964669]
11. Khoury P, Desmond R, Pabon A, et al. Clinical features predict responsiveness to imatinib in platelet-derived growth factor receptor-alpha-negative hypereosinophilic syndrome. Allergy 2016 Jun;71(6):803-810. [doi: 10.1111/all.12843] [Medline: 26797802]
12. Chen YYK, Khoury P, Ware JM, et al. Marked and persistent eosinophilia in the absence of clinical manifestations. J Allergy Clin Immunol 2014 Apr;133(4):1195-1202. [doi: 10.1016/j.jaci.2013.06.037] [Medline: 23987798]
13. Schuster B, Zink A, Eyerich K. Medical algorithm: diagnosis and treatment of hypereosinophilic syndrome. Allergy 2020 Nov;75(11):3003-3006. [doi: 10.1111/all.14368] [Medline: 32416612]
14. Rosenberg HF, Dyer KD, Foster PS. Eosinophils: changing perspectives in health and disease. Nat Rev Immunol 2013 Jan;13:9-22. [doi: 10.1038/nri3341] [Medline: 23154224]
15. Davies EL, Bulto LN, Walsh A, et al. Reporting and conducting patient journey mapping research in healthcare: a scoping review. J Adv Nurs 2023 Jan;79(1):83-100. [doi: 10.1111/jan.15479] [Medline: 36330555]
16. Mackey T, Baur C, Eysenbach G. Advancing infodemiology in a digital intensive era. JMIR Infodemiology 2022;2(1):e37115. [doi: 10.2196/37115] [Medline: 37113802]
17. Hindelang M, Sitaru S, Zink A. Transforming health care through chatbots for medical history-taking and future directions: comprehensive systematic review. JMIR Med Inform 2024;12:e56628. [doi: 10.2196/56628] [Medline: 39207827]
18. Hindelang M, Wecker H, Biedermann T, et al. Continuously monitoring the human machine? - a cross-sectional study to assess the acceptance of wearables in Germany. Health Informatics J 2024;30(2):14604582241260607. [doi: 10.1177/14604582241260607] [Medline: 38900846]
19. Pearce E, Raj H, Emezienna N, et al. The use of social media to express and manage medical uncertainty in dyskeratosis congenita: content analysis. JMIR Infodemiology 2024;4. [doi: 10.2196/46693] [Medline: 38224480]
20. Mavragani A. Infodemiology and infoveillance: scoping review. J Med Internet Res 2020;22(4). [doi: 10.2196/16206] [Medline: 32310818]
21. Li J, He Z, Zhang M, et al. Estimating rare disease incidences with large-scale internet search data: development and evaluation of a two-step machine learning method. JMIR Infodemiology 2023;3:e42721. [doi: 10.2196/42721]
22. Hwee J, Huynh L, Du S, et al. Hypereosinophilic syndrome in Europe: retrospective study of treatment patterns, clinical manifestations, and healthcare resource utilization. Ann Allergy Asthma Immunol 2023 Jun;130(6):768-775. [doi: 10.1016/j.anai.2023.02.022] [Medline: 36863663]
23. Kamiński M, Tizek L, Zink A. 'Dr. Google, what Is that on my skin?'-internet searches related to skin problems: Google Trends data from 2004 to 2019. Int J Environ Res Public Health 2021 Mar 4;18(5):2541. [doi: 10.3390/ijerph18052541] [Medline: 33806391]
24. Schielein MC, Tizek L, Seifert F, et al. Health care of chronic inflammatory skin diseases: do affected individuals seek dermatological care? Hautarzt 2019 Nov;70(11):875-882. [doi: 10.1007/s00105-019-04481-6] [Medline: 31560083]
25. Ziehfreund S, Tizek L, Arkema EV, et al. Identifying sarcoidosis trends using web search and real-world data in Sweden: a retrospective longitudinal study. Sci Rep 2024 Aug 20;14(1):19260. [doi: 10.1038/s41598-024-69223-8] [Medline: 39164281]

XSL•FO
RenderX

26. Sitaru S, Tizek L, Buters J, et al. Assessing the national burden of allergic asthma by web-search data, pollen counts, and drug prescriptions in Germany and Sweden. World Allergy Organ J 2023 Feb;16(2):100752. [doi: 10.1016/j.waojou.2023.100752] [Medline: 36896457]
27. Wecker H, Ziehfreund S, Hindelang M, et al. Change of perspective: impact of COVID-19 pandemic on axial spondyloarthritis-related web searches in Germany. Sci Rep 2024 Feb 19;14:4044. [doi: 10.1038/s41598-024-54047-3] [Medline: 38374132]
28. Ranpariya V, Chu B, Fathy R, et al. Dermatology without dermatologists? analyzing Instagram influencers with dermatology-related hashtags. J Am Acad Dermatol 2020 Dec;83(6):1840-1842. [doi: 10.1016/j.jaad.2020.05.039] [Medline: 32416205]
29. Wecker H, Ziehfreund S, Sitaru S, et al. Dilemmas of nomenclature: web search analysis reveals European preferences in atopic skin diseases. Clin Transl Allergy 2024 Apr;14(4):e12355. [doi: 10.1002/clt2.12355] [Medline: 38613147]
30. Wecker H, Ziehfreund S, Sitaru S, et al. Burden of atopic dermatitis in Europe: a population-centred approach leveraging web search data in 21 European countries. J Eur Acad Dermatol Venereol 2024 Aug;38(8):1637-1648. [doi: 10.1111/jdv.19989] [Medline: 38641980]
31. Schuster B, Tizek L, Schielein MC, et al. Retracing the COVID-19 pandemic in Germany from a public perspective using Google search queries related to “coronavirus”. Gesundheitswesen 2021 May;83(5):e9-e14. [doi: 10.1055/a-1398-5417] [Medline: 33862647]
32. Brandwatch. The social suite of the future. URL: www.brandwatch.com [accessed 2025-01-24]
33. South A, Michael S, Massicotte P. rnaturalearthhires: high resolution world vector map data from natural earth used in rnaturalearth. URL: https://ropensci.r-universe.dev/rnaturalearthhires [accessed 2025-03-27]
34. South A, Michael S, Massicotte P. rnaturalearthdata: world vector map data from natural earth used in ‘rnaturalearth’. URL: https://ropensci.r-universe.dev/rnaturalearthdata [accessed 2025-03-27]
35. Massicotte P, South A, Hufkens K. rnaturalearth: world map data from natural earth. URL: https://ropensci.r-universe.dev/rnaturalearth [accessed 2025-03-27]
36. Pebesma E. Simple features for R: standardized support for spatial vector data. R J 2018;10(1):439-446. [doi: 10.32614/RJ-2018-009]
37. Pebesma E, Bivand R. Spatial Data Science: With Applications in R: Chapman and Hall/CRC; 2023. [doi: 10.1201/9780429459016]
38. Wickham H, François R, Henry L, et al. dplyr: a grammar of data manipulation. 2023. URL: https://CRAN.R-project.org/package=dplyr [accessed 2025-03-27]
39. Wickham H. Ggplot2: Elegant Graphics for Data Analysis: Springer; 2016. [doi: 10.1007/978-3-319-24277-4]
40. Wickham H, Bryan J. readxl: Read Excel Files. 2023. URL: https://CRAN.R-project.org/package=readxl [accessed 2025-05-16]
41. Hurrell JM, Genta RM, Dellon ES. Prevalence of esophageal eosinophilia varies by climate zone in the United States. Am J Gastroenterol 2012 May;107(5):698-706. [doi: 10.1038/ajg.2012.6] [Medline: 22310220]
42. Neumann G, Kawaoka Y. Seasonality of influenza and other respiratory viruses. EMBO Mol Med 2022 Apr 7;14(4):e15352. [doi: 10.15252/emmm.202115352] [Medline: 35157360]
43. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
44. Hendriksz C. Rare disease impact report: insights from patients and the medical community. URL: https://www.academia.edu/18842593/Rare_Disease_Impact_Report_Insights_from_patients_and_the_medical_community [accessed 2024-09-19]
45. Strobel MJ, Alves D, Roufosse F, et al. Insights from social media on the patient experience of living with rare eosinophil-driven diseases. J Patient Exp 2022;9. [doi: 10.1177/23743735221143953] [Medline: 36530646]
46. Klion AD. Approach to the patient with suspected hypereosinophilic syndrome. Hematology Am Soc Hematol Educ Program 2022 Dec 9;2022(1):47-54. [doi: 10.1182/hematology.2022000367] [Medline: 36485140]
47. Szymczyk A, Jaworski J, Podhorecka M. The challenge of diagnosing and classifying eosinophilia and eosinophil disorders: a review. Cent Eur J Immunol 2024;49(1):60-69. [doi: 10.5114/ceji.2024.136512] [Medline: 38812609]
48. Jackson DJ, Akuthota P, Andradas R, et al. Improving care in eosinophil-associated diseases: a charter. Adv Ther 2022 Jun;39(6):2323-2341. [doi: 10.1007/s12325-022-02110-8] [Medline: 35489014]
49. Schober A, Tizek L, Johansson EK, et al. Monitoring disease activity of pollen allergies: what crowdsourced data are telling us. World Allergy Organ J 2022 Dec;15(12):100718. [doi: 10.1016/j.waojou.2022.100718] [Medline: 36438191]
50. Mick A, Tizek L, Schielein M, et al. Can crowdsourced data help to optimize atopic dermatitis treatment? comparing web search data and environmental data in Germany. J Eur Acad Dermatol Venereol 2022 Apr;36(4):557-565. [doi: 10.1111/jdv.17875] [Medline: 34921466]
51. Hindelang M, Tizek L, Harders C, et al. Hybrid care potential of teledermatology: the importance of linking digital and physical practice and acceptance of online services: a cross-sectional study. Health Sci Rep 2024 Jul;7(7):e2241. [doi: 10.1002/hsr2.2241] [Medline: 38983681]
52. Pauer F, Litzkendorf S, Göbel J, et al. Rare diseases on the internet: an assessment of the quality of online information. J Med Internet Res 2017 Jan 18;19(1):e23. [doi: 10.2196/jmir.7056] [Medline: 28100442]

XSL•FO
RenderX

53. Morgan T, Schmidt J, Haakonsen C, et al. Using the internet to seek information about genetic and rare diseases: a case study comparing data from 2006 and 2011. JMIR Res Protoc 2014 Feb 24;3(1):e10. [doi: 10.2196/resprot.2916] [Medline: 24565858]
54. Stenberg U, Westfal L, Dybesland Rosenberger A, et al. A scoping review of health literacy in rare disorders: key issues and research directions. Orphanet J Rare Dis 2024;19:328. [doi: 10.1186/s13023-024-03332-5] [Medline: 39243094]
55. Tozzi AE, Mingarelli R, Agricola E, et al. The internet user profile of Italian families of patients with rare diseases: a web survey. Orphanet J Rare Dis 2013;8(76). [doi: 10.1186/1750-1172-8-76] [Medline: 23680013]
56. Nölke L, Mensing M, Krämer A, et al. Sociodemographic and health-(care-)related characteristics of online health information seekers: a cross-sectional German study. BMC Public Health 2015;15. [doi: 10.1186/s12889-015-1423-0]
57. Percheski C, Hargittai E. Health information-seeking in the digital age. J Am Coll Health 2011;59(5):379-386. [doi: 10.1080/07448481.2010.513406] [Medline: 21500056]
58. Lupton D. Young people's use of digital health technologies in the global north: narrative review. J Med Internet Res 2021 Jan 11;23(1):e18286. [doi: 10.2196/18286] [Medline: 33427684]

Abbreviations

CEL-NOS: chronic eosinophilic leukemia-not otherwise specified
EGPA: eosinophilic granulomatosis with polyangiitis
HES: hypereosinophilic syndrome
I-HES: idiopathic hypereosinophilic syndrome
L-HES: lymphocytic hypereosinophilic syndrome
M-HES: myeloproliferative hypereosinophilic syndrome

Edited by R Cuomo; submitted 20.11.24; peer-reviewed by Q Niu, S Wei; revised version received 24.01.25; accepted 28.02.25; published 26.05.25.

Please cite as:
Hindelang M, Sitaru S, Zink A
Tracking Public Interest in Rare Diseases and Eosinophilic Disorders in Germany: Web Search Analysis
JMIR Infodemiology 2025;5:e69040
URL: https://infodemiology.jmir.org/2025/1/e69040
doi:10.2196/69040

© Michael Hindelang, Sebastian Sitaru, Alexander Zink. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 26.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Communicating Antimicrobial Resistance on Instagram: Content Analysis of #AntibioticResistance

Elin Nilsson[1,2*], PhD; Emma Oljans[3,4*], PhD; Anna-Carin Nordvall[5*], PhD, Prof Dr; Mirko Ancillotti[6*], PhD

[1]Umeå School of Business, Economics, and Statistics, Umeå University, Handelshögskolan, Biblioteksgränd 6, Umeå, Sweden

[2]CERUM (Center for Regional Science), Umeå University, Umeå, Sweden

[3]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

[4]Department of Movement, Culture and Society, Swedish School of Sport and Health Sciences, Stockholm, Sweden

[5]Department of Business Studies, Uppsala University, Uppsala, Sweden

[6]Centre for Research Ethics and Bioethics, Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

[*]all authors contributed equally

Corresponding Author:
Elin Nilsson, PhD
Umeå School of Business, Economics, and Statistics, Umeå University, Handelshögskolan, Biblioteksgränd 6, Umeå, Sweden

Abstract

Background: Antimicrobial resistance (AMR) is a major global health issue heavily influenced by human behavior. Effective communication and awareness-raising are crucial in curbing AMR, with social network sites (SNSs) significantly shaping health behaviors. Despite their potential, current analyses of AMR on SNSs have focused mainly on top-down communication initiatives.

Objective: This study aims to examine AMR on Instagram (Meta Platforms), identifying key actors, content themes, and the nature of the communication to understand how AMR is portrayed and perceived.

Methods: Based on the sender-message-channel-receiver model, this study used content analysis to review publicly accessible posts on Instagram. The data refer to 24 months, focusing on the hashtag "#antibioticresistance." After cleaning the data, 610 posts (10% of the total 6105) were analyzed.

Results: Content creators were predominantly information drivers or professionals in science and health. Posts frequently featured text-dominated visuals or images of bacteria and laboratory tests. However, the AMR posts were found to be siloed, with limited engagement beyond specific interest groups. The study highlighted the neutrality and accuracy of the content but noted the challenge of reaching a broader audience.

Conclusions: While Instagram serves as a platform for accurate and informative AMR communication, the post of it remains confined to niche groups, limiting its broader impact. To enhance engagement, AMR discussions should be integrated into more general interest content, use visually compelling formats, and encourage institutional participation and interactive user engagement.

(JMIR Infodemiology 2025;5:e67825) doi:10.2196/67825

KEYWORDS

antimicrobial resistance; social media engagement; communication; Instagram; health communication; content analysis

Introduction

Background

Antimicrobial resistance (AMR) is a major global issue, causing substantial health burdens and societal costs. Every year, about 1.3 million people die from infections caused by antimicrobial-resistant bacteria, and this figure is predicted to increase to 10 million by 2050 [1,2].

The Role of Community Awareness in Combating AMR

Community knowledge and awareness raising are key factors in curbing resistance because AMR largely depends on human behavior [3-5]. In addition to medical use, which is the primary driver of AMR [6,7], there is plenty of individually and contextually induced health and lifestyle behavior impacting human health—but also animal health and the environment—such as food production and consumption [8], international travel [9], or the adoption of basic hygiene practices to avoid infection spread [10]. Previous research stresses the necessity of considering the contextual motivations and preferences of participants and that health communication needs to explore and engage with these [11,12]. The World Health Organization's (WHO's) Global Action Plan highlights the need to emphasize and improve AMR awareness and understanding through effective communication, education, and training [13].

XSL•FO
RenderX

The Influence of Social Network Sites on Health Behaviors

Social influences shape health and lifestyle behaviors to a high degree. Research has highlighted the influence exerted by social network sites (SNSs), which stand as structural determinants of health nowadays for their role in information exchange and diffusion of beliefs [14]. SNSs are web-based services that allow individuals to (1) construct a public or semipublic profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system [15]. SNSs are frequently used to address and communicate health issues [16]. They were used in a few intervention studies to raise awareness of AMR, for educational purposes, and to promote behavioral change [17-21]. International health bodies such as the WHO and the European Centre for Disease Prevention and Control also engage in SNSs—primarily during Antimicrobial Awareness Weeks—and promote their use in awareness-raising campaigns [22,23]. However, to date, the analysis of AMR on SNSs has been almost completely limited to top-down communication, with a primary focus on X (formerly Twitter) [24-26]. Recent studies have highlighted the effectiveness of SNSs in rapidly disseminating health-related evidence to professionals, particularly during public health emergencies, where timeliness and platform dynamics influence uptake and visibility [27]. At the same time, SNSs demonstrated considerable potential for reaching broader audiences with public health messages, influencing awareness, attitudes, and behaviors [28]. However, the very features that facilitate wide dissemination can also contribute to the amplification of misinformation. A systematic review found that a significant proportion of health content on popular SNSs is inaccurate or misleading, especially in emotionally charged or polarizing contexts [29]. A recent bibliometric study showed a sharp rise in research on how SNSs spread health misinformation during the COVID-19 pandemic, underlining the need for active monitoring and strategies to reduce harm [30]. Content analyses of platforms such as YouTube have also revealed large amounts of misleading health information, such as false claims about vitamin D as a treatment for COVID-19, highlighting the difficulties in ensuring reliable health communication online [31]. This dual potential underscores the importance of developing communication strategies that not only ensure the visibility of accurate and accessible information but also address the risks associated with misinformation, particularly in the context of AMR, where public understanding and behavior are crucial. It has been shown that institutional awareness events, such as the World Antimicrobial Resistance Awareness Week (WAAW)—the major global initiative dedicated to raising awareness and promoting action on AMR—can rapidly increase the number of posts, including among laypeople, although this increase tends to be short-lived, typically returning to baseline levels within 48 hours [25]. Another study focusing on Twitter looked into the types of influential users and showed that the discussion was primarily influenced by news sources, health professionals, and governmental health organizations [24]. A study analyzing the information shared by Instagram (Meta Platforms) users of oral and topical antibiotics for treating acne vulgaris claimed the potential use of the SNS in elucidating patients' behavior and attitudes [32].

Among SNSs, Instagram is a platform in steady growth [33], among the most popular in size (fourth SNS globally, with 1478 million active users in 2022) [34] and impact on users' decision-making [35]. Content-wise, Instagram is a favorite SNS for social purposes [36]; therefore, it holds the potential for spreading knowledge and raising awareness of AMR, or on the contrary, to contribute to misinformation. Furthermore, considering that Instagram is an image-based platform of the first choice for entertainment purposes and co-creating with brands via social media [36], where topics such as food, wellness, and travel are amongst the most typical, the influence on health and lifestyle behavior of relevance for AMR is also worth attention.

To explore how Instagram content may shape awareness and behaviors related to AMR, it is useful to apply a theoretical model that captures the dynamics of message dissemination and reception.

Understanding Berlo's Sender-Message-Channel-Receiver Communication Model

The sender-message-channel-receiver (SMCR) communication model, developed by Berlo in 1960 [37,38], outlines 4 essential components in the communication process: the sender, the message, the channel, and the receiver. This model has been extensively used in communication research, including as a framework for understanding how SNS users disseminate and receive information [39-41].

The sender refers to the originator of the communication, whose characteristics, such as communication skills, knowledge, and attitudes, influence the effectiveness of the message delivered. In SNS contexts, the sender is any individual or organization posting content, with their social status or credibility impacting how their message is perceived.

The message itself consists of the information being transmitted. It can take various forms, including text, images, or videos, depending on the channel used.

The channel is the medium through which the message is transmitted, such as visual or auditory platforms. On Instagram, which serves as the primary communication channel in this context, users interact primarily through visual and textual content. In addition, online communities can also be considered vehicles or channels for displaying information. Typically, hashtags are not only part of the message but have a broad spectrum of functionality, including being a powerful tool for marketing promotion of products and services, acting as an evaluation marker that can set the interpretation model of the message, and activating networks of associations [42]. Finally, the receiver is the target audience of the message, whose characteristics—such as attitudes, social background, and previous knowledge—play a crucial role in determining how the message is interpreted.

XSL•FO
RenderX

Studying AMR Communication on Instagram

This study aims to examine the communication about AMR on Instagram, identifying key actors, content themes, and the nature of the communication to understand how AMR is portrayed and perceived using Berlo's SMCR communication model. Thus, the study contributes to addressing a broader challenge: while AMR is recognized as a pressing global health threat, little is known about how it circulates on SNSs like Instagram or how effectively these messages engage the general public. By analyzing how messages are framed, who produces them, and which publics they appear to reach, this study offers insights that can inform the development of more inclusive and effective digital communication strategies in the AMR field.

Methods

Study Design

This research uses a descriptive design and is grounded in inductive reasoning. The method of data collection and analysis is content analysis, which, in this study, integrates both quantitative and qualitative approaches [43,44]. The study incorporates quantitative elements but is primarily qualitative, guiding the choice to pursue trustworthiness as a measure of quality [45,46]. To foster transferability, a detailed description of the Methods and Results is offered. The study was planned following Bengtsson's recommendations to enhance the credibility of content analysis [47]. In the planning phase, the scope, relevance, and breadth of the aim, the size and characteristics of the sample and unit of analysis, the choice of data collection method, the choice of analysis method, and the ethical implications were thoroughly considered.

Data Collection

The data refers to the 24 months between December 1, 2020, and November 30, 2022. The timeframe was selected to include twice the WAAW, taking place in November. The feed posts (reels and stories were not considered) were collected retrospectively using an automated scripting tool for web scraping [48]. It was decided to use "#antibioticresistance" in the search, a widely used and recommended hashtag to engage in conversations about AMR [49], because it was deemed as a simpler but still scientifically proper expression of the notion, to which the public is relatively more used with respect to alternatives such as "antimicrobial resistance" or "AMR" [50].

The search returned 7380 hits. After cleaning the data, there were 6105 posts. Posts were excluded according to the following criteria: when non-English languages dominated the posts, duplicates, and reposts. To ensure the representativeness of the sample, hence enhancing the credibility and transferability of the study, a 10% sample ratio was applied (N=610). To prevent over-reporting content generated in concomitance with special AMR-related events, stratified random sampling was applied. For each month, 10% of the posts were randomly selected. The sample size was determined by the presumption that the data were sufficient based on the pretesting training sessions and pilot-testing results (78 posts) and consistency with previous sample sizes of Instagram studies [51,52]. This was later confirmed during the analysis by confidence in having reached, or closely approached, data saturation [53].

Data Analysis

Since there was no codebook of AMR in general on SNSs, it was created based on previous analyses of AMR in SNSs and a scoping review of AMR communication. The creation of the codebook was based on the SMCR model, and the coding process was adapted from Cohen et al [52]. To enhance credibility and dependability, investigator triangulation measures were implemented. A stepwise approach alternating individual and team sessions was used to enhance intercoder consistency and reliability [54]. As a pretest, 4 coders independently analyzed 15 different posts each in training sessions, followed by an iterative process of consensus coding and updating the original codebook. In a successive pilot coding round, the 4 coders analyzed together a random sample of 18 posts in order to confirm consistency in categorization. The results of the pilot test were included in the final sample. Afterwards, each coder independently analyzed one-quarter of the posts. Each coder's analysis was double-checked by another coder with full visibility of the original coding, and discrepancies were discussed and harmonized.

Content analysis was used in the study [55,56]. The unit of analysis comprised the caption, image, hashtags, and the content creator—their profiles were observed to categorize them according to their declared interests and visibility. Throughout the analysis, coders went back and forth between the post and the categories to enhance intracoder reliability and minimize the risk of misinterpretation [54]. During the coding process, posts were evaluated on the language used in captions, the emotional connotations of the images, and the context provided by hashtags. To assess the tone of the Instagram posts related to AMR, each post was categorized based on predefined sentiment categories: neutral, emotive, or promotional. For instance, posts categorized as emotive often included language that invoked fear or urgency about AMR, while neutral posts presented factual information without emotional framing. In the codebook, the application of the SMCR model generated 5 overarching categories, which are described together with the subcategories in Textbox 1. The categories in Textbox 1 were established through a rigorous content analysis of Instagram posts containing the hashtag "#antibioticresistance." Each category reflects a distinct type of content creator, based on their users' descriptions, bios, and linked websites. The professional background, communication style, and intended audience contributed to the categorization also. By categorizing content creators and their communication styles, this taxonomy provides valuable insights into the landscape of AMR discourse on Instagram, revealing both the strengths and limitations of current engagement strategies.

XSL•FO
RenderX

Textbox 1. Description of the categories and subcategories.

Sender: Content creator

Based on users' descriptions, bios, and linked websites. The content of the post contributes to the categorization, too. The number of followers and "likes" per post was noted.

- Ecology and animal care: Users who connected antibiotic resistance (AR) to animal rights and animal care.
- Science and health: Users working in biomedicine, biotechnology, and medicine who connected AR to science and health.
- Information driver: Users who focus on spreading knowledge and raising awareness of AR.
- Naturopathist: Users who promote a holistic approach to wellness, often focused on lifestyle and diet.
- Pharmacy and veterinary: Users who connected AR to the pharmacy and veterinary fields.
- Education: Education institutions and students.
- Motivators and art: Wellness motivators and illustrators posting on AR.

Receiver: Audience

Inferred from the post content (image, caption, and hashtags) and considering the presumable interest of the content creator, that is, when the post creator caters or not specifically to their closest audience.

- General audience: The content does not appeal to any specific audience.
- Customers: The post, implicitly or explicitly, aims to sell products or services.
- Followers: The content is directed to the creator's Instagram followers.
- Peers: The post is primarily directed to the creator's peers, for example, vegan community and health care professionals.
- Patients: The post seeks to communicate with patients.
- Students: The post seeks to communicate with students.
- Local community: The post is directed to a local community.
- Health organizations: The post is directed toward health care organizations and policymakers.

Message: Purpose

- The manifest primary reason for which the post is made.
- Information and awareness: Spreading knowledge and raising awareness of AR.
- Advertising: Selling products or services.
- Propaganda: Promoting ideological or political points of view through biased or partial communication.
- Infotainment: Sharing humorous, relatable content, often through comics or illustrations to inform about AR.

Message: Tone

The gist or attitude of the post, including the image, caption, and hashtags.

- Neutral: A communication style tendentially tends to be information- and fact-based.
- Emotive: A communication style involving emotional content or potentially triggering an emotional response
- Promotional: A style involving marketing-style communication strategies

Message and channel: Centrality of AMR

An appraisal of AR importance within the post (image, caption, and hashtags).

- Main: AR is the most important subject in the post.
- Relevant: AR is an important subject in the post.
- Collateral: AR is a subject of secondary importance in the post.
- Irrelevant: The post is about something else, unrelated to AR (AR only appears in the hashtags).

In addition, other descriptive characteristics of the posts were noted: (1) the number of followers of the creator of the posts (sender); (2) the image content and type (message): text (the visual section of the post, whether static or a carousel, is dominated by words), visual object (the visual section of the post, whether static or a carousel, is dominated by a visual object, eg, images, pictures, photos, graphics, and illustrations),

XSL•FO
RenderX

video, or audio; (3) the hashtags (channel); (4) the number of "likes" obtained by the posts (receiver).

Ethical Considerations

This study analyzed publicly accessible Instagram posts tagged with "#antibioticresistance" using a custom web scraping script that accessed content viewable without login, mimicking human browsing behavior. The tool did not access private or password-protected data. While platform terms of service do not always clearly address noncommercial academic research, the study acknowledged the complex and evolving legal and ethical landscape of automated data collection [57]. No personally identifiable information was collected, stored, or quoted. Usernames, profile information, and other potentially identifying content were excluded to minimize the risk of reidentification. The posts included a mix of location-specific and nonlocation-specific content, and we did not systematically categorize them based on geographic references. The analysis focused on aggregated patterns in publicly shared communication, not on individual users.

Although ethical review was not required for the use of public data, the study followed established ethical guidance for social media research, including transparency in reporting, minimization of harm, and attention to user expectations of privacy [58,59]. The overall aim was to contribute to the understanding of public health communication, not to evaluate or profile individuals.

Results

The results show that the sender categories could be divided into 7 main content creator types, with the biggest group being the information driver (n=186), followed by the professionals (n=147). Table 1 details the content creator types and general characteristics of their posts. The most common hashtags used are "bacteria," "AMR," "antimicrobialresistance," "antibiotic," "medicine," and "health." Typically, images contained text messages or pictures and illustrations—often of bacteria and test samples. The captions often described how antibiotic resistance (AR) and antimicrobial resistance (AMR) spread, why it needs to be stopped, and connected health consequences.

XSL•FO
RenderX

Table . Content creator types and general characteristics of their posts.

Category	Posts, n (%)	Top hashtags[a] (n)	Top visual content[b]	Likes		Followers	
				Total[c], n	Average[d]	Total[c], n	Average[d]
Ecology and animal care	56 (9.2)	Vegan (13), plantbased (10), antibiotics (9), deforestation (7), factoryfarming (6), animalagriculture (5), eatmoreplants (5), animalagricultureisdestroyingtheplanet (4), biodiversityloss (4), eatplants (4), straightfromthefarm (4)	Food, meat, becoming vegan, and animals (cow and fish birds).	1566	28	329,578	5885
Science and health	147 (24.1)	Bacteria (27), AMR[e] (25), microbiology (23), science (18), antibiotic (17), antimicrobialresistance (16), biology (13), medicine (12). Healthcare (11), lab (11), research (10)	Pills, bacteria, and lab tests.	25,440	173	3,315,184	22,552
Information driver	185 (30.3)	amr (92), antimicrobialresistance (79), superbugs (72), health (69), antibiotic (59), bacteria (59), medicine (57), stopsuperbugs (45), keepantibioticsworking (41), microbiology (37), antibioticstewardship (35), onehealth (35), science (35), healthcare (33), pharmacy(33), beantibioticsaware (30)	Pills, AMR facts, bacteria, doctor, and patient.	28,908	156	1,127,311	6094
Naturopathist	66 (10.8)	Antibiotics (29), health (12) antimicrobialresistance (11) bacteria (10), infection (8), beantibioticsaware (7), medicine (7), healthcare (5), healthy (5) wellness (5) microbiology (5) antibioticawareness (5)	Information, pills, hands, and animals.	3131	47	1,349,163	20,442

XSL•FO
RenderX

Category	Posts, n (%)	Top hashtags[a] (n)	Top visual content[b]	Likes		Followers	
				Total[c], n	Average[d]	Total[c], n	Average[d]
Pharmacy and veterinary	64 (10.5)	Antibiotics (24), antimicrobialresistance (15) pharmacy (14), antibiotic (13), antibioticstewardship (11), pharmacist (11) amr (10), bacteria (10), medical (7), doctor (7), infectiousdisease (7), resistance (7)	Dissertation, pills, bacteria, and hands.	1784	28	263,038	4110
Education	44 (7.2)	Antibiotics (10), bacteria (10), medicine (9), microbiology (8), amr (7), antibiotic (7) Science (7), antimicrobialresistance (6), publichealth (5), biology (4), pharmacy (4)	Research highlights, infographics research findings, and laboratory tests.	1269	29	326,894	7429
Motivators and art	48 (7.9)	Antibiotics (15), microbiology (11), bacteria (8), antibiotic (5), research (5) science (5), sciencecommunication (5) abstractart (4), art (4) artoftheday (4), podcast (4), ceramics (3)	Artwork, bacteria, selfies, and promoting hand hygiene products.	2970	62	45,771	954

[a]The top hashtags refer to the most frequent hashtags used per content creator category.

[b]The top image content refers to the most frequent pictures and illustrations used per content creator category.

[c]The total number of “likes” and “Followers”refers to the content creator categories, not to individual Instagram users.

[d]The average number of “likes” and “Followers”refers to the content creator categories, not to individual Instagram users.

[e]AMR: antimicrobial resistance.

In the following, each sender category (eg, “Ecology and animal care”) and their different sender subgroups (eg, “Vegans and vegetarians”) are described. For each sender category, the most relevant channel feature, namely the hashtags, and message features, such as characteristic visual content, are described. Accounting for the receiver, the number of “likes” is reported. For each sender subgroup, the emphasized message features include purpose, image content, tone, and centrality of AR. At the same time, as for the receiver analysis, the type of audience for which the post was made is shown.

Ecology and Animal Care

These users connected AR to environmental and animal rights issues. The average number of “likes” per post (n=28) is relatively low in this sender category (see Table 1). Their top hashtags were “vegan,” “plantbased,” “antibiotics,” “deforestation,” and “factoryfarming.” Pictures often showed animals such as cows, fish, and birds and food products. The category was divided into 3 subgroups: “Vegans and vegetarians,” “Farming and food companies,” and “Others in food and diet.”

“Vegans and vegetarians” informed and promoted diet change and animal welfare to a great extent. The tone was emotive for most posts (56%, 18/32 participants) and, sometimes, they slipped into propaganda. Almost half of the posts (44%, 14/32 participants) have AR as a centrality in the post. Animal welfare was also crucial to the next subgroup—“Farming and food companies.” They encouraged eating organic and focused on showing good animal care and food production. Relatively often (14%, 3/21 participants) their posts served commercial purposes and were directed to their customers. Foremost, 10 out of 21

XSL•FO
RenderX

participants (48%) adopted a neutral tone in their posts, where AR was relevant but not the main focus of the post. These 2 subgroups are those with the highest number of posts where the centrality of AR was deemed only collateral (24%, 5/21 participants). In the "Others in food and diet" subgroup, organizations and individuals that share a focus on food and diet without falling into any other subgroups were grouped. Further details are shown in Table 2.

Table . Ecology and animal care subgroups' characteristics.

Subgroup	Purpose, n (%)				Audience, n (%)			Image content, n (%)			Tone, n (%)			Centrality of AR[a], n (%)			
	I&A[b]	Propaganda	Advertising	Infotainment	General	Peers	Customers	Text	Visual object	Video or audio	Emotive	Neutral	Promotional	Main	Relevant	Irrelevant	Collateral
Vegans and vegetarians (n=32)	25 (78)	4 (12.5)	3 (9.4)	0 (0)	28 (87.5)	4 (12.5)	0 (0)	18 (56.3)	11 (34.3)	3 (9.4)	18 (56)	10 (31)	4 (13)	14 (44)	7 (22)	6 (18)	5 (16)
Farming and food companies (n=21)	9 (43)	1 (4.8)	10 (47.6)	1 (4.8)	13 (61.9)	0 (0)	8 (38)	10 (47.6)	10 (47.6)	1 (4.8)	8 (38)	10 (48)	3 (14)	4 (19)	10 (48)	2 (10)	5 (24)
Others in food and diet (n=3)	2 (67)	0 (0)	1 (33.3)	0 (0)	3 (100)	0 (0)	0 (0)	2 (66.7)	1 (33.3)	0 (0)	1 (33)	1 (33)	1 (33)	1 (33)	2 (67)	0 (0)	0 (0)

[a]AR: antibiotic resistance.

[b]I&A: information and awareness.

Science and Health

These sender categories showed their work within the AR field, either in research or care. Table 1 shows that the average number of "likes" per post (n=173) and users' followers (n=22,552) was the highest among all main categories. The category is divided into 4 subgroups: "Biomedicine labs," "Healthcare staff," "Researchers," and "Healthcare facilities." Their top hashtags were "bacteria," "AMR," "microbiology," "science," and "antibiotic." Pictures typically showed pills, laboratory tests, bacteria, and health care personnel.

"Biomedicine labs" encompasses collective profiles of laboratories and individuals creating laboratory-life content, thus focusing on scientific-specific aspects, such as AR mechanisms. This subgroup engages the most with its peers. The "Healthcare staff" subgroup consists of health care professionals. Their posts focused on proper antibiotic use and the threat posed by AR. Their audience was the public and, occasionally, a specific class of patients and their peers in health care. The "Researchers" subgroup comprises users who do research in academic or private settings. For the most part, they work in biomedicine and biotechnology. However, with respect to the "Biomedicine labs" subgroup, they focused more on individual work and achievements and, overall, gave a more personalized perspective to their posts and tried to reach a broader audience. The last subgroup, "Healthcare facilities," involves hospitals, clinics, ambulatories, etc, often providing private services. Their user profiles are mainly collective or of individual users communicating on behalf of or about the health care facility. Their communication was similar to that of the "Healthcare staff" subgroup but tended to engage local communities and promote their services more. Further details are shown in Table 3.

XSL•FO
RenderX

Table . Science and health professionals subgroups' characteristics.

Subgroups	Purpose, n (%)			Audience, n (%)							Image content, n (%)			Tone, n (%)			Centrality of AR[a], n (%)			
	I&A[b]	Advertising	Propaganda	General	Peers	Health organization	Patients	Followers	Community	Students	Text	Visual object	Video or audio	Neutral	Emotive	Promotion	Main	Relevant	Collateral	Irrelevant
Biomedical laboratories (n=37)	33 (89.2)	4 (10.8)	0 (0)	20 (54)	16 (43.2)	1 (2.7)	0 (0)	0 (0)	0 (0)	0 (0)	26 (70)	11 (29.7)	0 (0)	28 (76)	6 (16)	3 (8.1)	16 (43)	15 (41)	3 (8)	3 (8)
Health care staff (n=36)	31 (86.1)	3 (8.3)	2 (5.6)	28 (77.8)	5 (13.9)	0 (0)	3 (8.3)	0 (0)	0 (0)	0 (0)	17 (47.2)	18 (50)	1 (2.8)	26 (72)	5 (14)	5 (14)	25 (70)	5 (14)	4 (11)	2 (6)
Researches (n=49)	41 (83.7)	4 (8.2)	4 (8.2)	40 (8.2)	8 (16.3)	0 (0)	0 (0)	1 (2)	0 (0)	0 (0)	19 (38.8)	26 (53)	4 (8.2)	37 (76)	9 (18)	3 (6)	31 (63)	15 (31)	2 (4)	1 (2)
Health care facilities (n=25)	20 (80)	5 (20)	0 (0)	15 (60)	0 (0)	0 (0)	3 (12)	0 (0)	5 (20)	2 (8)	4 (16)	18 (72)	3 (12)	15 (60)	6 (24)	4 (16)	16 (64)	5 (20)	3 (12)	1 (4)

[a]AR: antibiotic resistance.

[b]I&A: information andawareness.

Information Driver

This main sender category of Instagram users is information-driven, and their manifest purpose is to spread knowledge and awareness. Their communication is effective, sometimes even entertaining, as humor is not disdained, and their content is tendentially short and easily accessible. As seen in Table 1, their top hashtags include "AMR," "antimicrobialresistance," "superbugs," "health," and "antibiotic." Pictures depicted pills and health care personnel. They have a comparatively high engagement from their audience in terms of number of "likes" (156 on average per post). This category has 3 subgroups: "AMR organizations and communities," "Journalists and reporters," and "Companies."

"AMR organizations and communities" involves nongovernmental organizations, communities, networks, and others focused on AMR. They mostly share information regarding AMR and AMR-related events (eg, the World AMR Awareness Week and conferences). This subgroup has the highest percentage of AMR-centered posts (84%, 66/79 participants), yet 14 out of 32 posts had a neutral tone (70%). The "Journalists and reporters" subgroup encompasses science news-based profiles, a few journals and newspapers or magazines, but also individual journalists and reporters. Usually, they reported recent scientific research results for the general audience but also for students and their followers and promoted AMR awareness and events. This subgroup tended more than others to use text in the image space of the posts (55 out of 77 participants). The last subgroup is "Companies." They primarily used the platform to promote their products (eg, tests, drugs, and supplements) but delivered their communication in terms of AMR information and awareness spreading. They often targeted potential customers and health organizations. This subgroup tended more than others to use pictures in the image space of the posts (25 out of 30 posts). Further details are shown in Table 4.

XSL•FO
RenderX

Table . Information driver subgroups’ characteristics.

Subgroups	Purpose, n (%)				Audience, n (%)							Image content, n (%)			Tone, n (%)			Centrality of AR[a], n (%)			
	I&A[b]	Advertising	Infotainment	Propaganda	General	Community	Patients	Students	Health organization	Followers	Customers	Visual object	Text	Audio or video	Neutral	Emotive	Promotion	Main	Relevant	Collateral	Irrelevant
AMR[c] organizations and communities (n=79)	68 (86)	7 (8.8)	3 (3.8)	1 (1.3)	67 (84.8)	5 (6.3)	4 (5.1)	2 (2.5)	1 (1.3)	0 (0)	0 (0)	41 (51.9)	38 (48.1)	0 (0)	55 (70)	21 (27)	3 (4)	66 (84)	9 (12)	2 (3)	2 (3)
Journalists and reporters (n=77)	71 (92.2)	4 (5.2)	2 (2.6)	0 (0)	55 (71.4)	0 (0)	0 (0)	9 (11.7)	2 (2.6)	8 (10.4)	3 (3.9)	21 (27.3)	55 (71.4)	1 (1.3)	48 (62)	21 (27)	8 (10)	48 (62)	14 (18)	10 (13)	5 (7)
Companies (n=30)	18 (60)	9 (30)	1 (3.3)	2 (6.7)	17 (56.7)	0 (0)	0 (0)	0 (0)	7 (23.3)	0 (0)	6 (20)	25 (83.3)	5 (16.7)	0 (0)	23 (76)	4 (14)	3 (10)	19 (62)	7 (24)	1 (3)	3 (10)

[a]AR: antibiotic resistance.

[b]I&A: information and awareness.

[c]AMR: antimicrobial resistance.

Naturopathist

These users promote a holistic approach to wellness, often focused on lifestyle and diet. An active lifestyle, keeping a balanced gut microbiota through wise food choices, and avoiding unnecessary antibiotic use were hot topics. The most common hashtags were “antibiotics,” “health,” “antimicrobialresistance,” “bacteria,” and “infection.” Pictures were often associated with food or health. This sender category includes the subgroups: “Natural healers,” “Body and nutrition,” and “Environment and health” subgroups.

Specific to the “Natural healers” was the latent message that the individual users were personally engaged in AMR. Relatively often, the posts were used to promote their products (eg, probiotics and natural antimicrobials) or services (eg, mindfulness and homeopathy consultancies). The second subgroup, “Body and nutrition,” included fitness enthusiasts who share information and tips about healthy eating. The “Environment and healthcare” subgroup involves individuals focusing on the environment, health, and sustainability. The latter two subgroups’ tone was coded as the most emotive with 3 out of 5 and 2 out of 4 posts being emotive. The last group’s communication was even characterized by anger. Further details are shown in Table 5.

XSL•FO
RenderX

Table . Naturopathist subgroups' characteristics.

Subgroups	Purpose, n (%)				Audience, n (%)						Image content, n (%)			Tone, n (%)			Centrality of AR[a], n (%)			
	I&A[b]	Advertising	Propaganda	Infotainment	General	Customers	Followers	Health organization	Patients	Community	Text	Visual object	Video or audio	Neutral	Emotive	Promotion	Main	Relevant	Collateral	Irrelevant
Natural healers (n=57)	43 (75.4)	10 (17.5)	3 (5.3)	1 (1.8)	43 (75.4)	9 (15.8)	3 (5.3)	1 (1.8)	1 (1.8)	0 (0)	29 (50.9)	27 (47.4)	1 (1.8)	41 (72)	11 (19)	5 (9)	29 (51)	16 (28)	7 (12)	5 (9)
Body and nutrition (n=5)	4 (80)	1 (20)	0 (0)	0 (0)	4 (80)	0 (0)	0 (0)	0 (0)	0 (0)	1 (20)	2 (40)	3 (60)	0 (0)	2 (4)	3 (60)	0 (0)	4 (80)	0 (0)	1 (20)	0 (0)
Environment and health (n=4)	3 (75)	1 (25)	0 (0)	0 (0)	3 (75)	1 (25)	0 (0)	0 (0)	0 (0)	0 (0)	4 (100)	0 (0)	0 (0)	1 (25)	2 (50)	1 (25)	2 (50)	0 (0)	2 (50)	0 (0)

[a]AR: antibiotic resistance.

[b]I&A: information and awareness.

Pharmacy and Veterinary

Users in this sender category focus on spreading knowledge about correct antibiotic use and AR and promoting their products and services for human and animal health. More than other categories, they include "One Health" in their posts. Their top hashtags involve "antibiotics," "antimicrobialresistance," "pharmacy," "antibiotic," and "antibioticstewardship," and the pictures primarily concern pills, bacteria, and hands (see Table 1).

The subgroups are "Pharmacists and vets"—pharmacy and veterinary medicine professionals—and "Pharmacy and veterinary companies." The latter subgroup has the highest promotional tone with 4 out of 15 of the posts (27%). Further details are shown in Table 6.

Table . Pharmacy and veterinary subgroups' characteristics.

Subgroups	Purpose			Audience				Image content			Tone			Centrality of AR[a]			
	I&A[b]	Infotainment	Advertising	General	Peers	Customers	Health organization	Text	Visual object	Video or audio	Neutral	Emotive	Promotion	Main	Relevant	Irrelevant	Collateral
Pharmacists and vets (n=49)	43 (87.8)	5 (10.2)	1 (2)	38 (77.6)	6 (12.2)	4 (8.2)	1 (2)	30 (61.2)	18 (36.7)	1 (2)	32 (65)	14 (29)	3 (6)	39 (80)	6 (12)	3 (6)	1 (2)
Pharmacy and veterinary companies (n=15)	9 (60)	0 (0)	6 (40)	9 (60)	0 (0)	3 (20)	3 (20)	8 (53.3)	7 (46.7)	0 (0)	9 (60)	2 (13)	4 (27)	8 (53)	4 (27)	2 (13)	1 (7)

[a]AR: antibiotic resistance.

[b]I&A: information and awareness.

XSL•FO
RenderX

Education

This sender category top hashtags were “antibiotics,” “bacteria,” “medicine,” “microbiology,” and “AMR.” Their top images comprehended research highlights, infographics, research findings, and laboratory tests. The category encompasses “Students” and “Education institutes.” The subgroup “Students” includes upper secondary education students, undergraduates, and graduate students (PhD students were placed in the Science and Health category).

Users in the “Students” subgroup shared what they learned about AR and promoted proper antibiotic use. Also, users in “Education institutes” created posts to spread awareness, besides promoting their courses and activities and highlighting research findings. It is the category whose posts were coded as using the most neutral tone (91%, 21/23 posts). Further details are shown in Table 7.

Table . Education subgroups’ characteristics.

Subgroups	Purpose			Audience			Image content		Centrality of AR[a]			
	I&A[b]	Infotainment	Advertising	General	Peers	Students	Visual object	Text	Main	Irrelevant	Relevant	Collateral
Students (n=21)	19 (90.5)	2 (9.5)	0 (0)	17 (80.9)	4 (19)	0 (0)	16 (76.2)	5 (23.8)	15 (71)	3 (14)	2 (10)	1 (5)
Educational institutes (n=23)	18 (78.3)	0 (0)	5 (21.7)	18 (78.3)	0 (0)	5 (21.7)	14 (60.9)	9 (39.1)	15 (65)	0 (0)	7 (30)	1 (4)

[a]AR: antibiotic resistance.

[b]I&A: information and awareness.

Motivators and Art

This sender category’s top hashtags were “antibiotics,” “microbiology,” “bacteria,” “antibiotic,” and “research.” Their top pictures included artwork, bacteria, and selfies (see Table 1). The category involves users who are less focused on AMR. Its subgroups are “Motivators” and “Art.” The former involves users who try to influence and provide feed content of interest, including AR. The latter involves illustrators and art communities that have created content for AR.

The “Motivators” subgroup often resorted to personal narratives, connected AR to life experiences, and promoted healthy habits. The “Art” subgroup used relatively more pictures in the image space of the posts, and 4 out of 13 (31%) used an emotive (sad or scary) tone. Further details are shown in Table 8.

Table . Motivators and art subgroups’ characteristics.

Subgroups	Purpose				Audience		Image content			Tone			Centrality of AR[a]			
	I&A[b]	Infotainment	Advertising	Propaganda	General	Followers	Visual object	Text	Video or audio	Neutral	Emotive	Promotion	Main	Relevant	Irrelevant	Collateral
Motivators (n=35)	30 (85.7)	3 (8.6)	2 (5.7)	0 (0)	31 (88.6)	4 (11.4)	21 (60)	11 (31.4)	3 (8.6)	26 (74)	8 (23)	1 (3)	19 (54)	9 (26)	4 (11)	3 (9)
Art (n=13)	10 (77)	1 (7.7)	1 (7.7)	1 (7.7)	12 (92.3)	1 (7.7)	9 (69.2)	4 (30.8)	0 (0)	9 (69)	4 (31)	0 (0)	5 (38)	6 (46)	1 (8)	1 (8)

[a]AR: antibiotic resistance.

[b]I&A: information and awareness.

Discussion

Principal Findings

The main findings of this study reveal that while Instagram serves as a platform for disseminating accurate and informative content about AMR, the discussions are largely confined to niche groups, limiting broader audience engagement. The findings showed a heterogeneous and siloed user landscape, where top-down communication is delivered by those who are knowledgeable or have a linked purpose, often of a commercial nature with minimal interaction beyond specific interest communities, suggesting a need for more inclusive and visually compelling content to enhance public awareness and engagement on this critical health issue. These results confirm previous studies [24-26].

Senders and Receivers

Although users created posts that were accessible to a general audience, the broader public rarely engaged with the content or showed visible interest. Virtually no content was created by someone not already engaged in the AMR field or not promoting specific viewpoints, services, or goods. The fragmented and siloed landscape of AMR on Instagram is also confirmed in the

XSL•FO
RenderX

analysis of user engagement patterns, which reveals that "likes" of posts and interactions occur predominantly within specific user groups, with each subgroup gravitating towards content tailored to their interests and priorities. This is shown in the images, the way posts are formulated and that most "likes" derive from other users within the same main category. For instance, a post by a "Science and health" user related to AMR and infection prevention would gain higher engagement rates among users in the same category and individuals with a vested interest in public health issues. In a nutshell, AMR is of no interest to the Average Joe of Instagram. Only a few posts had health organizations among their recipients and very few health authorities and policymakers, highlighting the platform's limitations as a forum for broader policy discussion. The challenge of informational homophily and its siloing effect for which science and health content reach almost exclusively already engaged audiences has been long known, and it was also detected in this study, indicating challenges in overcoming audience segmentation on SNSs [60,61]. It is also possible that algorithmic filtering limits the visibility of AMR content to users outside niche communities. Since content exposure is shaped by previous behavior and engagement patterns, users who are not already interacting with health-related content may be less likely to encounter posts about AMR, even if they might find the topic relevant or engaging [62].

The major actors in the AMR discussion on Instagram are the "Science and health" and "Information driver" content creators. These 2 groups are highly engaged in the AMR field, either possessing extensive professional expertise or being well-informed on the subject. Both groups want to raise awareness, but they do it either by practicing stewardship AMR (Science and health) or mobilizing individual and societal actions (Information drivers). They, together with "Pharmacy and veterinary" and "Education" users, have a specific responsibility deriving from being (and being perceived as) specialists and experts to communicate about AMR in a truthful, nonbiased, and educated manner, as whenever they engage in health and science communication through SNSs, there is the potential for impacting public health and individual behavior and public trust in general [63,64]. These senders are perceived as experts, and their social status as professionals or well-informed individuals lends authority to their messages. This aligns with the SMCR model, where the sender's characteristics—like their knowledge and credibility—determine the effectiveness of communication.

Messages

The purpose of 3% of the posts (18/611) was coded as propaganda, meaning that the content promoted ideological or political points of view through biased or partial communication. For the majority, this type of communication was adopted by users who stressed their concern for animal health and the environmental consequences of AMR. Therefore, AMR and antibiotic use are not ideologically or politically invested in themselves, at least not yet. Such communication happened in conjunction with animal and environmental themes, which tend instead to be polarizing, especially in social media, which are designed to monetize on disagreement among users [65].

Overall, the content of the posts was accurate and conveyed in a neutral tone. Narratives and images aimed at triggering an emotional response were also used but without compromising the veracity of the content, thus to be considered, for the most part, as a way to attract an audience and motivate behavior change through the mediating role of self-relevant emotions (primarily fear) [66]. Noteworthily, almost no content highlighted the role and responsibilities of institutions about AMR.

Posts were text-dominated to an extent above average for Instagram users. This was predictable, considering that they aimed at spreading information and raising awareness about AMR. A 2020 study on the Centers for Disease Control and Prevention's Instagram posts showed similar results about the role of text in the visual section of the analyzed posts [67]. A recent research study by Charani et al [68] on the message content produced by key actors in global health about the visual depiction of AMR found that the current narrative is one of power imbalances, where women and children from low-income and middle-income countries are presented with less dignity, respect, and power than those from high-income countries. On a positive note, none of this was detected in this study. Therefore, the problem of degradingly representing AMR stands with international health bodies but is not shared by other Instagram content creators, not even in the posts whose purpose was deemed as propaganda or that adopted an emotive tone. Charani et al [68] has also shown that imagery in global health communication plays a crucial role not just in transmitting facts but in shaping public understanding, emotional responses, and awareness of infectious diseases, including AMR. Their analysis of visual practices across global health documents highlights how images, when ethically and contextually used, can embed health issues in the public imagination and evoke empathy [68]. In our study, many Instagram posts featured images such as stylized bacteria, pills, or laboratory scenes. However, these visuals were often used in an illustrative or aesthetic way, with limited contextual or emotional framing. Given Instagram's visual-first design, future AMR communication on the platform could benefit from adopting more purposeful and ethically grounded imagery to broaden public engagement and improve risk communication.

Channels

In this study, hashtags are considered as channels, serving multiple functions beyond grouping content. Hashtags, in this context, were also a vehicle for marketing, activism, and community engagement, allowing senders to align their messages with specific networks of associations and broadening the scope and impact of their communication. Different sender categories used hashtags strategically based on their intended message and audience. For example, users in the "Ecology and animal care" category used hashtags like "vegan," "plantbased," and "deforestation" to align AR with environmental and animal rights issues. Here, hashtags extended the message beyond health care, invoking broader themes that resonate with specific "green" values and networks. Similarly, users in the "Science and health" category frequently used hashtags such as "bacteria," "AMR," and "microbiology," which are directly related to scientific work and communication. These hashtags helped cater

XSL•FO
RenderX

scientific content to targeted peers within the scientific community, as well as an informed public. Hashtags also served a commercial function, particularly in the “Pharmacy and veterinary” and “Information driver” categories. Hashtags like “pharmacy,” “antibioticstewardship,” and “health” not only connected the posts to public health but were also marketing tools for promoting products and services, often with a neutral tone to maintain credibility.

Rauschnabel et al [69] identified 10 different motivations for using hashtags (in order of frequency): amusing, organizing, designing, conforming, trendgaging, bonding, inspiring, reaching, summarizing, and endorsing. This research mainly showed organizing and reaching use, namely, to structure and organize the content of posting and to meet the conventions of specific groups of interest, respectively. To a minor extent, also conforming to use, that is, showing the desire to meet the conventions of specific groups of interest was observed.

Limitations

The study findings should be interpreted considering several methodological choices. A key strength lies in the use of stratified random sampling based on time, which ensured a balanced representation of content across the 24-month period while maintaining analytic feasibility. A 10% sample was drawn (N=610), consistent with established practices in media and communication research where full-population analysis is often impractical [70]. While this enabled the identification of dominant patterns and themes, rare or emerging discourses may have been underrepresented. Temporal stratification helped mitigate this by capturing variation in posting behavior over time [71]. To enhance credibility and dependability, coding was conducted iteratively, including pilot testing, coder training, and cross-checking. These strategies, along with attention to data saturation, supported the trustworthiness of findings. The study focused exclusively on Instagram and the hashtag “#antibioticresistance.” While this defined a clear and relevant dataset, AMR-related discussions may also occur under other hashtags or on other SNSs such as X or Facebook. Platform-specific user bases and engagement formats may influence how AMR is communicated and perceived. Future studies should broaden the scope to include additional platforms and terms. Automated accounts were not identified or excluded in this study. Since bots can inflate engagement metrics, future work should consider detection methods to ensure accurate interpretations of user interaction.

Finally, while the SMCR model provided a useful structure, it does not fully capture the participatory, multimodal nature of SNSs. Future work may benefit from combining it with frameworks better suited to networked communication.

Conclusions

The findings of this study provide an overview of how AMR is communicated on Instagram, particularly through the lens of the hashtag “#antibioticresistance.” It is evident that while Instagram serves as a platform for disseminating valuable information about AMR, the engagement with this content is largely confined to niche communities rather than reaching a broader audience. This conclusion synthesizes the key findings, discusses the implications for public health communication, and suggests pathways for enhancing engagement and awareness around AMR on social media platforms.

The content analysis of 611 Instagram posts revealed several critical insights into the landscape of AMR communication. First and foremost, the primary intention behind these posts was to inform the public about AMR, with 497 out of 611 of the posts (81.3%) categorized as information-driven. The predominant content creators were classified as “Information drivers” and “Science and health” professionals, who are engaged in raising awareness about AMR. Their posts were characterized by a neutral tone, with a significant amount of text-based content aimed at educating the audience about the implications of antibiotic resistance.

Despite the accuracy and neutrality of the content being commendable, the study highlighted a concerning trend of the isolated nature of the conversations of AMR within specific interest groups. The majority of engagement, as indicated by the number of likes and interactions, occurred within these niche communities, suggesting that the broader Instagram audience, or average individuals, remains largely disengaged from AMR discourse. This finding aligns with existing research that emphasizes the challenges of reaching diverse audiences in health communication, particularly in the context of social media.

The implications of these findings for public health communication are significant. AMR represents a pressing global health crisis, and effective communication strategies are essential for raising awareness and fostering behavioral change among the general population. Given the reliance on social media as a primary source of information for many individuals, particularly younger demographics, it is critical that AMR discussions are made more accessible and engaging. The findings of this study highlight the need for more inclusive, relatable, and visually engaging content that can resonate with a broader audience. To address the challenges identified in this study and promote communication that is actionable and impactful for diverse audiences, the following recommendations are proposed: (1) integrate AMR discussions into more general interest posts, such as stories related to daily life, popular culture, and trending topics in order to attract the attention of nonexperts; (2) use more visually engaging content that cuts through the noise; (3) increase institutional engagement; (4) promote inclusive narratives; (5) improve engagement strategies and campaigns that require user participation, such as challenges, question-and-answer sessions, and live discussions, to boost engagement and interaction or series of educational posts; and (6) highlight the individual and collective responsibility for AMR. Ultimately, addressing the challenges of siloing and promoting a more inclusive dialogue will be essential in mobilizing collective action against this critical global health issue.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Drug-resistant infections: a threat to our economic future. : World Bank Group; 2017 URL: https://documents1.worldbank.org/curated/en/323311493396993758/pdf/final-report.pdf [accessed 2025-07-31]
2. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 2022 Feb;399(10325):629-655. [doi: 10.1016/S0140-6736(21)02724-0]
3. Stålsby Lundborg C, Tamhankar AJ. Understanding and changing human behaviour--antibiotic mainstreaming as an approach to facilitate modification of provider and consumer behaviour. Ups J Med Sci 2014 May;119(2):125-133. [doi: 10.3109/03009734.2014.905664] [Medline: 24735112]
4. Kosiyaporn H, Chanvatik S, Issaramalai T, et al. Surveys of knowledge and awareness of antibiotic use and antimicrobial resistance in general population: a systematic review. PLoS ONE 2020;15(1):e0227973. [doi: 10.1371/journal.pone.0227973] [Medline: 31945117]
5. Ancillotti M, Eriksson S, Veldwijk J, Nihlén Fahlquist J, Andersson DI, Godskesen T. Public awareness and individual responsibility needed for judicious use of antibiotics: a qualitative study of public beliefs and perceptions. BMC Public Health 2018 Oct 3;18(1):1153. [doi: 10.1186/s12889-018-6047-8] [Medline: 30285689]
6. Holmes AH, Moore LSP, Sundsfjord A, et al. Understanding the mechanisms and drivers of antimicrobial resistance. The Lancet 2016 Jan;387(10014):176-187. [doi: 10.1016/S0140-6736(15)00473-0]
7. Costelloe C, Metcalfe C, Lovering A, Mant D, Hay AD. Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis. BMJ 2010 May 18;340:c2096. [doi: 10.1136/bmj.c2096] [Medline: 20483949]
8. Ancillotti M, Nilsson E, Nordvall AC, Oljans E. The status quo problem and the role of consumers against antimicrobial resistance. Front Sustain Food Syst 2022;6:6. [doi: 10.3389/fsufs.2022.834022]
9. Frost I, Van Boeckel TP, Pires J, Craig J, Laxminarayan R. Global geographic trends in antimicrobial resistance: the role of international travel. J Travel Med 2019 Dec 23;26(8):taz036. [doi: 10.1093/jtm/taz036] [Medline: 31115466]
10. Maillard JY, Bloomfield SF, Courvalin P, et al. Reducing antibiotic prescribing and addressing the global problem of antibiotic resistance by targeted hygiene in the home and everyday life settings: a position paper. Am J Infect Control 2020 Sep;48(9):1090-1099. [doi: 10.1016/j.ajic.2020.04.011] [Medline: 32311380]
11. Mickelsson M, Usai T, Chinofunga D, Oljans E. Health communication for AMR behaviour change: Zimbabwean students' relationships with the microbial world. JAC Antimicrob Resist 2023 Dec;5(6):dlad133. [doi: 10.1093/jacamr/dlad133] [Medline: 38089463]
12. Papadimou D, Malmqvist E, Ancillotti M. Socio-cultural determinants of antibiotic resistance: a qualitative study of Greeks' attitudes, perceptions and values. BMC Public Health 2022 Jul 28;22(1):1439. [doi: 10.1186/s12889-022-13855-w] [Medline: 35902816]
13. Global action plan on antimicrobial resistance. : World Health Organization; 2015 URL: https://www.who.int/publications/i/item/9789241509763 [accessed 2025-07-31]
14. Zhang J, Centola D. Social networks and health: new developments in diffusion, online and offline. Annu Rev Sociol 2019 Jul 30;45(1):91-109 [FREE Full text] [doi: 10.1146/annurev-soc-073117-041421]
15. boyd DM, Ellison NB. Social network sites: definition, history, and scholarship. J Comput Mediat Commun 2007 Oct;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
16. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520. [Medline: 25083128]
17. Zowawi HM, Abedalthagafi M, Mar FA, et al. The potential role of social media platforms in community awareness of antibiotic use in the Gulf cooperation council states: luxury or necessity? J Med Internet Res 2015 Oct 15;17(10):e233. [doi: 10.2196/jmir.3891] [Medline: 26471079]
18. Parveen S, Garzon-Orjuela N, Amin D, McHugh P, Vellinga A. Public health interventions to improve antimicrobial resistance awareness and behavioural change associated with antimicrobial use: a systematic review exploring the use of social media. Antibiotics (Basel) 2022 May 16;11(5):669. [doi: 10.3390/antibiotics11050669] [Medline: 35625313]
19. Wright R, Ferguson C, Bodrick M, Balkhy H, Jackson D, Davidson PM. Social media and drug resistance in nursing training: using a Twitterchat to develop an international community of practice for antimicrobial resistance. J Clin Nurs 2020 Jul;29(13-14):2723-2729. [doi: 10.1111/jocn.15185] [Medline: 31944460]
20. Thornber K, Huso D, Rahman MM, et al. Raising awareness of antimicrobial resistance in rural aquaculture practice in Bangladesh through digital communications: a pilot study. Glob Health Action 2019;12(sup1):1734735. [doi: 10.1080/16549716.2020.1734735] [Medline: 32153258]
21. Tarín-Pelló A, Marco-Crespo E, Suay-García B, Galiana-Roselló C, Bueso-Bordils JI, Pérez-Gracia MT. Innovative gamification and outreach tools to raise awareness about antimicrobial resistance. Front Microbiol 2022;13:977319. [doi: 10.3389/fmicb.2022.977319] [Medline: 36187952]

XSL•FO
RenderX

22. Huttner B, Saam M, Moja L, et al. How to improve antibiotic awareness campaigns: findings of a WHO global survey. BMJ Glob Health 2019;4(3):e001239. [doi: 10.1136/bmjgh-2018-001239] [Medline: 31179029]
23. Earnshaw S, Mancarella G, Mendez A, et al. European Antibiotic Awareness Day: a five-year perspective of Europe-wide actions to promote prudent use of antibiotics. Euro Surveill 2014 Oct 16;19(41):20928. [doi: 10.2807/1560-7917.es2014.19.41.20928] [Medline: 25345519]
24. Andersen B, Hair L, Groshek J, Krishna A, Walker D. Understanding and diagnosing antimicrobial resistance on social media: a yearlong overview of data and analytics. Health Commun 2019 Feb;34(2):248-258. [doi: 10.1080/10410236.2017.1405479] [Medline: 29206493]
25. Dyar OJ, Castro-Sánchez E, Holmes AH. What makes people talk about antibiotics on social media? A retrospective analysis of Twitter use. J Antimicrob Chemother 2014 Sep;69(9):2568-2572. [doi: 10.1093/jac/dku165] [Medline: 24862092]
26. Zucco R, Lavano F, Anfosso R, Bianco A, Pileggi C, Pavia M. Internet and social media use for antibiotic-related information seeking: findings from a survey among adult population in Italy. Int J Med Inform 2018 Mar;111:131-139. [doi: 10.1016/j.ijmedinf.2017.12.005] [Medline: 29425624]
27. Roberts-Lewis S, Baxter H, Mein G, et al. Examining the effectiveness of social media for the dissemination of research evidence for health and social care practitioners: systematic review and meta-analysis. J Med Internet Res 2024 Jun 5;26:e51418. [doi: 10.2196/51418] [Medline: 38838330]
28. Gough A, Hunter RF, Ajao O, et al. Tweet for behavior change: using social media for the dissemination of public health messages. JMIR Public Health Surveill 2017 Mar 23;3(1):e14. [doi: 10.2196/publichealth.6313] [Medline: 28336503]
29. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187. [doi: 10.2196/17187] [Medline: 33470931]
30. Adebesin F, Smuts H, Mawela T, Maramba G, Hattingh M. The role of social media in health misinformation and disinformation during the COVID-19 pandemic: bibliometric analysis. JMIR Infodemiology 2023 Sep 20;3:e48620. [doi: 10.2196/48620] [Medline: 37728981]
31. Quinn EK, Fenton S, Ford-Sahibzada CA, et al. COVID-19 and vitamin D misinformation on Youtube: content analysis. JMIR Infodemiology 2022;2(1):e32452. [doi: 10.2196/32452] [Medline: 35310014]
32. Reddy PS, DeBord LC, Gupta R, Kapadia P, Mohanty A, Dao H Jr. Antibiotics for acne vulgaris: using Instagram to seek insight into the patient perspective. J Dermatolog Treat 2021 Mar;32(2):188-192. [doi: 10.1080/09546634.2019.1631432] [Medline: 31190574]
33. Growth of monthly active users of selected social media platforms worldwide from 2019 to 2021. Statista. 2022. URL: https://www.statista.com/statistics/1219318/social-media-platforms-growth-of-mau-worldwide [accessed 2025-07-31]
34. Most popular social networks worldwide as of January 2022, ranked by number of monthly active users. Statista. 2022. URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users [accessed 2025-07-31]
35. Krallman A, Pelletier MJ, Adams FG. @Size vs #impact: social media engagement differences amongst Facebook, Twitter, and Instagram. In: Celebrating America's Pastimes: Baseball, Hot Dogs, Apple Pie and Marketing?: Cham: Springer International Publishing; 2016. [doi: 10.1007/978-3-319-26647-3_112]
36. Pelletier MJ, Krallman A, Adams FG, Hancock T. One size doesn't fit all: a uses and gratifications analysis of social media platforms. JRIM 2020 Jun 1;14(2):269-284. [doi: 10.1108/JRIM-10-2019-0159]
37. Berlo DK. Communication: An Introduction to Theory and Practice: Bolt, Rinehart, and Winston; 1960.
38. Berlo DK. Communication as process: review and commentary. Ann Int Commun Assoc 1977 Dec;1(1):11-27. [doi: 10.1080/23808985.1977.11923667]
39. Zhang N, Tseng HY, Thai J, et al. Health communication needs for COVID-19 prevention and control among college students. J Am Coll Health 2024 Dec;72(9):3204-3209. [doi: 10.1080/07448481.2022.2155060] [Medline: 36595622]
40. Chen S, Xiao L, Kumar A. Spread of misinformation on social media: what contributes to it and how to combat it. Comput Human Behav 2023 Apr;141:107643. [doi: 10.1016/j.chb.2022.107643]
41. Le QH, Phan Tan L, Hoang TH. Brand posts and brand co-creation in higher education communities: a social communication process theory. J Mark High Educ :1-24. [doi: 10.1080/08841241.2023.2239756]
42. Budnik E, Gaputina V, Boguslavskaya V. Dynamic of hashtag functions development in new media: hashtag as an identificational mark of digital communication in social networks. : Association for Computing Machinery Presented at: Proceedings of the XI International Scientific Conference Communicative Strategies of the Information Society; Oct 25, 2019. [doi: 10.1145/3373722.3373795]
43. Berg BL, Lune H. Qualitative Research Methods for the Social Sciences, 8th edition: Pearson Education, Inc., United States; 2012.
44. Krippendorff K. Content Analysis: An Introduction to Its Methodology: Sage publications; 2018.
45. Stahl NA, King JR. Expanding approaches for research: understanding and using trustworthiness in qualitative research. J Dev Educ 2020;44(1):26-28 [FREE Full text]
46. Lincoln YS, Guba EG. Naturalistic Inquiry: Sage Publications; 1985. [doi: 10.1016/0147-1767(85)90062-8]
47. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]

XSL•FO
RenderX

48. Pelechová Z. Instagram hashtag scraper. Apify. URL: https://apify.com/apify/instagram-hashtag-scraper [accessed 2025-07-31]
49. Concerned about antimicrobial resistance: spread the word using this social media toolkit. World Organisation for Animal Health. 2023. URL: https://www.woah.org/app/uploads/2023/05/amr-advocacy-toolkit-ed18724.pdf [accessed 2025-07-31]
50. Reframing resistance: how to communicate about antimicrobial resistance effectively.: Wellcome Trust; 2019 URL: https://wellcome.org/reports/reframing-antimicrobial-resistance-antibiotic-resistance [accessed 2025-07-31]
51. Lucibello KM, Vani MF, Koulanova A, deJonge ML, Ashdown-Franks G, Sabiston CM. #quarantine15: A content analysis of Instagram posts during COVID-19. Body Image 2021 Sep;38:148-156. [doi: 10.1016/j.bodyim.2021.04.002] [Medline: 33892438]
52. Cohen R, Irwin L, Newton-John T, Slater A. #bodypositivity: A content analysis of body positive accounts on Instagram. Body Image 2019 Jun;29:47-57. [doi: 10.1016/j.bodyim.2019.02.007] [Medline: 30831334]
53. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
54. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 1;19:1609406919899220. [doi: 10.1177/1609406919899220]
55. Stemler S. An overview of content analysis. PARE 2000;7(1):17. [doi: 10.7275/z6fm-2e34]
56. Stemler SE. Emerging trends in the social and behavioral sciences:1-14. [doi: 10.1002/9781118900772]
57. Weiger C, Smith KC, Cohen JE, Dredze M, Moran MB. How Internet Contracts Impact Research: Content Analysis of Terms of Service on Consumer Product Websites. JMIR Public Health Surveill 2020 Dec 2;6(4):e23579 [FREE Full text] [doi: 10.2196/23579] [Medline: 33263555]
58. Chiauzzi E, Wicks P. Digital Trespass: Ethical and Terms-of-Use Violations by Researchers Accessing Data From an Online Patient Community. J Med Internet Res 2019 Feb 21;21(2):e11985 [FREE Full text] [doi: 10.2196/11985] [Medline: 30789346]
59. Moreno MA, Goniu N, Moreno PS, Diekema D. Ethics of social media research: common concerns and practical considerations. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):708-713. [doi: 10.1089/cyber.2012.0334] [Medline: 23679571]
60. Kline KN. A decade of research on health content in the media: the focus on health challenges and sociocultural context and attendant informational and ideological problems. J Health Commun 2006;11(1):43-59. [doi: 10.1080/10810730500461067] [Medline: 16546918]
61. Scheufele DA. Beyond the choir? The need to understand multiple publics for science. Environmental Communication 2018 Nov 17;12(8):1123-1126. [doi: 10.1080/17524032.2018.1521543]
62. Dutta S, R A, E P. Breaking the bubble: a case study on the echo chamber effect in Instagram. J Inf Technol Teach Cases 2025;0. [doi: 10.1177/20438869251326279]
63. Engebretsen M. The role, impact, and responsibilities of health experts on social media. A focus group study with future healthcare workers. Front Commun 2024;9:9. [doi: 10.3389/fcomm.2024.1296296]
64. Stellefson M, Paige SR, Chaney BH, Chaney JD. Evolving role of social media in health promotion: updated responsibilities for health education specialists. Int J Environ Res Public Health 2020 Feb 12;17(4):1153. [doi: 10.3390/ijerph17041153] [Medline: 32059561]
65. Brossard D, Scheufele DA. The chronic growing pains of communicating science online. Science 2022 Feb 11;375(6581):613-614. [doi: 10.1126/science.abo0668] [Medline: 35143290]
66. Oh SH, Lee SY, Han C. The effects of social media use on preventive behaviors during infectious disease outbreaks: the mediating role of self-relevant emotions and public risk perception. Health Commun 2021 Jul;36(8):972-981. [doi: 10.1080/10410236.2020.1724639] [Medline: 32064932]
67. Kim Y, Kim JH. Using photos for public health communication: A computational analysis of the Centers for Disease Control and Prevention Instagram photos and public responses. Health Informatics J 2020 Sep;26(3):2159-2180. [doi: 10.1177/1460458219896673] [Medline: 31969051]
68. Charani E, Shariq S, Cardoso Pinto AM, et al. The use of imagery in global health: an analysis of infectious disease documents and a framework to guide practice. Lancet Glob Health 2023 Jan;11(1):e155-e164. [doi: 10.1016/S2214-109X(22)00465-X] [Medline: 36463917]
69. Rauschnabel PA, Sheldon P, Herzfeldt E. What motivates users to hashtag on social media? Psychology and Marketing 2019 May;36(5):473-488. [doi: 10.1002/mar.21191]
70. Neuendorf KA. The Content Analysis Guidebook, 2nd edition: SAGE Publications; 2017. [doi: 10.4135/9781071873045]
71. Ruths D, Pfeffer J. Social sciences. Social media for large studies of behavior. Science 2014 Nov 28;346(6213):1063-1064. [doi: 10.1126/science.346.6213.1063] [Medline: 25430759]

Abbreviations

AMR: antimicrobial resistance
SMCR: sender-message-channel-receiver

XSL•FO
RenderX

SNS: social network site
WAAW: World Antimicrobial Resistance Awareness Week
WHO: World Health Organization

Edited by T Mackey; submitted 22.10.24; peer-reviewed by D Golemi-Kotra, N Idowu, S Pena-Fernandez; revised version received 08.05.25; accepted 27.06.25; published 20.08.25.

Please cite as:
Nilsson E, Oljans E, Nordvall AC, Ancillotti M
Communicating Antimicrobial Resistance on Instagram: Content Analysis of #AntibioticResistance
JMIR Infodemiology 2025;5:e67825
URL: https://infodemiology.jmir.org/2025/1/e67825
doi:10.2196/67825

© Elin Nilsson, Emma Oljans, Anna-Carin Nordvall, Mirko Ancillotti. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 20.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

#GenderAffirmingHormoneTherapy and Health Information on TikTok: Thematic Content Analysis

Julia Rose Beatini[1*], MD; Nora Yanyi Sun[2*]; Julianna K Coleman[3], MD; Maetal E Haas-Kogan[3], AB; Andrea Pelletier[3,4], MPH, MD; Deborah Bartz[3,4], MPH, MD; Alex Sogomon Keuroghlian[3,5,6], MPH, MD

[1]Thomas Jefferson University, Philadelphia, PA, United States
[2]Harvard University, Cambridge, MA, United States
[3]Harvard Medical School, Boston, MA, United States
[4]Department of Obstetrics and Gynecology, Brigham and Women's Hospital, Boston, MA, United States
[5]Department of Psychiatry, Massachusetts General Hospital, 55 Fruit Street, Wang Building 8th Floor, Boston, MA, United States
[6]The National LGBTQIA+ Health Education Center, The Fenway Institute, Boston, MA, United States
[*]these authors contributed equally

Corresponding Author:
Alex Sogomon Keuroghlian, MPH, MD
Harvard Medical School, Boston, MA, United States

Abstract

Background: Transgender and gender diverse people often turn to online platforms for information and support regarding gender-affirming hormone therapy (GAHT); however, analysis of this social media content remains scarce.

Objective: We characterized GAHT-related videos on TikTok to highlight the implications relevant to GAHT prescribers.

Methods: We used a web scraper to identify TikTok videos posted under the hashtags #genderaffirminghormonetherapy and #genderaffirminghormones as of November 2023. We identified recurrent themes via qualitative content analysis and assessed health education videos with the Patient Education Materials Assessment Tool for Audiovisual Materials (PEMAT-A/V) scale and a modified Currency, Relevance, Authority, Accuracy, and Purpose (CRAAP) test.

Results: Out of 69 videos extracted, 71% (49/69) were created by GAHT users, 24.6% (17/69) were created by health care workers, and 21.7% (15/69) were created to provide health education. Themes included physical changes on testosterone, GAHT access, and combating misinformation and stigma surrounding GAHT. Health education videos scored highly on PEMAT-A/V items assessing understandability (mean 88.3%, SD 11.3%) and lower on actionability (mean 60.0%, SD 45.8%). On the CRAAP test, videos scored highly on the relevance, authority, and purpose domains but lower on the currency and accuracy domains.

Conclusions: Discussions of GAHT on TikTok build community among transgender and gender diverse users, provide a platform for digital activism and resistance against legislation that limits GAHT access, and foster patient-provider dialogue. Educational videos are highly understandable and are created by reliable sources, but they vary in terms of currency and quality of supporting evidence, and they lack in actionability.

(JMIR Infodemiology 2025;5:e66845) doi:10.2196/66845

KEYWORDS

transgender; gender diverse; transgender and gender diverse; TGD; gender fluid; online platform; social media; gender affirming; hormone therapy; gender-affirming hormone therapy; GAHT; social media content; media information; social media analysis; TikTok; web scraper; hashtag; themes; qualitative content analysis; patient education materials assessment; PEMAT; Currency, Relevance, Authority, Accuracy, and Purpose; CRAAP; audiovisual materials; qualitative

Introduction

TikTok (ByteDance) is a short-form video-sharing application that boasts 97.6 million active users in the United States. Since TikTok's spike in popularity in 2020, transgender and gender diverse (TGD) content creators have used the platform as a space to document and share their experiences with others. While TikTok has the potential to disseminate health information and improve access to gender-affirming care, it has come under scrutiny for spreading misinformation, bias, and hate speech [1].

Despite the lack of systematic analysis, the spread of information about gender-affirming care within the TGD TikTok community has been cited as an example of community-engaged knowledge exchange and peer-to-peer support [2]. Furthermore, TikTok has the potential to improve access to health information among communities that experience health inequities as the result of

XSL•FO
RenderX

discrimination, because having positive impressions of knowledgeable professionals on social media may help decrease medical mistrust and enhance access to care offline [2]. Lowering medical mistrust among TGD communities is crucial, given that 24% of respondents in the 2022 US Transgender Survey reported avoiding medical care due to fear of mistreatment by a provider [3].

While TikTok videos have the potential to improve access to health information, peer support, and trust in medical professionals, TikTok may also be used to spread misinformation, disinformation, and hateful rhetoric against lesbian, gay, bisexual, transgender, queer, intersex, and all asexually and gender diverse (LGBTQIA+) people on the platform [4]. For example, the use of gender-affirming hormone therapy (GAHT), which involves administering hormones like estrogen and testosterone or puberty blockers to alter gendered physical characteristics among TGD youth, has increasingly been attacked; its controversy has led to online hate speech and, in several instances, threats of violence against hospitals and individual providers online [5,6].

Previous studies have queried TikTok to explore attitudes toward and experiences with other types of medical care, using qualitative methods to determine the content and tone of posts about medical interventions ranging from contraceptive methods to erectile dysfunction treatment [7-9]. Others have focused on analyzing the quality and accuracy of health information reported on the platform [10-15]. Their analyses yielded insights into misconceptions about care, the prevalence of inaccurate factual claims about treatment, and salient elements of individual experiences with care, all of which have the potential to inform how clinicians treat and counsel their patients. However, there have been no analyses of users' attitudes towards, experience with, or knowledge about GAHT.

The aim of this study is to explore popular TikTok content posted under the hashtags #genderaffirminghormones and #genderaffirminghormonetherapy. Using previously validated methods, we (1) describe the demographic characteristics, attitudes, and affiliations of video content creators; (2) perform a qualitative analysis of video content to identify content themes; and (3) assess the understandability, actionability, and reliability of information presented in a subset of educational videos. In doing so, we aim to better understand the degree to which TikTok is a vehicle for sharing valuable information about GAHT and treatment access versus a potentiator of misinformation and harmful biases.

Methods

Data Extraction

We used the web-scraping application Apify (Apify Technologies s.r.o.) TikTok scraper to download all TikTok videos posted publicly under the hashtags #genderaffirminghormonetherapy and #genderaffirminghormones as of November 17, 2023; Apify provided all videos as MP4 files. While videos that fit the inclusion criteria may be available under alternate hashtags, only these two hashtags were selected for this study, as it was infeasible to scrape the vast content created under broader hashtags; additionally, users who are actively seeking information on GAHT would most likely search these two hashtags. The scrape included a total of 86 videos. We applied the following exclusion criteria: (1) non-English language video, (2) country codes in the European Union or China (based on differences in data usage agreements in these regions), (3) GAHT not mentioned in the video, (4) duplicate video, or (5) video posts removed following the scrape.

Descriptive Analysis of the Content

For all eligible videos, we recorded the date posted and video duration, and TikTok engagement statistics, including the number of video views, likes, comments, shares, and number of creator fans. Through discussion and consensus, the first and second authors determined five main categories of videos after reviewing all videos (personal experience, health education, politics, creator opinion, and humor) and categorized each video. The sum, median, and IQR of engagement statistics were calculated for each type of video to best characterize the distribution of engagement; these metrics were selected given that specific videos may go viral on the platform, thereby skewing the data.

The first author then determined content creator demographics, including self-described gender identity and sexual orientation, for each video via the exploration of content creators' public profiles, including the user, bios, current or previous videos, video captions, and comment responses [16]. Only explicitly stated identities from content creators were included to avoid assumptions about their identities; if no such statements were available, we marked the field as "unknown." We selected "not applicable" if the account belonged to an organization rather than an individual. We similarly obtained creators' GAHT user status, health care worker status, and organizational affiliations from information on their public profiles. Finally, the first and second authors individually rated each video on whether the creator displayed a positive (eg, supportive or encouraging), negative (eg, dismissive or transphobic), neutral (eg, purely informative), or ambiguous or mixed attitude (authors could not ascertain or agree on the creator's intentions) toward GAHT. Through discussion, the authors came to a consensus on the final ratings of all videos.

Quality Rating of the Content

We further analyzed a subset of health education videos to assess the understandability, actionability, and reliability of information presented. "Health education" videos contained at least one message about which creators aimed to inform viewers [17,18].

The first and second authors independently used the Patient Education Materials Assessment Tool for Audiovisual Materials (PEMAT-AV) to assess information understandability (ie, accessibility of the information presented for the layperson, including statements like "The material uses common, everyday language"; items 1 - 13) and actionability (ie, the feasibility of implementing the information presented, including statements like "The material breaks down any action into manageable, explicit steps"; items 14 - 17) of patient educational materials. Each item in the scale is rated as "agree" (1 point), "disagree"

XSL•FO
RenderX

(0 points), or "not applicable." The first and second author then compared each rating they assigned until they came to a consensus on all items. Finally, scores were calculated as a percentage of the possible points obtained for all items, excluding those rated as not applicable. Higher percentages suggest higher levels of understandability and actionability, with a threshold of >75% used to indicate "high quality" [19-21].

The senior author, who is a physician-scientist specialized in gender-affirming care, also used a modified Currency, Relevance, Authority, Accuracy, and Purpose (CRAAP) test to assess the reliability and accuracy of health information presented in educational videos [22]. The CRAAP test is used for the quantitative assessment of digital health information [23-25] by assessing five domains of information reliability: currency, relevance, authority, accuracy, and purpose. We adapted a previously published scale and scoring key by modifying the language and removing several items that were not applicable to audiovisual media [23]. Our adapted scale contained a total of 18 items (3 assessing currency, 4 assessing reliability, 3 assessing authority, 3 assessing accuracy, and 5 assessing purpose).

Overall scores ranged between 0 and 34, with higher scores suggesting higher reliability. Based on previous work using the CRAAP test, we considered a final score of <20% as unreliable, 20‐46% as reliable with caution, 46‐80% as good reliability, and >80% as excellent reliability [23]. We used Cronbach α to assess the interitem reliability of the scale (α=.88). The rubric for the modified CRAAP test is provided in Table 1.

XSL•FO
RenderX

Table . Modified Currency, Relevance, Authority, Accuracy, and Purpose (CRAAP)[a] test.

Item	Scoring key			
	0	1	2	3
Currency (3 items, 5 points)				
Date created	>5 y	1 - 5 y	<1 y	N/A[b]
Information outdated or debunked	Yes	No	N/A	N/A
Embedded links or suggested resources still accessible	None listed	No longer accessible	Still accessible	N/A
Relevance (4 items, 4 points)				
Information answers a central question	No	Yes	N/A	N/A
Information identifies an intended audience	No	Yes	N/A	N/A
Information appropriate for the needs of the intended audience	No	Yes	N/A	N/A
Information avoids overgeneralization	No	Yes	N/A	N/A
Authority (3 items, 5 points)				
Identity of the author or source	None identified	Medication user or patient	Expert in the field	N/A
Author's credentials	None	Lived experience	Licensed medical professional	N/A
Author qualified to discuss the topic[c]	No	Yes	N/A	N/A
Accuracy (3 items, 7 points)				
Derivation of information	Unclear	Individual lived experience	Professional experience	Evidence-based review
Is the information supported by evidence[d]	No	Yes	N/A	N/A
What kind of evidence supports the claim	None	Individual lived experience	Expert/consensus opinion	Published evidence-based guidelines
Purpose (5 items, 8 points)				
Purpose of information	Advertisement	Persuading or entertaining	Informing	Teaching
Intentions or purpose clear	No	Yes	N/A	N/A
Nature of information	Propaganda	Opinion	Facts	N/A
Point of view appears objective and impartial	No	Yes	N/A	N/A
Political, ideological, cultural, or religious biases	Yes	No	N/A	N/A

[a]This modified CRAAP Test is developed from the original test by Sarah Blakeslee, which is licensed for adaptations under a Creative Commons Attribution 4.0 International License.

[b]N/A: not applicable

[c]Creators were considered qualified to discuss a topic if they either had confirmed credentials indicating relevant subject matter expertise and licensure relevant to the clinical topic, or lived experience in a patient testimonial.

[d]Information presented in videos was considered supported if contemporary scientific evidence was congruent with the information.

Ethical Considerations

This study involves the analysis of publicly available data from TikTok; data from private accounts were not assessed as part of this project. Investigators only had access to creators' account names while viewing videos during the initial scoring and coding of data. Any identifiable information, including account names, was removed from the data prior to dissemination. This study

XSL•FO
RenderX

received a Not Human Subjects Research Determination from the Harvard Longwood Campus Institutional Review Board.

Results

Descriptive Analysis of the Content

Our search identified 84 videos posted between February 2022 (the earliest video searching the TikTok hashtag recalled) and October 2023; while data collection was performed in November 2023, no new videos under the hashtag were posted in November 2023. Fifteen videos were determined ineligible for analysis based on non-English language (n=6), country code in the European Union or China (n=1), GAHT not mentioned in the video (n=6), duplicate (n=1), or video post removed following the scrape (n=1). The remaining 69 videos were included in our analysis. The median (IQR) video duration was 56 (23 - 73) seconds. Video characteristics and creator demographics are summarized in Table 2.

XSL•FO
RenderX

Table . Video characteristics and creator demographics of the 69 TikTok videos analyzed.

Video characteristics	n (%), N=69
Video type	
Personal experience	45 (65.2)
Health education	15 (21.7)
Politics	4 (5.8)
Creator opinion	4 (5.8)
Humor	1 (1.4)
Attitude toward GAHT[a]	
Positive	58 (84.1)
Negative	2 (2.9)
Neutral	5 (7.2)
Mixed or ambiguous	4 (5.8)
Medication referenced	
Testosterone	41 (59.4)
Estrogen	11 (15.9)
Antiandrogens	1 (1.4)
Not specified	17 (24.6)
Creator gender identity	
Nonbinary	29 (42.0)
Cisgender woman	16 (23.2)
Trans masculine	14 (20.3)
Transgender man	9 (13.0)
Transgender woman	6 (8.7)
Trans feminine	1 (1.4)
Unknown	1 (1.4)
N/A[b]	3 (4.3)
Creator sexual orientation	
Bisexual	10 (14.5)
Queer	10 (14.5)
Lesbian	1 (1.4)
Gay	1 (1.4)
Pansexual	1 (1.4)
Unknown	43 (62.3)
N/A	3 (4.3)
Health care professional	
Yes	17 (24.6)
No or unsure	52 (75.4)
GAHT user	
Yes	49 (71.0)
No or unsure	20 (28.9)
Creator affiliation	
None	65 (94.2)

XSL•FO
RenderX

Video characteristics	n (%), N=69
Health care organization	3 (4.3)
Religious organization	1 (1.4)

[a]GAHT: gender-affirming hormone therapy.

[b]N/A: not applicable.

Common themes that appeared in videos are described in Table 3. Self-identified health care workers created 50% of videos about estrogen and antiandrogen regimens, 46% of videos about GAHT access and legality, and 33% of videos about physical changes on testosterone. They created relatively smaller proportions of videos about testosterone regimens (24%), combating misinformation or social stigma around GAHT (11%), and physical changes due to testosterone (10%). None posted antitrans rhetoric.

Table . Content themes and subthemes appearing in the 69 TikTok videos analyzed.

Content themes and subthemes	n (%), N=69
Physical changes on testosterone	29 (42.0)
Voice changes	20 (29.0)
Facial or body hair growth	11 (15.9)
Clitoral growth	7 (10.1)
Skin changes (acne or oiliness)	6 (8.7)
Body odor changes	3 (4.3)
Body composition changes	6 (8.7)
Physical changes on estrogen	2 (2.9)
Breast development	4 (2.9)
Pre-post therapy photo reveal	17 (24.6)
Testosterone regimens	13 (18.8)
Medication formulations and routes of administration	2 (2.9)
Medication safety or monitoring	2 (2.9)
Concomitant use of estrogen-containing medications	6 (8.7)
Estrogen and antiandrogen regimens	4 (5.8)
Medication formulations and routes of administration	3 (4.3)
Medication safety or monitoring	13 (18.8)
GAHT[a] access and legality	6 (8.7)
Reviews new state guidelines or policy proposals	4 (5.8)
Reviews the process of obtaining medical clearance or prescription	3 (4.3)
Promotes a health care practice offering GAHT	1 (1.4)
Solicits advice on obtaining a prescription	1 (1.4)
Offers resources for funding GAHT	9 (13.0)
Combating misinformation and social stigma about GAHT	6 (8.7)
Validates TGD[b] identities	5 (7.2)
Emphasizes mental health benefits of access to GAHT	2 (2.9)
Normalizes the use of GAHT	29 (42.0)

[a]GAHT: gender-affirming hormone therapy.

[b]TGD: transgender and gender diverse.

Regarding engagement, videos had a total of 446,318 views, 43,743 likes, 1184 comments, and 438 shares. The sums, medians, and IQRs of engagement measures across each video type are described in Table 4. The top three most viewed videos each discussed physical changes on testosterone therapy and accounted for 46% of total views, 55 of total likes, 32% of total comments, and 56% of total shares. The single most-viewed video (104,800 views) depicted a nurse practitioner discussing

XSL•FO
RenderX

labial changes due to testosterone with the aid of a plastic anatomic model.

Table . Number of views, likes, comments, shares, and creator fans for the 69 TikTok videos analyzed.

Engagement statistics by video type	All, n=69	Personal experience, n=45	Health education, n=15	Politics, n=4	Creator opinion, n=4	Humor, n=1
View count						
Total	440,711	272,602	138,843	18,165	9808	1293
Median (IQR)	847 (283 - 2649)	843 (313 - 2600)	423 (276 - 4315)	1451 (235 - 5758)	2122 (1148 - 3424)	N/A[a]
Like count						
Total	43,743	26,431	12,281	3671	1051	309
Median (IQR)	77 (21 - 248)	77 (28 - 222)	25 (11 - 242)	149 (19 - 1048)	185 (106 - 341)	N/A
Comment count						
Total	1169	585	246	226	96	16
Median (IQR)	6 (1-16)	6 (2-13)	2 (0 - 9)	34 (0 - 91)	23 (19 - 28)	N/A
Share count						
Total	437	198	101	118	20	0
Median (IQR)	0 (0 - 3)	0 (0 - 1)	0 (0 - 6)	8 (0 - 37)	5 (2-9)	N/A
Creator fans						
Total	488,055	185,202	129,321	162,177	10,857	498
Median (IQR)	1871 (593 - 3448)	1871 (759 - 3047)	189 (39 - 16,648)	3429 (2257 - 41,717)	2269 (1924 - 4640)	N/A

[a]N/A: not applicable.

Quality Rating of the Content

Health education videos averaged 88.3% (SD 11.1%; median 85.7%) on PEMAT-A/V understandability items, and 60.0% (SD 45.8%; median 100%) on actionability items. Together, the weighted mean (SD) PEMAT-A/V score for all items was 81.7% (18.7%). Videos averaged 74.5% (SD 18.4%) in total on the CRAAP test, with mean (SD) component scores as follows: currency, 58.7% (19.2%); relevance, 93.3% (SD 20.0%); authority, 82.7% (SD 35.3%); accuracy, 61.0% (SD +29.3%); and purpose 81.7% (SD 14.1%).

Discussion

Principal Findings and Comparison With Previous Works

This study characterized the creators and content of 69 TikTok videos related to GAHT. The most common videos were those made by TGD content creators who shared personal experiences on GAHT. For example, many videos were part of weekly or monthly series in which the creator reviewed the physical effects of their medications or applied or injected medication on camera while providing updates to their followers. Attitudes among users were overwhelmingly positive, with a few instances of ambivalence or mixed attitudes reflecting a preference for one mode of GAHT administration over another.

The large proportion of personal experience videos, which also had high engagement from viewers, reflects the longstanding popularity of TGD video blogs (vlogs) across other social media sites, including Reddit and YouTube [26-28]. As in the current study, prior work has noted an increased frequency of videos created by testosterone users relative to estrogen/anti-androgen users, which may be partially due to the shorter onset time of visible bodily changes with testosterone use [29,30]. Prior YouTube-based ethnographic research suggests these vlogs simultaneously function as opportunities for creators to reflect on and visualize their own gender affirmation journeys, as well as digital diaries to share their narratives with others and engage in broader dialogue [28]. The prevalence of and engagement with personal experience videos in this dataset suggest that TikTok provides a similar space for TGD GAHT users to continually reaffirm their personal identities and engage with the community through digital narrative-sharing.

Our findings suggest that TikTok also functions as a platform for digital activism and resistance. Nearly a third of videos dealt with issues related to GAHT access, newly imposed restrictions on GAHT use, and disinformation about GAHT in the media. In these videos, creators reviewed the process of obtaining a prescription, described sources of funding, or detailed how users could continue to access GAHT despite restrictive legislation passed in Florida and Utah during the timeframe studied. Many used TikTok’s duet function to respond directly to disinformation circulating in the media and among politicians. These videos represent a timely means of intracommunity resistance against restrictive legislation by providing GAHT users with steps to continue GAHT. As legislatures in the United States and abroad attempt to ban and restrict access to care, social media platforms including TikTok may play an increasingly important role in community organizing and harm reduction.

XSL•FO
RenderX

Our results suggest that TikTok provides space for dialogue between GAHT users and health care workers, which may help reduce medical mistrust and facilitate the safe administration of GAHT. Content made by health care workers constituted nearly a quarter of all videos and was created with the intention of sharing health information relevant to GAHT users. These videos showed high levels of engagement, and often directly addressed viewers' comments or private messages. The ability for health care providers to respond directly to GAHT users' questions about hormones supports its potential as a space for effective digital knowledge mobilization, as prior work suggests [2]. Furthermore, health care providers also benefit from this interaction by gaining a deeper understanding of common questions and concerns among TGD people interested in GAHT.

Our analysis of health education videos suggests that information contained in educational videos is of high understandability, low actionability, and moderate reliability. An average PEMAT understandability score of 88.3% across educational videos suggests a high level of accessibility to viewers. The lower average actionability score of 60.0% may reflect the fact that the goal of many health education videos was to explain a particular phenomenon (eg, the mechanisms behind expected physical changes due to testosterone) rather than to guide patients' decisions about treatment.

High scores on the relevance, authority, and purpose components of the CRAAP test (>80%) suggest that educational information regarding GAHT is well-suited to the needs of the intended audience, created by reliable sources, and shared with the purpose of informing viewers about their health. The lower scores on the accuracy and currency components of the scale reflect the finding that while many content creators used their professional experience to support their claims, few cited evidence-based guidelines or provided viewers with further reading or updates as new information emerged. While the content analyzed was overall understandable, it did not reliably contain the level or depth of detail present in evidence-based clinical practice guidelines. It may therefore be important for providers to work with patients to contextualize information about GAHT found on the platform.

Clinical Implications

Our findings have several clinical implications. Providers should be aware that patients may use TikTok as a source of health knowledge, and that this information varies in depth and accuracy. GAHT prescribers may consider incorporating routine screening questions about patients' consumption of medication-related social media content and using these to either augment, contextualize, or correct information found online. As much as GAHT users may use TikTok as a space to seek information from health care providers, it can also better inform providers about TGD patients' needs and priorities. These needs are apparent where videos made by providers and users diverge thematically. For instance, providers created a high proportion of videos surrounding drug safety and monitoring, whereas GAHT users focused more on desired medication effects, suggesting a potential need for comprehensive counseling on expected and adverse effects.

Furthermore, videos created by providers also differed from user-created videos in that few of them directly addressed a nonbinary audience, despite nonbinary content creators being the most well-represented demographic in the videos analyzed, suggesting that there may be a lack of awareness surrounding GAHT-related needs of nonbinary people. In fact, only one health education video created by a provider explicitly addressed a nonbinary audience, and nearly a quarter used language that reinforced a binary gender paradigm (eg, "this video is for anyone transitioning male to female"). Thus, TikTok may also offer cisgender providers an opportunity to better understand the unique needs of diverse groups seeking out GAHT, which may allow for a more patient-centered and culturally responsive approach to counseling.

Limitations

There are several limitations to this study. First, the content analyzed is limited to what appears under the specific search terms #genderaffirminghormonetherapy and #genderaffirming hormones. While there is likely broader discourse on TikTok surrounding this topic, broadening the search terms—for example, to #hormonetherapy or #testosterone—would have yielded results that are not specific or relevant to TGD communities. While the videos analyzed were created over a 20-month period, they represent a single snapshot in time within the rapidly changing landscape of social media.

We were unable to compare quantitative statistics between different video types due to the small sample size for some video types and the non-normal distribution of engagement statistics with videos that were outliers with regard to viewership. Future studies should conduct rigorous statistical comparisons on video metrics in a larger sample size.

Additionally, TikTok has been criticized for its tendency toward "collaborative filtering," a method of predicting users' interests based on their previous views and activity in the app. By using physiognomic data, some argue that TikTok is more likely to recommend creators who look like the platform's white and able-bodied top influencers, and less likely to recommend creators who belong to underrepresented minority groups, which can also be referred to as "shadow banning" [31-33]. In this context, it is important to consider that some perspectives may be systematically privileged over others.

Similarly, the collected data may be vulnerable to bias towards more positive experiences. Users may be more willing to share positive experiences given the nature of the community formed under these hashtags; those with negative experiences may be less willing to share their experiences. However, we attempted to mitigate bias in content selection by using a web scraper rather than a TikTok account for data collection.

Our analysis involved collecting self-reported demographic characteristics from public profiles, which can be falsely reported for a variety of reasons, including stigma or safety concerns. Thus, it is possible that the number of TGD content creators and GAHT users within the dataset was underreported. Moreover, we chose to limit our focus to video content to analyze the dialogue between creators and viewers, though

XSL•FO
RenderX

future research may also include the comment sections of such videos.

Finally, while an abundance of clinical evidence supports the efficacy of GAHT, there remains debate nationally and internationally on certain aspects of GAHT. Videos were rated based on the contemporary understanding of GAHT, with the authorship team comparing information presented in videos to the majority consensus in the academic field; however, we acknowledge that our ratings are limited by a lack of consensus in the clinical community on certain topics related to GAHT.

Conclusions

This study evaluated the discourse around GAHT on TikTok to better understand the extent to which it is being used as a tool for building community and disseminating health knowledge. Overall, our results suggest that TikTok allows GAHT users to document their experiences, connect with other community members, and advocate for GAHT as legislation restricts access to treatment. TikTok also provides a space for direct user-provider dialogue, whereby users can have questions answered by health professionals with a high level of information understandability. Health professionals should be aware that patients may use TikTok as a source of information and should be ready to explore these sources of knowledge with patients, as they vary in terms of currency and quality of supporting evidence. Health care workers may utilize social media platforms such as TikTok as an opportunity for bidirectional learning and health knowledge dissemination between clinicians and GAHT users.

Conflicts of Interest

ASK declares royalties as editor of a McGraw Hill textbook on transgender and gender diverse health care and of an American Psychiatric Association textbook on gender-affirming psychiatric care. The authors declare no competing financial interests.

References

1. Social media safety index 2023. GLAAD. 2023. URL: https://glaad.org/publications/social-media-safety-index-2023/ [accessed 2024-05-14]
2. MacKinnon KR, Kia H, Lacombe-Duncan A. Examining TikTok's potential for community-engaged digital knowledge mobilization with equity-seeking groups. J Med Internet Res 2021 Dec 9;23(12):e30315. [doi: 10.2196/30315] [Medline: 34889739]
3. James SE, Herman JL, Durso LE, et al. Early insights: a report of the 2022 US Transgender Survey. : National Center for Transgender Equality; 2024 URL: https://transequality.org/sites/default/files/2024-02/2022%20USTS%20Early%20Insights%20Report_FINAL.pdf [accessed 2024-04-11]
4. Brewster J, Arvanitis L, Pavilonis V, Wang M. Beware the 'New Google:' TikTok's search engine pumps toxic misinformation to its young users. NewsGuard. 2022. URL: https://www.newsguardtech.com/misinformation-monitor/september-2022/ [accessed 2025-04-11]
5. O'Reilly KB. Terrifying bomb threats against children's hospitals must stop. American Medical Association. 2022. URL: https://www.ama-assn.org/practice-management/physician-health/terrifying-bomb-threats-against-childrens-hospitals-must-stop [accessed 2024-05-14]
6. Hughes LD, Gamarel KE, Restar AJ, et al. Adolescent providers' experiences of harassment related to delivering gender-affirming care. J Adolesc Health 2023 Oct;73(4):672-678. [doi: 10.1016/j.jadohealth.2023.06.024] [Medline: 37589604]
7. Stoddard RE, Pelletier A, Sundquist EN, et al. Popular contraception videos on TikTok: an assessment of content topics. Contraception 2024 Jan;129:110300. [doi: 10.1016/j.contraception.2023.110300] [Medline: 37802460]
8. Wu J, Trahair E, Happ M, Swartz J. TikTok, #IUD, and user experience with intrauterine devices reported on social media. Obstet Gynecol 2023 Jan 1;141(1):215-217. [doi: 10.1097/AOG.0000000000005027] [Medline: 36473194]
9. Babar M, Loloi J, Patel RD, et al. Cross-sectional and comparative analysis of videos on erectile dysfunction treatment on YouTube and TikTok. Andrologia 2022 Jun;54(5):e14392. [doi: 10.1111/and.14392] [Medline: 35122283]
10. Abramson M, Feiertag N, Javidi D, Babar M, Loeb S, Watts K. Accuracy of prostate cancer screening recommendations for high-risk populations on YouTube and TikTok. BJUI Compass 2023 Mar;4(2):206-213. [doi: 10.1002/bco2.200] [Medline: 36816146]
11. Aragon-Guevara D, Castle G, Sheridan E, Vivanti G. The reach and accuracy of information on autism on TikTok. J Autism Dev Disord 2023 Aug 6. [doi: 10.1007/s10803-023-06084-6] [Medline: 37544970]
12. Bernstein A, Zhu M, Loloi J, et al. TikTok as a source of information regarding premature ejaculation: a qualitative assessment. Sex Med 2023 Apr;11(2):qfac020. [doi: 10.1093/sexmed/qfac020] [Medline: 36910705]
13. Dubin JM, Aguiar JA, Lin JS, et al. The broad reach and inaccuracy of men's health information on social media: analysis of TikTok and Instagram. Int J Impot Res 2024 May;36(3):256-260. [doi: 10.1038/s41443-022-00645-6] [Medline: 36402921]
14. Tam J, Porter EK, Lee UJ. Examination of information and misinformation about urinary tract infections on TikTok and YouTube. Urology 2022 Oct;168(1):35-40. [doi: 10.1016/j.urology.2022.06.030] [Medline: 35809700]
15. Wu J, Greene M, Happ M, Trahair E, Montoya M, Swartz JJ. Medication abortion on TikTok: misinformation or reliable resource? Am J Obstet Gynecol 2023 Jun;228(6):749-751. [doi: 10.1016/j.ajog.2023.01.023] [Medline: 36736679]

XSL•FO
RenderX

16. Paciente R, Pranoto EP, Woolard A, Munro E, Lombardi K. The queers are all right: a content analysis of LGBTQIA + mental health on TikTok. Cult Health Sex 2024 Jun;26(6):778-789. [doi: 10.1080/13691058.2023.2253882] [Medline: 37729467]
17. Schreier M. Qualitative Content Analysis in Practice: SAGE Publications Ltd; 2012. [doi: 10.4135/9781529682571]
18. Hung YH, Miles A, Trevino Z, et al. BIPOC experiences of racial trauma on TikTok: a qualitative content analysis. Contemp Fam Ther 2023 May 9;2023:1-11. [doi: 10.1007/s10591-023-09669-6] [Medline: 37361259]
19. Ashai S, Pace L, Malik RD. Assessment of stress urinary incontinence-related podcasts on stitcher. Urology 2023 Apr;174:79-85. [doi: 10.1016/j.urology.2023.01.044] [Medline: 36791811]
20. Charow R, Snow M, Fathima S, et al. Evaluation of the scope, quality, and health literacy demand of Internet-based anal cancer information. J Med Libr Assoc 2019 Oct;107(4):527-537. [doi: 10.5195/jmla.2019.393] [Medline: 31607810]
21. Mehra M, Brody PA, Kollapaneni SS, Sakhalkar O, Rahimi S. Evaluating the quality, readability, and activity of online information on brain arteriovenous malformations. Cureus 2023 Sep;15(9):e45984. [doi: 10.7759/cureus.45984] [Medline: 37900528]
22. Blakeslee S. The CRAAP test. LOEX Q 2004;31(3):4 [FREE Full text]
23. Garcia M, Daugherty C, Ben Khallouq B, Maugans T. Critical assessment of pediatric neurosurgery patient/parent educational information obtained via the internet. J Neurosurg Pediatr 2018 May;21(5):535-541. [doi: 10.3171/2017.10.PEDS17177] [Medline: 29451451]
24. McNamee LA, Shakartzi H, Wasser TE, Li Y, Kim CS. Assessing the accuracy and quality of YouTube videos on early pregnancy loss. Heliyon 2022 Dec;8(12):e11947. [doi: 10.1016/j.heliyon.2022.e11947] [Medline: 36506362]
25. Shah V, Lava CX, Hakimi AA, Hoa M. Evaluating quality, credibility, and readability of online over-the-counter hearing aid information. Laryngoscope 2024 Jul;134(7):3302-3309. [doi: 10.1002/lary.31278] [Medline: 38280186]
26. Darwin H. Doing gender beyond the binary: a virtual ethnography. Symb Interact 2017 Aug;40(3):317-334 [FREE Full text] [doi: 10.1002/symb.316]
27. Stein A. Transitioning out loud and online. Contexts (Berkeley Calif) 2016 May;15(2):40-45. [doi: 10.1177/1536504216648150]
28. Raun T. Out Online: Trans Self-Representation and Community Building on YouTube: Routledge; 2016. [doi: 10.4324/9781315599229]
29. Raun T. Video blogging as a vehicle of transformation: exploring the intersection between trans identity and information technology. International Journal of Cultural Studies 2015 May;18(3):365-378. [doi: 10.1177/1367877913513696]
30. Coleman E, Radix AE, Bouman WP, et al. Standards of care for the health of transgender and gender diverse people, version 8. Int J Transgend Health 2022;23(Suppl 1):S1-S259. [doi: 10.1080/26895269.2022.2100644] [Medline: 36238954]
31. Faddoul M. A TikTok novelty: face-based filter bubbles. X. 2020. URL: https://twitter.com/MarcFaddoul/status/1232014908536938498 [accessed 2024-05-14]
32. Tillewein H, Mohon-Doyle K, Cox D. A critical discourse analysis of sexual violence survivors and censorship on the social media platform TikTok. Arch Sex Behav 2024 Nov;53(10):3785-3794. [doi: 10.1007/s10508-024-02987-2] [Medline: 39317917]
33. Delmonaco D, Mayworm S, Thach H, Guberman J, Augusta A, Haimson OL. "What are you doing, TikTok?": How marginalized social media users perceive, theorize, and "prove" shadowbanning. Proc ACM Hum-Comput Interact 2024 Apr 17;8(CSCW1):1-39. [doi: 10.1145/3637431]

Abbreviations

CRAAP: Currency, Relevance, Authority, Accuracy, and Purpose
GAHT: gender-affirming hormone therapy
LGBTQIA+: lesbian, gay, bisexual, transgender, queer, intersex, and all asexually and gender diverse
PEMAT-A/V: Patient Education Materials Assessment Tool for Audiovisual Materials
TGD: transgender and gender diverse

Edited by T Nguyen; submitted 12.10.24; peer-reviewed by C Olezeski, K MacKinnon, RC Dimitroyannis; revised version received 11.02.25; accepted 13.02.25; published 29.04.25.

Please cite as:
Beatini JR, Sun NY, Coleman JK, Haas-Kogan ME, Pelletier A, Bartz D, Keuroghlian AS
#GenderAffirmingHormoneTherapy and Health Information on TikTok: Thematic Content Analysis
JMIR Infodemiology 2025;5:e66845
URL: https://infodemiology.jmir.org/2025/1/e66845
doi:10.2196/66845

XSL•FO
RenderX

© Julia Rose Beatini, Nora Yanyi Sun, Julianna K Coleman, Maetal E Haas-Kogan, Andrea Pelletier, Deborah Bartz, Alex Sogomon Keuroghlian. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 29.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Experiences of Public Health Professionals Regarding Crisis Communication During the COVID-19 Pandemic: Systematic Review of Qualitative Studies

Tsuyoshi Okuhara[1], MPH, PhD; Marina Terada[1], MPH; Hiroko Okada[1], MPH, PhD; Rie Yokota[2], MPH, PhD; Takahiro Kiuchi[1], PhD, MD

[1]Department of Health Communication, The University of Tokyo, Tokyo, Japan

[2]Department of Medical Communication, Hoshi University, Tokyo, Japan

Corresponding Author:
Tsuyoshi Okuhara, MPH, PhD
Department of Health Communication
The University of Tokyo
7 Chome-3-1 Hongo
Tokyo, 113-8654
Japan
Phone: 81 3 5800 6549
Email: okuhara-ctr@umin.ac.jp

Abstract

Background: The COVID-19 pandemic emerged in the digital age and has been called the first "data-driven pandemic" in human history. The global response demonstrated that many countries had failed to effectively prepare for such an event. Learning through experience in a crisis is one way to improve the crisis management process. As the world has returned to normal after the pandemic, questions about crisis management have been raised in several countries and require careful consideration.

Objective: This review aimed to collect and organize public health professionals' experiences in crisis communication to the public during the COVID-19 pandemic.

Methods: We searched PubMed, MEDLINE, CINAHL, Web of Science, Academic Search Complete, PsycINFO, PsycARTICLES, and Communication Abstracts in February 2024 to locate English-language articles that qualitatively investigated the difficulties and needs experienced by health professionals in their communication activities during the COVID-19 pandemic.

Results: This review included 17 studies. Our analysis identified 7 themes and 20 subthemes. The 7 themes were difficulties in pandemic communication, difficulties caused by the "infodemic," difficulties in partnerships within or outside of public health, difficulties in community engagement, difficulties in effective communication, burnout among communicators, and the need to train communication specialists and establish a permanent organization specializing in communication.

Conclusions: This review identified the gaps between existing crisis communication guidelines and real-world crisis communication in the digital environment and clarified the difficulties and needs that arose from these gaps. Crisis communication strategies and guidelines should be updated with reference to the themes revealed in this review to effectively respond to subsequent public health crises.

Trial Registration: PROSPERO CRD42024528975; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=528975

International Registered Report Identifier (IRRID): RR2-10.2196/58040

(JMIR Infodemiology 2025;5:e66524) doi:10.2196/66524

KEYWORDS

COVID-19; health communication; infodemic; misinformation; social media; SARS-CoV-2; pandemic; infectious; digital age; systematic review; internet; public health; government; health professional; crisis communication; qualitative; disinformation; eHealth; digital health; medical informatics

XSL•FO
RenderX

Introduction

The COVID-19 pandemic claimed millions of lives. It resulted in a public health crisis and caused economic and social turmoil worldwide. No country, irrespective of region or wealth, was spared the devastating effects of the COVID-19 pandemic. Given that there were no available drugs or vaccines early in the pandemic, communication was an important means of containing the crisis. Even after vaccines were developed, communication to increase trust in the vaccines was central to ending the crisis. Therefore, communication is essential in dealing with a pandemic [1].

Before the COVID-19 outbreak, crisis communication guidelines had been published by the World Health Organization (WHO) [2-4] and crisis communication strategies had been studied by researchers [5-9]. However, when the COVID-19 pandemic started, public health organizations worldwide acknowledged their lack of preparation and training for effective communication during such chaos [10-15]. Furthermore, the communication technology infrastructure has become increasingly complex over the last few decades. Social media platforms now seamlessly connect people to both accurate and false information, which tends to flow to recipients faster than viruses spread [16]. During the COVID-19 pandemic, public health organizations worldwide experienced difficulties with the “infodemic” of misinformation on social media [17]. Before the pandemic, researchers had recognized the importance of management of misinformation and studied countermeasures [18-21]. However, the COVID-19 pandemic highlighted the inexperience of public health agencies in dealing with the influence of misinformation during an emergency [22,23]. Therefore, the COVID-19 pandemic presented public health agencies with unprecedented challenges and highlighted the need to update existing crisis management communication strategies. A crisis is an important opportunity for learning; learning through experience in a crisis is the only way to improve the crisis management process [24,25]. Now that the world has returned to normal following the pandemic, questions requiring reflection have been raised about the crisis management in each country. Therefore, studies are needed to collect and organize data on public health professionals’ experiences in crisis communication worldwide during the COVID-19 pandemic. This work is essential for updating crisis communication strategies to prepare for subsequent public health crises.

We conducted a systematic review of qualitative studies that focused on public health professionals’ experiences in crisis communication during the COVID-19 pandemic in diverse countries. We examined the difficulties that public health professionals experienced during the COVID-19 pandemic, the challenges they faced in overcoming those difficulties, and the needs to be met in future public health crises. We also discussed the gaps between existing crisis communication guidelines and real-world experiences in the COVID-19 pandemic that need to be bridged going forward.

Methods

Overview

This systematic review followed the guidelines provided in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [26]. In addition, we referred to the Sample, Phenomenon of Interest, Design, Evaluation, Research type tool for the synthesis of qualitative evidence [27]. The protocol was previously published [28] and registered with the international Prospective Register of Systematic Reviews (registration: CRD42024528975).

Literature Search

We searched the following databases on February 7, 2024: PubMed, MEDLINE, CINAHL, Web of Science, Academic Search Complete, PsycINFO, PsycARTICLES, and Communication Abstracts. We filtered our database searches to include articles published from January 1, 2020, to January 31, 2024. We used a combination of keywords with reference to previous studies to search the abstracts in these databases [29-31]: “((government*) OR (ministr*) OR (department*) OR (office*) OR (municipalit*) OR (prefecture*) OR (province*) OR (state*) OR (count*) OR (organization*) OR (institution*) OR (center*) OR (agenc*) OR (sector*) OR (authorit*)) AND ((covid-19) OR (coronavirus) OR (sars-cov-2)) AND ((interview*) OR (focus group*) OR (questionnaire*) OR (survey*)) AND ((communicat*) OR (messag*) OR (inform*) OR (recommend*) OR (announce*)) AND ((qualitative) OR (mix method)).”

Study Selection

We used Rayyan software (Qatar Computing Research Institute) [32] to screen the identified studies and automatically remove duplicates. Study inclusion and exclusion criteria are shown in Textboxes 1 and 2.

Titles and abstracts were independently screened to identify eligible studies using selection criteria established by the first author (TO) and the second author (MT). Then, the full texts of the remaining studies were screened independently by the first and second authors. Any disagreements during the screening process were discussed until consensus was reached, assisted by the third author (HO), as necessary.

XSL•FO
RenderX

Textbox 1. Study inclusion criteria.

- The study aim was to investigate public health professionals' experiences in crisis communication during the COVID-19 pandemic in the digital age with the infodemic of misinformation on social media.
- Regarding study content: qualitative studies of communications from governments and public health agencies to the public focusing on addressing the infodemic of misinformation on social media platforms.
- Regarding design: studies with qualitative data (eg, interviews, documents, and free-text questionnaire responses), those that used content analysis of qualitative data, reviews of qualitative studies, and mixed methods studies with qualitative results that met the study aim.
- Studies on individuals (irrespective of age, gender, ethnicity, or nationality), such as officials, health professionals, and researchers working for governments and public health agencies.
- Gray literature (information produced outside traditional publishing and distribution channels, such as conference proceedings and theses) if sufficient information was provided to confirm its eligibility (ie, full-length descriptions of research objectives, methods, results, discussion, and conclusions).
- Papers written in English and conducted from (and including) January 2020.

Textbox 2. Study exclusion criteria.

- Quantitative studies with quantitative data (eg, observational and interventional studies)
- Studies on journalists in media companies, patients, and the public
- Studies not published in full-text format
- Non–English-language papers
- Studies that did not meet the study aim that public health professionals' experiences in crisis communication during the COVID-19 pandemic in the digital age with the infodemic of misinformation on social media (eg, those on content analysis of media information, information searches by the public, COVID-19 patient management in hospitals, and patient-provider communication)

Quality Assessment

The Joanna Briggs Institute Critical Appraisal Checklist for Qualitative Research was used to assess the methodological quality of eligible studies [33,34]. This Joanna Briggs Institute checklist assesses the descriptive, interpretative, theoretical, and evaluative validity of qualitative studies. The 10 items of the checklist are evaluated as "yes," "no," "unclear," or "not applicable." The first (TO) and second (MT) authors independently performed quality assessments of the included studies. Any disagreements were discussed until consensus was reached, assisted by the third author (HO) as necessary.

Data Synthesis

Thematic synthesis was used to synthesize the collected data [35]. Thematic synthesis is recommended as a systematic method for synthesizing qualitative evidence [36]. In the first stage, free line-by-line coding of texts and quotations in the results and discussion sections of the included studies was conducted by TO. Next, 2 reviewers (TO and MT) independently grouped similar codes and generated data-driven descriptive themes. Consensus was reached through discussion, and the third reviewer (HO) was consulted when necessary. Finally, TO developed analytical themes by organizing the descriptive themes generated in the previous stage. This process of developing analytical themes involved repeated discussions among TO, MT, and HO.

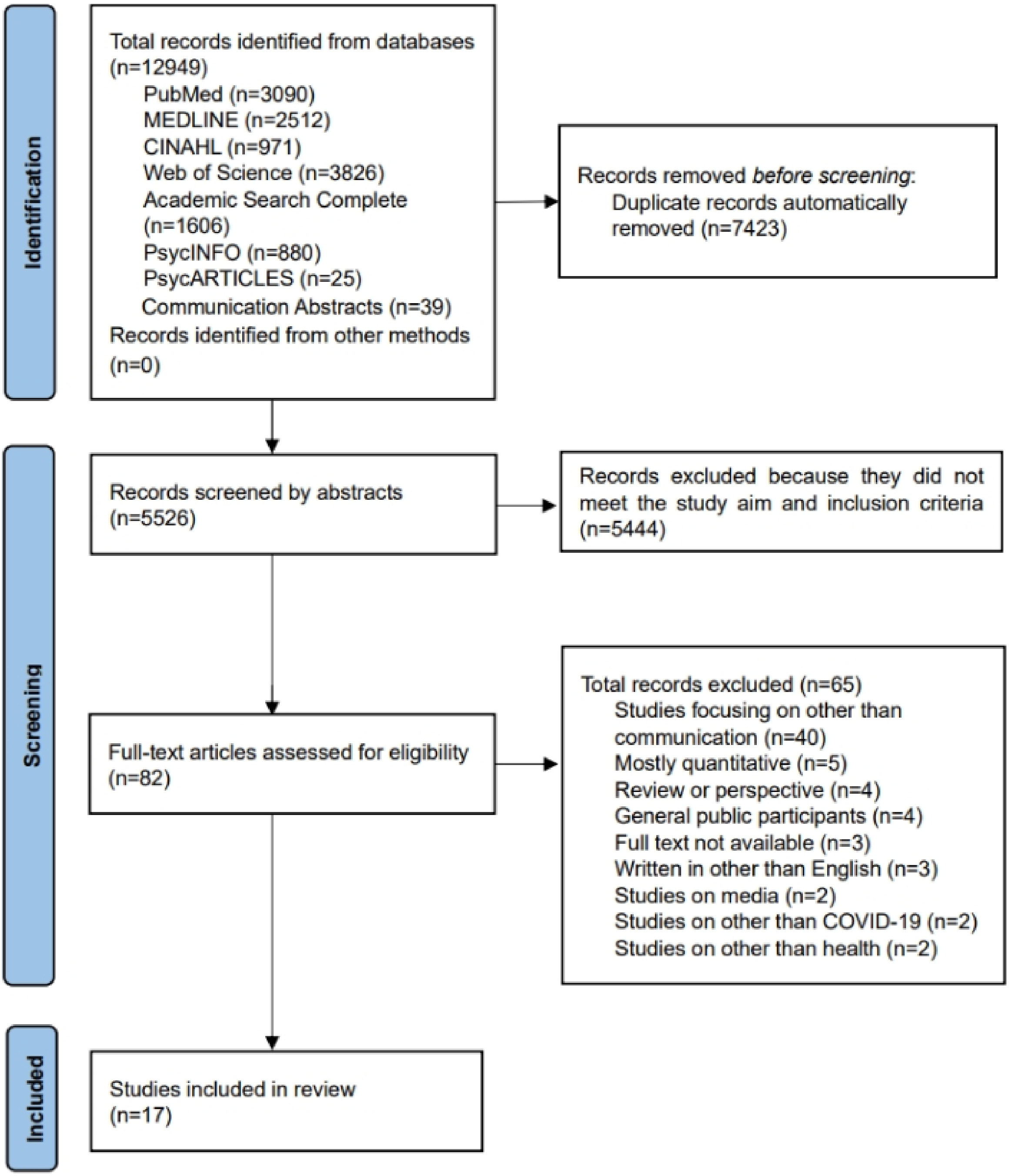
Results

Study Characteristics

Figure 1 shows the PRISMA flow diagram of the study selection. We included 17 studies in this review. Table 1 shows the characteristics of the included studies. A total of 5 studies were conducted in the United States, 4 in Canada, 2 in Switzerland, and 2 in Iran, and the other studies included participants from Europe, the Middle East, Asia, South America, and Africa. Participants' occupations included communication specialists, medical professionals, scientists, and officials in public health institutions and local municipalities. The median number of study participants was 20 (IQR 12.5-26), and 367 health professionals were represented overall. The time frame in which the data were collected was from March 2020 to December 2022. The included studies showed an overall good methodological quality; the median number of studies classified as "yes" was 8 (IQR 7-9). Results of the quality appraisal are shown in Multimedia Appendix 1 [11,15,37-51].

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) search process flowchart.



XSL•FO
RenderX

Table 1. Study characteristics of qualitative studies regarding crisis communication during the COVID-19 pandemic.

Study	Country	Settings	Participants, n	Period when the data were collected	Study aim
Atighechian et al [15], 2021	Iran	Universities, governments, and hospitals	Health professionals and experts, including university faculty members, policy makers, physicians, and nurses working in the infectious disease unit (n=19)	March 2020-June 2020	To identify the challenges of COVID-19-related information among people in point of experts' views
Nehushtan et al [41], 2023	Israel	14 municipalities	Officials in local municipalities, including chief executive officers, mayors, and officials responsible for health in emergencies (n=42)	October 2020-February 2021	To explore local municipalities' management of the COVID-19 pandemic
Sears et al [43], 2024	United States	One state	Public health workers, including sanitarian, educator, and administrative positions (n=11)	October 2020-March 2021	To gain an in-depth perspective of public health workers' experiences during the complex and dynamic climate the COVID-19 pandemic
Colman et al [37], 2021	Belgium, the Netherlands, United Kingdom, Sweden, and Germany	Academic or public health research institutions	Scientists with an official government advisory role during the pandemic (n=21)	December 2020-April 2021	To explore the views and experiences of scientists working on government advisory boards
Bravo et al [40], 2023	Paraguay, Uruguay, United States, Canada, Germany, Spain, New Zealand, Australia, and South Korea	Universities, governments, a research center, a health care center, and a nongovernmental organization	Experts with experience in health crisis management or risk communication (n=10)	December 2020-March 2021	To identify a framework for risk communication during health crises from the voices of international experts by using the COVID-19 pandemic as a case study
Rubinelli et al [11], 2023	Switzerland	Institutions responsible for communicating with the public at the national and cantonal levels	Individuals responsible for public institutional communication within key public health institutions (n=25)	January 2021-July 2021	To collect individual experiences of communicating the situation and protective measures to the public
Ort and Rohrbach [42], 2024	Switzerland	Key Swiss public health institutions at the federal and cantonal levels	Individuals responsible for public institutional communication within key public health institutions (n=25)	January 2021-July 2021	To explore public health institutions' challenges in implementing their COVID-19–related communication strategies
Pringle et al [50], 2022	Canada	Vancouver, the most diverse area	Communication specialists, medical professionals, officials in community service organizations, and volunteer community advocates (n=27)	May 2021-November 2021	To examine how community leaders, advocates, and public health communication specialists have approached community engagement
Engdawork et al [45], 2024	Ethiopia	A capital city	Stakeholders in the local government and private sectors engaged in social interventions to prevent COVID-19 (n=21)	September 2021-October 2021	To investigate the effectiveness of structural interventions during the earlier period of the pandemic in promoting adoption of preventive actions, challenges encountered during implementation, and draw lessons for future pandemic responses in low- and middle-income settings

XSL•FO
RenderX

Study	Country	Settings	Participants, n	Period when the data were collected	Study aim
Dubé et al [38], 2022	Canada	One province	Communication specialists in charge of developing health authorities’ COVID-19 communication and health care professionals actively engaged in public discussion in traditional and social media (n=11)	September 2021-December 2021	To explore how communication specialists working in health and governmental institutions and health care professionals have communicated about COVID-19
Lowe et al [39], 2022	Canada	11 jurisdictions	Public health officials, frontline health care workers, health scholars (social, epidemiological, policy, and clinical researchers), and health care worker union leaders (n=34)	September 2021-December 2021	To assess COVID-19 pandemic public health messaging for its potential to encourage or undermine public trust and adherence
Ittefaq [44], 2023	United States	Three states	Communication officials working in local health departments (n=14)	February 2022-April 2022	To explore challenges in information dissemination on social media, and factors contributing to burnout among communication officials
Kamruzzaman et al [49], 2023	Bangladesh	Three divisions that reported the highest COVID-19 cases	Health professionals, including district-level health education officers, residential medical officers, and pertinent national specialists (n=14)	February 2022-May 2022	To understand how the social context influences risk communication and community response during the COVID-19 pandemic
Bates et al [46], 2023	United States	One county	Public health professionals working at city health departments and a county health department (n=7)	March 2022-May 2022	To determine how public health officials perceived misinformation and political polarization during the pandemic, and to learn more about strategies county health officials used to combat misinformation
Strand et al [48], 2023	United States	Midwestern states	Public health professionals in local and state public health departments, universities, and health care organizations (n=48)	Summer of 2022	To describe the lived experiences of public health professionals working during the COVID-19 pandemic and to provide lessons learned and best practices to inform preparation for a future infectious disease pandemic
Bazrafshan et al [47], 2023	Iran	Provincial and national public health institutions	Public health professionals across provincial and national health authorities (n=20)	October 2022-December 2022	To develop a conceptual framework for health risk communication and infodemic management during epidemics and health emergencies
Johnston et al [51], 2024	South Africa and Zambia	18 community health organizations	Individuals working in community health organizations with engagement in health education and information services (n=18)	Not mentioned	To investigate the strategies, challenges, and needs of community health organizations involved in public COVID-19 education to understand their role in public health crises in relation to communicating health information

XSL•FO
RenderX

Data Synthesis

Our analysis identified 242 free codes, which were organized into 41 descriptive themes: 7 analytical themes and 20 subthemes. Multimedia Appendix 2 [11,15,37-51] shows the analytical themes and subthemes, the studies that contributed to those themes, and direct quotations from the included studies to support those themes.

Difficulties in Pandemic Communication

Gap Between Scientific Uncertainty and Expectations of Certainty

The COVID-19 pandemic revealed a gap between the normal reality of scientific uncertainty and political and public expectations of certainty, which made public health communication difficult [11,37-40]. The traditional scientific method of generating, evaluating, and acting on evidence was incompatible with the urgency of the pandemic [11,37]. However, participants were required by policy makers and citizens to provide rapid, definitive conclusions and explanations based on uncertain evidence in an uncertain situation [11,37,38] (quotation 1). This demand contrasted with the "slowness" of science [37]. Changes were unfolding rapidly in terms of scientific knowledge, the spread of the infection, and political, economic, and social conditions, and this required several changes in public health policies over a short time [37,38] (quotation 2). The gap between the uncertainty of science and unrealistic expectations of certainty resulted in public criticism of public health professionals and difficulties in public health communication [11,37-40] (quotation 3).

Communication Challenges in a "Slow Disaster"

Participants described the characteristics of the COVID-19 pandemic as a "slow disaster" [40]. Most disasters are short-lived, but the nature of the COVID-19 pandemic meant that they had to continuously deal with changing circumstances [11,40,41]. In the early stages of the pandemic, citizens cooperated with public health recommendations [11]. However, over time, their patience waned, their trust in public health professionals declined, and compliance worsened [11,40,41] (quotation 4). In addition, health professionals experienced difficulty in using communication to encourage citizens to adopt preventive behaviors amid fatigue from a pandemic with no seeming end in sight [11,40-42] (quotation 5).

Difficulties Caused by the Infodemic

Difficulties in Public Health Activities Due to Misinformation

Misinformation about the severity and mortality of COVID-19 and the safety of vaccines spread on social media, and affected citizens' attitudes and behaviors [11,15,42-47] (quotations 6 and 7). Participants were forced to devote significant resources to identifying and correcting misinformation [11,15] (quotations 8 and 9) but did not have effective measures to counter the sensational communication strategies used by purveyors of misinformation [11,39,43] (quotation 10). It was also more difficult to persuade people who had acquired a skeptical attitude through misinformation than it was to simply convey correct information [11,42,43,45].

Countering Misinformation

During the COVID-19 pandemic, participants learned several strategies to deal with misinformation. The first was the importance of timely communication; it was crucial that participants disseminated messages before misinformation spread [11,15,47] (quotation 11). Second, participants recognized that social listening improved their understanding of citizens' psychosocial aspects and information needs, as well as the quality of information they provided [11] (quotations 12 and 13). Third, participants had to recognize and address fear and anxiety among citizens [39] (quotation 14). Finally, participants recognized the importance of actively using social media to disseminate accurate information and guide people to reliable information sources [11,38,39,46,47] (quotation 15). However, the lack of human resources with expertise in using social media made it difficult to counter misinformation using these platforms [11] (quotation 16).

Difficulties in Partnerships Within and Outside Public Health

Tensions Within the Community of Public Health Experts

Participants recognized the importance of public health agencies partnering with epidemiologists, data scientists, sociologists, communication scholars, and other professionals with unique expertise for developing and implementing pandemic communication strategies [11,15,37,38,40,42,47-49]. This was because pandemic communication had to incorporate consideration of the social, economic, and political context that unfolded along with the health crisis [15,37,40,47-49] (quotation 17). However, there were coordination difficulties, especially in the early stages of the pandemic. Expert committees tended to be dominated by biomedical and virological researchers and often excluded sociologists and anthropologists [37] (quotation 18). A reason cited for the limited effectiveness of communication to citizens was that the strategies used lacked an understanding of people's sociocultural beliefs [38,49].

Tensions Between Public Health and Politics

The conflict of interest between health care and the economy was a major factor that characterized the communication difficulties during the COVID-19 pandemic [11,15,37-39,42,43,46,48,49] (quotation 19). The conflict between safeguarding public health and maintaining the economy abrogated the coherence of policy decisions and messages to the public and led to public confusion and distrust of public health [15,37-39,48]. This conflict of interest between health care and the economy also created tensions between public health professionals and political leaders who wanted to maintain their political popularity [37,48]. At the policy-making level, some political leaders did not accept or use the scientific evidence provided by public health experts [37,39,48,49]. Moreover, political leaders sometimes used and abused public health professionals to evade their own responsibilities in communicating with the public [37,46] (quotation 20). At the policy practice level, public health professionals were sometimes obstructed by political leaders from recommending preventive

XSL•FO
RenderX

behaviors and vaccination for citizens, rather than receiving political support [43,46] (quotations 21 and 22).

Difficulties in Coordination Between Public Health and Mass Media

Participants recognized the importance of close collaboration with the mass media [11,15,37,42,43,47]. They understood the influence of mass media in shaping public opinion and journalists' commitment to scientifically accurate and balanced reporting [15,37,47]. However, they recognized that during the COVID-19 pandemic, the mass media often engaged in misleading reporting, as well as pitting public health professionals against each other and politicians against public health professionals, quoting out of context, and linking public health professionals to specific political decisions [11,37,42,43]. In addition, some participants perceived that the biased discussion and criticism of public health activities in mass media coverage led to a decline in people's trust in public health [43] (quotations 23 and 24).

Difficulties in Community Engagement

Need to Tailor Communication to Community Realities

Some participants recognized that many of the COVID-19 recommendations were not consistent with community realities [45,50,51]. For example, small living quarters, large families, and essential travel by public transportation to buy food and work affected compliance with preventive behaviors such as social distancing [45,51]. Some citizens had to prioritize other essential living activities over infection prevention behaviors. For example, people of lower economic status had to go out to earn their living even during lockdown periods [45,51] (quotations 25 and 26). Compliance with COVID-19 prevention recommendations meant that many citizens faced economic hardship, food insecurity, domestic violence, and mental health problems [45].

Need to Consider Local Cultural Factors

Cultural factors such as a given community's dominant religion could also pose a barrier to compliance with the COVID-19 prevention recommendations [38,39,45,50] (quotations 27 and 28). However, participants recognized that cultural factors could act as both facilitators and inhibitors of public health activities [50]. They adapted their communication strategies to reflect community-specific sociocultural factors and incorporated ideas such as using culturally significant meeting places (eg, local religious centers) [45,50] (quotations 29 and 30).

Need for Bottom-Up and 2-Way Communication

Participants identified that the effectiveness of communication from health professionals to communities was inhibited by its 1-way nature [45,47,49,51] (quotation 31). Community groups and leaders were involved in implementing infection prevention programs; however, they had little involvement in planning and designing feasible programs [45]. Community participation tended to be lower when information was distributed from public health agencies to communities in a top-down manner. These top-down communication strategies, which lacked collaboration with the community, inhibited acceptance of recommendations for preventive behaviors [45,47,49]. This suggested that bottom-up and 2-way communication that involved the community were required to foster community engagement [45,47,49] (quotation 32).

Need to Build Trust With Communities

Participants generally responded that a trusting relationship between public health and the community was a factor in increasing community engagement [39,40,45,46,48,50,51]. They noted the importance of building trust with local political, religious, business, and agricultural leaders, along with schools, newspapers, radio stations, and other local organizations [40,48,50,51] (quotations 33 and 34). Existing local networks were especially important in developing grassroots communication activities [46] (quotation 35). In addition to trusting relationships with organizations, participants stated the importance of one-on-one trust relationships between health professionals and local residents [39,46] (quotation 36). However, in areas where public health outreach services had been reduced in the years before the pandemic, it was difficult to quickly rebuild trust between public health professionals and the community during the pandemic [50].

Need for Communication Through Community Channels

Communication through community-specific communication channels, such as local television and radio stations, social media platforms, and connections with trusted individuals, were emphasized as ways to increase community engagement [11,38,40,41,44-46,49-51] (quotation 37). Formal and informal communications were developed, including traditional media campaigns and disseminating messages via social media [11,40,41,45,50,51] (quotation 38). Participants noted that the key communication channels, including newspapers, radio, and social media, varied by community resident group [11,40,44,45,50,51] (quotation 39). For those groups using social media in particular, attempts were made to increase their engagement by encouraging their participation in communication activities [45,46] (quotation 40).

Difficulties in Effective Communication

Need for Uniformity and Promptness in Communication

Participants identified the absence of reliable sources of information known to citizens as an impediment to effective communication [11,15] (quotation 41). The plethora of available information sources, including mass media and social media, created confusion among citizens [11,15,43] (quotation 42). Furthermore, the importance of rapid information dissemination was crucial in communication regarding a hitherto unknown infectious disease [11,15,39,44,47] (quotations 43 and 44). However, participants faced a dilemma whereby prioritizing the speed of communication did not allow sufficient time to create effective messages. For example, translation into multiple languages was time-consuming [11,44] (quotations 45 and 46). In addition, it took time to crunch the vast amount of information and create concise, clear messages [11,44,45] (quotation 47).

Need for Understandable and Persuasive Communication

Participants emphasized the importance of efforts to ensure the public understood messages [11,37-40,42,45,48] (quotation 48).

XSL•FO
RenderX

These messages needed to have a clear purpose, use plain language and illustrations, and be persuasive to be easily understood and accepted by all citizens [11,38-40,45,48]. However, participants experienced difficulties in creating messages that addressed the various levels of citizens' individual health literacy [38,39,42,45] (quotation 49). Understandable communication was also important for politicians and policy makers who did not necessarily have basic scientific knowledge [37].

Need for Communication to Empower People

Participants noted the harms associated with health authorities generating stigma for certain populations [38-40]. For example, they accused young adults of often failing to follow recommendations for social distancing, and therefore, transmitting the virus, or of prolonging the pandemic by not being vaccinated [38,39] (quotation 50). They stressed that effective communication strategies should emphasize helping people make better informed decisions rather than punishing them with blame or fear or offering temporary reassurance [38,40] (quotations 51 and 52).

Burnout Among Communicators

Difficulties With Information Overload and Requests

Participants indicated that they felt like they were drowning in an overwhelming influx of information related to COVID-19 [11,41,51]. They tried to extract relevant information from this torrent; however, they did not know how to do so [51] (quotation 53). In addition, they were under intense pressure from the community to share the latest information about the novel virus [11,44] (quotation 54). Furthermore, public health professionals were expected to respond to constant media requests for updated information [11] (quotation 55).

Lack of Trust in Public Health

Participants experienced a lack of public trust, which led to communication difficulties [11,15,39,40,44,48,51]. A major contributing factor to this was discrepancies in the information disseminated by the government, municipalities, public health agencies, and professionals [15,39,48] (quotations 56 and 57). The confusion caused by these discrepancies increased people's distrust and decreased their willingness to accept infection prevention recommendations [15,40,44] (quotations 58 and 59).

Attacks on Public Health Professionals by Citizens

Participants experienced criticism and attacks from citizens despite their best efforts to overcome the aforementioned difficulties [11,37,43,44,46,48] (quotation 60). Daily criticism and attacks from citizens through social media, email, and community face-to-face meetings accelerated burnout among participants [11,43,44,46,48] (quotations 61 and 62). Accordingly, they sought ways to prevent burnout, including learning to set emotional boundaries for criticism [43] (quotation 63). They noted that rare words of gratitude from citizens empowered them [37,43] (quotation 64).

Need to Train Communication Specialists and Establish a Permanent Organization

Need to Train Communication Specialists

There was a notable lack of human resources with communication expertise during the COVID-19 pandemic [11,42,44,45,47]. In the early stages of the pandemic, public health agencies made efforts to increase the number of communications personnel by reorganizing their human resources [11]. However, securing a sufficient number of communications personnel, relative to the overwhelming volume of information that needed to be addressed, was difficult [11,42,45,47] (quotation 65). Personnel who had been moved to communications duties from other departments often lacked basic communication skills and competencies [11,42,45]. In addition, even those who had been previously trained in communications lacked the experience and ability to communicate effectively in the emergency pandemic situation [11,42] (quotation 66).

Need to Establish a Permanent Organization Specializing in Communication

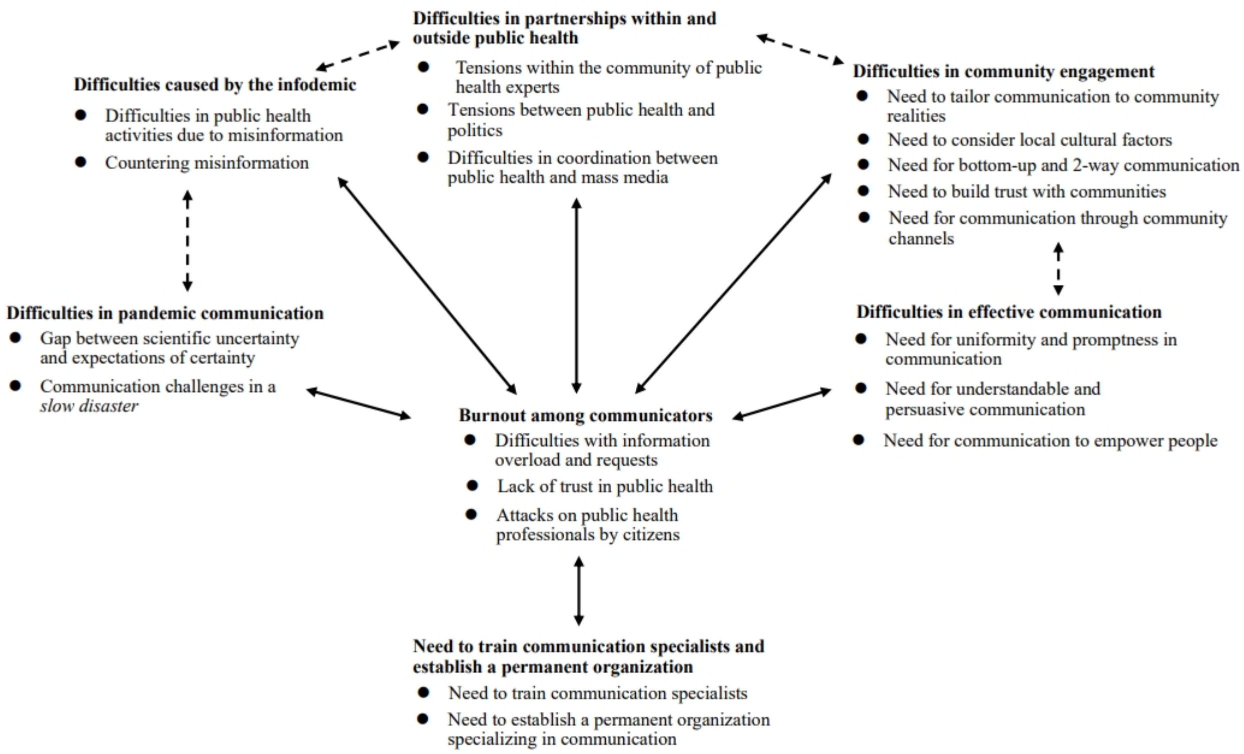
Participants identified the rigidity within existing organizational structures as a problem. They emphasized that the many procedures, time-consuming approval processes, and inflexible and rigid protocols in the organizations inhibited rapid and effective public health communication [11,50]. They agreed on the importance of establishing a permanent organization specializing in public health communication [40,41] and noted that such an organization should train communication specialists, accumulate methods for effective communication strategies, build cross-functional partnerships with other organizations, and establish a structure to respond quickly in further public health crises [47] (quotation 67).

Conceptual Model

Figure 2 shows the conceptual model developed from the above results. Themes 3.2.1 to 3.2.5 were interrelated, and the difficulties experienced by communicators resulted in their burnout (theme 3.2.6). The difficulties and needs indicated in themes 3.2.1 to 3.2.6 indicated the need for future training of communication specialists and establishing a permanent organization specializing in communication (theme 3.2.7). It was assumed that training experts and establishing organizations would reduce difficulties and enable effective communication in subsequent public health crises.

XSL•FO
RenderX

Figure 2. Conceptual model developed from qualitative studies regarding crisis communication during the COVID-19 pandemic.



Discussion

Principal Findings

This systematic review of qualitative studies examined the difficulties, challenges, and needs experienced by public health professionals during the COVID-19 pandemic and identified 7 themes. The theme of difficulties in pandemic communication encompassed difficulties stemming from scientific uncertainty and the "slow disaster." Public health crisis communication inherently involves uncertainty [52,53], and the WHO and the Centers for Disease Control and Prevention (CDC) recommended explicitly communicating information about uncertainties [1,54,55]. Researchers in crisis communication argued that communicating uncertainty increased rather than decreased public trust [56,57]. However, this systematic review revealed that risk communication in the real world is not as simple as the above recommendation suggests. Uncertainty reduction theory indicates that humans are intrinsically motivated to reduce uncertainty [58]. Therefore, communicating uncertainty creates a conflict with people's demand for certainty. However, when people's trust in their government and communicators is stronger, they tend to more successfully accept uncertainty [59]. People's trust in government and public health agencies may offer a clue to resolving communication difficulties associated with uncertainty. Furthermore, neither the WHO nor CDC guidelines contained details on how to deal with communication difficulties stemming from a slow disaster [1,54,55]. Coping with pandemic fatigue was one of the difficulties stemming from the slow disaster. Although previous studies have examined factors associated with pandemic fatigue during the COVID-19 pandemic [60,61], much remains unknown about pandemic fatigue. Further research should consider effective communication strategies for a slow disaster.

The communication difficulties in the COVID-19 pandemic were characterized by the destructive impact of the infodemic. A survey conducted in the United Kingdom in 2020 showed that 46% of the public had been exposed to fake news about COVID-19 and 40% said they could not tell the difference between truth and lies [62]. Previous studies have examined effective debunking methods for misinformation [18-21]. The CDC also developed public health infodemic surveillance systems in the wake of the COVID-19 pandemic [63]. Furthermore, there are more than 100 laws against disseminating misinformation in different countries worldwide [64]. A multifaceted approach is needed to prepare for future public health infodemics, including surveillance, communication, and legal regulation.

The WHO guidelines to address COVID-19 emphasized the importance of collaboration within public health agencies and with external partners [55]. However, this systematic review found that, in reality, tensions in and outside of public health agencies hindered an effective crisis response. During noncrisis periods, governments, public health agencies, researchers, and media are often siloed, making crisis-related coordination and information sharing difficult [65]. In addition, political and economic interests that conflict with public health policies hinder an effective pandemic response [66]. Such partnership failures, which were experienced in past epidemics and pandemics, were repeated in the COVID-19 pandemic. Addressing this is a crucial challenge going forward.

Existing public health organization guidelines emphasized the importance of community engagement strategies [1,54,55].

XSL•FO RenderX

Many studies have shown that community-based cultural factors were related to preventive behaviors and mortality rates during the COVID-19 pandemic [67-70]. Furthermore, language and cultural barriers prevented access to information, understanding of messages, and compliance with recommendations during the pandemic [71,72]. This systematic review showed that top-down, 1-way communication to the community hindered effective pandemic responses, despite the importance of a bottom-up approach that involves community stakeholders and residents in decision-making having been officially emphasized [73]. Communication in public health crises requires adapting communication strategies to the cultural, social, and demographic background of the local community to gain support among the target population [74]. To achieve this, it is important to break away from top-down, 1-way communication and adopt a 2-way, bottom-up approach that includes dialogue with the community [75].

During epidemics and pandemics, it is important that information from public health agencies is not overtaken by competing misinformation [25]. The first message that an audience receives shapes their subsequent attitudes [76]. Therefore, quick dissemination of information based on partial evidence is better than delayed dissemination of information based on complete evidence [1,55,77] because prompt communication is an essential principle of risk communication [54]. However, this systematic review revealed that the speed of communication hindered the effectiveness of communication during the COVID-19 pandemic. Public health professionals experienced difficulty in securing time for translation, pretesting, and creating easy-to-understand messages as they were under pressure to communicate quickly. The COVID-19 pandemic highlighted the difficulty of following existing crisis communication guidelines in a real-world crisis response.

Many public health professionals experienced burnout during the course of the pandemic. The main factors contributing to burnout were information overload that exceeded limited human resources, along with criticism and attacks on public health professionals from the public. The lack of public trust in public health also contributed to attacks against health professionals. The degree of trust in public institutions was associated with the rate of COVID-19 infection and the associated mortality rate [78]. A 2022 report by the Organisation for Economic Co-operation and Development highlighted that public trust was a key insight from the evaluation of responses to the pandemic, which pointed to the importance of building trust over a long period before a crisis occurs [73]. Building public trust and preventing burnout among public health professionals are essential for preparing for future public health crises.

The aforementioned 6 themes suggested the seventh theme, the need to train communication specialists and establish permanent organizations specializing in communication. These measures are necessary to address the aforementioned issues brought to light by the COVID-19 pandemic. COVID-19 showed that many countries had failed to learn the lessons of past global infections (eg, severe acute respiratory syndrome and influenza A virus subtype) and had failed to prepare for a future public health crisis [73,79]. Even now, many countries are still not prepared for future public health crises [80]. Another public health crisis occurring is not a matter of "if" but of "when" [81]. The best way to manage a crisis is to prevent one [25], and the second-best way to manage a crisis is to prepare for one [82]. All public health institutions and professionals must learn from the difficulties, challenges, and needs identified in this systematic review and update their strategies and guidelines to implement more effective communication in the next public health crisis.

Future Directions for Practitioners

The results of this systematic review suggest the following practice implications, which may help to prepare for the next public health crisis. (1) The scientific process is accompanied by uncertainty; however, politicians and citizens seek certainty. It is necessary to increase trust in public health organizations and address the communication difficulties associated with uncertainty, to address pandemic fatigue, and to develop effective communication strategies for future slow-onset disasters. (2) More research and practice are needed to manage misinformation in public health crises, including surveillance and communication strategies for "prebunking" and debunking information. (3) Partnerships between stakeholders at both the policy-making and communication practice levels are needed to manage public health crises. Such partnerships are important for enabling the creation and transmission of consistent messages, and avoiding confusion among citizens and distrust in public health. (4) It is necessary to build trusting relationships between public health organizations and communities before a crisis occurs and to enable bottom-up communication during crises. (5) It is also necessary to address the trade-off between communication promptness and effectiveness and conduct communication with the aim of empowerment. (6) Measures are needed to prevent burnout among health professionals during a crisis. (7) To address these issues and support an effective response to future public health crises, it is necessary to train more communications specialists, establish permanent organizations specializing in communication, and update strategies and guidelines.

Limitations

This systematic review had several limitations. First, we conducted a rigorous literature search and qualitative synthesis with 2 or more reviewers. However, we could not completely rule out the possibility that some relevant literature had not been included. Second, we did not weight the interpretation of study results according to the quality appraisal of the included studies; however, the included studies showed an overall good methodological quality. Third, because this was a systematic review of previous studies, our interpretations were limited by the data that were reported in the included studies. Fourth, participants in the included studies had various occupational backgrounds such as policy makers, officials in local municipalities, frontline health care workers, and scientists. A strength of this review was that it reflected the experiences of participants from diverse backgrounds; however, it was also limited by not differentiating experiences at the policy-making level from those at the policy implementation level on the front line. Fifth, another strength was that we included studies from various countries in Europe, the Middle East, Asia, Africa, and

XSL•FO
RenderX

North and South America; however, a limitation was that we did not make any economic or cultural distinctions. Finally, because all crises are novel and involve contextual differences, the generalizability of the findings and implications of this study to future crises is limited. Despite these limitations, this review has the important implications mentioned earlier, in that it identified the gaps between existing crisis communication guidelines and real-world crisis communication and the difficulties and needs that arise from those gaps.

Conclusions

This systematic review of qualitative studies identified the following issues that need to be addressed to prepare for subsequent public health crises. Despite the importance of collaboration within and outside public health and community engagement being highlighted in existing crisis communication guidelines, there was insufficient preparation and response to the COVID-19 pandemic. Although prompt communication is an essential principle for crisis response, the trade-off between promptness and the effectiveness of communication should be addressed. Difficulties specific to "slow disasters" and "infodemics" characterized the challenges encountered during the COVID-19 pandemic. Information overload, a shortage of human resources, and a lack of trust in public health contributed to burnout among health professionals. Public health professionals need to address the difficulties and needs identified in this systematic review by training communication specialists and establishing permanent organizations specializing in communication. One health professional described the difficulties resulting from the lack of preparation during the COVID-19 pandemic as "we are building the plane while we are flying" [44]. Of course, airplanes must be built before they fly, and in the case of a public health crisis, preparations must be made before the crisis arises.

Acknowledgments

This work was supported by the Japan Society for the Promotion of Science KAKENHI (20K10397).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed in this study protocol.

Authors' Contributions

TO was responsible for the conceptualization, methodology, data analysis, writing of the original draft, and funding acquisition. MT, HO, and RY were responsible for data analysis. TK was responsible for supervision. All authors contributed to reviewing and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Quality appraisal of included studies.
[DOC File , 54 KB - infodemiology_v5i1e66524_app1.doc]

Multimedia Appendix 2
Themes and illustrative quotes from qualitative studies regarding crisis communication during the COVID-19 pandemic.
[DOCX File , 32 KB - infodemiology_v5i1e66524_app2.docx]

Multimedia Appendix 3
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 34 KB - infodemiology_v5i1e66524_app3.pdf]

References

1. Communicating risk in public health emergencies: a WHO guideline for emergency risk communication (ERC) policy and practice. World Health Organization. 2018 Jan 10. URL: https://www.who.int/publications/i/item/9789241550208 [accessed 2025-02-19]
2. Effective media communication during public health emergencies: a WHO handbook. World Health Organization. 2012 Jun 16. URL: https://www.who.int/publications/i/item/effective-media-communication-during-public-health-emergencies-a-who-handbook [accessed 2025-02-19]
3. Outbreak communication: best practices for communicating with the public during an outbreak. World Health Organization. 2005 Jun 13. URL: https://www.who.int/publications/i/item/outbreak-communication-best-practices-for-communicating-with-the-public-during-an-outbreak [accessed 2025-02-19]

XSL•FO
RenderX

4. Communication for behavioural impact (COMBI): a toolkit for behavioural and social communication in outbreak response. World Health Organization. 2012. URL: https://www.asset-scienceinsociety.eu/pages/toolkit-behavioural-and-social-communication-outbreak-response [accessed 2025-02-19]
5. Reynolds B, W Seeger M. Crisis and emergency risk communication as an integrative model. J Health Commun 2005;10(1):43-55. [doi: 10.1080/10810730590904571] [Medline: 15764443]
6. Eckert S, Sopory P, Day A, Wilkins L, Padgett D, Novak J, et al. Health-related disaster communication and social media: mixed-method systematic review. Health Commun 2018 Dec;33(12):1389-1400. [doi: 10.1080/10410236.2017.1351278] [Medline: 28825501]
7. Winters M, Jalloh MF, Sengeh P, Jalloh MB, Conteh L, Bunnell R, et al. Risk communication and Ebola-specific knowledge and behavior during 2014-2015 outbreak, Sierra Leone. Emerg Infect Dis 2018 Feb;24(2):336-344 [FREE Full text] [doi: 10.3201/eid2402.171028] [Medline: 29350151]
8. Quinn P. Crisis communication in public health emergencies: the limits of 'legal control' and the risks for harmful outcomes in a digital age. Life Sci Soc Policy 2018 Feb 06;14(1):4 [FREE Full text] [doi: 10.1186/s40504-018-0067-0] [Medline: 29404722]
9. Takamura N, Orita M, Taira Y, Matsunaga H, Yamashita S. Experiences of crisis communication during radiation emergency and risk communication for recovery of the community in Fukushima. J Radiat Res 2021 May 05;62(Supplement_1):i95-100 [FREE Full text] [doi: 10.1093/jrr/rraa113] [Medline: 33978182]
10. Ortiz dos Santos F. Myths and misconceptions on COVID-19: ‘Congo check’ and ‘Talato’ verification experiences. Front Commun 2021 Mar 31;6 [FREE Full text] [doi: 10.3389/fcomm.2021.627214]
11. Rubinelli S, Häfliger C, Fiordelli M, Ort A, Diviani N. Institutional crisis communication during the COVID-19 pandemic in Switzerland. A qualitative study of the experiences of representatives of public health organizations. Patient Educ Couns 2023 Sep;114:107813 [FREE Full text] [doi: 10.1016/j.pec.2023.107813] [Medline: 37247524]
12. Suntayakorn C. Policy and the politics of COVID-19 communication responses: a comparative case study of two border provinces in Thailand. Asian Aff Am Rev 2022;49(4):217-244. [doi: 10.1080/00927678.2022.2126674]
13. Harish V, Samson TG, Diemert L, Tuite A, Mamdani M, Khan K, et al. Governing partnerships with technology companies as part of the COVID-19 response in Canada: a qualitative case study. PLOS Digit Health 2022 Dec;1(12):e0000164 [FREE Full text] [doi: 10.1371/journal.pdig.0000164] [Medline: 36812643]
14. Kajal F, Lall D, Chattu VK, Kumar S, Yadav A, Agarwal S, et al. Covid-19 lockdown governance in Uttar Pradesh, India: a call for equity? Cogent Soc Sci 2024;10(1):1-16. [doi: 10.1080/23311886.2023.2299084]
15. Atighechian G, Rezaei F, Tavakoli N, Abarghoian M. Information challenges of COVID-19: a qualitative research. J Educ Health Promot 2021 Jul 30;10:279 [FREE Full text] [doi: 10.4103/jehp.jehp_1271_20] [Medline: 34485576]
16. Briand SC, Cinelli M, Nguyen T, Lewis R, Prybylski D, Valensise CM, et al. Infodemics: a new challenge for public health. Cell 2021 Dec 09;184(25):6010-6014 [FREE Full text] [doi: 10.1016/j.cell.2021.10.031] [Medline: 34890548]
17. The Lancet Infectious Diseases. The COVID-19 infodemic. Lancet Infect Dis 2020 Aug;20(8):875 [FREE Full text] [doi: 10.1016/S1473-3099(20)30565-X] [Medline: 32687807]
18. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
19. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
20. Prike T, Ecker UK. Effective correction of misinformation. Curr Opin Psychol 2023 Dec;54:101712 [FREE Full text] [doi: 10.1016/j.copsyc.2023.101712] [Medline: 37944323]
21. van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. Nat Med 2022 Mar;28(3):460-467. [doi: 10.1038/s41591-022-01713-6] [Medline: 35273402]
22. Khan S, Mishra J, Ahmed N, Onyige CD, Lin KE, Siew R, et al. Risk communication and community engagement during COVID-19. Int J Disaster Risk Reduct 2022 May;74:102903 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.102903] [Medline: 35313476]
23. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
24. Schiffino N, Taskin L, Donis C, Raone J. Post-crisis learning in public agencies: what do we learn from both actors and institutions? Policy Stud 2016 May 29;38(1):59-75. [doi: 10.1080/01442872.2016.1188906]
25. Coombs WT. Ongoing Crisis Communication: Planning, Managing, and Responding. Thousand Oaks, CA: SAGE Publications; 2012.
26. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
27. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qual Health Res 2012 Oct;22(10):1435-1443. [doi: 10.1177/1049732312452938] [Medline: 22829486]

XSL•FO
RenderX

28. Okuhara T, Terada M, Okada H, Kiuchi T. Experiences of governments and public health agencies regarding crisis communication during the COVID-19 pandemic in the digital age: protocol for a systematic review of qualitative studies. JMIR Res Protoc 2024 Jun 27;13:e58040 [FREE Full text] [doi: 10.2196/58040] [Medline: 38935414]
29. Whitehead HS, French CE, Caldwell DM, Letley L, Mounier-Jack S. A systematic review of communication interventions for countering vaccine misinformation. Vaccine 2023 Jan 27;41(5):1018-1034 [FREE Full text] [doi: 10.1016/j.vaccine.2022.12.059] [Medline: 36628653]
30. Xia S, Nan X. Motivating COVID-19 vaccination through persuasive communication: a systematic review of randomized controlled trials. Health Commun 2024 Jul 31;39(8):1455-1478. [doi: 10.1080/10410236.2023.2218145] [Medline: 37254940]
31. Panickar R, Aziz Z, Mohd Sani N, Kamarulzaman A. The use of technology in vaccine safety communication: a systematic review of randomised controlled trials. Patient Educ Couns 2023 Jul;112:107707. [doi: 10.1016/j.pec.2023.107707] [Medline: 36989861]
32. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
33. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015 Sep;13(3):179-187. [doi: 10.1097/XEB.0000000000000062] [Medline: 26262565]
34. Hannes K, Lockwood C, Pearson A. A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. Qual Health Res 2010 Dec;20(12):1736-1743. [doi: 10.1177/1049732310378656] [Medline: 20671302]
35. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45 [FREE Full text] [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
36. Harden A, Thomas J. Thematic synthesis. The Cochrane Collaboration. URL: https://training.cochrane.org/resource/thematic-synthesis [accessed 2025-02-19]
37. Colman E, Wanat M, Goossens H, Tonkin-Crine S, Anthierens S. Following the science? Views from scientists on government advisory boards during the COVID-19 pandemic: a qualitative interview study in five European countries. BMJ Glob Health 2021 Sep;6(9):e006928 [FREE Full text] [doi: 10.1136/bmjgh-2021-006928] [Medline: 34580072]
38. Dubé È, Labbé F, Malo B, Pelletier C. Public health communication during the COVID-19 pandemic: perspectives of communication specialists, healthcare professionals, and community members in Quebec, Canada. Can J Public Health 2022 Dec;113(Suppl 1):24-33 [FREE Full text] [doi: 10.17269/s41997-022-00697-7] [Medline: 36131219]
39. Lowe M, Harmon SH, Kholina K, Parker R, Graham JE. Public health communication in Canada during the COVID-19 pandemic. Can J Public Health 2022 Dec;113(Suppl 1):34-45 [FREE Full text] [doi: 10.17269/s41997-022-00702-z] [Medline: 36329358]
40. Bravo P, Martinez-Pereira A, Fernández-González L, Dois A. What is needed to effectively communicate risk during a health crisis? A qualitative study with international experts based on the COVID-19 pandemic. BMJ Open 2023 May 09;13(5):e067531 [FREE Full text] [doi: 10.1136/bmjopen-2022-067531] [Medline: 37160386]
41. Nehushtan H, Barel H, Donchin M, Bar-Zeev Y. 'A war of a different kind': local municipalities' experiences in managing the COVID-19 pandemic in Israel-a qualitative study. BMJ Open 2023 Mar 15;13(3):e068226 [FREE Full text] [doi: 10.1136/bmjopen-2022-068226] [Medline: 36921952]
42. Ort A, Rohrbach T. (Dis)harmony in times of crisis? An analysis of COVID-related strategic communication by Swiss public health institutions. Public Health 2024 Mar;228:112-118 [FREE Full text] [doi: 10.1016/j.puhe.2023.12.033] [Medline: 38354580]
43. Sears JS, Maziarz L, Wilken L, Bussard M, Knippen K. Public health in the public eye: experiences of Ohio's public health workforce during COVID-19. Health Promot Pract 2024 Mar;25(2):227-234. [doi: 10.1177/15248399221124598] [Medline: 36154309]
44. Ittefaq M. "Beating your head against the wall": burnout on the rise among local health department communication officials. Health Secur 2023;21(6):440-449. [doi: 10.1089/hs.2023.0002] [Medline: 37962563]
45. Engdawork K, Amdework E, Assefa S, Ayode D, Tadele G. Experiences and lessons from structural interventions against COVID-19 in Addis Ababa, Ethiopia. Front Sociol 2024 Jan 26;9:1305549 [FREE Full text] [doi: 10.3389/fsoc.2024.1305549] [Medline: 38344393]
46. Bates G, Titi M, Dickson-Gomez J, Young S, Keval A, Meurer J. Navigating misinformation and political polarization of COVID-19: interviews with Milwaukee, Wisconsin county public health officials. Front Public Health 2023 Dec 21;11:1215367 [FREE Full text] [doi: 10.3389/fpubh.2023.1215367] [Medline: 38179566]
47. Bazrafshan A, Sadeghi A, Bazrafshan MS, Mirzaie H, Shafiee M, Geerts J, et al. Health risk communication and infodemic management in Iran: development and validation of a conceptual framework. BMJ Open 2023 Jul 30;13(7):e072326 [FREE Full text] [doi: 10.1136/bmjopen-2023-072326] [Medline: 37518081]
48. Strand MA, Sidhu S, Hohman A, Jansen R, McDonough S. Qualitative analysis of a roundtable exercise "reflecting on the COVID-19 experience of public health professionals". J Public Health Manag Pract 2023;29(4):E128-E136. [doi: 10.1097/PHH.0000000000001703] [Medline: 36727794]

XSL•FO
RenderX

49. Kamruzzaman M, Rahman A, Reidpath DD, Akhter S. Risk communication and community engagement in the context of COVID-19 response in Bangladesh: a qualitative study. Front Public Health 2023 Jan 05;11:1267446 [FREE Full text] [doi: 10.3389/fpubh.2023.1267446] [Medline: 38249367]
50. Pringle W, Sachal SS, Dhutt GS, Kestler M, Dubé È, Bettinger JA. Public health community engagement with Asian populations in British Columbia during COVID-19: towards a culture-centered approach. Can J Public Health 2022 Dec;113(Suppl 1):14-23 [FREE Full text] [doi: 10.17269/s41997-022-00699-5] [Medline: 36329357]
51. Johnston JS, Zhang Aluri K, Job N, Kuhnert KL, Prober C, Ward V, et al. Exploring the role of community health organizations in promoting public health during a health crisis: a qualitative study of COVID-19 responses in South Africa and Zambia. Glob Health Promot 2024 Mar;31(1):65-74 [FREE Full text] [doi: 10.1177/17579759231205854] [Medline: 37909401]
52. van der Bles AM, van der Linden S, Freeman AL, Mitchell J, Galvao AB, Zaval L, et al. Communicating uncertainty about facts, numbers and science. R Soc Open Sci 2019 May 08;6(5):181870 [FREE Full text] [doi: 10.1098/rsos.181870] [Medline: 31218028]
53. Seeger MW. Best practices in crisis communication: an expert panel process. J Appl Commun Res 2006;34(3):232-244. [doi: 10.1080/00909880600769944]
54. Crisis and emergency risk communication (CERC). Center for Disease Control and Prevention. URL: https://emergency.cdc.gov/cerc/manual/index.asp [accessed 2024-08-22]
55. COVID-19 global risk communication and community engagement strategy – interim guidance. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/covid-19-global-risk-communication-and-community-engagement-strategy [accessed 2025-02-19]
56. Rogers MB, Amlôt R, Rubin GJ, Wessely S, Krieger K. Mediating the social and psychological impacts of terrorist attacks: the role of risk perception and risk communication. Int Rev Psychiatry 2007 Jun;19(3):279-288. [doi: 10.1080/09540260701349373] [Medline: 17566905]
57. Han PK, Scharnetzki E, Scherer AM, Thorpe A, Lary C, Waterston LB, et al. Communicating scientific uncertainty about the COVID-19 pandemic: online experimental study of an uncertainty-normalizing strategy. J Med Internet Res 2021 Apr 22;23(4):e27832 [FREE Full text] [doi: 10.2196/27832] [Medline: 33769947]
58. Berger CR, Calabrese RJ. Some explorations in initial interaction and beyond: toward a developmental theory of interpersonal communication. Hum Commun Res 1975 Dec;1(2):99-112 [FREE Full text] [doi: 10.1111/j.1468-2958.1975.tb00258.x]
59. Liu BF, Bartz L, Duke N. Communicating crisis uncertainty: a review of the knowledge gaps. Public Relat Rev 2016 Sep;42(3):479-487. [doi: 10.1016/j.pubrev.2016.03.003]
60. Lai DW, Jin J, Yan E, Lee VW. Predictors and moderators of COVID-19 pandemic fatigue in Hong Kong. J Infect Public Health 2023 May;16(5):645-650 [FREE Full text] [doi: 10.1016/j.jiph.2023.03.003] [Medline: 36913768]
61. Taylor S, Rachor GS, Asmundson GJ. Who develops pandemic fatigue? Insights from latent class analysis. PLoS One 2022 Nov 10;17(11):e0276791 [FREE Full text] [doi: 10.1371/journal.pone.0276791] [Medline: 36355709]
62. Half of UK adults exposed to false claims about coronavirus. Ofcom. 2020 Apr 09. URL: https://www.ofcom.org.uk/about-ofcom/latest/features-and-news/half-of-uk-adults-exposed-to-false-claims-about-coronavirus [accessed 2024-08-22]
63. Chiou H, Voegeli C, Wilhelm E, Kolis J, Brookmeyer K, Prybylski D. The future of infodemic surveillance as public health surveillance. Emerg Infect Dis 2022 Dec;28(13):S121-S128 [FREE Full text] [doi: 10.3201/eid2813.220696] [Medline: 36502389]
64. Yadav K, Erdoğdu U, Siwakoti S, Shapiro JN, Wanless A. Countries have more than 100 laws on the books to combat misinformation. How well do they work? Bull At Sci 2021 May 11;77(3):124-128. [doi: 10.1080/00963402.2021.1912111]
65. Nelson S, Abimbola S, Jenkins A, Naivalu K, Negin J. Information sharing, collaboration, and decision-making during disease outbreaks: the experience of Fiji. J Decis Syst 2021 May 13;31(1-2):171-188. [doi: 10.1080/12460125.2021.1927486]
66. Abraham T. Risk and outbreak communication: lessons from alternative paradigms. Bull World Health Organ 2009 Aug;87(8):604-607 [FREE Full text] [doi: 10.2471/blt.08.058149] [Medline: 19705010]
67. Huynh TL. Does culture matter social distancing under the COVID-19 pandemic? Saf Sci 2020 Oct;130:104872 [FREE Full text] [doi: 10.1016/j.ssci.2020.104872] [Medline: 32550745]
68. Chen C, Frey CB, Presidente G. Culture and contagion: individualism and compliance with COVID-19 policy. J Econ Behav Organ 2021 Oct;190:191-200 [FREE Full text] [doi: 10.1016/j.jebo.2021.07.026] [Medline: 34566218]
69. Gupta M, Shoja A, Mikalef P. Toward the understanding of national culture in the success of non-pharmaceutical technological interventions in mitigating COVID-19 pandemic. Ann Oper Res 2022;319(1):1433-1450 [FREE Full text] [doi: 10.1007/s10479-021-03962-z] [Medline: 33558781]
70. Gelfand MJ, Jackson JC, Pan X, Nau D, Pieper D, Denison E, et al. The relationship between cultural tightness-looseness and COVID-19 cases and deaths: a global analysis. Lancet Planet Health 2021 Mar;5(3):e135-e144 [FREE Full text] [doi: 10.1016/S2542-5196(20)30301-6] [Medline: 33524310]
71. Brekke JP. Informing hard-to-reach immigrant groups about COVID-19—reaching the Somali population in Oslo. J Refug Stud 2022 Mar;35(1):641-661. [doi: 10.1093/jrs/feab053]

XSL•FO
RenderX

72. Hourani L, Lambert S, Peeler R, Lane B, Borst C. Graphic novels: a new stress mitigation tool for military training: developing content for hard-to-reach audiences. Health Commun 2017 May;32(5):541-549. [doi: 10.1080/10410236.2016.1140265] [Medline: 27548497]
73. First lessons from government evaluations of COVID-19 responses: a synthesis. Organization for Economic Cooperation and Development. 2022 Jan 21. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/01/first-lessons-from-government-evaluations-of-covid-19-responses-a-synthesis_3638f4ec/483507d6-en.pdf [accessed 2025-02-19]
74. Lanham A, Lubari E, Gallegos D, Radcliffe B. Health promotion in emerging collectivist communities: a study of dietary acculturation in the South Sudanese community in Logan City, Australia. Health Promot J Austr 2022 Jan;33(1):224-231 [FREE Full text] [doi: 10.1002/hpja.491] [Medline: 33848390]
75. Warren GW, Lofstedt R. Risk communication and COVID-19 in Europe: lessons for future public health crises. J Risk Res 2021 Jul 08;25(10):1161-1175. [doi: 10.1080/13669877.2021.1947874]
76. Covello VT. Best practices in public health risk and crisis communication. J Health Commun 2003;8 Suppl 1:5-8; discussion 148-51. [doi: 10.1080/713851971] [Medline: 14692565]
77. Rubinelli S, Purnat TD, Wilhelm E, Traicoff D, Namageyo-Funa A, Thomson A, et al. WHO competency framework for health authorities and institutions to manage infodemics: its development and features. Hum Resour Health 2022 May 07;20(1):35 [FREE Full text] [doi: 10.1186/s12960-022-00733-0] [Medline: 35525924]
78. COVID-19 National Preparedness Collaborators. Pandemic preparedness and COVID-19: an exploratory analysis of infection and fatality rates, and contextual factors associated with preparedness in 177 countries, from Jan 1, 2020, to Sept 30, 2021. Lancet 2022 Apr 16;399(10334):1489-1512 [FREE Full text] [doi: 10.1016/S0140-6736(22)00172-6] [Medline: 35120592]
79. Maxmen A. Has COVID taught us anything about pandemic preparedness? Nature 2021 Aug;596(7872):332-335. [doi: 10.1038/d41586-021-02217-y] [Medline: 34389832]
80. Global health security index 2021 report and data. Global Health Security Index. URL: https://ghsindex.org/report-model/ [accessed 2025-02-19]
81. Wucker M. The Gray Rhino: How to Recognize and Act on the Obvious Dangers We Ignore. New York City, NY: St. Martin's Publishing Group; 2016.
82. Frandsen F, Johansen W. Organizational Crisis Communication: A Multivocal Approach. Thousand Oaks, CA: SAGE Publications; 2016.

Abbreviations

CDC: Centers for Disease Control and Prevention
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WHO: World Health Organization

Edited by T Purnat; submitted 15.09.24; peer-reviewed by M Sanjeev, N Pelizzari; comments to author 11.10.24; revised version received 26.10.24; accepted 23.01.25; published 14.03.25.

Please cite as:
Okuhara T, Terada M, Okada H, Yokota R, Kiuchi T
Experiences of Public Health Professionals Regarding Crisis Communication During the COVID-19 Pandemic: Systematic Review of Qualitative Studies
JMIR Infodemiology 2025;5:e66524
URL: https://infodemiology.jmir.org/2025/1/e66524
doi:10.2196/66524
PMID:

©Tsuyoshi Okuhara, Marina Terada, Hiroko Okada, Rie Yokota, Takahiro Kiuchi. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 14.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating the Content and Quality of Videos Related to Hypertrophic Scarring on TikTok in China: Cross-Sectional Study

Jiangkun Wang[1,2], MM; Kai Xu[1], MD, PhD; Juanjuan Wu[1], MSN; Wen Liang[1], MM; Weiming Qiu[1], MD, PhD; Song Wang[1,2], MD, PhD

[1]Department of Burns and Plastic Surgery, General Hospital of Central Theater Command, Wuhan, China

[2]School of Medicine, Wuhan University of Science and Technology, Wuhan, China

Corresponding Author:
Song Wang, MD, PhD
Department of Burns and Plastic Surgery, General Hospital of Central Theater Command
No.627, Wuluo Road
Wuhan, 430060
China
Phone: 86 18502386879
Email: m13896055663@163.com

Abstract

Background: Hypertrophic scars (HTSs) are a predominant condition after burns and trauma, and it causes severe physiological and psychological problems. TikTok (Douyin in Chinese), a popular platform for sharing short videos, has shown the potential to spread health information, including information related to HTSs. Educating the public to obtain correct information is important to reduce the incidence of physiological and psychological problems caused by HTSs. However, the quality and reliability of HTS-related video content on TikTok in mainland China have not been thoroughly studied.

Objective: This study aims to evaluate the content and quality of short videos related to HTSs on the Chinese version of TikTok (Douyin) and explore the factors related to their quality, providing valuable insights for health information dissemination.

Methods: We collected a sample of 153 TikTok videos in Chinese related to HTSs and categorized them according to video source and content. We evaluated the video content using a coding schema, and a hexagonal radar schema was used to intuitively display the spotlight and weight of each aspect of the videos. We evaluated quality using 4 standardized tools: the modified DISCERN (mDISCERN) questionnaire, the *Journal of the American Medical Association*, the Global Quality Scale (GQS), and the Health on the Net Foundation Code of Conduct. We also explored the potential relationship between video quality and characteristics.

Results: The analysis showed that health care professionals uploaded all videos about treating HTSs, which matched the hexagonal radar model analysis findings. The quality assessment scores for the *Journal of the American Medical Association*, GQS, mDISCERN, and the Health on the Net Foundation Code of Conduct had median values of 1 (IQR 1-2), 2 (IQR 2-3), 2 (IQR 2-3), and 3 (IQR 3-4), respectively, indicating a need to improve the quality and reliability of videos on HTSs. In addition, high-quality videos were more popular, based on metrics such as likes, comments, favorites, and shares (P<.001). Interestingly, the time when the videos were uploaded positively correlated with GQS and mDISCERN scores (r=0.393; P<.001 and r=0.273; P<.001), while the video length did not significantly correlate with evaluation scores (P=.78, P=.20, P=.07, and P=.04).

Conclusions: The quality of TikTok videos related to HTSs is generally moderate. Users should exercise caution when seeking information on HTSs from TikTok. It is advisable to choose videos uploaded by health care professionals from the burn department and the burn plastic surgery department, and in the Chinese context, those produced in first-tier cities and emerging first-tier cities.

(JMIR Infodemiology 2025;5:e64792) doi:10.2196/64792

KEYWORDS

hypertrophic scars; health education; TikTok; social media; information quality

XSL•FO
RenderX

Introduction

Background

Hypertrophic scars (HTSs) are a common fibrotic skin condition that can develop from various sources, including acute or chronic wounds, deep burns, and surgical incisions [1]. The overall incidence of HTSs ranges from 4% to 16%, but among patients with burns, the prevalence can be as high as 70% [1,2]. A study of Chinese college students reported an incidence rate of HTSs at 5.2% [3]. In high-income countries, around 100 million individuals are affected by HTSs [4]. HTSs can negatively impact a person's appearance and lead to impaired skin function, joint deformities, and decreased mobility, significantly affecting mental and physical well-being [5]. The annual global cost for HTS care is estimated at nearly US $20.8 billion, with the United States spending about US $4 billion on treatment yearly [6]. The market for HTS and keloid scar treatments is projected to grow, potentially reaching US $37.9 billion by 2026, with a compound annual growth rate of 9.9% [7]. Making lifestyle changes, such as minimizing intense physical labor, avoiding spicy foods, reducing alcohol consumption, and limiting time spent in hot baths, may help lower the risk of developing HTSs [2]. Early detection, diagnosis, and effective treatment are essential for improving patient outcomes and addressing the physiological and psychological issues related to HTSs. Therefore, educating the public about accurate and reliable health information is crucial in reducing the incidence of problems associated with HTSs.

Health Information in the Digital Era

The rapid advancement of internet technology has transformed how we share and communicate health information [8,9]. Remarkably, around 80% of individuals worldwide rely on online resources to inform themselves about health matters [10,11]. This shift enhances health communication and unprecedentedly empowers patients in their education and decision-making [12]. With easy access to information online, patients are no longer passive recipients; they have become proactive seekers, fully engaged in their health outcomes [13]. In recent years, health education videos designed to inform viewers, individually or collectively, have surged in popularity [14]. Unlike traditional text, videos on social media platforms present information more digestibly and effectively motivate users toward healthier behaviors through compelling visuals [15,16]. Short video platforms have the potential to spread health education widely, but patients may encounter challenges in using these technologies. Patients' primary concern when searching for online health information is the quality of the information [17]. The rise of numerous content creators and the lack of regulation on these platforms often lead to concerns about the trustworthiness of the medical information shared [18]. When searching for online health information, the quality of the information is a primary concern for patients [17]. For many nonprofessionals, evaluating the quality of online health information sources, especially for patients with lower health literacy levels, is not easy [19]. Due to the varying quality of short video content, patients often struggle to distinguish between true and false information. This can lead to the spread of misleading information, potentially impacting patients' understanding of their own medical conditions [20]. It is important to recognize that the content creators on video platforms could be (1) patients themselves or (2) health professionals. Experience sharing by patients who have experienced similar conditions can be a handy educational mechanism for other patients or their caregivers [21]. Health professionals, experts in their fields, can also offer helpful advice. However, more must be done to verify the authenticity of the sources (patients or health professionals) producing these videos. This is one of the primary reasons why a lot of misinformation or unverified health information is propagated on social media and video platforms. Therefore, establishing robust oversight of online health videos is crucial to ensure that patients receive reliable and accurate health information.

This Study

TikTok (ByteDance), or Douyin, its Chinese name, is the leading video social media app in China, captivating a vast audience. Focusing on diverse content, such as food, travel, and education, it has attracted over 750 million daily active users, making it a vital platform for engagement and discovery [22]. During the COVID-19 pandemic, TikTok videos about the SARS-CoV-2 garnered 93.1 billion views by July 2020 [17,19]. In addition, videos tagged with #cancer have amassed over 1.1 billion views worldwide [23]. As a platform for disseminating health information, social media has significant differences in video quality and information accuracy. Previous studies have shown that videos published by health care professionals typically have higher scientific validity and credibility [24]. However, the popularity of these videos is often limited by the social influence of the publishers. Videos posted by social media influencers with many followers may attract more viewers and interactions, even if they lack professionalism. In addition, Ming et al [25] found that erroneous information is commonly present in health education videos released by for-profit organizations. This further highlights the necessity of evaluating video quality and authenticity. Previous research has examined the quality of videos on various themes on TikTok, revealing differences in video quality. For instance, videos about *Helicobacter pylori* infection [11], breast cancer [26], liver cancer [20], and inflammatory bowel disease [27] are generally considered unsatisfactory in quality. In contrast, videos related to plastic surgery are deemed satisfactory in quality and reliability [28]. We found many videos about HTSs on Douyin, the Chinese version of TikTok, but the quality of the information presented is yet to be evaluated. To address this research gap, we assessed the content, quality, and reliability of HTS-related videos on TikTok. We examined the relationship between the quality of video content and audience engagement, focusing specifically on interactive indicators, such as likes, comments, favorites, and shares.

Methods

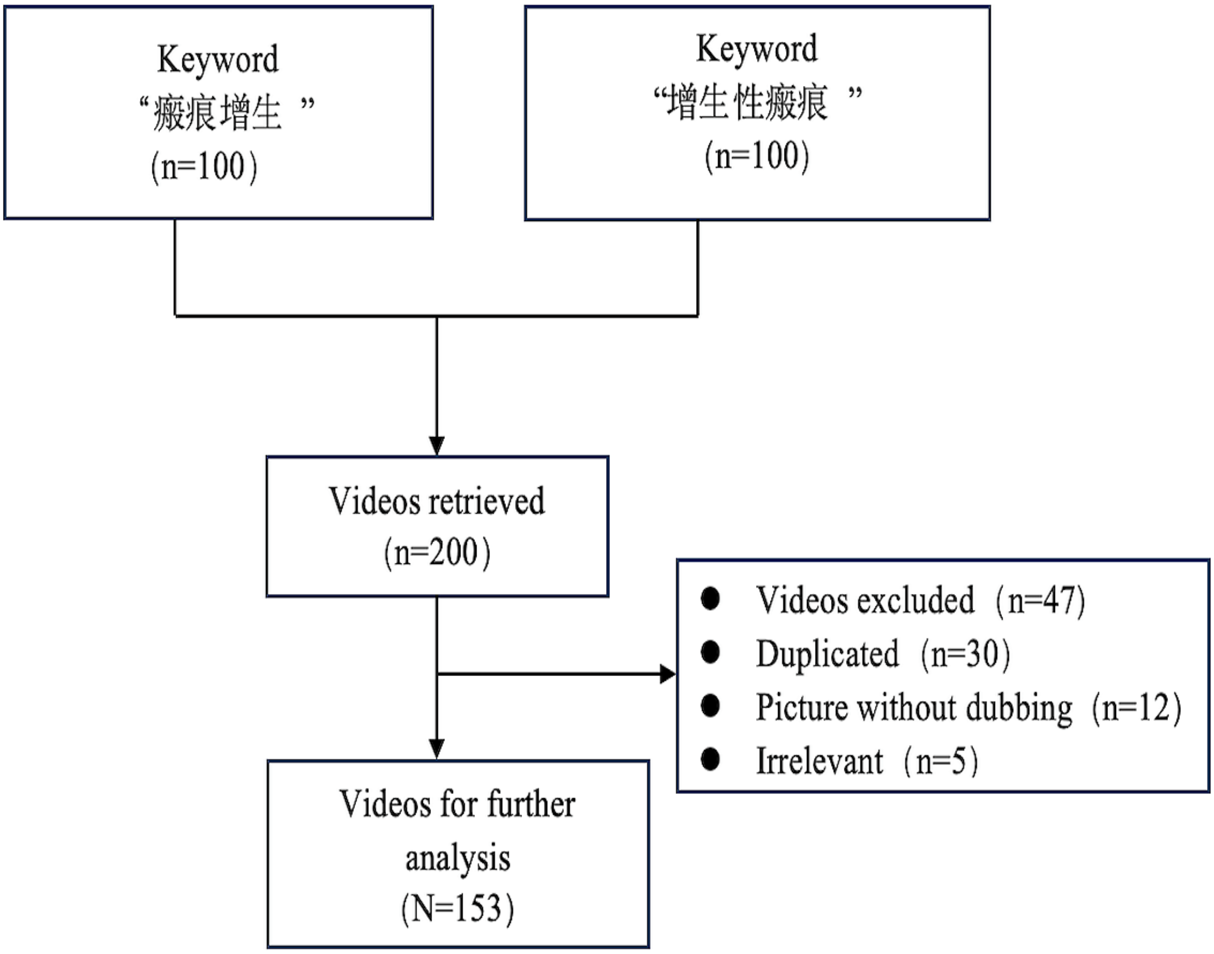
Search Strategy and Data Extraction

In this cross-sectional study, we used the keywords "瘢痕增生" ("scar hyperplasia" in Chinese) and "增生性瘢痕" ("hypertrophic scars" in Chinese) to search on the Chinese version of TikTok ("Douyin") on February 28, 2024, with the

XSL•FO
RenderX

default sorting option of "overall ranking." To avoid bias caused by personalized recommendations, we used newly registered accounts to conduct searches. We did not apply any filtering conditions to restrict the search. Consumers seeking general health videos typically do not scroll very far when searching online; they usually browse only the first few pages of search results. Furthermore, videos that rank low in the search results of the "overall ranking" mode often have little relevance to the topic [29]. Considering the aforementioned situation, we selected the top 100 videos for further analysis of the search results. Subsequently, we excluded non-Chinese, irrelevant, repetitive, and silent videos, resulting in 153 videos selected for the final data analysis (Figure 1).

Figure 1. Search strategy for short videos on hypertrophic scars.

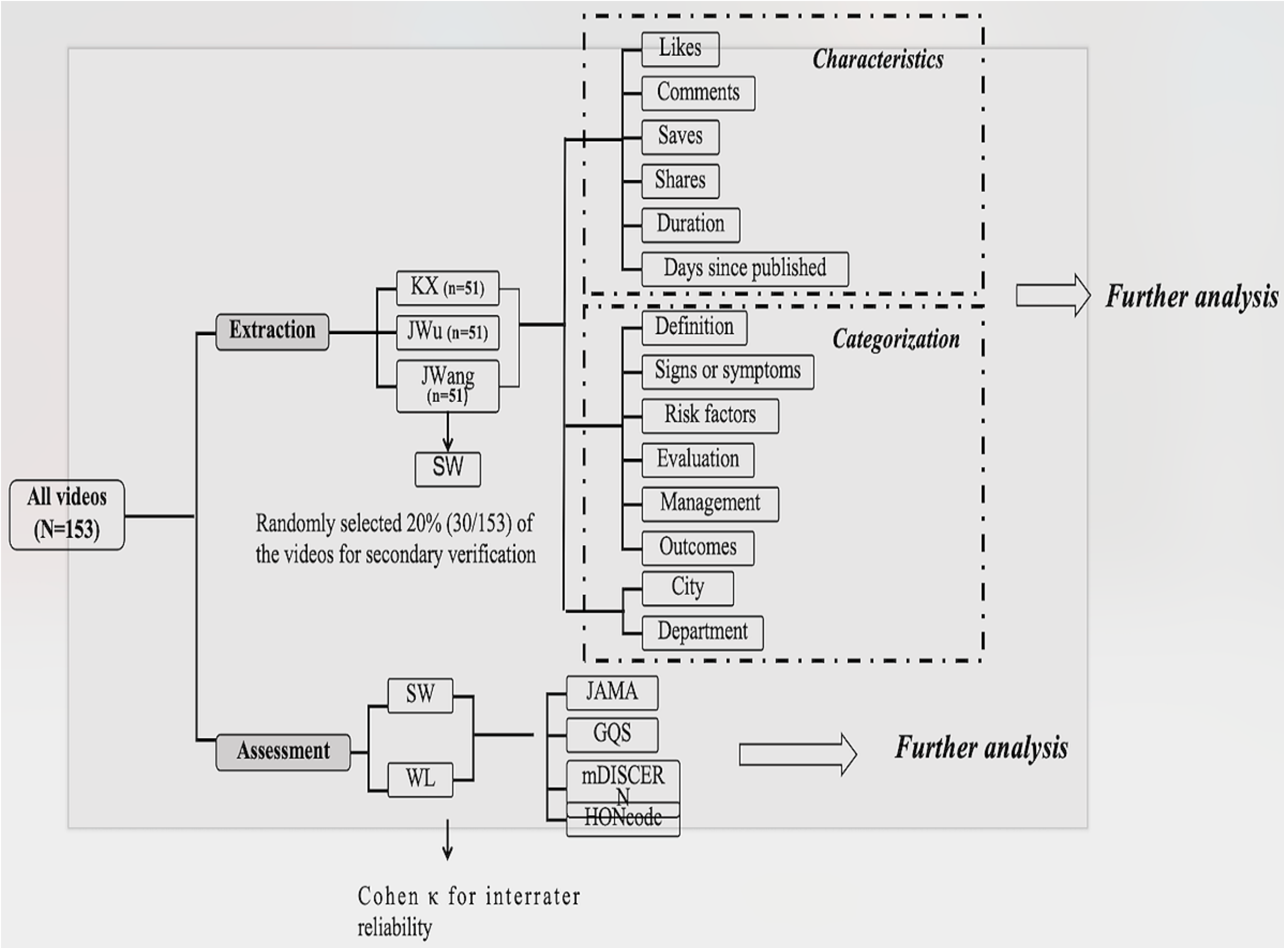


Exclusion criteria were non-Chinese videos and repetitive videos—which refers to videos with the same content but different sources; we based our judgments on the video descriptions and main content. We excluded silent videos, which are defined as content that consists only of images or text, with no voice or background sound. We determined whether a video was silent by overseeing each clip to ensure no audio. Moreover, we excluded irrelevant videos, which refers to videos that do not pertain to the themes of "scar hyperplasia" or "hypertrophic scars." Examples included advertisements, entertainment videos, or content related to other health topics. We based our judgments on the video descriptions and main content.

We extracted data directly from the public information provided by the TikTok platform, as it lacks a bulk data export function. Consequently, we manually recorded the relevant data for each video. Team members used browser tools, including screenshots and text-copying functions, to transfer video information into Microsoft Excel spreadsheets for further classification and analysis. Three team members (J Wu, KX, and J Wang) completed the data extraction, each responsible for a specific portion of the videos. To ensure the accuracy of the data entry, we developed a unified operations manual, and cross-checking was conducted by another team member (SW) after the data entry was finished. In addition, we randomly selected 20% (30/153) of the videos for secondary verification, which resulted in a data consistency rate of over 95% (Figure 2).

XSL•FO
RenderX

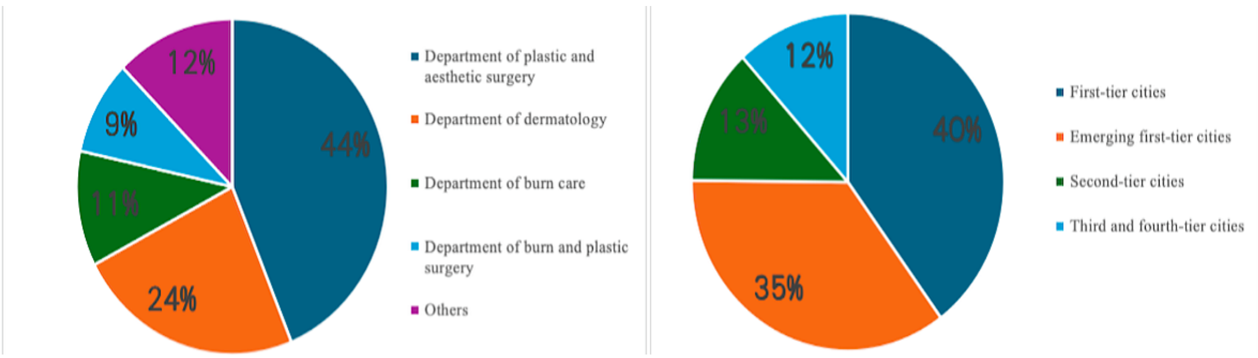
Figure 2. Flowchart for data extraction and analysis. GQS: Global Quality Scale; HONcode: Health on the Net Foundation Code of Conduct; *JAMA*: *Journal of the American Medical Association*; mDISCERN: modified DISCERN.



Video Classification

The content of the videos was classified through manual review. Three authors (J Wu, KX, and J Wang) independently watched each video and categorized it into 1 of the following 6 groups based on its content: the definition, signs and symptoms, risk factors, evaluation, management, or outcomes. All videos were provided by health care professionals; we further classified them according to department categories, specifically including plastic and aesthetic surgery, dermatology, burn care, burn and plastic surgery, and a general category termed "other departments," which includes various departments, such as traditional Chinese medicine and pediatric surgery. In addition, to comprehensively consider the distribution of video resources, we categorized the videos based on city administrative levels, including first-tier cities, emerging first-tier cities, second-tier cities, third-tier cities, and fourth-tier cities, to reflect regional differences more accurately (Figures 2 and 3, Tables 1 and 2) [30].

Figure 3. Percentage of videos on health care professionals from different departments and city tiers.



XSL•FO
RenderX

Table 1. Scoring criteria and weight for categorization of cities.

Scoring criteria	Weight
Commercial resource concentration	0.19
Big Brand Favorability Index	0.32
Commercial Core Index	0.40
Commercial support maturity	0.28
Urban hub connectivity	0.2
Transportation connectivity	0.33
Talent Mobility Index	0.27
Industry Synergy Index	0.20
Commercial resource regional centrality	0.2
Urban population activity	0.22
Consumption activity	0.38
Social activity	0.31
Nighttime activity	0.31
New economy competitiveness	0.20
Corporate leadership	0.36
New Consumption Index	0.32
Industry Chain Ecosystem Index	0.32
Future flexibility	0.19
Innovation Atmosphere Index	0.32
Talent Attraction Index	0.39
City Size Index	0.29

Table 2. The 2024 China city classification rankings.

City classification	List of cities
First-tier cities	Shanghai, Beijing, Shenzhen, and Guangzhou
Emerging first-tier cities	Chengdu, Hangzhou, Chongqing, Suzhou, Wuhan, Xi’an, Nanjing, Changsha, Tianjin, Zhengzhou, Dongguan, Wuxi, Ningbo, Qingdao, and Hefei
Second-tier cities	Foshan, Shenyang, Kunming, Jinan, Xiamen, Fuzhou, Wenzhou, Changzhou, Dalian, Shijiazhuang, Nanning, Harbin, Jinhua, Nanchang, Changchun, Nantong, Quanzhou, Guiyang, Jiaxing, Taiyuan, etc.
Third-tier cities	Urumqi, Linyi, Haikou, Huzhou, Yangzhou, Yancheng, Luoyang, Tangshan, Jining, Langfang, Taizhou, Ganzhou, Hohhot, Zhenjiang, Wuhu, Shantou, Handan, Jiangmen, Zibo, Yinchuan, etc.
Fourth-tier cities	Zhoushan, Qingyuan, Quzhou, Zhumadian, Deyang, Yibin, Longyan, Rizhao, Hongzhi, Anshan, Maoming, Binzhou, Qinhuangdao, Jilin, Kaifeng, etc.

The weights of the primary and secondary dimensions of the ranking were determined through scoring by the expert committee of the New First-Tier Cities Research Institute, while the indicators mentioned after the secondary dimensions were calculated using the principal component analysis method. The indicators for each subdimension of the ranking were primarily derived from data collected throughout 2023 or up to early 2024.

In this study, to determine the geographic location of video creators, we mainly obtained relevant information through 2 channels. First channel was user profile information, where the TikTok account profile of the video uploader usually voluntarily disclosed location information, such as city name or workplace. We manually checked the account home page of each uploader and recorded the geographic location mentioned in their profile (such as a hospital in Beijing). The second channel was the certification identification and employer information, where the TikTok platform usually included the publisher’s employer and department information for certified accounts. For example, the authentication information may have included “burn and plastic surgery department of a hospital in Shanghai” or “dermatology department of a hospital in Guangzhou,” based on which we could determine the geographic location of the uploader.

XSL•FO
RenderX

Assessment of Video Content, Quality, and Reliability

We used the 6 questions developed by Goobie et al [31] to assess video content, focusing on disease definition, signs and symptoms, risk factors, evaluation, management, and outcomes. The hexagonal radar chart is a unique statistical tool that can display data from 6 different fields at the same time. Each dataset is mapped onto a separate axis, and the data points on each axis are connected by continuous lines to form a hexagonal outline. The main goal of this chart design is to visually represent the focus and impact weight of a specific subject, such as video content, across 6 core dimensions [19]. By doing so, the hexagonal radar chart simplifies the comprehension of complex data and provides a clear and user-friendly visual representation for both users and researchers [12,32].

The videos' reliability and quality were assessed using 4 standardized evaluation tools: modified DISCERN (mDISCERN; Table 3), Global Quality Scale (GQS; Table 4), the *Journal of the American Medical Association* (*JAMA*; Table 5), and the Health on the Net Foundation Code of Conduct (HONcode; Table 6). We used a multifaceted approach to assess the quality and reliability of the educational content of the videos collected, mainly based on the following considerations:

- Each benchmark focuses on different dimensions. mDISCERN assesses the information quality, GQS evaluates overall content quality, *JAMA* evaluates the reliability of the video, and the HONcode examines the ethics and credibility of health information.
- By integrating multiple benchmarks, it is possible to more comprehensively capture the differences in quality and reliability dimensions in short videos.
- The multibenchmark method improves the objectivity of evaluation and avoids bias that may arise from a single benchmark.

Table 3. Description of modified DISCERN (mDISCERN) for evaluating the quality of the videos with information on hypertrophic scars.

mDISCERN	Scores (1 point is given for every yes and 0 points for no)
Is the video clear, concise, and understandable?	0-1
Are reliable sources of information used? (ie, publication cited and speaker is a specialist)	0-1
Is the information presented balanced and unbiased?	0-1
Are additional sources of information listed for patient reference?	0-1
Are areas of uncertainty or controversy mentioned?	0-1

Table 4. Description of the Global Quality Scale (GQS) for evaluating the quality of the videos with hypertrophic scars information.

GQS	Scores (range from 1=poor quality to 5=excellent flow and quality)
The information is of poor quality, and the flow of the site is poor. Most information is missing and not useful for patients at all.	1
The information is generally of poor quality and flow. Some information is listed, but many important topics are missing, and it is of very limited use to patients.	2
Moderate quality and suboptimal flow: Some vital information is adequately discussed, but other topics are poorly discussed and somewhat useful for patients.	3
Good quality and flow: Most relevant information is listed, but some topics still need to be covered. It is useful for patients.	4
The information is of excellent quality and has excellent flow. It is beneficial for patients.	5

Table 5. Description of the *Journal of the American Medical Association (JAMA)* for evaluating the quality of the videos with hypertrophic scar information.

JAMA benchmark criteria	1 point for each criterion, with a total score of 4 points
Authorship	Author and contributor credentials and their affiliations should be provided.
Attribution	All copyright information should be clearly listed, and references and sources for content should be stated.
Currency	The initial date of posted content and subsequent updates to the content should be provided.
Disclosure	Conflicts of interest, funding, sponsorship, advertising, support, and video ownership should be fully disclosed.

XSL•FO
RenderX

Table 6. Description of the Health on the Net Foundation Code of Conduct (HONcode) for evaluating the quality of videos with information about hypertrophic scars.

HONcode	Detail
Authority	Any medical or health advice given in the video must come from a qualified health professional unless it is clearly stated that the information does not come from a qualified health source.
Complementarity	The information provided in the video must be designed to support the patient's HTS[a] self-management, but it is not meant to replace the patient-physician relationship.
Privacy policy	The information in the video maintains the right to confidentiality and respect of the individual patient featured.
Referenced and dated	Each video contains references to source data on information presented or contains a specific HTML link to source information.
Justifiability	Each video containing claims on the benefits or performance of specific skills and behaviors, interventions, treatments, products, etc must be supported by evidence through references or HTML links.
Transparency	The video must provide the viewer with contact information or a URL to more information.
Financial disclosure	Any individual or organization that contributes funds, services, or material in the posted video must be clearly identified in the video or video description.
Advertising policy	If an advertisement supports funding to the video or the video's developers, it must be clearly stated. Included advertising must be clearly differentiable to the viewer. There should be a clear difference between the advertising material and the educational material in the video.

[a]HTS: hypertrophic scar.

mDISCERN is the most commonly used quality research tool [33]. This method has been widely used to evaluate information quality on video-sharing platforms [34]. Considering that the video studied belongs to the medical category, mDISCERN is based on the following 5 aspects: clarity, relevance, traceability, robustness, and fairness. The mDISCERN has 5 questions that need answers as "yes" or "no." A score of 1 indicates yes, 0 indicates no, and the maximum score is 5 [35].

GQS was used to assess the overall content quality of the videos in this study. The GQS is a commonly used 5-point scale comprising 5 criteria ranging from 1 to 5, with higher scores indicating better quality [36-38].

JAMA was used to evaluate the reliability of the video [39]. The rating is according to the 4 predetermined issues: authorship, attribution, currency, and disclosure. There is 1 point for each criterion, with a total score of 4 points [40].

The HONcode consists of 8 issues that are predetermined for the rating: authority, complementarity, privacy policy, reference and date, justifiability, transparency, financial disclosure, and advertising policy [41,42]. The details of the scoring criteria are mentioned subsequently. First, any medical or health advice given in the video must come from a qualified health professional unless it is clearly stated that the information does not come from a qualified health source. Second, the information provided in the video must be designed to support the patient's HTS self-management, but it is not meant to replace the patient-physician relationship. Third, the information in the video maintains the right to confidentiality and respect of the individual patient featured. Fourth, each video contains references to source data on the information presented or contains a specific HTML link to source information. Fifth, each video containing claims on the benefits or performance of specific skills or behaviors, interventions, treatments, products, etc must be supported by evidence through references or HTML links. Sixth, the video must provide the viewer with contact information or a URL to more information. Seventh, any individual or organization that contributes funds, services, or material in the posted video must be clearly identified in the video or video description. Eighth, if an advertisement supports funding to the video or the video's developers, it must be clearly stated. Included advertising must be differentiable to the viewer: There should be a clear difference between the advertising material and the educational material in the video. There is 1 point for each criterion, with a total score of 8 points.

Although *JAMA* and HONcode are commonly used to evaluate formal or long-format medical content (such as websites, journals, or organizational publications), their application has been extended to user-generated content on the TikTok short video platform [43]. In this study, we had a detailed discussion on the benchmark before scoring and adjusted the scope of application of the scoring criteria. For the "disclosure" rating item, we focused on whether the video identified the publisher's identity and affiliation rather than detailing funding sources or advertising disclosures. Regarding the "citation source" standard, many videos did not provide explicit references and often used vague terms, such as "research shows" or "experts say." To tackle this issue, we reached the following consensus: (1) videos that do not provide any source explanation will receive a score of 0; (2) content that mentions vague references (like "research shows") but fails to specify the source will receive a score of 0.5, indicating a partially satisfied score; and (3) videos that list their sources or include relevant reference information in the video description will receive a score of 1. We also focused on evaluating whether the core medical information of the video was accurately conveyed based on the video duration limit rather than comprehensive coverage. In addition, during the rating process, we considered the background information of the video creator (such as certification marks or institutional affiliations) to help evaluate the credibility of the references.

The videos were evaluated by 2 qualified physicians (SW and WL) who have extensive experience in scar treatment. Before scoring the videos, the 2 evaluators reviewed the mDISCERN,

XSL•FO
RenderX

GQS, *JAMA*, and HONcode scoring guidelines and conducted detailed discussions to prevent cognitive bias. The final score for each video was calculated by averaging the scores given by the 2 evaluators. If there was a significant difference between the scores of the 2 experts, the final score was determined through discussion with the third arbitrator (KX; Figure 2). In the evaluation process of 153 videos, Cohen κ values rated by experts showed high consistency (κ>0.80). Therefore, no case involved the third arbitrator.

Ethical Considerations

All information used in this study came from publicly published TikTok (Douyin in Chinese) videos. This study did not involve clinical data, human specimens, or animal experiments, nor did it involve personal privacy. No personal data identifying the uploader's identity, such as username or profile picture, were recorded or stored during the research process. Data analysis only focused on video content and interaction metrics (such as likes, comments, and shares). The study strictly abided by the terms of use of the TikTok platform and did not obtain the platform's undisclosed data through any technical means. The research content did not involve any potential harm to user interests or platform rules and was only used for academic purposes. Therefore, this study did not require an ethics review.

Statistical Analyses

The data were analyzed using SPSS Statistics (version 29; IBM Corp). Continuous variables were presented as medians with IQRs, while categorical variables were presented in terms of numbers and percentages. Cohen κ was used to measure interrater reliability between the 2 evaluators. According to the criteria set by Landis and Koch [44], a κ value > 0.8 indicates almost perfect agreement, a value between 0.6 and 0.8 indicates substantial agreement, a value between 0.4 and 0.6 indicates moderate agreement, and a value <0.4 indicates poor agreement. Spearman correlation analysis was conducted to assess the relationships between quantitative variables. A significance level of P<.001 was considered statistically significant.

Results

Video Characteristics

A total of 153 videos about scar hyperplasia and HTSs were found on Chinese TikTok, all posted by health care professionals. In terms of departmental distribution of video uploads, professionals from the department of plastic and aesthetic surgery contributed the highest proportion of video content, accounting for 67 videos (43.8%), followed by dermatology (n=36, 23.5%), burn care (n=17, 11.1%), burn and reconstructive surgery (n=14, 9.2%), and "other departments," which included traditional Chinese medicine and pediatric surgery (n=19, 12.4%). Further analysis by city tier revealed significant differences in video publication volume. Health care professionals in first-tier cities were the most active, accounting for 61 (39.9%) of the video uploads, followed by new first-tier cities (n=54, 35.3%), second-tier cities (n=20, 13.1%), and third- and fourth-tier cities (n=18, 11.8%; Figure 3). The general characteristics of the videos are presented in Tables 7-9.

The median time since upload was 212 (IQR 54-321) days, and the average video duration was 43 (IQR 33-58, SD 36) seconds. All videos received a maximum of 21,000 likes (median 72, IQR 31-189), 1230 comments (median 9, IQR 4-32), 7580 favorites (median 21, IQR 7-66), and 2292 shares (median 20, IQR 9-63). Table 8 describes the critical features of the videos uploaded by health care professionals from different departments. Notably, videos posted by dermatologists stood out in several engagement metrics, specifically with higher numbers of likes (median 112.5, IQR 44.5-254), comments (median 12, IQR 5.75-46.25), saves (median 36.5, IQR 13.75-94.75), and shares (median 29, IQR 11.75-70.75). This phenomenon may reflect the public's interest and preference for educational dermatology videos. Further analysis (Table 9), which focused on the essential characteristics of videos uploaded by health care professionals from different city tiers, revealed a notable phenomenon. Although some cities may not have the overall resource advantage, emerging first-tier cities' videos showed unique appeal in user engagement. The median numbers of likes, comments, saves, and shares were 94 (IQR 37.75-183), 12 (IQR 5-55.5), 24.5 (IQR 8-78.75), and 26.5 (IQR 10.25-66), respectively. This finding suggested that video content dissemination strategies should focus more on regional characteristics and alignment with user needs. In addition, the shortest video was 13 seconds long, the longest was 282 seconds long, and the first video was uploaded 1047 days before our search. In contrast, the most recent video was uploaded the day before data collection.

XSL•FO
RenderX

Table 7. Characteristics of hypertrophic scar videos (N=153).

Parameters	Values
Video source, n (%)	
Health care professionals	153 (100)
Department classification, n (%)	
Department of plastic and aesthetic surgery	67 (43.8)
Department of dermatology	36 (23.5)
Department of burn care	17 (11.1)
Department of burn and plastic surgery	14 (9.2)
Other departments	19 (12.4)
City classification, n (%)	
First-tier cities	61 (39.9)
Emerging first-tier cities	54 (35.3)
Second-tier cities	20 (13.1)
Third- and fourth-tier cities	18 (11.8)
Likes, median (IQR)	22 (31-189)
Comments, median (IQR)	9 (4-32)
Saves, median (IQR)	21 (7-6)
Shares, median (IQR)	20 (9-3)
Duration (s), median (IQR)	44 (33-8)
Days since published, median (IQR)	159 (54-21)
JAMA[a] score, median (IQR)	1 (1-2)
GQS[b] score, median (IQR)	2 (2-3)
mDISCERN[c] score, median (IQR)	2 (2-3)
HONcode[d] score, median (IQR)	3 (3-4)

[a]*JAMA*: *Journal of the American Medical Association*.
[b]GQS: Global Quality Scale.
[c]mDISCERN: modified DISCERN.
[d]HONcode: Health on the Net Foundation Code of Conduct.

XSL•FO
RenderX

Table 8. Characteristics of hypertrophic scars in videos across different departments.

Variable	Department of plastic and aesthetic surgery (n=67), median (IQR)	Department of dermatology (n=36), median (IQR)	Department of burn care (n=17), median (IQR)	Department of burn and plastic surgery (n=14), median (IQR)	Others (n=19), median (IQR)	Overall (n=153), median (IQR)
Likes	52 (30.5-61.5)	112.5 (44.5-254)	70 (34-189)	62 (37.25-146.25)	96 (18.5-234)	22 (31-189)
Comments	8 (4-19)	12 (5.75-46.25)	7 (2-33)	8.5 (3-27)	10 (3.5-40)	9 (4-32)
Saves	17 (8-56.5)	36.5 (13.75-94.75)	22 (6-50)	14.5 (7-35.25)	45 (5-97.5)	21 (7-66)
Shares	14 (8-55.5)	29 (11.75-70.75)	22 (6-50)	16.5 (7.75-47)	32 (8.5-96.5)	20 (9-63)
Duration (s)	45 (33.5-59)	41.5 (30-55.25)	45 (5-97.5)	43.5 (29.5-55.75)	53 (42.5-58.5)	44 (33-58)
Days since published	16.5 (39.5-334)	143.5 (71.5-281)	162 (88-196)	258 (69.25-464)	145 (28-260)	159 (54-321)
JAMA[a] score	1 (1-1.5)	1.75 (1-2)	1.5 (1-2)	1 (1-1)	1 (1-1.5)	1 (1-2)
GQS[b] score	2 (2-3)	2 (2-3)	2 (2-2.5)	2 (2-2.75)	2 (2-3)	2 (2-3)
mDISCERN[c] score	2 (2-3)	2 (2-3)	3 (2-3)	2 (2-3)	2 (2-2)	2 (2-3)
HONcode[d] score	3 (3-4)	3 (3-3)	3 (3-4)	3.25 (3-4)	3 (3-3.5)	3 (3-3.75)

[a]*JAMA*: *Journal of the American Medical Association*.

[b]GQS: Global Quality Scale.

[c]mDISCERN: modified DISCERN.

[d]HONcode: Health on the Net Foundation Code of Conduct.

Table 9. Characteristics of hypertrophic scars in videos across different city tiers.

Variable	First-tier cities (n=61), median (IQR)	Emerging first-tier cities (n=54), median (IQR)	Second-tier cities (n=20), median (IQR)	Third- and fourth-tier cities (n=18), median (IQR)	Overall (n=153), median (IQR)
Likes	79 (34-220)	94 (37.75-183)	53.5 (29.5-108.75)	51.5 (24-144.25)	22 (31-189)
Comments	8 (4-33)	12 (5-45.5)	10.5 (3-15)	10.5 (4-16.5)	9 (4-32)
Saves	22 (8-76)	24.5 (8-78.75)	9 (4-52.5)	13 (8-46)	21 (7-66)
Shares	20 (9-63)	26.5 (10.25-66)	18 (7.75-47.25)	10 (7-40)	20 (9-63)
Duration (s)	41 (31-55)	43 (30.25-54.5)	50 (38.25-60.5)	58 (37-102)	44 (33-58)
Days since published	188 (80-317)	119.5 (54.75-325.5)	156.5 (76.75-329)	25 (14-168)	159 (54-321)
JAMA[a] score	1 (1-2)	1 (1-2)	1 (1-1)	1.25 (1-2)	1 (1-2)
GQS[b] score	2 (2-3)	2 (2-3)	2 (2-2)	2 (2-3)	2 (2-3)
mDISCERN[c] score	2 (2-3)	2 (1-3)	2 (2-3)	2.5 (2-3)	2 (2-3)
HONcode[d] score	3 (3-4)	3 (3-4)	3 (3-4)	3 (3-4)	3 (3-4)

[a]*JAMA*: *Journal of the American Medical Association*.

[b]GQS: Global Quality Scale.

[c]mDISCERN: modified DISCERN.

[d]HONcode: Health on the Net Foundation Code of Conduct.

Analysis of Video Content

According to the hexagonal radar chart, the most frequently discussed topic in all videos was managing HTSs, which appeared in approximately 72.5% (111/153) of the videos. This was followed by the symptoms and definitions of HTSs mentioned in 47.7% (73/153) and 24.2% (37/153) of the videos. However, the outcomes and risk factors of HTSs should have been addressed, with only 11.8% (18/153) and 19.6% (30/153) of the videos discussing these aspects. The least mentioned topic was the evaluation of HTSs, with only 9.2% (14/153) of the videos adequately covering evaluation, while 90.8% (139/153) of the videos provided little to no information on this aspect (Figure 4). An in-depth analysis of the hexagonal radar chart structures presented by various city tiers revealed a common phenomenon—regardless of city tier, the video content predominantly focused on managing HTSs, while the evaluation of HTSs was notably less addressed (Figure 5). This finding aligned with our overall evaluation of the video content, further confirming the distribution bias of video resources toward specific topics.

XSL•FO
RenderX

Figure 4. Hexagonal radar charts of the content of videos on hypertrophic scars.

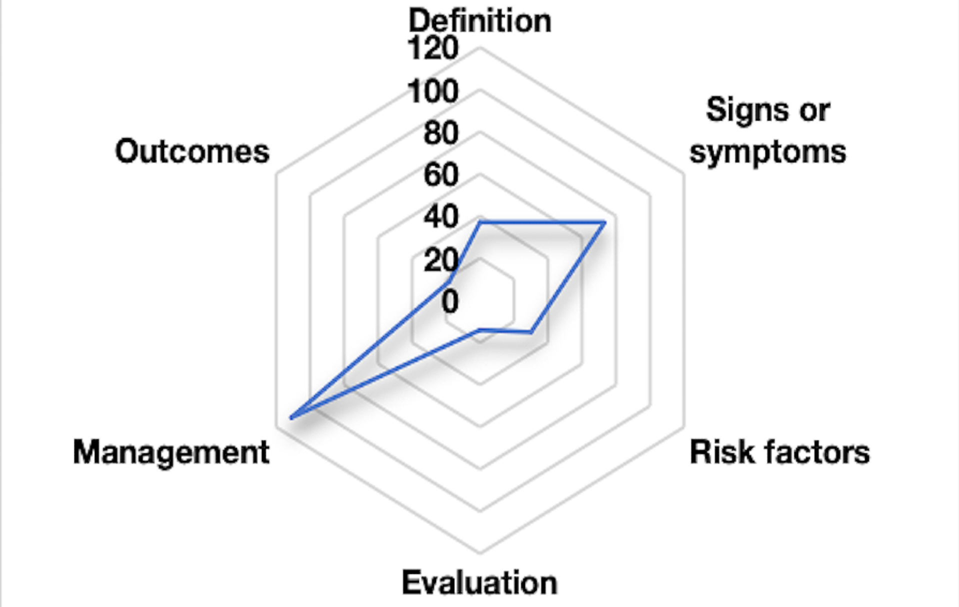
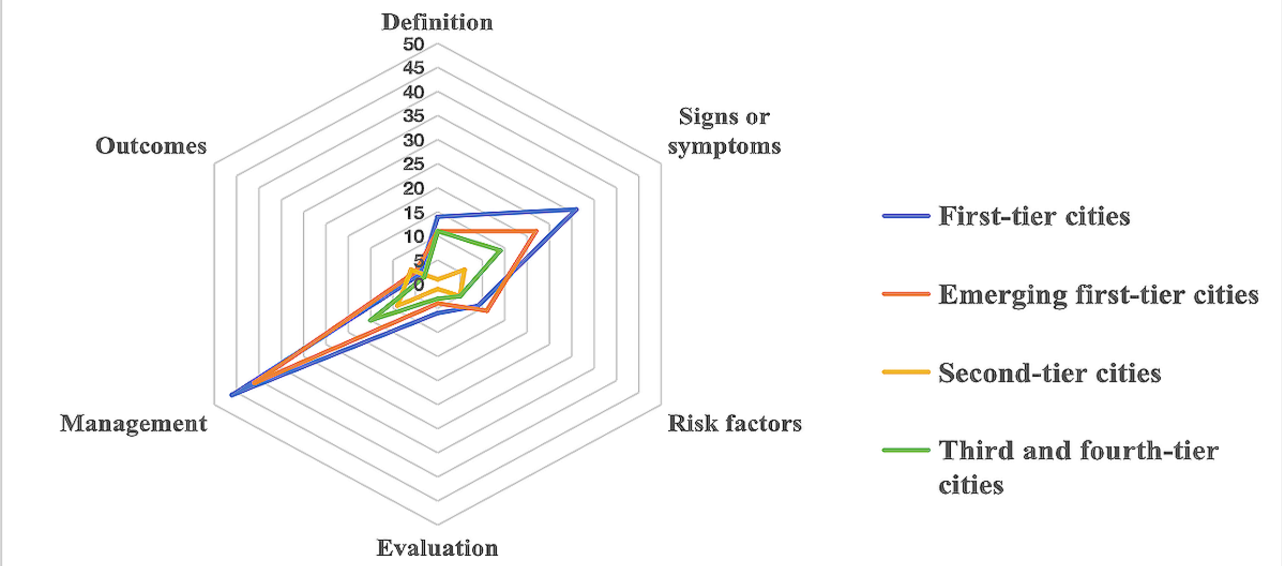


Figure 5. Hexagonal radar charts of the content of different city-tier videos on hypertrophic scars.



Assessment of Video Quality

We found that the median *JAMA* score for all uploaded TikTok videos was 1 (IQR 1-2). When we used the mDISCERN score to assess the usability and reliability of the videos, the median score was 2 (IQR 2-3). Specifically, the median GQS score for the overall quality of the TikTok videos was 2 (IQR 2-3), while the median HONcode score was 3 (IQR 3-4).

To explore whether health care professionals from different departments and cities influenced the quality and reliability of videos, we conducted a detailed categorization based on departmental affiliations and city tiers. The results showed slight differences in the quality scores, specifically *JAMA*, GQS, mDISCERN, and HONcode, among videos uploaded from 1 to 3(IQR 1.75-2.25), indicating an overall low quality of the videos. There was little variation in video quality ratings among different city levels; HONcode ratings were primarily concentrated at 3 (IQR 3-4) points, suggesting moderate overall content quality. Further analysis revealed that videos related to burn plastic surgery and burn surgery had relatively high-quality ratings, with HONcode score of 3.25 and an mDISCERN rating of 3. These departments exhibited stronger professionalism and scientific content. However, the videos suffered from insufficient interactivity, indicated by fewer likes and shares, resulting in a lower dissemination effect compared to dermatology videos.

For new first-tier cities, the video quality ratings (GQS: 2 and mDISCERN: 2) were comparable to those in first-tier cities. In addition, the median upload time for these videos was shorter (119.5 days compared to 188 days), indicating that the content was timelier. In contrast, videos from third- and fourth-tier cities achieved HONcode ratings of 3 and mDISCERN ratings of 2, showing no significant disadvantage in terms of quality. However, these cities had lower upload volumes and interaction metrics, averaging 51.5 likes and 10.5 comments. This could be attributed to limited medical resources and a smaller number of creators in those areas.

XSL•FO
RenderX

Correlation Analysis

The nonnormal distribution of the data led us to use Pearson correlation analysis to investigate the relationships between different video variables and all evaluation scores (Table 10). We found that each video variable positively correlated with the scores obtained from the 4 evaluation methods. Notably, likes, comments, favorites, and shares were the only variables that showed significant correlations with all evaluation scores ($P<.001$), indicating that higher-quality videos tended to be more appreciated by viewers. Specifically, the number of days since video upload was significantly positively correlated only with GQS scores ($r=0.393$; $P<.001$) and mDISCERN scores ($r=0.273$; $P<.001$). In contrast, video duration did not significantly correlate with the evaluation scores (Table 10).

Table 10. Pearson correlation analysis between the video variables and all evaluation scores.

Variables	*JAMA* [a]	GQS[b]	mDISCERN[c]	HONcode[d]
Likes				
r	0.514[e]	0.740[e]	0.394[e]	0.287[e]
P value	<.001	<.001	<.001	<.001
Comments				
r	0.403[e]	0.613[e]	0.438[e]	0.426[e]
P value	<.001	<.001	<.001	<.001
Saves				
r	0.504[e]	0.736[e]	0.424[e]	0.293[e]
P value	<.001	<.001	<.001	<.001
Shares				
r	0.470[e]	0.701[e]	0.413[e]	0.301[e]
P value	<.001	<.001	<.001	<.001
Days since published				
r	0.123	0.393[e]	0.273[e]	0.098
P value	.11	<.001	<.001	.23
Duration				
r	0.023	0.105	0.149	0.169
P value	.78	.20	.07	.04

[a]*JAMA*: *Journal of the American Medical Association*.

[b]GQS: Global Quality Scale.

[c]mDISCERN: modified DISCERN.

[d]HONcode: Health on the Net Foundation Code of Conduct.

[e]The correlation is significant at a significance level of .01 (2-tailed).

In addition, we used Spearman correlation analysis to reveal the relationships between different video variables. We observed a positive correlation between the following variables: likes and comments ($\rho=0.777$; $P<.001$), likes and saves ($\rho=0.941$; $P<.001$), likes and shares ($\rho=0.904$; $P<.001$), likes and uploads ($\rho=0.534$; $P<.001$), comments and saves ($\rho=0.781$; $P<.001$), comments and shares ($\rho=0.820$; $P<.001$), comments and uploads ($\rho=0.404$; $P<.001$), saves and shares ($\rho=0.897$; $P<.001$), saves and uploads ($\rho=0.499$; $P<.001$), and shares and uploads ($\rho=0.564$; $P<.001$). Meanwhile, there was no significant relationship between video duration and other variables (Table 11).

XSL•FO
RenderX

Table 11. Spearman correlation analysis between the video variable.

Variable	Likes	Comments	Saves	Shares	Days since published	Duration
Likes						
ρ	1	0.777[a]	0.941[a]	0.904[a]	0.534[a]	0.072
P value	—[b]	<.001	<.001	<.001	<.001	.38
Comments						
ρ	0.777[a]	1	0.781[a]	0.820 [a]	0.404 [a]	0.122
P value	<.001	—	<.001	<.001	<.001	.13
Saves						
ρ	0.941[a]	0.781[a]	1	0.897[a]	0.499[a]	0.105
P value	<.001	<.001	—	<.001	<.001	.20
Shares						
ρ	0.904[a]	0.820[a]	0.897[a]	1	0.564[a]	0.072
P value	<.001	<.001	<.001	—	<.001	.38
Days since published						
ρ	0.534[a]	0.404[a]	0.499[a]	0.564[a]	1	0.020
P value	<.001	<.001	<.001	<.001	—	.80
Duration						
ρ	0.072	0.122	0.105	0.072	0.020	1
P value	.38	.13	.20	.38	.80	—

[a]The correlation is significant at a significance level of .01 (2-tailed).

[b]Not applicable.

Discussion

Principal Findings

Health problems are crucial and need daily attention, accurate assessment, and timely intervention. With the increasing popularity of the mobile internet, it has become one of the most popular ways to obtain health and medical information. A survey shows that 70% of internet users rely on the internet as their primary source of health information [45]. In this cross-sectional study, we used *JAMA* GQS, mDISCERN, and HONcode tools to evaluate the quality and reliability of HTS-related videos on the Chinese version of TikTok (Douyin). The results showed that the quality and reliability of HTS-related videos from TikTok were generally moderate. From the perspective of video sources, HTS-related videos were mainly released by health professionals. TikTok has strict verification rules to protect users' interests, information security, and content reliability, and it requires only certified institutions or individuals to share medical-related videos on the platform [46]. In terms of video content, the video integrity was insufficient. Most (111/153, 72.5%) of the videos were related to HTS management. From the perspective of video classification, compared with other departments and cities, the videos uploaded by health professionals in burn departments and burn plastic-surgery departments, and videos produced in first-tier and emerging first-tier cities, were of slightly higher quality.

Users should exercise caution when seeking information on HTSs from TikTok. It is advisable to choose videos uploaded by health care professionals from burn departments and burn plastic surgery departments, and in the Chinese context, those produced in first-tier and emerging first-tier cities.

Analysis of Overall Video Quality and Correlation

Our research uncovered an interesting phenomenon—only a few (1/153, 0.7%) videos thoroughly covered all aspects of HTSs, offering authoritative and practical guidance. Most (111/153, 72.5%) videos focused mainly on treatment methods, with symptom descriptions coming next and preventive measures mentioned less frequently. This may be related to the format of the TikTok platform, where video lengths vary; however, according to the latest statistics, the average length of popular videos is about 40 seconds [47]. This characteristic requires creators to present health information within a limited time frame, thereby affecting the depth and coverage of the video, and encourages users to create multiple videos on the same topic, each focusing on different aspects [48]. Our findings support this observation. Because a single video cannot cover all 6 core aspects of HTSs due to time constraints, users tend to split these into multiple videos presented as a series [46]. However, social media platforms usually recommend videos based on algorithms or randomness, making it difficult for users to access comprehensive health information systematically [49].

XSL•FO
RenderX

In the evaluation process of 153 videos, *JAMA*, GQS, mDISCERN, and HONcode scale values rated by experts showed high consistency (Cohen κ>0.80). Most videos on this platform did not receive high scores based on evaluations using *JAMA*, GQS, mDISCERN, and HONcode scales. This suggests that short videos about HTSs have poor quality and reliability. According to the recommendation algorithm of TikTok, people may primarily watch recently uploaded videos, and longer videos might cause viewers to lose patience and interest, leading to video skips. In addition, this mechanism determines that videos with more likes are more likely to be recommended; therefore, popular videos with lower quality have become more popular, further exacerbating the gap between video quality and popularity. We also found that videos from third- and fourth-tier cities received higher scores; however, this result only partially reflects the quality and reliability of their video content. The main reason is the relatively limited sample size from third- and fourth-tier cities, which may introduce some statistical bias. Therefore, caution is needed when interpreting these scores to avoid misinterpretation or misleading conclusions. To address this issue, we recommend that short video platforms introduce professional certification for experts and use unique markers to improve the trustworthiness of medical video content and reduce the spread of misinformation. The review standards for content uploaders on short video platforms are not yet comprehensive and strict. A significant number of nonprofessionals are still posting medical and scientific videos, which, to some extent, affects the accuracy and authority of the content [19]. Therefore, platforms should improve their verification and management procedures to ensure the professionalism and reliability of medical videos.

Our research discovered a potential link between video attributes and evaluation scores. We found a positive relationship between video length and evaluation scores, indicating that longer videos may improve quality by offering more informative content. However, this correlation was not statistically significant (P>.05). Previous studies have suggested that high-quality videos are often longer, which is consistent with our findings [50,51]. Excessively long videos might decrease viewer interest, resulting in fewer views, likes, and user engagement. This decrease in interest may stem from reduced viewer motivation despite the comprehensive content [52]. Therefore, publishers should consider video length carefully to maintain viewer interest and effectiveness of dissemination while upholding content quality. In addition, metrics, such as likes, comments, favorites, and shares, can gauge video popularity. Our analysis found significant positive correlations (P<.001) between these metrics and evaluation scores, indicating that high-quality videos are more likely to receive viewer approval. This finding aligns with the research conducted by Kong et al [12], which evaluated the quality of TikTok videos focused on diabetes health education. Their study found that higher-quality videos tend to receive greater recognition from audiences, evidenced by increased praise and sharing rates. In addition, our research revealed a positive correlation between the upload time of videos and their quality ratings, such as the GQS and mDISCERN scores (P<.001). This suggests that audiences prefer more timely and relevant content. Similar to the findings by Kong et al [12], we observed that the upload timing of videos is positively correlated with user engagement. However, in contrast to the work of Kong et al [12] and other studies that examined YouTube videos as sources of health information, our research found that TikTok videos received lower overall quality ratings. Specifically, in this study, the median *JAMA* score for TikTok videos was 1 (IQR 1-2), while YouTube videos typically received higher ratings in comparable studies (Kong et al [12] reported a median score of 2.5). This disparity may be attributed to the distinct characteristics of each platform. The short video format of TikTok, usually limited to 40 seconds, restricts the depth of content, whereas YouTube allows for longer videos that are more likely to adhere to *JAMA* and HONcode standards for comprehensive information. Furthermore, we found that interaction metrics for TikTok videos, such as likes and shares, were significantly correlated with GQS and mDISCERN ratings (r=0.740 and r=0.394, respectively; P<.001). This supports the conclusion made by Kong et al [12] that high-quality videos tend to engage audiences more actively. In addition, our study revealed variations in interaction metrics among medical professionals from different departments and cities, with videos uploaded from first-tier cities showing higher rates of likes and shares (P<.05). This finding has not been extensively explored in research on other platforms, suggesting that user behavior on TikTok may be influenced by unique regional and professional factors.

Analysis of Evaluation Tools

This study comprehensively used *JAMA*, GQS, mDISCERN, and HONcode to evaluate the quality and reliability of TikTok short videos, mainly based on the considerations mentioned subsequently. First, each benchmark focuses on different dimensions. *JAMA* evaluates authorship and transparency, GQS evaluates overall content quality, mDISCERN focuses on information reliability, and the HONcode examines the ethics and credibility of health information. Second, by integrating multiple benchmarks, it is possible to more comprehensively capture the differences in quality and reliability dimensions in short videos. Third, the multibenchmark method improves the objectivity of evaluation and avoids bias that may arise from a single benchmark. However, we also recognize that these benchmarks were not originally designed for short videos and may pose applicability challenges. For example, *JAMA* and HONcode standards are commonly used to evaluate formal or long-format medical content (websites, journals, or organizational publications). However, this study attempts to extend their application to user-generated content on the TikTok short video platform. These videos mainly focus on visual effects and have a duration between 33 and 58 seconds, so they may not fully meet the requirements of *JAMA* standards for content depth and information transparency. To overcome this challenge, 2 scoring experts had a detailed discussion on the benchmark before scoring and adjusted the scope of application of the scoring criteria. For example, the "disclosure" rating item focuses on whether the video identifies the publisher's identity and affiliation rather than detailing funding sources or advertising disclosures. The rating experts also focus on evaluating whether the core medical information of the video is accurately conveyed based on the video duration limit rather than comprehensive coverage. By adjusting the scope of

XSL•FO
RenderX

application of these benchmarks (such as simplifying citation source standards), they still have reference value in evaluating the accuracy and credibility of core medical information in short videos. In addition, the review study by Li et al [53] indicates that mDISCERN is the most commonly used evaluation tool for health information videos. However, the review mainly focuses on long-format health education content and needs to explore the applicability of mDISCERN, specifically on short video platforms. Secondly, mDISCERN's single benchmark may need to be able to cover the multidimensional quality assessment needs in short videos. Therefore, our study attempts to compensate for the dimensions that a single benchmark may overlook, such as video transparency and overall content quality, by combining other benchmarks, such as *JAMA* and GQS. Therefore, this study chooses to comprehensively use multiple benchmarks to evaluate the quality and reliability of short videos from different perspectives. Future research should further optimize and develop evaluation tools for short videos to enhance their applicability and scientific validity.

Limitations and Future Directions

This research is the first to use 4 evaluation tools (*JAMA*, GQS, mDISCERN, and HONcode) to comprehensively evaluate the quality and reliability of high-frequency videos about HTSs on the TikTok platform. The study also includes an in-depth analysis of the relationship between video characteristics (likes, comments, favorites, and shares) and video quality. However, there are limitations to this study. First, the sample is limited to videos uploaded on the Chinese TikTok platform, which may limit the generalizability of the findings to other languages (such as English) and platforms (such as BiliBili). Despite focusing on Chinese TikTok, the research aligns with studies on videos from various platforms. Given the prevalence of HTS as a health issue, the findings may offer insights for video content in other languages and platforms (such as international versions of TikTok and YouTube). Second, there is a lack of standardized methods for evaluating health information video content on TikTok [46]. The study used 4 standardized evaluation tools due to their proven effectiveness in assessing video quality on media platforms and their previous use in studies evaluating TikTok video quality [54,55]. However, these assessments are somewhat subjective. Despite 2 raters confirming the scores and using Cohen κ to quantify interrater reliability, subjective differences cannot be ignored [20]. This highlights the need for the development of more suitable scoring standards. Third, limiting the analysis scope to verified accounts may result in certain limitations, such as not including videos published by ordinary users (such as patients) or unverified accounts. These videos may contain patients' firsthand experiences or other nonprofessional information, which can impact the comprehensiveness of research conclusions. Future research should consider expanding the scope of analysis and adopting broader validation criteria to cover a more diverse range of video sources. Fourth, although we only selected videos uploaded by medical professionals certified by the platform, we cannot completely rule out the ambiguity of the author's identity information. For example, some uploaders may not be the video's actual creators or information providers but may only participate in video publishing. This uncertainty may affect the accuracy of *JAMA*'s benchmark "authorship" score, potentially leading to bias in research results. Fifth, there are inherent issues with viewing TikTok as a platform for disseminating health information. TikTok's recommendation algorithm tends to push videos that easily attract attention rather than the most scientifically sound ones. This mechanism may lead to the dissemination of misleading or incomplete information. The subject of this study is limited to short videos related to HTSs on the TikTok platform, and all videos are uploaded by medical professionals. Although this choice ensures the scientific and credible nature of the video content, it also limits the generalizability of the research results. Videos related to HTSs uploaded by other groups, such as ordinary users or unverified health influencers, were not included in the analysis, which may limit the applicability of the research results to the broader dissemination of short videos related to health. In addition, the data source of this study is limited to Chinese TikTok (Douyin). The platform culture, user behavior, and regulatory policy may differ from the international version of TikTok or other social media platforms. Therefore, the research results may not directly apply to short video platforms in other countries or regions. In addition, TikTok's video duration limit (usually within 40 s) poses a challenge to the comprehensiveness of health information. Although short videos on the platform can attract viewers to understand a topic quickly, their structure and depth often need to be improved, which may not meet the audience's needs for complex health issues. Therefore, the effectiveness and limitations of TikTok as a tool for disseminating health information need to be further explored.

In addition, there are other evaluation criteria, such as the Video Popularity Index, that can be considered for assessing the quality of health-related information [56]. We recommend that future research incorporate various evaluation methods and platforms to assess the quality of video information more accurately.

The rapid advancement of internet technology and the rising health standards have led to the growing popularity of internet-based health promotion methods. Patients have shifted from being passive recipients to seeking health information actively [20]. With the widespread use of electronic devices, such as smartphones, and the flourishing multimedia technology, visual social media has become a crucial channel for accessing health information. However, the quality of video content varies greatly, leading to significant challenges. Some videos are misleading and provide inaccurate information to viewers, prompting professionals to advocate for stricter regulations. The Chinese government recently issued guidelines for media platforms to publish scientifically accurate health information, a move with global implications [57]. Enhancing the quality of health promotion videos has become a pressing issue requiring all stakeholders' attention. A high-quality health promotion video should be scientifically accurate, appeal to a broad audience, and be easily understandable while eliminating any misleading content. Therefore, rigorous evaluation of video quality is essential to ensure the dissemination of reliable information. Future research should focus on constructing and optimizing platforms to better cater to the public's health information needs.

XSL•FO
RenderX

Conclusions

This research gathered 153 videos about HTSs from TikTok, a popular short video–sharing social media platform in China, and comprehensively evaluated their information quality. The findings revealed that the videos lacked reliable sources and content quality. Overall, videos on the topic of HTSs produced by health care professionals from the burn department and burn plastic surgery department as well as those from first-tier and emerging first-tier Chinese cities demonstrated more significant insights regarding quality and reliability. They provide audiences with more reliable medical information. Therefore, people may prefer content from these departments and cities when seeking information about HTSs. As video-sharing platforms become increasingly popular sources of health information, it is essential to improve regulation and quality control. Users should be cautious when seeking health care management information on short video platforms. To ensure access to accurate information on hypertrophic scarring, we recommend referring to professional and authoritative sources and platforms to safeguard health effectively.

Acknowledgments

This work received support from the Natural Science Foundation of China (grant 82102340), the Postdoctoral Fund of Central Theater General Hospital (grant BSH022), and the Yuying Fund of Central Theater General Hospital (grant ZZYHL202101).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SW and KX developed and planned the study. SW and WL conducted video reviews and provided ratings. J Wu, KX, and J Wang collected and analyzed the data. J Wang initially drafted the manuscript, which was then reviewed and edited by J Wu. SW and WQ revised the manuscript for content. All authors contributed to writing, revising, and editing the manuscript and approved the final draft for submission.

Conflicts of Interest

None declared.

References

1. Finnerty CC, Jeschke MG, Branski LK, Barret JP, Dziewulski P, Herndon DN. Hypertrophic scarring: the greatest unmet challenge after burn injury. Lancet 2016 Oct 01;388(10052):1427-1436 [FREE Full text] [doi: 10.1016/S0140-6736(16)31406-4] [Medline: 27707499]
2. Ogawa R. The most current algorithms for the treatment and prevention of hypertrophic scars and keloids: a 2020 update of the algorithms published 10 years ago. Plast Reconstr Surg 2022 Jan 01;149(1):79e-94e [FREE Full text] [doi: 10.1097/PRS.0000000000008667] [Medline: 34813576]
3. Liu AH, Sun XL, Liu DZ, Xu F, Feng SJ, Zhang SY, et al. Epidemiological and clinical features of hypertrophic scar and keloid in Chinese college students: a university-based cross-sectional survey. Heliyon 2023 Apr;9(4):e15345 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e15345] [Medline: 37123979]
4. Marshall CD, Hu MS, Leavitt T, Barnes LA, Lorenz HP, Longaker MT. Cutaneous scarring: basic science, current treatments, and future directions. Adv Wound Care (New Rochelle) 2018 Feb 01;7(2):29-45 [FREE Full text] [doi: 10.1089/wound.2016.0696] [Medline: 29392092]
5. Obaidi N, Keenan C, Chan RK. Burn scar management and reconstructive surgery. Surg Clin North Am 2023 Jun;103(3):515-527. [doi: 10.1016/j.suc.2023.01.012] [Medline: 37149387]
6. Aarabi S, Longaker MT, Gurtner GC. Hypertrophic scar formation following burns and trauma: new approaches to treatment. PLoS Med 2007 Sep;4(9):e234 [FREE Full text] [doi: 10.1371/journal.pmed.0040234] [Medline: 17803351]
7. Atiyeh B, Ibrahim A. Nonsurgical management of hypertrophic scars: evidence-based therapies, standard practices, and emerging methods: an update. Aesthetic Plast Surg 2020 Aug;44(4):1345-1347. [doi: 10.1007/s00266-020-01766-3] [Medline: 32766896]
8. Sillence E, Blythe JM, Briggs P, Moss M. A revised model of trust in internet-based health information and advice: cross-sectional questionnaire study. J Med Internet Res 2019 Nov 11;21(11):e11125 [FREE Full text] [doi: 10.2196/11125] [Medline: 31710297]
9. Yigzaw KY, Wynn R, Marco-Ruiz L, Budrionis A, Oyeyemi SO, Fagerlund AJ, et al. The association between health information seeking on the internet and physician visits (the seventh Tromsø Study - part 4): population-based questionnaire study. J Med Internet Res 2020 Mar 05;22(3):e13120 [FREE Full text] [doi: 10.2196/13120] [Medline: 32134387]
10. Onder ME, Zengin O. YouTube as a source of information on gout: a quality analysis. Rheumatol Int 2021 Jul;41(7):1321-1328 [FREE Full text] [doi: 10.1007/s00296-021-04813-7] [Medline: 33646342]

XSL•FO
RenderX

11. Tan W, Liu Y, Shi Z, Zheng B, Feng L, Wang J, et al. Information quality of videos related to Helicobacter pylori infection on TikTok: cross-sectional study. Helicobacter 2024;29(1):e13029. [doi: 10.1111/hel.13029] [Medline: 37823482]
12. Kong W, Song S, Zhao YC, Zhu Q, Sha L. TikTok as a health information source: assessment of the quality of information in diabetes-related videos. J Med Internet Res 2021 Sep 01;23(9):e30409 [FREE Full text] [doi: 10.2196/30409] [Medline: 34468327]
13. Yang X, Xue X, Shi Z, Nan S, Lian C, Ji Z, et al. The reliability, functional quality, understandability, and actionability of fall prevention content in YouTube: an observational study. BMC Geriatr 2022 Aug 09;22(1):654 [FREE Full text] [doi: 10.1186/s12877-022-03330-x] [Medline: 35945509]
14. Shah N, Mathur VP, Kathuria V, Gupta T. Effectiveness of an educational video in improving oral health knowledge in a hospital setting. Indian J Dent 2016;7(2):70-75 [FREE Full text] [doi: 10.4103/0975-962X.184646] [Medline: 27433049]
15. Feng B, Malloch YZ, Kravitz RL, Verba S, Iosif A, Slavik G, et al. Assessing the effectiveness of a narrative-based patient education video for promoting opioid tapering. Patient Educ Couns 2021 Feb;104(2):329-336 [FREE Full text] [doi: 10.1016/j.pec.2020.08.019] [Medline: 32900605]
16. Rus HM, Cameron LD. Health communication in social media: message features predicting user engagement on diabetes-related Facebook pages. Ann Behav Med 2016 Oct;50(5):678-689. [doi: 10.1007/s12160-016-9793-9] [Medline: 27059761]
17. Chen K, Zhou L, Zhao R, Tang Y. Assessing the quality of hearing aids-related videos on TikTok. Front Public Health 2022;10:901976 [FREE Full text] [doi: 10.3389/fpubh.2022.901976] [Medline: 35812482]
18. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. Health Informatics J 2015 Sep;21(3):173-194 [FREE Full text] [doi: 10.1177/1460458213512220] [Medline: 24670899]
19. Song S, Xue X, Zhao YC, Li J, Zhu Q, Zhao M. Short-video apps as a health information source for chronic obstructive pulmonary disease: information quality assessment of TikTok videos. J Med Internet Res 2021 Dec 20;23(12):e28318 [FREE Full text] [doi: 10.2196/28318] [Medline: 34931996]
20. Zheng S, Tong X, Wan D, Hu C, Hu Q, Ke Q. Quality and reliability of liver cancer-related short Chinese videos on TikTok and Bilibili: cross-sectional content analysis study. J Med Internet Res 2023 Jul 05;25:e47210 [FREE Full text] [doi: 10.2196/47210] [Medline: 37405825]
21. Chandrasekaran R, Konaraddi K, Sharma SS, Moustakas E. Text-mining and video analytics of COVID-19 narratives shared by patients on YouTube. J Med Syst 2024 Feb 15;48(1):21. [doi: 10.1007/s10916-024-02047-1] [Medline: 38358554]
22. TikTok revenue and usage statistics (2024). Business of Apps. URL: https://www.businessofapps.com/data/tik-tok-statistics/ [accessed 2024-12-08]
23. Zenone M, Ow N, Barbic S. TikTok and public health: a proposed research agenda. BMJ Glob Health 2021 Nov;6(11):e007648 [FREE Full text] [doi: 10.1136/bmjgh-2021-007648] [Medline: 34819326]
24. Subramanian T, Araghi K, Akosman I, Tuma O, Hassan A, Lahooti A, et al. Quality of spine surgery information on social media: a DISCERN analysis of TikTok videos. Neurospine 2023 Dec;20(4):1443-1449 [FREE Full text] [doi: 10.14245/ns.2346700.350] [Medline: 38171310]
25. Ming S, Han J, Li M, Liu Y, Xie K, Lei B. TikTok and adolescent vision health: content and information quality assessment of the top short videos related to myopia. Front Public Health 2022;10:1068582 [FREE Full text] [doi: 10.3389/fpubh.2022.1068582] [Medline: 36684892]
26. Liu H, Peng J, Li L, Deng A, Huang X, Yin G, et al. Assessment of the reliability and quality of breast cancer related videos on TikTok and Bilibili: cross-sectional study in China. Front Public Health 2023;11:1296386 [FREE Full text] [doi: 10.3389/fpubh.2023.1296386] [Medline: 38317686]
27. He Z, Wang Z, Song Y, Liu Y, Kang L, Fang X, et al. The reliability and quality of short videos as a source of dietary guidance for inflammatory bowel disease: cross-sectional study. J Med Internet Res 2023 Feb 09;25:e41518 [FREE Full text] [doi: 10.2196/41518] [Medline: 36757757]
28. Zhang J, Han P, Tang Y, Xi W, Xiao X, Yang F. Popular science and education of cosmetic surgery in China: quality and reliability evaluation of Douyin short videos. Health Expect 2023 Jun;26(3):1221-1226 [FREE Full text] [doi: 10.1111/hex.13737] [Medline: 36807969]
29. Cole C. Looking for information: a survey of research on information seeking, needs, and behavior (4th edition). Donald O.Case and Lisa M.Given. Bingley, UK: Emerald Group Publishing, 2016. 528 pp. $82.95 (hardcover). (ISBN: 9781785609688). J Assoc Inf Sci Technol 2016 Dec 21;68(9):2284-2286. [doi: 10.1002/asi.23778]
30. The 2024 charm ranking of new first-tier cities has been released: Wuxi returns, and Kunming drops out again. Zhicheng Data Platform. URL: https://www.datayicai.com/report/detail/1000009 [accessed 2024-07-18]
31. Goobie GC, Guler SA, Johannson KA, Fisher JH, Ryerson CJ. YouTube videos as a source of misinformation on idiopathic pulmonary fibrosis. Ann Am Thorac Soc 2019 May;16(5):572-579. [doi: 10.1513/AnnalsATS.201809-644OC] [Medline: 30608877]
32. Xue X, Yang X, Xu W, Liu G, Xie Y, Ji Z. TikTok as an information hodgepodge: evaluation of the quality and reliability of genitourinary cancers related content. Front Oncol 2022;12:789956 [FREE Full text] [doi: 10.3389/fonc.2022.789956] [Medline: 35242704]

XSL•FO
RenderX

33. Zhang Y, Sun Y, Xie B. Quality of health information for consumers on the web: a systematic review of indicators, criteria, tools, and evaluation results. J Assoc Inf Sci Technol 2015 Apr 29;66(10):2071-2084. [doi: 10.1002/asi.23311]
34. Bashir K, Nazim M. The impact of artificial intelligence on employee performance and satisfaction: a study from academic sector. SSRN Journal Preprint posted online January 7, 2022 [FREE Full text] [doi: 10.2139/ssrn.4951891]
35. Singh AG, Singh S, Singh PP. YouTube for information on rheumatoid arthritis--a wakeup call? J Rheumatol 2012 May;39(5):899-903. [doi: 10.3899/jrheum.111114] [Medline: 22467934]
36. Karakoyun A, Yildirim A. YouTube videos as a source of information concerning Behçet's disease: a reliability and quality analysis. Rheumatol Int 2021 Dec 29;41(12):2117-2123. [doi: 10.1007/s00296-021-05009-9] [Medline: 34590188]
37. Yang S, Zhan J, Xu X. Is TikTok a high-quality source of information on thyroid cancer? Endocrine 2023 Aug;81(2):270-276. [doi: 10.1007/s12020-023-03332-8] [Medline: 36840912]
38. Bernard A, Langille M, Hughes S, Rose C, Leddin D, Veldhuyzen van Zanten S. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. Am J Gastroenterol 2007 Sep;102(9):2070-2077. [doi: 10.1111/j.1572-0241.2007.01325.x] [Medline: 17511753]
39. Ferhatoglu MF, Kartal A, Ekici U, Gurkan A. Evaluation of the reliability, utility, and quality of the information in sleeve gastrectomy videos shared on open access video sharing platform YouTube. Obes Surg 2019 May;29(5):1477-1484. [doi: 10.1007/s11695-019-03738-2] [Medline: 30706318]
40. Harwood RH, Goldberg SE, Brand A, van Der Wardt V, Booth V, Di Lorito C, et al. Promoting activity, independence, and stability in early dementia and mild cognitive impairment (PrAISED): randomised controlled trial. BMJ 2023 Aug 29;382:e074787 [FREE Full text] [doi: 10.1136/bmj-2023-074787] [Medline: 37643788]
41. Chan C, Sounderajah V, Daniels E, Acharya A, Clarke J, Yalamanchili S, et al. The reliability and quality of YouTube videos as a source of public health information regarding COVID-19 vaccination: cross-sectional study. JMIR Public Health Surveill 2021 Jul 08;7(7):e29942 [FREE Full text] [doi: 10.2196/29942] [Medline: 34081599]
42. Stellefson M, Chaney B, Ochipa K, Chaney D, Haider Z, Hanik B, et al. YouTube as a source of chronic obstructive pulmonary disease patient education: a social media content analysis. Chron Respir Dis 2014 May;11(2):61-71 [FREE Full text] [doi: 10.1177/1479972314525058] [Medline: 24659212]
43. D'Ambrosi R, Bellato E, Bullitta G, Cecere AB, Corona K, De Crescenzo A, SICSeG (Italian Society of Shoulder, Elbow Surgery) Research Committee. TikTok content as a source of health education regarding epicondylitis: a content analysis. J Orthop Traumatol 2024 Mar 23;25(1):14 [FREE Full text] [doi: 10.1186/s10195-024-00757-3] [Medline: 38521890]
44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
45. Prestin A, Vieux SN, Chou WY. Is online health activity alive and well or flatlining? Findings from 10 years of the health information national trends survey. J Health Commun 2015;20(7):790-798. [doi: 10.1080/10810730.2015.1018590] [Medline: 26042588]
46. Gong X, Dong B, Li L, Shen D, Rong Z. TikTok video as a health education source of information on heart failure in China: a content analysis. Front Public Health 2023;11:1315393 [FREE Full text] [doi: 10.3389/fpubh.2023.1315393] [Medline: 38146471]
47. TikTok video duration by followers. Statista. URL: https://www.statista.com/statistics/1372569/tiktok-video-duration-by-number-of-views/ [accessed 2024-12-11]
48. Li Y, Guan M, Hammond P, Berrey LE. Communicating COVID-19 information on TikTok: a content analysis of TikTok videos from official accounts featured in the COVID-19 information hub. Health Educ Res 2021 Jul 12;36(3):261-271 [FREE Full text] [doi: 10.1093/her/cyab010] [Medline: 33667311]
49. Bai G, Fu K, Fu W, Liu G. Quality of internet videos related to pediatric urology in mainland China: a cross-sectional study. Front Public Health 2022;10:924748 [FREE Full text] [doi: 10.3389/fpubh.2022.924748] [Medline: 35784238]
50. Du R, Zhang Y, Wang M, Lu N, Hu Y. TikTok and Bilibili as sources of information on Helicobacter pylori in China: a content and quality analysis. Helicobacter 2023 Oct;28(5):e13007. [doi: 10.1111/hel.13007] [Medline: 37452727]
51. Gaş S, Zincir Ö, Bozkurt AP. Are YouTube videos useful for patients interested in botulinum toxin for bruxism? J Oral Maxillofac Surg 2019 Sep;77(9):1776-1783. [doi: 10.1016/j.joms.2019.04.004] [Medline: 31077674]
52. Ergenç M, Uprak TK. YouTube as a source of information on Helicobacter pylori: content and quality analysis. Helicobacter 2023 Aug;28(4):e12971. [doi: 10.1111/hel.12971] [Medline: 36942858]
53. Li B, Liu M, Liu J, Zhang Y, Yang W, Xie L. Quality assessment of health science-related short videos on TikTok: a scoping review. Int J Med Inform 2024 Jun;186:105426. [doi: 10.1016/j.ijmedinf.2024.105426] [Medline: 38531256]
54. Fullard AC, Johnston SM, Hehir DJ. Quality and reliability evaluation of current internet information regarding mesh use in inguinal hernia surgery using HONcode and the DISCERN instrument. Hernia 2021 Oct 14;25(5):1325-1330. [doi: 10.1007/s10029-021-02406-8] [Medline: 33852079]
55. Karataş ME, Karataş G. Evaluating the reliability and quality of the upper eyelid blepharoplasty videos on YouTube. Aesthetic Plast Surg 2022 Apr 30;46(2):754-759. [doi: 10.1007/s00266-021-02504-z] [Medline: 34331096]
56. Kartal A, Kebudi A. Evaluation of the reliability, utility, and quality of information used in total extraperitoneal procedure for inguinal hernia repair videos shared on WebSurg. Cureus 2019 Sep 04;11(9):e5566 [FREE Full text] [doi: 10.7759/cureus.5566] [Medline: 31695985]

XSL•FO
RenderX

57. Guidance on establishing and improving the release and dissemination mechanism of health science knowledge through all media. National Health Commission of the People's Republic of China. URL: http://www.nhc.gov.cn/cms-search/xxgk/%20getManuscriptXxgk.htm?id=1c67c12c86b44fd2afb8e424a2477091 [accessed 2024-06-03]

Abbreviations

GQS: Global Quality Scale
HONcode: Health on the Net Foundation Code of Conduct
HTS: hypertrophic scar
***JAMA*:** *Journal of the American Medical Association*
mDISCERN: modified DISCERN

Edited by T Mackey; submitted 26.07.24; peer-reviewed by F Medina, B Eapen, R Chandrasekaran; comments to author 17.10.24; revised version received 15.12.24; accepted 19.03.25; published 29.04.25.

Please cite as:
Wang J, Xu K, Wu J, Liang W, Qiu W, Wang S
Evaluating the Content and Quality of Videos Related to Hypertrophic Scarring on TikTok in China: Cross-Sectional Study
JMIR Infodemiology 2025;5:e64792
URL: https://infodemiology.jmir.org/2025/1/e64792
doi:10.2196/64792
PMID:

©Jiangkun Wang, Kai Xu, Juanjuan Wu, Wen Liang, Weiming Qiu, Song Wang. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 29.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Social Media and the Evolution of Vaccine Preferences During the COVID-19 Pandemic: Discrete Choice Experiment

Robbie Maris[1*], MMS; Zack Dorner[2*], PhD; Stephane Hess[3*], PhD; Steven Tucker[4], PhD

[1]Centre for Education Policy and Equalising Opportunities, University College London, 20 Bedford Way, London, United Kingdom

[2]Department of Environmental Management, Lincoln University, Lincoln, New Zealand

[3]Choice Modelling Centre, Institute for Transport Studies, University of Leeds, Leeds, United Kingdom

[4]Waikato Management School, University of Waikato, Hamilton, New Zealand

[*]these authors contributed equally

Corresponding Author:
Robbie Maris, MMS
Centre for Education Policy and Equalising Opportunities, University College London, 20 Bedford Way, London, United Kingdom

Abstract

Background: Vaccine information and misinformation are spread through social media in ways that may vary by platform. Understanding the role social media plays in shaping vaccine preferences is crucial for policymakers and researchers.

Objective: This study aims to test whether social media use is associated with changes in vaccine preferences during the COVID-19 pandemic in New Zealand, and whether trust in sources of information has a moderating role.

Methods: Our data consist of a balanced panel of 257 web-based respondents in New Zealand in August 2020, October-November 2020, and March-April 2021. We use a novel approach with stated choice panel data to study transitions between different vaccine preference groups. We analyze the associations between these transitions and social media use. We classify respondents as resistant (never chose a vaccine), hesitant (chose a vaccine between 1 and 5 times), and provaccine (chose a vaccine 6 out of 6 times) in each wave of data.

Results: We found a positive or neutral association between social media use and vaccine uptake. Facebook, Twitter (pre-2022), and TikTok users who are provaccine are less likely to become hesitant or resistant. Facebook and Instagram users who are hesitant are more likely to become pro. Some social media platforms may have a more positive association with vaccine uptake preferences for those who do not trust the government.

Conclusions: The paper contributes to the wider literature, which shows social media can be associated with reinforcing both pro and antivaccination sentiment, and these results depend on where individuals get their information from and their trust in such sources.

(JMIR Infodemiology 2025;5:e66081) doi:10.2196/66081

KEYWORDS

health information sources; social media; COVID-19; SARS-COV-2; respiratory; infectious; pulmonary; pandemic; vaccination; stated preferences; attitude; perspective; discrete choice experiment; misinformation; global pandemic; sentiment analysis; social listening; public health

Introduction

The COVID-19 pandemic has had substantial global impacts, with over 6 million deaths reported to the World Health Organization to date, alongside other impacts such as 5% - 20% of cases leading to long COVID and widespread economic disruption [1-3]. Vaccination is a vital tool for combating the pandemic, both now and into the future [4]. Despite the clear benefits of vaccination against COVID-19 and other diseases, many people are hesitant or resistant [5-7]. A rapidly emerging body of literature focuses on the characteristics associated with COVID-19 vaccine uptake [7-10]. Indeed, it is important to use data from this period to study what influences vaccination preferences, given the unprecedented global attention on COVID-19 vaccination, the speed of vaccine development, and the potential for future pandemics [11]. Given worries that vaccination hesitancy could be on the rise, it is doubly important to learn from the COVID-19 pandemic [12].

The role of information, who provides that information, and trust in that information have all been shown to be key factors in vaccination uptake [6,13]. Social media platforms are increasingly used as a source of information [14]. The nature of social media means that information sources and types vary greatly, with everything from conspiracy theories to scientific information. Both types of information have been shown to disseminate in a relatively similar manner, with some differences

XSL•FO
RenderX

by platform [15,16]. There is evidence that users tend to seek out information that accords with their pre-existing beliefs, regardless of whether it is factual or not [17]. Thus, users end up in echo-chambers reinforcing their own beliefs regarding vaccination and other socially important topics, particularly when issues are controversial [16,18,19]. Indeed, prior work suggests antivaccine sentiment generates more engagement than provaccine sentiment [20] and that exposure to misinformation can decrease intent to vaccinate [12,21,22]. However, there are still relatively few papers that use panel data to investigate the relationship between social media use and vaccine preferences, and our understanding of this relationship is still limited.

A range of studies use survey methods to investigate the association between social media use and willingness to vaccinate against COVID-19. These studies show a mixed picture. Several multi-national studies of European nations and the United States of America show either a negative or nonsignificant association between social media and willingness to vaccinate, depending on the country [23,24]. Similarly, Park et al [8] conducted a survey in the United States and found the lowest vaccine acceptance levels among those who rely on social media for information. In another USA-based survey, Al-Ugdah et al [25] find a positive association between social media users and willingness to vaccinate, but the opposite when no medical or government source of information is used alongside it. Bendau et al [26] found that Germans who report the use of social media to garner information on vaccination have higher acceptance than those who do not. Wang et al [27] report a similar finding in China. These studies all use cross-sectional surveys, but do not have a panel sample recording changes over time.

There are fewer studies using panel surveys, tracking changes in intention to vaccinate and their relationship with social media use. Romer and Jamieson [28] conducted a panel survey in the United States in March and July 2020. They find that social media use is associated with conspiracy beliefs and vaccine resistance, with underlying political ideology playing a key role. Beliefs and preferences around health behaviors were relatively stable over the survey period. In their panel survey, Theocharis et al [29] find that COVID-19 conspiracy beliefs change over time. Twitter reduces the propensity to believe in this conspiracy, whereas other platforms (Facebook, YouTube, Messenger, and WhatsApp) increase these beliefs. They do not look specifically at intention to vaccinate.

In this paper, we investigate whether social media use is positive or negative for COVID-19 vaccination uptake, both on average and in terms of polarization over time. We conducted a web-based stated preferences panel survey of New Zealanders over 3 waves between August 2020 and April 2021, while vaccines were being developed, approved by regulators, and initially rolled out. New Zealand provides an interesting case study due to the high levels of government trust at the time [30-32], Thus, New Zealand has the potential to show a best-case scenario for the impact of social media on vaccine uptake.

We collected data on respondents' social media use over the previous 6 months, by platform, in the third survey wave. This 6-month period roughly covers the time between the first survey wave and the third. To test for polarization, we model transitions between types of vaccine preferences between waves using a Logit model, to test whether social media use is associated with a change or a reinforcement of earlier choices. We then model the association of social media use with likely vaccine uptake overall (positive or negative on average) using a pooled partial proportional odds (PPO) model with time fixed effects [33]. Finally, we test how trust in government, friends, and family could moderate the effect of platform use. We contribute to the literature by furthering our understanding of the potential role of social media and trust in information on shaping vaccine preferences. Additionally, we demonstrate a novel approach to analyzing stated preference data, we use panel data to explore changes over time, and we examine preferences during the COVID-19 pandemic in New Zealand—an informative context to study the interactions between social media use, trust in information, and vaccine preferences.

Methods

Data

We use the New Zealand data from Hess et al [7], which is a web-based stated choice survey on COVID-19 vaccines, undertaken on 20 - 29 August, 2020 in the case of New Zealand. Additionally, we repeated the survey for the panel from 26 October to 18 November, 2020 (wave 2) and 23 March to 24 April, 2021 (wave 3). This element of the research design was important to be able to track changes over time. The panel was recruited using the Qualtrics survey company, and was initially representative of the New Zealand population of 18 years and older for age and gender (reweighting as needed for analysis, described at the end of this section). Qualtrics recruited into the survey from a pool of potential New Zealand-based participants who are available for the purpose of collecting representative data in web-based surveys, with a small financial incentive. We checked the survey responses across the panel, and there were no discernible signs of poor quality responses (eg, very short completion time).

The time period of the 3 waves captures a key moment for information gathering and preference formation regarding vaccines. Development and testing were being undertaken during waves 1 and 2. The Pfizer-BioNTech vaccine was approved by the United States Food and Drug Administration in December 2020, with Medsafe in New Zealand following on 3 February 2021. As the only vaccine initially offered in New Zealand, the rollout of Pfizer-BioNTech to a select group had begun by wave 3, with more advanced rollouts underway in other countries.

Additionally, over this time period, trust in the government increased and was high due to their strong lockdown policies, which were successful in eliminating community COVID-19 cases within New Zealand [30-32]. Demonstrating this high support, Prime Minister Jacinda Ardern's Labor Party won re-election in October 2020, increasing their vote share to the highest ever for a single political party since New Zealand introduced its proportional representation electoral system in 1996, at 50% of the vote.

XSL•FO
RenderX

The main part of the survey is a stated choice survey, using a technique often called a discrete choice experiment [34]. Respondents were presented with 6 hypothetical, but realistic, choices regarding a COVID-19 vaccine. In each choice, respondents were asked to choose their preferred option out of 2 vaccines (each available as paid or free with a wait time) and no vaccine, giving them 4 vaccine choices and one no vaccine. The vaccines varied by risk of infection, serious illness, protection duration, mild or severe side effects, and population coverage. We used the Qualtrics web-based survey platform. It randomly assigned each respondent one of 6 blocks of choice sets. Respondents saw the same block of questions as initially assigned to them in each survey wave.

The choice sets were generated using a D-efficient design [35] from the NGene software package (ChoiceMetrics) [36]. They were calibrated such that an overall view of each respondent's vaccine preferences over key vaccine attributes could be established, within credible ranges. Thus, they give us an overall view of COVID-19 vaccine preferences, rather than asking about specific vaccines that were in development. Each of the 4 question blocks is balanced such that a raw count of the number of vaccines chosen out of 6 is roughly comparable between respondents. For more details on the questions, see Hess et al [7].

Discrete, stated choice surveys are widely used in health due to their effectiveness in providing an accurate and fine-scale measure of individuals' preferences [7,34]. While stated survey measures of sentiment, willingness, or intention to vaccinate are also valid approaches used in the literature (as covered in the introduction), the strengths of a discrete choice experiment mean it is highly suited to understanding vaccine preferences within a specific case study using hypothetical but realistic vaccine options. This is particularly important as the vaccines were still under development and their attributes were not clear. Hence, we could also not measure real vaccine uptake during this period [22].

For each wave, we sum the number of times respondent i selects a vaccine option across the 6 choice scenarios. We then classify individuals into 3 vaccine uptake categories, which we denote as resistant (vaccine chosen 0 times), hesitant (vaccine chosen 1 to 5 times), and provaccine (vaccine chosen 6 of 6 times). As noted in the previous paragraph, there may be some differences between the 4 choice blocks that could drive different counts between individuals, but these should not be significant and are unlikely to affect our classification of respondents into 3 categories (this classification approach reduces the influence of any minor measurement error across the blocks). Respondents also saw the same choice block in each wave, so differences between waves, within individuals, are not due to changes in questions.

In the New Zealand survey of Hess et al [7] in wave 3 only, we asked about the frequency of social media use, by platform, in the last 6 months. This period roughly equates with the time between the first and third survey waves, so that we can measure the association between use and changes in vaccination uptake preferences between survey waves. Platforms included are Facebook, Instagram, Twitter, and TikTok. We use dummy coding, with anyone using the platform at least once a week being coded as a user.

We asked participants about their trust in various sources for information about COVID-19 vaccines. We test their moderating effect on social media and vaccine uptake. They were measured on a 5-point Likert scale, and we coded them as dummy variables of trust (1) versus neutral or distrust (0).

We pool our data across the 3 waves, including social media use by type, a rich set of covariates, and time fixed effects. The original wave 1 sample included was representative by age and gender; however, due to attrition, the wave 3 sample needs some reweighting. We reweight for all analyses on the basis of age, gender, and ethnicity from the 2018 census [37,38]. We drop 110 observations with missing data on the variables we use in the modeling. As we collected social media use in wave 3 only, we are left with a balanced panel across the 3 waves of 257.

Ethical Considerations

This project was granted ethics approval for research by the Waikato Management School Ethics Committee (application WMS 20/68). The panel was recruited using the Qualtrics survey company, and respondents were provided a small financial incentive by Qualtrics to participate in the sample. Participants provided their informed consent to participate in the study and had the option to withdraw their data at any time. All data were anonymized by Qualtrics before being sent to the research team to protect the privacy and confidentiality of our research participants.

Empirical Modeling

We are interested in understanding three main points from the data as follows: (1) the association between social media use (by platform) and an individual becoming less (or more) provaccine over time; (2) the overall association of social media use (by platform) with likely vaccine uptake, and (3) the moderating influence of trust in information sources and vaccine uptake.

To understand the first point, we estimate simple Logit models for each transition between waves. This parsimonious modeling approach allows us to estimate the association between social media use and an increase/decrease in likely vaccine uptake over time, at an individual level. Our dependent variable is a binary indicator Lijk=1 for when individual i transitioned to a lower vaccine uptake group between waves j and k, and Hijk indicating if individual i transitioned to a higher vaccine uptake. We include the full set of control variables. We model the transitions separately for the different starting groups in wave j because these groups are fundamentally different from each other. For example, we model Lijk for provaccine individuals, for each wave. This allows us to understand whether social media is associated with staying provaccine or becoming less provaccine between waves j and k. If we had not separated these groups out, we would have implicitly treated transitions from being provaccine to hesitant the same as moving from hesitant to resistant.

For the latter 2 points, we estimate a PPO model [33] on the ordinal dependent variable of stated vaccine uptake of resistance

XSL•FO
RenderX

through to provaccine. This model can give us an overall estimate of the association between social media platforms and vaccine uptake. However, it will not give us as clear a picture of transition over time as the Logit approach outlined above.

The PPO is a specific case of a generalized ordered logit (gologit). It allows us to take into account the ordinal nature of our outcome variable, but relax the proportional odds assumption of the more commonly used ordered logit, when needed [39,40]. The proportional odds assumption states that if we estimate a series of cumulative logit models by successively collapsing the ordinal variable into a binary variable (with different cutoffs), the odds ratios for each regressor will be equal across all models (within the limits of random error). When this assumption is violated, ordinal logit models may produce misleading and biased results [40]. Brant [41] developed a test to determine whether the proportional odds assumption holds for a set of variables. In our case, we find the proportional odds assumption to be violated for several covariates and the overall model. Hence, the PPO model allows us to relax the proportional odds assumption when needed, but keep it for the covariates where it is not violated.

An alternative model we could have estimated, which also does not require the proportional odds assumption to hold, is the multinomial logit model [39]. This model runs a series of logit regressions on every possible binary combination of the categorical variable (in our case, {1 vs 2, 1 vs 3, 2 vs 3}). However, this approach fails to account for the inherent ordering of Uit and computes a number of unnecessary parameters [40].

Using maximum likelihood, our PPO model estimates the probability that individual i in wave t is resistant, hesitant, or pro, represented by Uit∈{1,2,3}, respectively:

$$P(Uit>k)=exp(\alpha k+Xit\beta 1k+Zit\beta 2)1+[exp(\alpha k+Xit\beta 1k+Zit\beta 2)],k=(1,2).$$

In our case, we can think of the PPO model as essentially 2 sets of logits, modeling the probability of being in categories 2 and 3 over 1 for the first set of coefficients (k=1), and category 3 over 1 and 2 for the second set of coefficients (k=2). Of course, both logit models are estimated simultaneously, and both are used to calculate predicted probabilities of being in any given category, as represented in the above equation. The coefficients β1k for covariates Xit vary over k. The coefficients β2 for covariates Zit do not vary over k when they do not violate the proportional odds assumption, reducing the number of coefficients to estimate. Covariates are placed into either Xit or Zit, depending on the results of Brant tests. These include time-invariant variables, such as social media use, and wave time fixed effects.

To investigate our aim (2) (stated above), we estimate a separate PPO for each social media platform, plus the full set of controls. For aim (3), we interact social media use with trust in different sources of information on vaccines (friends or family, the government, and social media).

Results

Descriptive Statistics

We start by presenting summary statistics in Table 1 for the variables we use in our modeling (pooled across the 3 waves). In terms of social media use, most of the sample uses at least one social media platform out of the 4 we questioned them on (188/257, 73.2%). A majority of respondents use Facebook (177/257, 68.9%), with use of Instagram, Twitter, and TikTok being significantly lower at 30.4% (n=78), 15.2% (n=39), and 7% (n=18), respectively. Most respondents trust the government (199/257, 77.4%) and a minority trust their friends and family (84/257, 32.7%) as sources of vaccine information. As discussed in the Data section, our New Zealand sample is an informative context to study vaccine preferences, given the relatively high levels of trust in government, as COVID-19 vaccines were rolled out. Moreover, most respondents distrust social media generally (195/257, 75.9%) as a source of vaccine information. Across the 3 waves, we classify 76.1% (587/771) of responses as belonging to the provaccine uptake group, 18.3% (141/771) of responses as belonging to the vaccine-hesitant uptake group, and 5.6% (43/771) of responses as belonging to the resistant uptake group.

XSL•FO
RenderX

Table . Summary statistics.

Variable	Values		Total	Observations
	N (%)	Mean (SD)		
Resistant	5.6	—[a]	257	771
Hesitant	18.3	—	257	771
Pro	76.1	—	257	771
Social media user	73.2	—	257	257
Facebook user	68.9	—	257	257
Instagram user	30.4	—	257	257
Twitter user[b]	15.2	—	257	257
TikTok user	7.0	—	257	257
Female	42.9	—	257	257
Male	57.1	—	257	257
Maori and Pacific	5.4	—	257	257
Trusts family or friends	32.7	—	257	257
Trusts government	77.4	—	257	257
Distrusts social media	75.9	—	257	257
University-educated	43.2	—	257	257
Income (NZ $, 000s)[c]	—	46.1 (35)	257	771
Age (years)	—	52.0 (15.7)	257	257

[a]Not applicable.

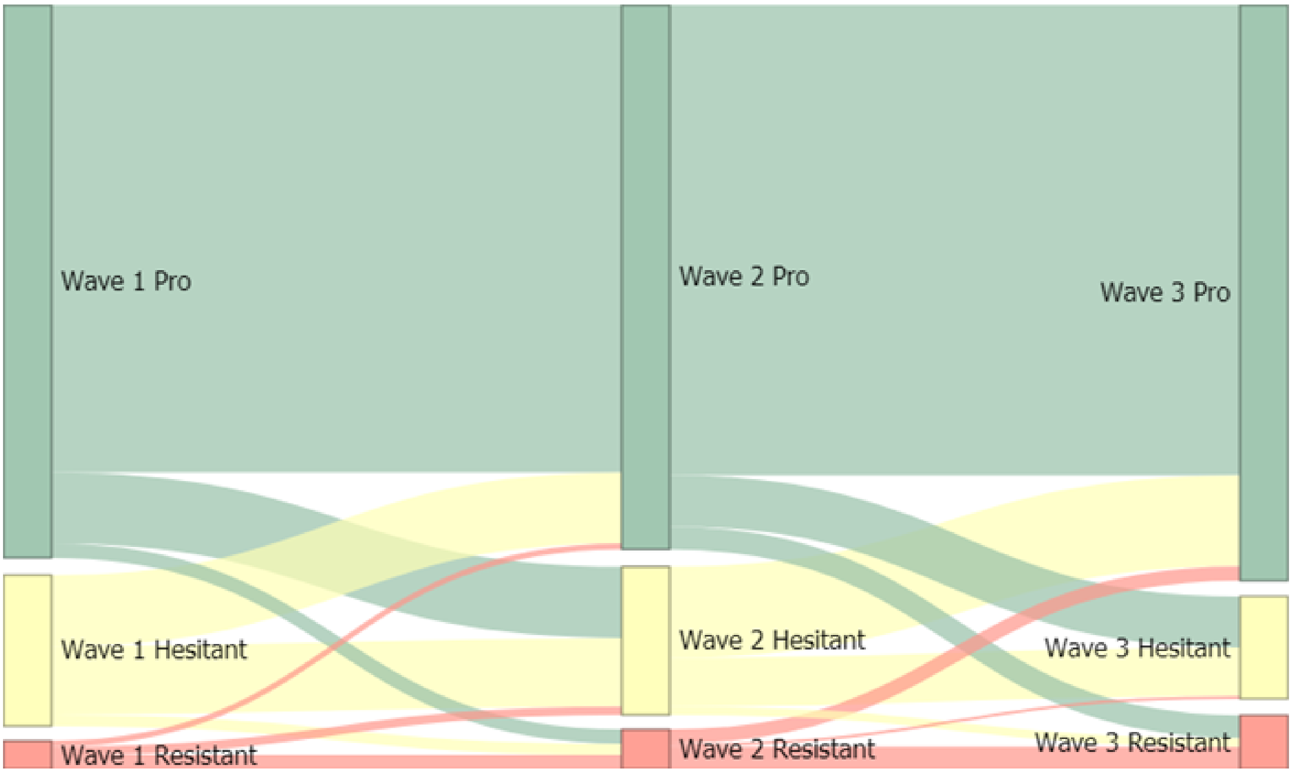
[b]This work was undertaken before the change of ownership and approach at Twitter in 2022. Most users of less-common social media platforms in our sample (TikTok, Twitter and Instagram) also used Facebook. Only 6% of social media users did not use Facebook.

[c] This variable is measured in New Zealand dollars (NZ $). At the modal time this variable was measured (August 25, 2020), the exchanged rate to US dollars (US $) was NZ $1 = US $0.65,

In Figure 1, we use a Sankey flow diagram to visually depict transitions between vaccine uptake groups across the 3 waves. Evidently, there are significant movements between vaccine uptake groups over time, particularly for the hesitant. Overall, we see movement away from hesitant by wave 3, towards provaccine and resistant. There is also a considerable proportion of respondents who remain stable in their vaccine uptake group membership across the 3 waves. Multimedia Appendix 1 contains the transition matrices.

XSL•FO
RenderX

Figure 1. Sankey diagram of individual transitions between vaccine uptake groups across waves.



Transition Models

In Table 2, we present the marginal effects at the means (MEMs) of social media use on the probability of provaccine individuals moving to a lower vaccine uptake group. These marginal effects show the change in probability for the dependent variable occurring when all other variables are set at their sample means. The first column shows the transition between waves 1 and 2, and the second column represents the transition between waves 2 and 3. Each row represents the MEMs from a model with a full set of controls, and one type of social media platform (or all social media platforms in the case of the first row). The full logit results, including controls, are in Multimedia Appendix 1.

Table . Marginal effects at the means (MEMs) of social media use on the probability that provaccine individuals decrease uptake between waves. Cluster robust SEs in parentheses. Results have been reweighted on age, gender, and ethnicity.

MEMs for platform use	Dep. variable: decrease in uptake	
	Waves 1 to 2 (n=194), mean (SE)	Waves 2 to 3 (n=191), mean (SE)
Social media use	−0.0187 (0.0678)	−0.109 (0.0879)
Facebook use	0.0161 (0.0583)	−0.156 (0.0829)[a]
Instagram use	−0.0484 (0.0549)	−0.132 (0.0827)
Twitter use	−0.164 (0.0566)[b]	−0.145 (0.0777)[a]
TikTok use	−0.151 (0.0548)[b]	−0.144 (0.0809)[a]

[a]$P<.1$
[b]$P<.01$

We can see that Twitter and TikTok are associated with a significantly lower probability of provaccine individuals decreasing uptake between waves 1 and 2 (P<.01). Twitter users are 16.4 percentage points less likely to decrease uptake, and TikTok users are 15.1 percentage points less likely. The MEMs are similar in size for these 2 platforms for the waves 2 to 3 transition (however, the MEMs are only marginally significant). The MEMs are significantly stronger for the waves 2 to 3 transition for general social media, Facebook, and Instagram use. For the waves 2 to 3 transition, provaccine Facebook users are 15.6 percentage points less likely to decrease vaccine uptake (significant at the 10% level). We do not estimate this model for the vaccine-hesitant group moving to resistant, as there are too few such transitions.

We estimate the increase in uptake from hesitant in Table 3. We find statistically significant MEMs of social media use on the probability of positive changes in uptake for the waves 1 to 2 transition. Hesitant social media, Facebook, and Instagram users are 29.6, 23.4, and 45.3 percentage points more likely to increase uptake (respectively) between waves 1 and 2. We find no significant MEMs for the waves 2 to 3 transition. We do not

XSL•FO
RenderX

include any other transition models due to sample size and rarity of transitions. For example, only 4 individuals went from hesitant to resistant between waves 1 and 2, making it impossible to model these transitions using our suite of covariates.

Table . Marginal effects at the means (MEMs) of social media use on the probability that vaccine-hesitant individuals increase uptake between waves. Cluster robust SEs in parentheses. Results have been reweighted on age, gender, and ethnicity.

MEMs for platform use	Dep. variable: increase in uptake	
	Waves 1 to 2 (n=53), mean (SE)	Waves 2 to 3 (n=52), mean (SE)
Social media use	0.296 (0.105)[a]	0.108 (0.157)
Facebook use	0.234 (0.112)[b]	0.0694 (0.151)
Instagram use	0.453 (0.150)[a]	−0.101 (0.266)
Twitter use	0.0264 (0.131)	0.166 (0.179)
TikTok use	0.107 (0.200)	0.0315 (0.265)

[a]*P*<.01

[b]*P*<.05

Partial Proportional Odds Model Results

In Table 4, we present the coefficients for the full PPO results of our base models with pooled data, wave fixed effects, and no interactions. Each pair of columns shows a separate model, varying only by the social media platform dummy included in the model. The first set of columns, (1), is social media use of the specified platform. The left column, labeled 1 versus 2, 3, models being hesitant or provaccine, over being resistant. Hence, a positive coefficient indicates a more provaccine orientation (more likely to be hesitant or provaccine than resistant). The right column, labeled 1, 2 versus 3, models being provaccine, over being hesitant or resistant, with positive coefficients again indicating a more provaccine orientation. Where coefficients are missing, these have been restricted to being the same across both columns, as the proportional odds assumption is not violated. As described in the Methods section, our modeling approach allows us to relax the assumption that covariates have the same impact on preferences across different transitions. Moreover, our results in Table 4 show how social media use and a range of individual characteristics correlate with our novel vaccine preference groupings derived from discrete choice experiment data.

XSL•FO
RenderX

Table . Full partial proportional odds (PPO) modeling results with time fixed effects, excluding interactions. Each pair of columns represents a separate model, varying only by type of Social media user, in the column heading. Cluster robust SEs in parentheses, clustered at the individual level. Results have been reweighted on age, gender, and ethnicity. 1 versus 2, 3 shows the model predicting the likelihood of being 2 (hesitant) or 3 (pro), over 1 (resistant), where a positive number shows more likely to be in 2 or 3. No coefficient in 1, 2 versus 3 means the model is restricted to assume the same coefficient across both columns, as the proportional odds assumption is not violated for that variable. Wave FE (fixed effect) is a fixed effect, coded as a dummy variable for that wave.

	Dependent variable: vaccine preference, where resistant is 1 and pro is 3									
	(1) Social media, mean (SE)		(2) Facebook, mean (SE)		(3) Instagram, mean (SE)		(4) Twitter, mean (SE)		(5) TikTok, mean (SE)	
	1 versus 2, 3	1, 2 versus 3 (n=257, Observations=771)	1 versus 2, 3	1, 2 versus 3 (n=257, Observations=771)	1 versus 2, 3	1, 2 versus 3 (n=257, Observations=771)	1 versus 2, 3	1, 2 versus 3 (n=257, Observations=771)	1 versus 2, 3	1, 2 versus 3 (n=257, Observations=771)
Social media user	0.265 (0.292)	—[a]	0.249 (0.287)	—	3.760 (0.764)[b]	1.274 (0.410)[b]	14.99 (0.436)[b]	1.024 (0.576)[c]	0.377 (0.587)	—
Income	0.00796 (0.00404)[c]	—	0.00815 (0.00410)[d]	—	0.0544 (0.0194)[b]	0.00855 (0.00438)[c]	0.00820 (0.00392)[d]	—	0.00835 (0.00398)[d]	—
University	−0.352 (0.348)	—	−0.358 (0.349)	—	−0.361 (0.304)	—	−0.250 (0.291)	—	−0.335 (0.358)	—
Age (years)	−0.0115 (0.00984)	—	−0.0118 (0.00975)	—	0.00885 (0.0101)	—	−0.00426 (0.00898)	—	−0.00967 (0.00851)	—
Male	0.0593 (0.309)	—	0.0652 (0.310)	—	0.284 (0.299)	—	−0.0757 (0.297)	—	0.00144 (0.277)	—
Māori and Pacific	−0.935 (0.515)[c]	—	−0.944 (0.515)[c]	—	−0.947 (0.446)[d]	—	−0.847 (0.442)[c]	—	−0.999 (0.512)[c]	—
Trust government	0.841 (0.298)[b]	—	0.853 (0.298)[b]	—	2.465 (0.562)[b]	0.895 (0.292)[b]	0.849 (0.263)[b]	—	0.820 (0.292)[b]	—
Trust family or friends	−0.126 (0.352)	—	−0.136 (0.349)	—	−1.399 (0.552)[d]	−0.110 (0.313)	−0.139 (0.302)	—	−0.178 (0.323)	—
Wave 2 FE	0.146 (0.259)	—	0.146 (0.259)	—	0.183 (0.275)	—	0.169 (0.263)	—	0.153 (0.260)	—
Wave 3 FE	−0.618 (0.360)[c]	0.340 (0.256)	−0.618 (0.360)[c]	0.340 (0.256)	−0.618 (0.406)	0.352 (0.243)	−0.584 (0.390)	0.338 (0.259)	−0.604 (0.363)[c]	0.348 (0.254)
Intercept	3.128 (0.883)[b]	0.726 (0.775)	3.153 (0.882)[b]	0.752 (0.770)	−0.255 (0.846)	−0.644 (0.655)	2.704 (0.665)[b]	0.390 (0.532)	3.229 (0.850)[b]	0.826 (0.670)

[a]Not applicable.
[b]*P*<.01.
[c]*P*<.1
[d]*P*<.05

Looking across the models, we find that Instagram and Twitter use have a statistically significant positive relationship with vaccine uptake. On the other hand, at least one of the 4 types of social media, Facebook and TikTok use, did not have significant effects. We present the MEMs for social media of these models in Table 5. Here, we see highly statistically significant associations for Instagram and Twitter. Specifically, Instagram is associated with an 11.8 percentage point decrease in the likelihood of being resistant and a 20.3 percentage point increase in the likelihood of being provaccine. The direction is the same for Twitter, albeit with lower marginal effects. It is worth noting again that Twitter and TikTok use is relatively uncommon in our sample (Table 1).

In terms of demographic controls in Table 5, we see that income has a positive and significant association with vaccine uptake. University education, age, and gender do not have a statistically significant association with uptake level. Māori or Pacific ethnicity has a negative association with vaccine uptake, at the 5% or 10% level. Towards the bottom of Table 4, we see that the survey wave fixed effects have little predictive power on vaccine uptake. Wave 3 is associated with a lower chance of being vaccine resistant, with significance at the 10% level for 3 of the 5 models.

XSL•FO
RenderX

Table . Average marginal effects (marginal effects at the mean) of social media use on being in different vaccine uptake categories (from the models in Table 5). Cluster robust standard errors in parentheses, calculated using the delta-method. Results have been reweighted on age, gender, and ethnicity.

Platform	Resistant, mean (SE)	Hesitant, mean (SE)	Pro, mean (SE)
Social media	−0.0120 (0.0140)	−0.0362 (0.0408)	0.0481 (0.0545)
Facebook	−0.0112 (0.0134)	−0.0338 (0.0396)	0.0450 (0.0527)
Instagram	−0.118 (0.0205)[a]	−0.0858 (0.0563)	0.203 (0.0591)[a]
Twitter	−0.0583 (0.0144)[a]	−0.0967 (0.0702)	0.155 (0.0699)[b]
TikTok	−0.0145 (0.0203)	−0.0485 (0.0704)	0.0630 (0.0904)

[a]*P*<.01.

[b]*P*<.05

Trust Interactions in PPO Models

Next, we look at the trust variables. For the base models in Table 5, we see that the coefficient on trust in government for information about vaccines is positive and strongly statistically significant, as expected. Trust in family or friends for vaccine information has a negative but statistically insignificant relationship with uptake, except for the Instagram model, where it is negative and significant for the 1 versus 2, 3 component of the model. We exclude distrust in social media from the model as it leads to within-sample predictions of negative probabilities for a handful of observations (a drawback of PPO models [40]).

Next, we interact with trust in 2 key information sources (government, and friends and family) on vaccines, with social media type in our base PPO models. This analysis is to test whether trust in information sources has a mediating role in how social media use is associated with vaccine uptake.

We might expect that social media use will be associated with a higher likelihood of vaccine uptake for those who trust the government for vaccine information (compared with not trusting), given the use of social media by the government to inform the public about COVID restrictions. Finally, we expect social media users will be less likely to take the vaccine if they trust their friends or family for vaccine information (compared with not trusting), given that friends or family may share vaccine information on social media and are a less reliable source of information. To test these hypotheses, we follow the same approach as before and re-estimate the models from Table 5, with just the interaction of interest. We present the MEMs for these interactions in Table 6.

Let us first consider general social media use and trust in government, in the top left of the table. We see that social media use for those who trust the government is associated with a 4.7% lower probability of being resistant, which is statistically significant at the 5% level. However, none of the other marginal effects is significant, nor is the difference between the marginal effects for those who trust the government for vaccine information and those who do not. There is a similar result for Facebook users.

XSL•FO
RenderX

Table . Marginal effects at the mean (MEMs) for the interactions between social media use and trust in sources of information regarding vaccines. Cluster robust SEs in parentheses, calculated using the delta method. Results have been reweighted on age, gender, and ethnicity.

MEMs	Trust in government			Trust in friends and family		
	Resistant	Hesitant	Pro	Resistant	Hesitant	Pro
Social media, mean (SE)						
Trust=0	−0.00527 (0.0420)	−0.00823 (0.0651)	0.0135 (0.107)	−0.0222 (0.0167)	−0.0668 (0.0481)	0.0891 (0.0638)
Trust=1	−0.0470 (0.0184)[c]	−0.00534 (0.0523)	0.0524 (0.0633)	0.0227 (0.0189)	0.0768 (0.0660)	−0.0995 (0.0840)
Difference significance	—[a]	—	—	Yes[b]	Yes[b]	Yes[b]
Facebook, mean (SE)						
Trust=0	0.0445 (0.0462)	−0.0824 (0.0796)	0.0379 (0.111)	−0.0209 (0.0157)	−0.0635 (0.0459)	0.0844 (0.0606)
Trust=1	−0.0255 (0.0150)[b]	−0.0274 (0.0560)	0.0530 (0.0617)	0.0229 (0.0191)	0.0757 (0.0647)	−0.0986 (0.0829)
Difference significance	—	—	—	Yes[b]	Yes[b]	Yes[b]
Instagram, mean (SE)						
Trust=0	−0.260 (0.0413)[d]	−0.00929 (0.0983)	0.269 (0.104)[c]	−0.0905 (0.0203)[d]	−0.107 (0.0632)[b]	0.197 (0.0696)[d]
Trust=1	−0.0455 (0.0189)[c]	−0.133 (0.0599)[c]	0.179 (0.0625)[d]	−0.0653 (0.0325)[c]	−0.175 (0.0708)[c]	0.240 (0.0875)[d]
Difference significance	Yes[d]	Yes[b]	—	—	—	—
Twitter, mean (SE)						
Trust=0	−0.117 (0.0315)[d]	−0.306 (0.0849)[d]	0.424 (0.0887)[d]	−0.0525 (0.0148)[d]	−0.00257 (0.0994)	0.0551 (0.0990)
Trust=1	−0.0393 (0.0104)[d]	−0.0104 (0.0751)	0.0496 (0.0763)	−0.0746 (0.0189)[d]	−0.196 (0.0770)[c]	0.270 (0.0815)[d]
Difference significance	Yes[d]	Yes[d]	Yes[d]	—	Yes[b]	Yes[b]
TikTok, mean (SE)						
Trust=0	−0.106 (0.0280)[d]	−0.302 (0.0588)[d]	0.408 (0.0693)[d]	−0.0522 (0.0138)[d]	−0.0571 (0.106)	0.109 (0.107)
Trust=1	−0.0364 (0.00972)[d]	0.127 (0.116)	−0.0906 (0.116)	−0.0384 (0.0267)	0.0245 (0.157)	0.0139 (0.156)
Difference significance	Yes[d]	Yes[d]	Yes[d]	—	—	—

[a]Not applicable.
[b]*P*<.1
[c]*P*<.05
[d]*P*<.01

Instagram, Twitter, and TikTok users have higher probabilities of being pro and lower probabilities of being hesitant or resistant, for both individuals who trust the government and those who do not. The MEMs of social media use on those who trust the government and those who do not are counterintuitive. For example, Twitter use is associated with a 11.7 percentage point decrease in the probability of being resistant, for those who do not trust the government. On the other hand, Twitter use is associated with only a 3.9 percentage point decrease in being resistant for those who do trust the government. This is a statistically significant difference. Instagram and TikTok use shows largely similar patterns to the Twitter example mentioned here. Thus, Twitter, Instagram, and TikTok use is associated more positively with vaccine uptake for those who do not trust the government, compared with those who do trust the government.

The right side of Table 6 shows the MEMs for interactions between social media use and trust in friends and family. We might expect those who trust friends and family for vaccine information to be more susceptible to misinformation on social media. We see this for social media users in general in the top rows; social media users who trust friends and family are more likely to be resistant and hesitant, and less likely to be pro, at the 10% level. This finding is true for Facebook users as well. There are no significant differences for Instagram or TikTok users. However, we see the opposite for Twitter users. Twitter users who trust their friends and family are more likely to be pro and less likely to be hesitant compared with those who do not trust their friends and family, again at the 10% level of significance.

XSL•FO
RenderX

Discussion

In this paper, we have demonstrated a novel approach to analyzing stated preference data. Specifically, we categorize individuals as resistant, hesitant, or pro, based on the number of times they chose a vaccine option in a set of 6 stated choices. Our stated choice method (discrete choice experiment) is a well-established means of understanding individual vaccine preferences; our means of analyzing such data adds another method to understanding vaccine orientation. We investigate the association between individual choices and social media use over time using both a Logit and PPO model, and we use the PPO model to understand the moderating role of trust in sources of information. Our data were collected during vaccine development and initial rollout, between August 2020 and April 2021. Our context in New Zealand had a high trust in the government at the time.

We find a positive or neutral association between social media use and vaccine uptake. In our transition modeling, we see some evidence that Facebook, Twitter, and TikTok users who are pro are less likely to become hesitant or resistant. Facebook and Instagram users who are hesitant are more likely to become pro. Over the 3 survey waves, we find a positive association between Instagram and Twitter and being more likely to uptake the vaccine.

Our results on trust in information sources provide some more interesting details. As expected, we find a strong positive association between trust in government for vaccine information and being more provaccination, shown in our main PPO results in Table 4. However, we provide some additional evidence that some social media platforms may have a more positive effect on vaccine uptake preferences for those who do not trust the government. Our Table 5 results show Instagram, Twitter, and TikTok use is more positively associated with being more provaccination for those who do not trust the government, compared with those who trust the government. This finding points to a potentially positive role for some social media platforms. Potentially good information from trusted, non-government sources could reach those who lack trust in the government but are still open to being convinced on the merits of vaccination.

These findings provide a marginally positive picture for social media and vaccine uptake. However, given the concerns from the wider literature outlined in our introduction, we urge caution in how our results are interpreted. While social media could help increase vaccine uptake, our data do not strongly refute the potential for social media to also decrease vaccine uptake. Social media still has just as significant potential to spread bad information about vaccination to susceptible individuals. Indeed, the spread of misinformation via social media and its impact on vaccine hesitancy has been a key concern in the literature [22].

There are several reasons why we may not have found a negative association between social media use and vaccine preferences. First, a limitation of our study is the dataset itself. The social media question is only in the third survey wave, and asks about social media use over the previous 6 months. It would have been preferable to include social media use in each survey wave and ask about social media use over a shorter timeframe. Future research would also ideally cover other social media platforms that may spread misinformation, such as Telegram and YouTube. However, our mixed results by platform align with previous work that shows different social media platforms have varying effects on the spread of vaccine misinformation [29]. Despite these limitations, we argue that our novel analysis of a stated choice survey and the panel nature of the dataset still demonstrate a useful contribution.

Second, we undertake a web-based survey, with potential for sample selection bias. It seems likely that individuals with conspiracy beliefs will be less likely to undertake a survey from University researchers. Hence, we may have missed those at the most extreme end of resistance. As such, a survey may not be the best approach to understand the extremely resistant. Data pulled directly from the use of social media platforms may help in this regard [19]. Of course, analyzing such data can give a better idea of social media users, but it does not include a control of non-social media users, and may not be as clear in providing information about individual vaccine preferences.

Third, in the context of high government trust at the time in New Zealand, we find low levels of vaccine resistance of 4 to 7 percent, which is relatively low by international comparison [7]. This low proportion makes it impossible to identify the reasons behind the increasing levels of resistance we observed over time, without a much larger sample size. Thus, we acknowledge that other contexts are likely to have much higher potential for misinformation on social media, having a greater impact on levels of hesitancy or resistance. For instance, previous work in the United States using panel data found that social media was associated with higher levels of vaccine hesitancy during the COVID-19 pandemic [28]. Furthermore, we do observe polarization, with increasing levels of individuals in both the resistant and pro categories over time.

We note that a limitation of our data is that it cannot be interpreted as causal, hence, we frame our findings as associations between social media use and vaccine preferences. We collected social media use for individuals in survey wave 3, over the last 6 months. This period coincides approximately with the length of time between the first and third survey waves. However, this is second-best panel data; first-best would include recent social media use in each wave, meaning we could track longitudinally how changes in social media use change vaccine preferences. This would allow us to be closer to finding causal relationships, though we are not sure how much social media use changes over time, and therefore, if such an approach would allow us enough variation. There may still be potential endogeneity issues with such an approach. We also point out that the transition modelling gets us closer to a causal interpretation, as opposed to cross-sectional associations, as we see whether changes in vaccine preferences are associated with social media use between observations.

Other future research could use a similar stated preference panel data method to track the evolution of vaccine preferences. However, it would be helpful if it included more detailed questions on social media use and other sources of vaccine

XSL•FO
RenderX

information. Our study highlights the need to collect such vaccine preference data at the crucial point where new vaccines are developed and rolled out, so that we can study how preferences form. Thus, opportunities should be explored to collect such data on new COVID vaccines and other new vaccines in development.

While our study was overall positive in terms of social media use and vaccine uptake, we contribute to a literature with a mixed picture, as outlined in the introduction. The concerning potential for polarization to continue to increase is still ever present. The downsides are large, including both worsening public health as well as increasingly extreme political conflict based on disagreement over basic facts. Whether naive or bad actors are spreading misinformation, it is a pressing issue to both better understand the problem and potential solutions.

Acknowledgments

We would like to thank John Gibson for his very helpful and insightful guidance at an early stage of this research project. We would like to thank the participants of the 2022 New Zealand Association of Economists (NZAE) Annual Conference for their thoughtful and useful suggestions. We would also like to thank the wider authorship of Hess et al [7] for their input into the development of the survey. We acknowledge the Waikato Management School for providing funding for this project. SH acknowledges the financial support by the European Research Council through the advanced grant 101020940-SYNERGY.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional modeling results.
[DOCX File, 33 KB - infodemiology_v5i1e66081_app1.docx]

References

1. Apergis E, Apergis N. The impact of COVID-19 on economic growth: evidence from a Bayesian Panel Vector Autoregressive (BPVAR) model. Appl Econ 2021 Dec 14;53(58):6739-6751. [doi: 10.1080/00036846.2021.1946479]
2. World Health Organisation. In the wake of the pandemic: preparing for long covid. : World Health Organisation; 2021 39 URL: https://iris.who.int/bitstream/handle/10665/339629/Policy-brief-39-1997-8073-eng.pdf [accessed 2025-05-20]
3. Coronavirus (COVID-19) dashboard. World Health Organisation. 2022. URL: https://covid19.who.int [accessed 2022-08-05]
4. Lavine JS, Bjornstad ON, Antia R. Immunological characteristics govern the transition of COVID-19 to endemicity. Science 2021 Feb 12;371(6530):741-745. [doi: 10.1126/science.abe6522] [Medline: 33436525]
5. Andre FE, Booy R, Bock HL, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. Bull World Health Organ 2008 Feb;86(2):140-146. [doi: 10.2471/blt.07.040089] [Medline: 18297169]
6. de Figueiredo A, Simas C, Karafillakis E, Paterson P, Larson HJ. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. The Lancet 2020 Sep;396(10255):898-908. [doi: 10.1016/S0140-6736(20)31558-0]
7. Hess S, Lancsar E, Mariel P, et al. The path towards herd immunity: predicting COVID-19 vaccination uptake through results from a stated choice study across six continents. Soc Sci Med 2022 Apr;298:114800. [doi: 10.1016/j.socscimed.2022.114800] [Medline: 35287066]
8. Park S, Massey PM, Stimpson JP. Primary source of information about COVID-19 as a determinant of perception of COVID-19 severity and vaccine uptake: source of information and COVID-19. J Gen Intern Med 2021 Oct;36(10):3088-3095. [doi: 10.1007/s11606-021-07080-1] [Medline: 34378115]
9. Ruiz JB, Bell RA. Predictors of intention to vaccinate against COVID-19: results of a nationwide survey. Vaccine (Auckl) 2021 Feb 12;39(7):1080-1086. [doi: 10.1016/j.vaccine.2021.01.010] [Medline: 33461833]
10. Zeballos Rivas DR, Lopez Jaldin ML, Nina Canaviri B, Portugal Escalante LF, Alanes Fernández AMC, Aguilar Ticona JP. Social media exposure, risk perception, preventive behaviors and attitudes during the COVID-19 epidemic in La Paz, Bolivia: a cross sectional study. PLoS ONE 2021;16(1):e0245859. [doi: 10.1371/journal.pone.0245859] [Medline: 33481945]
11. Hogan MJ, Pardi N. mRNA vaccines in the COVID-19 pandemic and beyond. Annu Rev Med 2022 Jan 27;73(1):17-39. [doi: 10.1146/annurev-med-042420-112725] [Medline: 34669432]
12. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. BMJ Glob Health 2020 Oct;5(10):e004206. [doi: 10.1136/bmjgh-2020-004206] [Medline: 33097547]

XSL•FO
RenderX

13. Thaker J, Cook C. Planning for a COVID-19 vaccination campaign in New Zealand: trust, affective and cognitive attitudes, and COVID-19 vaccine intention. Communication Research and Practice 2022 Jan 2;8(1):54-69. [doi: 10.1080/22041451.2021.2022325]
14. Schmidt AL, Zollo F, Del Vicario M, et al. Anatomy of news consumption on Facebook. Proc Natl Acad Sci U S A 2017 Mar 21;114(12):3035-3039. [doi: 10.1073/pnas.1617052114] [Medline: 28265082]
15. Cinelli M, Quattrociocchi W, Galeazzi A, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 6;10(1):16598. [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
16. Del Vicario M, Bessi A, Zollo F, et al. The spreading of misinformation online. Proc Natl Acad Sci U S A 2016 Jan 19;113(3):554-559. [doi: 10.1073/pnas.1517441113] [Medline: 26729863]
17. Cinelli M, Brugnoli E, Schmidt AL, Zollo F, Quattrociocchi W, Scala A. Selective exposure shapes the Facebook news diet. In: Iñiguez G, editor. PLOS ONE 2020;15(3):e0229129. [doi: 10.1371/journal.pone.0229129] [Medline: 32168347]
18. Bail CA, Argyle LP, Brown TW, et al. Exposure to opposing views on social media can increase political polarization. Proc Natl Acad Sci U S A 2018 Sep 11;115(37):9216-9221. [doi: 10.1073/pnas.1804840115] [Medline: 30154168]
19. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. Vaccine (Auckl) 2018 Jun 14;36(25):3606-3612. [doi: 10.1016/j.vaccine.2018.05.040] [Medline: 29773322]
20. Puri N, Coomes EA, Haghbayan H, Gunaratne K. Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. Hum Vaccin Immunother 2020 Nov 1;16(11):2586-2593. [doi: 10.1080/21645515.2020.1780846] [Medline: 32693678]
21. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021 Mar;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
22. Ruggeri K, Vanderslott S, Yamada Y, et al. Behavioural interventions to reduce vaccine hesitancy driven by misinformation on social media. BMJ 2024 Jan 16;384:e076542. [doi: 10.1136/bmj-2023-076542] [Medline: 38228339]
23. Brailovskaia J, Schneider S, Margraf J. To vaccinate or not to vaccinate!? predictors of willingness to receive Covid-19 vaccination in Europe, the U.S., and China. In: C-Y C, editor. PLOS ONE 2021;16(12):e0260230. [doi: 10.1371/journal.pone.0260230] [Medline: 34851986]
24. Mascherini M, Nivakoski S. Social media use and vaccine hesitancy in the European Union. Vaccine (Auckl) 2022 Mar 25;40(14):2215-2225. [doi: 10.1016/j.vaccine.2022.02.059] [Medline: 35249775]
25. Al-Uqdah L, Franklin FA, Chiu CC, Boyd BN. Associations between social media engagement and vaccine hesitancy. J Community Health 2022 Aug;47(4):577-587. [doi: 10.1007/s10900-022-01081-9] [Medline: 35332393]
26. Bendau A, Plag J, Petzold MB, Ströhle A. COVID-19 vaccine hesitancy and related fears and anxiety. Int Immunopharmacol 2021 Aug;97:107724. [doi: 10.1016/j.intimp.2021.107724] [Medline: 33951558]
27. Wang R, Qin C, Du M, Liu Q, Tao L, Liu J. The association between social media use and hesitancy toward COVID-19 vaccine booster shots in China: a web-based cross-sectional survey. Hum Vaccin Immunother 2022 Nov 30;18(5):2065167. [doi: 10.1080/21645515.2022.2065167] [Medline: 35671384]
28. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. Soc Sci Med 2020 Oct;263:113356. [doi: 10.1016/j.socscimed.2020.113356] [Medline: 32967786]
29. Theocharis Y, Cardenal A, Jin S, et al. Does the platform matter? social media and COVID-19 conspiracy theory beliefs in 17 countries. New Media & Society 2023 Dec;25(12):3412-3437. [doi: 10.1177/14614448211045666]
30. Beattie A, Priestley R. Fighting COVID-19 with the team of 5 million: Aotearoa New Zealand government communication during the 2020 lockdown. Soc Sci Humanit Open 2021;4(1):100209. [doi: 10.1016/j.ssaho.2021.100209] [Medline: 34585139]
31. Lieberoth A, Lin SY, Stöckli S, et al. Stress and worry in the 2020 coronavirus pandemic: relationships to trust and compliance with preventive measures across 48 countries in the COVIDiSTRESS global survey. R Soc Open Sci 2021 Feb 10;8(2):200589. [doi: 10.1098/rsos.200589] [Medline: 33972837]
32. Sibley CG, Greaves LM, Satherley N, et al. Effects of the COVID-19 pandemic and nationwide lockdown on trust, attitudes toward government, and well-being. Am Psychol 2020;75(5):618-630. [doi: 10.1037/amp0000662] [Medline: 32496074]
33. Peterson B, Harrell FE. Partial proportional odds models for ordinal response variables. Appl Stat 1990;39(2):205. [doi: 10.2307/2347760]
34. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26(8):661-677. [doi: 10.2165/00019053-200826080-00004] [Medline: 18620460]
35. Rose JM, Bliemer MCJ. Stated choice experimental design theory: the who, the what and the why. In: Handbook of Choice Modelling: Edward Elgar Publishing; 2014:152-177. [doi: 10.4337/9781781003152.00013]
36. NGene 1.2. ChoiceMetrics. 2018. URL: http://www.choice-metrics.com/ [accessed 2025-05-12]
37. Pacific peoples ethnic group. Stats NZ. 2018. URL: https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries/pacific-peoples [accessed 2022-05-19]
38. Māori population estimates: at 30 june 2021. Stats NZ. 2021 Jun 30. URL: https://www.stats.govt.nz/information-releases/maori-population-estimates-at-30-june-2021#:~:text=New%20Zealand's%20estimated%20M%C4%81ori%20ethnic,17.1%20percent%20of%20national%20population) [accessed 2022-05-19]

XSL•FO
RenderX

39. Long JS, Freese J. Regression Models for Categorical Dependent Variables Using Stata: Stata Press Publication; 2001.
40. Williams R. Understanding and interpreting generalized ordered logit models. J Math Sociol 2016 Jan 2;40(1):7-20. [doi: 10.1080/0022250X.2015.1112384]
41. Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 1990 Dec;46(4):1171-1178. [doi: 10.2307/2532457] [Medline: 2085632]

Abbreviations

MEM: marginal effects at the mean
PPO: partial proportional odds

Edited by R Cuomo; submitted 03.09.24; peer-reviewed by N Yiannakoulias, SH Raza; revised version received 22.04.25; accepted 22.04.25; published 28.05.25.

Please cite as:
Maris R, Dorner Z, Hess S, Tucker S
Social Media and the Evolution of Vaccine Preferences During the COVID-19 Pandemic: Discrete Choice Experiment
JMIR Infodemiology 2025;5:e66081
URL: https://infodemiology.jmir.org/2025/1/e66081
doi:10.2196/66081

© Robbie Maris, Zack Dorner, Stephane Hess, Steven Tucker. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 28.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Social Media Posts on Lifestyle Behaviors: Sentiment and Content Analysis

Yan Yee Yip[1,2], MPharm, MClinPharm; Mohd Ridzwan Yaakub[3], BSc, MSc, PhD; Mohd Makmor-Bakry[1], BPharm, MClinPharm, PhD; Muhammad Iqbal Abu Latiffi[3], BSc, MSc; Wei Wen Chong[1], BPharm, PhD

[1]Centre for Quality Management of Medicines, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, Kuala Lumpur, Malaysia

[2]Centre for Clinical Epidemiology, Institute for Clinical Research, National Institutes of Health, Ministry of Health Malaysia, Setia Alam, Shah Alam, Malaysia

[3]Center for Artificial Intelligence Technology, Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor Darul Ehsan, Malaysia

Corresponding Author:
Wei Wen Chong, BPharm, PhD
Centre for Quality Management of Medicines, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, Kuala Lumpur, Malaysia

Abstract

Background: There has been an increase in the prevalence of noncommunicable diseases in Malaysia. This can be prevented and managed through the adoption of healthy lifestyle behaviors, including not smoking, avoiding alcohol consumption, maintaining a balanced diet, and being physically active. The growing importance of using social media to deliver information on healthy behaviors has led health care professionals (HCPs) to lead these efforts. To ensure effective delivery of information on healthy lifestyle behaviors, HCPs should begin by understanding users' current opinions about these behaviors and whether the users are receptive to recommended health practices. Nevertheless, there has been limited research conducted in Malaysia that aims to identify the sentiments and content of posts, as well as how well users' perceptions align with recommended health practices.

Objective: This study aims to examine social media posts related to various lifestyle behaviors, by using a combination of sentiment analysis to analyze users' sentiments and manual content analysis to explore the content of the posts and how well users' perceptions align with recommended health practices.

Methods: Using keywords based on lifestyle behaviors, posts originating from X (formerly known as Twitter) and published in Malaysia between November and December 2022 were scraped for sentiment analysis. Posts with positive and negative sentiments were randomly selected for content analysis. A codebook was developed to code the selected posts according to content and alignment of users' perceptions with recommended health practices.

Results: A total of 3320 posts were selected for sentiment analysis. Significant associations were observed between sentiment class and lifestyle behaviors (χ^2_6=67.64; P<.001), with positive sentiments higher than negative sentiments for all lifestyle behaviors. Findings from content analysis of 1328 posts revealed that most of the posts were about users' narratives (492/1328), general statements (203/1328), and planned actions toward the conduct of their behavior (196/1328). More than half of tobacco-, diet-, and activity-related posts were aligned with recommended health practices, whereas most of the alcohol-related posts were not aligned with recommended health practices (63/112).

Conclusions: As most of the alcohol-related posts did not align with recommended health practices, the findings reflect a need for HCPs to increase their delivery of health information on alcohol consumption. It is also important to ensure the ongoing health promotion of the other 3 lifestyle behaviors on social media, while continuing to monitor the discussions made by social media users.

(JMIR Infodemiology 2025;5:e65835) doi:10.2196/65835

KEYWORDS

consumer health information; content analysis; chronic illness; health promotion; healthy lifestyle; lifestyle; lifestyle risk reduction; internet; primary prevention; sentiment analysis; social media

XSL•FO
RenderX

Introduction

Noncommunicable diseases (NCDs) have become a significant global health challenge, accounting for 74% of all deaths worldwide [1]. NCDs have also become a growing public health concern in middle-income countries within the Southeast Asia region. Malaysia is one of the countries in the region that has been significantly impacted by NCDs, with over 67% premature NCD-related mortality and over 70% disease burden [2]. Approximately, 2.5% of Malaysian adults, which accounts for over half a million people were affected by all 4 major NCDs, which are diabetes, hypertension, hypercholesterolemia, and obesity [3].

The World Health Organization (WHO) has identified 4 key modifiable risk factors associated with an elevated risk of NCDs, which are tobacco smoking, harmful use of alcohol, unhealthy diet, and physical inactivity [1,4]. The adoption of healthy lifestyle behaviors, including not smoking, avoiding alcohol consumption, maintaining a balanced diet, and being physically active can reduce modifiable risk factors, effectively preventing and managing NCDs. However, national surveys have indicated that the actual adoption of healthy behaviors among Malaysians remains low. For example, over 84% of Malaysian adults were inactive in sports, with half of the population leading a sedentary lifestyle, spending more than 2 hours sitting while awake [3].

In this regard, it is important to deliver information on healthy lifestyle behaviors, with health care professionals (HCPs) being ideally positioned to lead these efforts. Various technologies, such as mHealth applications, wearable devices [5], and social media platforms, can support the delivery of health information. Social media platforms have been widely used for health information delivery as these platforms are accessible to larger populations at a lower cost [6,7]. To effectively promote healthy lifestyle behaviors on social media, HCPs could begin by understanding users' current opinions on lifestyle behaviors and whether they are receptive to recommended health practices [8]. This could be achieved by examining social media posts discussing on lifestyle behaviors. X (formerly known as Twitter) is one of the microblog-based social media platform that allows users to freely express their opinions through posts, previously referred to as tweets. As of January 2024, approximately 5.71 million social media users in Malaysia were on X, accounting for 16.5% of the country's population [9], which highlights the growing popularity of X among Malaysians.

When users express their opinions in writing on social media, a range of emotions may be conveyed. Sentiment analysis is the process of classifying this textual data based on the emotions conveyed within the text as positive, negative, or neutral sentiments [10]. It can be conducted through manual annotation of posts or computational approaches [11]. Computational approaches in sentiment analysis are preferred as they are more cost-efficient, and can leverage large amounts of publicly accessible and concise real-time data across different regions and demographics [12]. Methodologies of computational approaches include lexicon-based sentiment analysis, which uses pre-existing dictionaries containing words with pre-assigned sentiment scores of positive, negative, or neutral. Lexicon-based sentiment analysis is effective when limited labelled training data is available with a strong association of sentiments with specific words. The usage of this approach has been documented in numerous studies that analyze sentiments regarding lifestyle behaviors such as the examination of policies on electronic cigarettes [13], vegan-related posts [14], and organic food posts [15].

Lexicon-based sentiment analysis, however, may have limited coverage in terms of vocabulary and often misses sarcasm or irony [16]. Positive sentiments may not necessarily translate to good health practices and vice versa. For example, the sentence "I love tobacco" showed positive sentiments, but the actual context is related to the user's preference towards unhealthy lifestyle behaviors. In order to further understand the topics communicated on social media and how users' perceptions are aligned with recommended health practices, lexicon-based sentiment analysis can be supported with manual content analysis. A codebook can be used to manually assign labels to each post, which will provide a more in-depth analysis of the posts [17].

The examination of social media posts across multiple lifestyle behaviors of tobacco smoking, alcohol consumption, diet, and physical activity could facilitate effective comparisons of findings across these different behaviors. The use of such findings would enable HCPs to use social media to deliver information on healthy behaviors by targeting areas where the lifestyle behaviors are not aligned with recommended health practices. Analyses that are focused within a geographic location would provide opportunities for HCPs to prioritize region-targeted health information delivery on social media. Such health information could also potentially be replicated in other countries with similar digital and health ecosystems.

Nevertheless, there have been limited studies that used the combined approaches of lexicon-based sentiment and content analysis to examine social media posts on lifestyle behaviors. The majority of the available studies were focused on other health-related issues such as the examination of users' perceptions on diabetes [18] and marijuana usage [19]. A study by Kasson et al [20] have used both sentiment analysis and manual content analysis to examine users' vaping behaviors. However, the study was confined to vaping behaviors during the e-cigarette or vaping use-associated lung injury outbreak and did not address other lifestyle behaviors. In addition, there is a lack of studies examining the opinions of social media users in Malaysia on lifestyle behaviors, despite the increasing burden of NCDs and the rising prevalence of unhealthy lifestyle behaviors in the country.

Therefore, this study used a combination of lexicon-based sentiment analysis and manual content analysis to understand the discussions on lifestyle behaviors among social media users in Malaysia. This study had three objectives: (1) to determine the sentiments of social media users in Malaysia regarding lifestyle behaviors, (2) to identify the content of posts and ascertain if users' perceptions were aligned with recommended health practices, and (3) to explore the associations between the alignment of users' perceptions with recommended health practices and sentiment class.

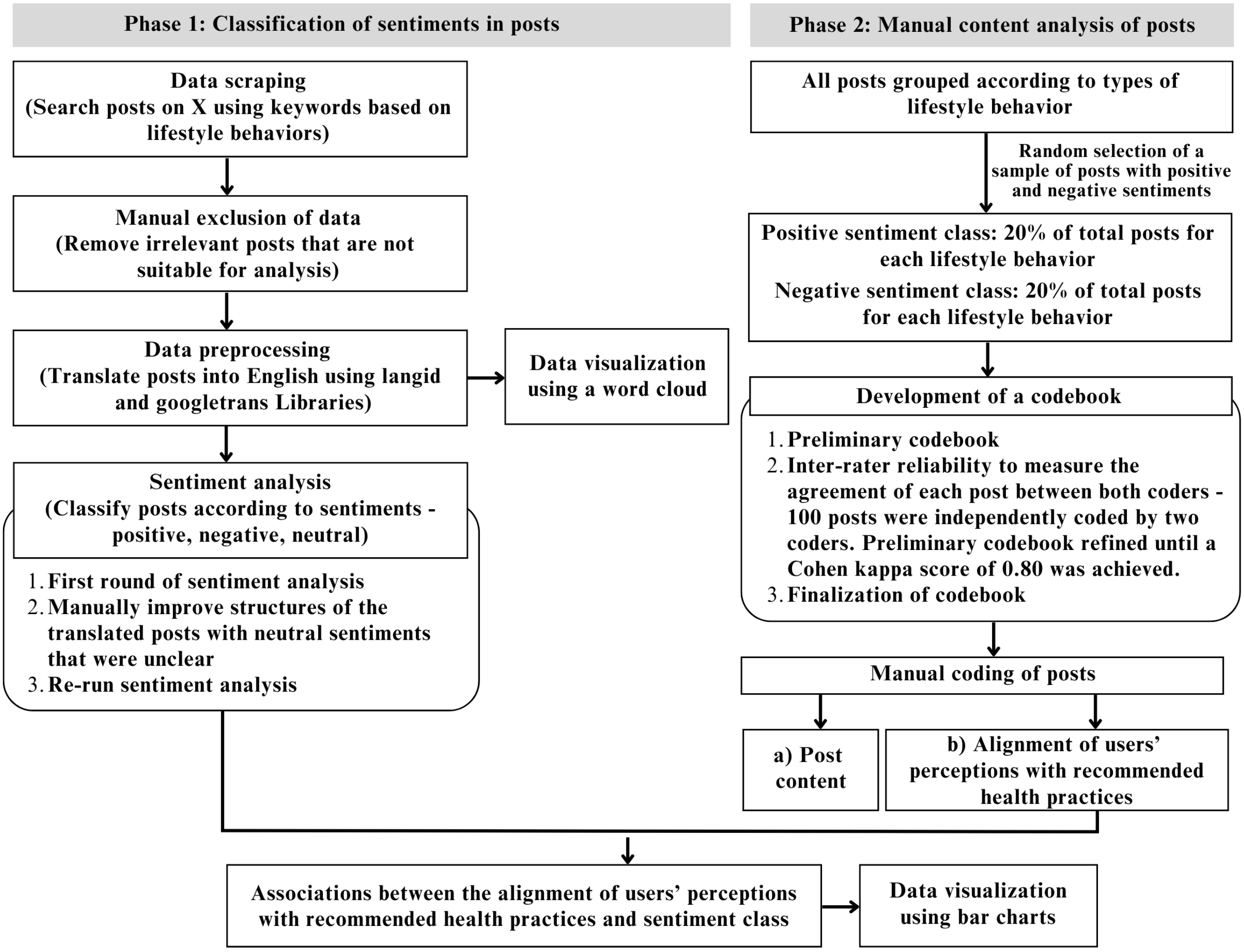
XSL•FO
RenderX

Methods

Overview

Figure 1 shows the overall study methods. In the classification of sentiments in posts, data was scraped from X. Following the manual exclusion of irrelevant posts, the data was cleaned, preprocessed and analyzed for sentiments. Data visualization was subsequently conducted. In the manual content analysis of posts, a random selection of posts with positive and negative sentiments for each lifestyle behavior was manually coded to identify the content of posts and the alignment of users' perceptions with recommended health practices. Associations between the alignment of users' perceptions and sentiment class were subsequently explored.

Figure 1. Overall study methods.



Classification of Sentiments in Posts

Data Scraping

The automated process of extracting large amounts of data from X is known as data scraping. All posts with keywords related to the 4 lifestyle behaviors aimed at reducing the 4 key modifiable risk factors for NCDs were scraped. These keywords are related to tobacco and its derivative products, alcohol, dietary, and physical activity. The 4 sets of keywords are provided in Multimedia Appendix 1. The keywords were derived from published systematic reviews on the management of lifestyle behaviors using social media [21,22]. Additional keywords commonly used locally were added upon discussion with all researchers.

Posts spanning 2 consecutive months were selected. This time frame was deemed to be appropriate, as similar studies analyzing health-related sentiments on the X platform have also utilized data across a 2-month period [23,24]. The selection of the 2 consecutive months was conducted by initially scraping all posts from January to December 2022 according to each month. The 2 months with the highest number of posts were from November to December 2022. Post based in Malaysia were determined using longitude and latitude metadata (4.2105°N, 101.9758°E). In terms of language, posts in Malay and English were scraped. Malay is the national language of Malaysia, whereas English is widely spoken and understood by Malaysians. They also use a mix of both languages, resulting in multilingual data.

Data scraping was conducted separately for each set of keywords using the SNScrape library on Python. In addition to posts, other X metadata such as timestamp, X username, number of reposts, language, and location were also collected.

Manual Exclusion of Data

All posts were manually screened by 2 researchers to exclude those not suitable for analysis, with discrepancies resolved among the research team. The exclusion steps were as follows:

XSL•FO
RenderX

First, the exclusion of posts not from Malaysia—During data scraping, longitude and latitude data retained posts located within the coordinates but posted outside of Malaysia, such as parts of Singapore and Thailand. The “location” metadata was therefore used to manually exclude these posts.

Second, the exclusion of irrelevant posts—Irrelevant posts include posts with different definitions (eg, “exercising” your rights), posts not related to health care (eg, religious restrictions on alcohol), indecipherable posts and posts made by bots. Bots were verified by manually checking the user’s profile for any unusual activity patterns that exhibited automated behavior (eg, high frequency of posts without breaks).

Data Preprocessing

Before conducting sentiment analysis, data preprocessing was carried out to ensure that the text data was cleaned, transformed, and prepared for analysis.

The model selected for sentiment analysis was the Valence Aware Dictionary and Sentiment Reasoner (VADER) on Python. It is a lexicon-based sentiment analysis tool [25] that is suitable for analyzing social media posts, generating results with high classification success [26]. As VADER is trained for sentiment analysis in English, all posts in Malay and mixed languages were preprocessed and translated into English using the langid and googletrans libraries on Python.

Other data preprocessing steps (tokenization, lower casing of texts, removal of stop words, html links, numbers, punctuations, emojis, and acronyms) were not executed. This is attributed to the unique advantages of VADER, in which assessment scores would account for capitalism, punctuations, emojis, English acronyms (eg, “LOL”), and colloquialisms (eg, “meh”) [27-29].

Following data preprocessing, a word cloud used to provide a visual representation of the words in the overall X dataset.

Sentiment Analysis

Computational, lexicon-based sentiment analysis was conducted to determine users’ sentiments. This approach was selected over manual annotations and other types of computational methods as it is cost-effective and does not require training for large datasets. As it relies on a predefined lexicon, it does not require significant computational resources. For general sentiment analysis, lexicon-based approaches often performs well enough to capture the overall sentiment trends in social media data [25].

The posts that have undergone data preprocessing were then analyzed for polarity using VADER. VADER classifies posts into positive, neutral, and negative sentiments. Positive sentiments have a compound score of ≥0.05, neutral sentiments have a score between >−0.05 and<0.05 and negative sentiments have a score of ≤−0.05 [25].

Computer-assisted translation tools may limit the extent of translation in posts that contain local dialects and slang. The translated posts may become indecipherable, causing them to be classified as having “neutral” sentiments. For example, “x” in Malay, which means “no” in English may not have undergone translation, resulting in sentiments not being classified accurately. Following the first round of sentiment analysis, the structures of the translated posts that were unclear and yielded neutral sentiments were manually improved. The sentiment analysis was then re-run to enhance the robustness of sentiment classification and to reduce the inaccurate labeling of posts.

Manual Content Analysis of Posts With Positive and Negative Sentiments

Sentiment analysis classifies texts according to the emotions conveyed [10]. However, there was no further elaboration on the content and whether users’ perceptions were aligned with recommended health practices.

Therefore, a sample of posts with positive and negative sentiments were randomly selected for manual content analysis. Stratified sampling was conducted by dividing the posts according to the type of lifestyle behavior and sentiment class (positive or negative). For each type of lifestyle behavior and sentiment class, 20% of the total posts for the particular lifestyle behavior were subjected to random selection. This would allow the posts for each lifestyle behavior to have an equal number of positive and negative sentiments. The random sample of posts was generated by using the random number equation in Microsoft Excel relative to the ID number attached to each post. This approach was adopted from a previous content analysis study on X, which also manually coded a random sample of 20% of total posts [30].

A preliminary codebook with 2 categories was developed through discussions among the research team to classify the content of posts and the alignment of users’ perceptions with recommended health practices. This codebook is partially adapted from the codes used by Miller et al [31]. The recommended health practices are based on WHO’s health recommendations [32]. In brief, WHO advocates healthy practices, including abstaining from smoking and alcohol consumption, maintaining a balanced diet, and engaging in regular physical activities. The codes were mutually exclusive. Using the preliminary codebook, 100 posts (25 posts for each type of lifestyle behavior) were independently coded by 2 coders. Interrater reliability was conducted to measure the agreement of each post between both coders. The preliminary codebook was refined until a Cohen kappa score of 0.80 was achieved. The remaining posts were then coded independently using the finalized codebook by both coders.

Table 1 provides a brief description of the codes for the finalized codebook, with a more comprehensive codebook provided in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Brief description of codes.

Category	Definition
Post content (Topical content dealing with the lifestyle behavior mentioned in the post)	
Self-narrative of current lifestyle behaviors	Narration of self's current lifestyle behaviors.
Narrative of others' current lifestyle behaviors	Talked about other people's current lifestyle behaviors.
Planned action related to lifestyle behaviors	A planned action that will be conducted by the person who wrote the post.
Recommendations related to lifestyle behaviors	A recommendation by the person who wrote the post, providing instruction, advice, or suggestion to others.
Direct question	Direct question used in a post.
General statement	General statement that is not under any of the other categories above.
Alignment of users' perceptions with recommended health practices (Whether users' perceptions in posts are aligned with WHO's health recommendations)	
Aligned with recommended health practices	Users agreed with the conduct of recommended health practices, that included not smoking, avoiding alcohol consumption, maintaining a balanced diet and being physically active.
Not aligned with recommended health practices	Users were not agreeable with the conduct of recommended health practices (eg, consumed oily food, refused to exercise).
Users' perceptions cannot be defined	The perceptions of the user could not be defined or linked with health practices.

Data Analysis

Data analysis was performed using descriptive statistics with all variables expressed in frequencies and percentages. The Pearson chi-square test was used to compare the associations between the categorical variables, with P values <.05 considered to be statistically significant. IBM SPSS Statistics version 26.0 was used for data analysis.

Findings were also visualized using a word cloud and bar charts. In addition, examples of posts were provided to describe the study findings.

Ethical Considerations

This study was approved by the Medical Research and Ethics Committee, Ministry of Health Malaysia (NMRR ID-23 - 00293-CIM [IIR]) on March 23, 2023, and the Research Ethics Committee, Universiti Kebangsaan Malaysia (UKM PPI/111/8/JEP-2023 - 174) on April 13, 2023.

As this study relied solely on publicly available social media data on X and did not involve direct interaction with individuals, informed consent was not applicable. No compensation was offered or provided, as the study did not involve direct participation of human participants.

No identifiable private user information was collected or analyzed. All data used in the analysis were publicly available and did not contain personally identifiable information.

Results

Overview of X Dataset

Figure 2 shows the flowchart of the selection of the X dataset.

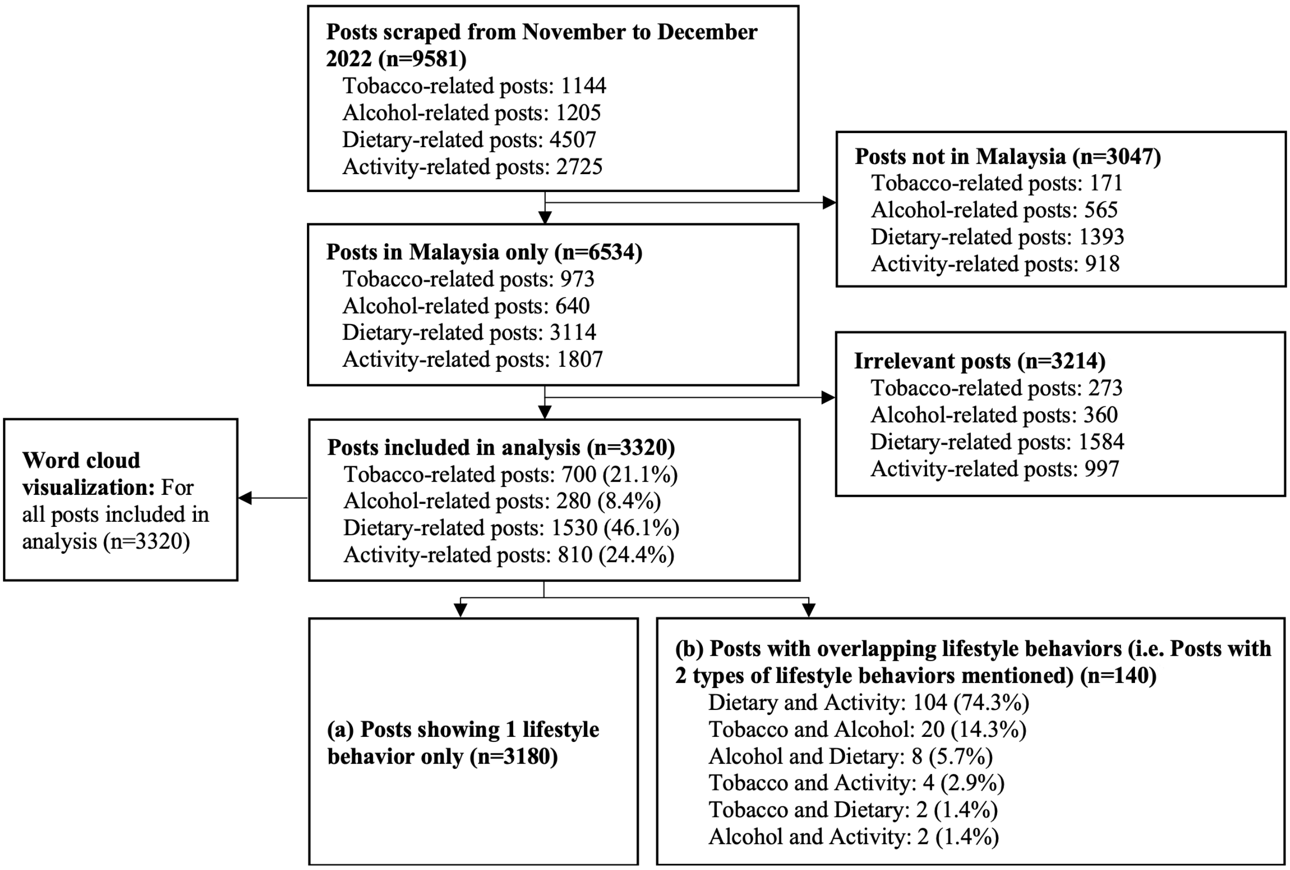
A total of 9581 posts were scraped from November to December 2022. Following the exclusion of 3047 posts that were not in Malaysia and 3214 irrelevant posts, 3320 posts across 4 types of lifestyle behaviors were retained for sentiment analysis. Almost half of the posts were dietary-related (1530/3320, 46.1%), followed by activity-related (810/3320, 24.4%) and tobacco-related (700/3320, 21.1%) posts. Alcohol-related posts were present in only one-tenth of the posts (280/3320, 8.4%).

As data scraping was conducted separately for each lifestyle behavior, a post may appear more than two times across different behaviors. Out of the 3320 posts, 3180 (95.8%) posts showed 1 lifestyle behavior only. There were 140 posts with two types of lifestyle of behaviors mentioned, with three-quarters (104/140, 74.3%) of posts mentioning both dietary- and activity-related behaviors.

A word cloud was used to visualize the overall X dataset (n=3320; see Multimedia Appendix 3). Overall, the 5 terms most commonly mentioned by users were "diet," "rice," "eat," "sugar," and "smoke."

XSL•FO
RenderX

Figure 2. Flowchart of selection of X dataset. Examples of irrelevant posts include (1) tobacco-related posts: gaming-type of posts (eg, smoke mentioned in a game), music band names (eg, Cigarettes After Sex), tweet mentioning terms as a location (eg, Hookah Island, Vape Shop); (2) alcohol-related posts: banning alcohol due to religion with no links to health (eg, at Qatar for the World Cup), sarcasm-based (eg, "you must be drunk," which translates to "you must be kidding me"); (3) dietary-related posts: nephew in Malay (eg, anak buah), nonhealth posts (eg, fruit on the trees or plants, sugar daddy); (4) activity-related posts: posts mentioning terms as a location (eg, Sports Direct), nonhealth posts (eg, who will be Sports Minister?, "exercising" your rights). With regard to posts overlapping lifestyle behaviors, as data scraping was conducted separately for each lifestyle behavior, a post may appear more than 2 times across different behaviors (eg, the post talks about both smoking and dietary habits).



Findings From Classification of Sentiments in Posts

Table 2 presents the frequency distribution of sentiment analysis, with examples for each lifestyle behavior provided in Multimedia Appendix 4.

The overall percentage of positive sentiments almost doubled that of negative sentiments (1874/3320, 56.5%) vs (1027/3320, 30.9%). Results have shown a significant association between sentiment class and lifestyle behaviors (χ^2_6=67.64; P<.001), with positive sentiments being higher than negative sentiments for all lifestyle behaviors. The trends for dietary- and activity-related posts were similar, with both showing approximate percentages of 60% for positive sentiments and 27% for negative sentiments. This was followed by alcohol-related posts with positive sentiments of 54.7% (153/280). Less than half of tobacco-related posts (314/700, 44.9%) had positive sentiments, with the percentage differences between sentiment classes slightly below 2%.

XSL•FO
RenderX

Table . Frequency distribution of sentiment analysis before and after structure improvement of posts using Valence Aware Dictionary and Sentiment Reasoner (VADER; n=3320).

Lifestyle behaviors	Sentiment count before structure improvement of posts, n (%)			Sentiment count after structure improvement of posts, n (%)[a]			Total, n
	Positive	Neutral	Negative	Positive	Neutral	Negative	
Tobacco-related posts	254 (36.3)	197 (28.1)	249 (35.6)	314 (44.9)	84 (12)	302 (43.1)	700
Alcohol-related posts	120 (42.9)	95 (33.9)	65 (23.2)	153 (54.7)	41 (14.6)	86 (30.7)	280
Dietary-related posts	724 (47.3)	455 (29.7)	351 (23)	916 (59.9)	197 (12.9)	417 (27.2)	1530
Activity-related posts	378 (46.7)	246 (30.4)	186 (22.9)	491 (60.6)	97 (12)	222 (27.4)	810
Total	—[b]	—	—	1874 (56.5)	419 (12.6)	1027 (30.9)	3320

[a]Sentiment count after structure improvement of posts were used for analysis. A Pearson chi-square test was conducted to test the associations between sentiment class and lifestyle behaviors (χ^2_6=67.64, *P*<.001).

[b]Not applicable.

Findings From Manual Content Analysis of Posts

A total of 1328 posts with an equal number of positive and negative sentiments for each lifestyle behavior were selected for manual content analysis. They comprised of 280 tobacco-related posts (140 posts for each sentiment class), 112 alcohol-related posts (56 posts for each sentiment class), 612 dietary-related posts (306 posts for each sentiment class), and 324 activity-related posts (162 posts for each sentiment class).

Prior to the manual content analysis of all 1328 posts, 100 posts were first subjected to interrater reliability testing. The Cohen kappa scores for both categories of post content and the alignment of users' perceptions with recommended health practices were 0.807 and 0.801, respectively.

The frequency of posts is tabulated in Table 3. Overall, the content with the 3 highest number of posts were self-narratives of current lifestyle behaviors (492/1328, 37%), general statements (203/1328, 15.3%) and planned actions (196/1328, 14.8%). Self-narratives were the most popular content for all types of lifestyle behaviors except for tobacco-related posts, in which the majority were narratives of others' current behaviors (96/280, 34.3%). Question-based posts were the least popular content for tobacco-, alcohol-, and activity-related posts, with less than 10% present in all types. Users' perceptions in more than half of the posts were aligned with recommended health practices (769/1328, 57.9%). Similar proportions were observed for all types of lifestyle behaviors except for alcohol-related posts, in which posts not aligned with recommended health practices were double those aligned with recommended health practices (63/112, 56.3% vs 33/112, 29.4%).

Figure 3 shows the frequency of posts that demonstrated the alignment of users' perceptions with recommended health practices according to the sentiment classification. A total of 3 main findings were observed. First, in dietary- and activity-related posts, significant associations between sentiment class and alignment of users' perceptions with recommended health practices were observed (χ^2_2=30.98, *P*<.001 and χ^2_2=24.16, *P*<.001; respectively). In both positive and negative sentiment classes, the percentages of posts aligned with recommended health practices were significantly higher than those not aligned with recommended health practices and those with undefined user perceptions, with percentages ranging from 49.3% to 80.2%. Posts with positive sentiments that aligned with recommended healthy practices showcased users' optimism to stay healthy (eg, "I'm ready to cut sugar. Let's go" [D-919-positive]). Meanwhile, negative sentiments that aligned with recommended healthy practices highlighted users' worries to stay healthy (eg, "Feel the weight.. rise suddenly. Sad. Have to fix it" [P-522-negative]).

Second, in tobacco-related posts, there was no significant association between sentiment class and alignment of users' perceptions with recommended health practices (χ^2_2=5.76; *P*=.06). Among posts with positive sentiments, the percentage of posts that aligned with recommended health practices was similar to those not aligned with recommended health practices, with a percentage difference of 7.1%. When users posted about tobacco with positive emotions, the likelihood of their perceptions aligning with the recommended health practices of not smoking (eg, "Please pray that I can stop smoking…" [T-563-positive]) or aligning with hazardous smoking practices (eg, "My kind of chill with cigar" [T-412-positive]) were similar. Despite the lack of significant association, the percentage of posts with negative sentiments that aligned with recommended health practices was noticeably higher than those that did not align with recommended health practices (86/140, 61.4% vs 41/140, 29.3%).

Third, a lack of significant association between sentiment class and alignment of users' perceptions with recommended health practices was also observed in alcohol-related posts (χ^2_2=4.62; *P*=.10). Although the findings were not statistically significant, the percentage of posts not aligned with recommended health practices was higher than those aligned with recommended health practices for both positive (37/56, 66.1% vs 12/56, 21.4%) and negative (26/56, 46.4% vs 21/56, 37.5%) sentiment classes.

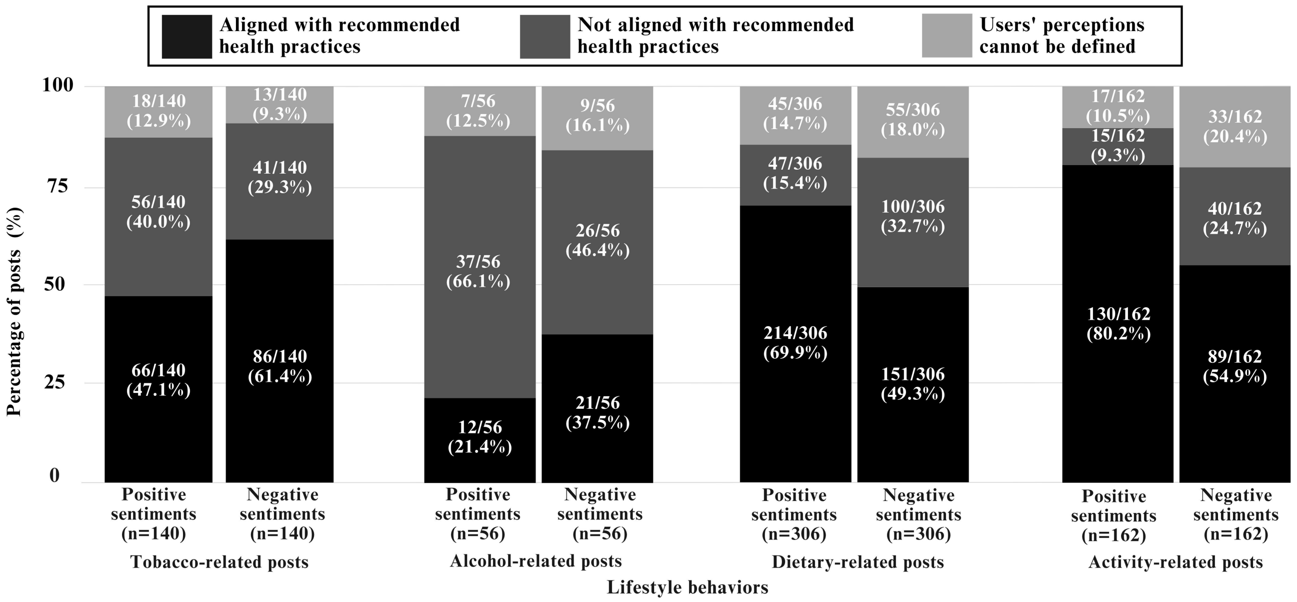
XSL•FO
RenderX

Positive sentiments leading to alcohol consumption included celebratory posts (eg, "I'm gonna have so much wine this weekend" [A-55-positive]), whereas negative sentiments involved users coping with worries ("I am going to drown my sorrows in alcohol and pick things back up tomorrow" [A-198-negative]). Examples of posts selected for manual content analysis are provided in Multimedia Appendix 5.

Table . Frequency of posts according to post content and alignment of users' perceptions with recommended health practices (n=1328).

Category	Lifestyle behaviors				Total posts (N=1328)
	Tobacco-related (n=280)	Alcohol-related (n=112)	Dietary-related (n=612)	Activity-related (n=324)	
Post content, n (%)					
Self-narrative of current lifestyle behaviors	70 (25)	36 (32.1)	243 (39.7)	143 (44.1)	492 (37)
Narrative of others' current lifestyle	96 (34.3)	17 (15.2)	40 (6.5)	22 (6.8)	175 (13.2)
Planned action related to lifestyle behaviors	26 (9.3)	15 (13.4)	97 (15.8)	58 (17.9)	196 (14.8)
Recommendations related to lifestyle behaviors	35 (12.5)	13 (11.6)	83 (13.6)	46 (14.2)	177 (13.3)
Direct question	17 (6.1)	6 (5.4)	45 (7.4)	17 (5.3)	85 (6.4)
General statement	36 (12.8)	25 (22.3)	104 (17)	38 (11.7)	203 (15.3)
Alignment of users' perceptions with recommended health practices, n (%)					
Aligned with recommended health practices	152 (54.3)	33 (29.4)	365 (59.7)	219 (67.6)	769 (57.9)
Not aligned with recommended health practices	97 (34.6)	63 (56.3)	147 (24)	55 (17)	362 (27.3)
Users' perceptions cannot be defined	31 (11.1)	16 (14.3)	100 (16.3)	50 (15.4)	197 (14.8)

Figure 3. Stacked bar charts of alignment of users' perceptions with recommended health practices stratified according to sentiment classification. A Pearson chi-square test was conducted for each lifestyle behavior to test the associations between sentiment class and alignment of users' perceptions with recommended health practices (tobacco: χ^2_2=5.76, P=.06; alcohol: χ^2_2=4.62, P=.10; dietary: χ^2_2=30.98, P<.001; activity: χ^2_2=24.16, P<.001).



XSL•FO
RenderX

Discussion

Overview

To the best of our knowledge, this is the first study in the region that examined discussions on X across multiple lifestyle behaviors. This study is also the first of its kind that used dual approaches of lexicon-based sentiment analysis and manual content analysis of posts to examine users' sentiments, post content and the alignment of users' perceptions with recommended health practices. Positive sentiments were significantly higher than negative sentiments for all 4 lifestyle behaviors. In dietary- and activity-related posts, users exhibited twice as many positive sentiments as negative ones. The majority of the sampled posts were self-narratives of current lifestyle behaviors. More than half of the sampled tobacco-, dietary-, and activity-related posts were aligned with WHO's recommended health practices, with contrasting results in alcohol-related posts.

Principal Findings

Data scraping has shown that dietary-related topics were the lifestyle behaviors most frequently discussed. The usage of a more extensive set of search terms in scraping dietary-related posts covered a variety of nutrition-based topics involving individuals across all age groups. This resulted in dietary-related discourses among young children and adolescents (eg, formula milk, vegetables, and fruits consumption). Users frequently mentioned "rice," which is attributed to Malaysians' staple diet, with the average Malaysian adult consuming 82.3 kilograms of rice annually [33]. In contrast, alcohol-related topics were the least discussed lifestyle behaviors among users. Alcohol-related discussions in Malaysia were largely anchored on themes related to cultural and religious beliefs. Alcohol consumption among Malaysians is generally lower as behaviors are influenced by compartmentalization among the three main races in Malaysia [34]. Malays who are Muslims are not allowed to consume alcohol as it is forbidden in Islam [35], whereas no restrictions were imposed on the Chinese and Indian communities [34]. The prohibition of alcohol consumption in certain communities was hypothesized as one of the reasons for the lower frequency of alcohol-related posts, compared to other lifestyle behaviors.

Findings indicate that positive sentiments significantly outweighed negative sentiments for all lifestyle behaviors. In dietary- and activity-related posts, positive sentiments were found to be twice as many as negative sentiments, which is consistent with the sentiment analysis findings of Shaw et al [36] who analyzed over 1.5 million posts on dietary and exercise topics. In our study, more than half of the sampled posts were either self-narratives or planned actions for self-implementation. It could be postulated that the posts with positive sentiments were driven by self-determination theory (SDT), a comprehensive theory of human motivation and personality that focuses on individuals' intrinsic tendencies for growth. SDT assumes the importance of autonomous motivation, which is a type of self-emanating motivation that is consistent with users' innate values to engage in behaviors or pursue a goal [37,38]. The field of autonomous motivation has been extensively studied in the context of dietary and exercise lifestyle behaviors [39-41]. Individuals who are autonomously motivated have a sense of self-control over their actions (eg, choose to exercise regularly), leading to an increase in positive sentiments, personal fulfilment and enjoyment in the actions pursued [40].

In tobacco-related posts, positive sentiments were found to be higher than negative sentiments, albeit with a small percentage difference. This suggests that in tobacco-related discourses, users tend to either feel positive emotions (eg, satisfaction, happiness, and trust) or negative emotions (eg, dissatisfaction, unhappiness, and worry). Mixed sentiments were found to be prevalent in discussions related to vaping among both the scientific community [42], and the general public [43]. In Malaysia, the Health Ministry has proposed the Generational End Game plan, which would ban tobacco sales for those born after 2005. The bill was first tabled at the country's parliamentary discussions in July 2022 and has yet to be finalized at the time when the social media posts were scraped from X [44]. Such uncertainties towards health policy changes have generated both positive and negative reactions, with the issue being debated constantly throughout the year. Users either praised the government's efforts to mitigate smoking behaviors or expressed concerns about such "untested" plans [45]. The negative reactions may also stemmed from users' awareness of the adverse effects of smoking, with over 90% of male lung cancer patients in Malaysia having a significant history of smoking [46]. In addition, 3500 out of 10,000 annual deaths were linked to smoking [47].

As discussed, the predominance of self-narratives in posts related to diet, physical activity and alcohol consumption is likely due to users' autonomous motivation and self-awareness to perform a behavior. Conversely, most posts related to tobacco were found to be linked to narratives of other users. As tobacco smoking has been associated with social stigmatization due to its negative health impact on others, users may have been more reluctant to post from a first-person perspective. Instead, users opt to openly discuss the smoking habits of others. In addition, sharing experiences in a third-person perspective may be preferred by users to maintain anonymity. During the 20th century, smokers were often viewed as "mysterious" or "cool," but this social status has slowly diminished over the past two decades [48,49]. In Malaysia, this was propelled by smoking reduction strategies, such as the ban on smoking in public eateries implemented in 2019, that socially impacted users' impression toward smoking [50].

In recent years, there has been a growing trend of health influencers using online platforms to actively share their dietary and fitness regimens. Previous studies have shown that social media users who were exposed to this information delivered by health influencers as well as content from other social media users, were more likely to be receptive to adopting healthy practices, such as maintaining a balanced diet and being physically active [51-53]. The results were consistent with the findings of this study, which showed a significantly higher percentage of dietary- and activity-related posts by social media users that were aligned with recommended health practices. Nevertheless, HCPs must remain active in advocating positive lifestyle behaviors on social media. Although almost one-fifth of posts for these two lifestyle behaviors were on planned

XSL•FO
RenderX

actions, this may not always translate into actions by the population. This is a caveat of much research that relies on social media or self-reported data on social media. It is often unclear whether individuals actually follow through on what they post about, highlighting the intention-behavior gap [54]. Findings from the Malaysian National Health and Morbidity Survey (NHMS) survey conducted in 2023 have shown that the actual adoption of healthy practices was still lacking among the Malaysian public. Almost 95.1% Malaysian adults did not meet the recommended daily intake of fruits and vegetables, consuming only two servings of fruit or vegetables daily instead of the recommended five servings daily [3] The prevalence of physical inactivity among Malaysian adults was at 29.9% [3], which was also considerably higher than other Asian countries, including China and India [55,56].

There was a lack of significant association between sentiment class and alignment with recommended health practices in both tobacco- and alcohol-related posts. Despite the smaller number of sampled alcohol-related posts, it is interesting to note that users' perceptions with recommended health practices had contrasting outcomes compared to the other three lifestyle behaviors. In more than half of the alcohol-related posts with positive and negative sentiments, users' perceptions were not aligned with recommended health practices. Most users perceived alcohol consumption as a casual and an affordable social activity and did not acknowledge the potential health risks involved. A survey conducted in Thailand, a country of similar income setting, had previously mentioned the popularity of alcohol being a social activity among urban communities [57]. In Malaysia, alcoholic beverages were available for purchase at neighborhood convenience stores, which allowed for easy purchases of takeaway alcohol [35]. This further downplayed users' awareness of the negative consequences of alcohol consumption [58].

Assessment of posts made by social media users on X allows HCPs to identify priority areas for social media-based health information delivery on this platform. As most alcohol-related posts do not align with health recommendations, it is postulated that greater emphasis should be placed on strategies to limit alcohol consumption among users in Malaysia. The WHO has proposed collaborative efforts with HCPs and journalists to improve targeted public health messaging to the public. A guide was recently developed for journalists to facilitate media reporting to communities on the harms of alcohol consumption [58]. While the other 3 lifestyle behaviors were mostly aligned with recommended health practices, it remains essential for HCPs to continuously deliver information advocating healthy behaviors. Online approaches allow HCPs to deliver information beyond geographic barriers, reaching a wider audience in diverse community settings. Therefore, health information can be adopted by users in countries with similar cultural beliefs, including countries within the Southeast Asia region.

Strengths and Limitations

The strengths of this study included its comprehensive coverage of 4 lifestyle behaviors aimed at reducing the 4 key modifiable risk factors under WHO's health priority [4]. This allowed for the simultaneous analysis of posts across different lifestyle behaviors. Unlike most studies that focus on global contexts, this research uniquely focused on the Malaysian context. It provides insights into the cultural and social dynamics that influence discussions around lifestyle behaviors in this specific region. Notably, the inclusion of alcohol-related posts in analysis shed light on culturally and socially nuanced discussions within the region. This is particularly valuable in regions like Malaysia, where religious and cultural factors strongly influence alcohol consumption. In addition, the retrospective examination of social media posts utilized approaches of lexicon-based sentiment analysis and manual content analysis. The dual approach provided real-time and spontaneous insights into users' opinions on lifestyle behaviors while addressing limitations of single-method studies. Findings from this study could assist HCPs in prioritizing the delivery of region-specific health information through social media.

Nevertheless, a few limitations should be considered. First, potential bias may exist during the selection of social media dataset. In self-selection bias, users who choose to share their opinions on social media may not represent the broader population. The study may be subjected to data selection bias as it included only social media posts in Malay and English, excluding other spoken languages in Malaysia, such as Chinese and Tamil. Nevertheless, sentiment analysis studies on X in Malaysia have largely concentrated on data scraped in Malay and English [59,60]. Demographic bias may be present due to the overrepresentation or underrepresentation of certain groups on X. For instance, the majority of users in Malaysia who post on social media are aged between 25 and 34 years old [9]. The limitations in the availability of metadata on X also prevented the collection of demographic data such as age, gender and race, as most users did not disclose this information in their profiles. Population bias may occur in geotagged posts utilizing longitude and latitude metadata. Previously literature has indicated that only 1% of users would geotag their location in posts [61]. Nevertheless, this is the most effective method to scrape posts that are published within a specific location.

Second, there are limitations in the study design and methods used for sentiment analysis and manual content analysis. The study is cross-sectional in nature and provides a snapshot of discussions at a specific time. Therefore, temporal bias may exist, making it challenging to track changes in sentiments or behaviors over time. In addition, posts that were collected for sentiment analysis and manual content analysis over two consecutive months may not accurately reflect year-round sentiments and discussions, as findings may vary due to the presence of health-related events occurring at certain times of the year. The events may include a change in legislations, prominent public health campaigns or disease outbreaks. The quality of the dataset was ensured by verifying that there were no notable health-related occurrences between November and December 2022. In addition, previous studies analyzing users' sentiments and content have similarly explored health data over two consecutive months [23,24]. The manual content analysis of social media posts can be time-consuming due to the involvement of large datasets, therefore, only 20% of the total posts were randomly selected using stratified sampling. This percentage was previously utilized in a content analysis study

XSL•FO
RenderX

by Mathieson et al [30]. While analyzing the full dataset would provide more comprehensive findings, the randomized sample offers a reliable snapshot for identifying the thematic content without the substantial time and resource demands of manual analysis for the entire dataset. Furthermore, prior to sentiment analysis, the computer-assisted translation of posts from Malay to English may have led to inaccuracies due to the usage of local dialects, sarcasm or slangs. To enhance sentiment labeling, the structures of translated posts with neutral sentiments that were unclear were manually refined, and sentiment analysis was repeated.

Third, the study results should be interpreted with caution, regarding posts on alcohol consumption due to the smaller sample size of 112 posts. A power analysis indicated that this sample size is adequate for detecting effects, with a power of 0.82. Furthermore, we acknowledge the presence of potential interactions in posts with overlapping lifestyle behaviors (eg, a post that talks about diet and physical activity). In sentiment analysis, the conduct of Pearson chi-square tests also did not account for potential confounding factors or interactions in posts with overlapping lifestyle behaviors. To account for this limitation, we compared the proportions between sentiment count for posts showing 1 lifestyle behavior only (n=3180), and sentiment count for posts across 4 types of lifestyle behaviors (n=3320). The proportions of sentiment counts for both were similar to each other. In addition, while many of the Pearson chi-square associations were significant, these may not imply causality and thus may not inform categorically that the observed sentiments result in practicing different lifestyle behaviors or the direction of the relationship.

Implications and Further Research

The findings from this study could help HCPs to prioritize the delivery of health information on lifestyle behaviors using social media tailored to the targeted region, which is Malaysia. Given the low number of alcohol-related posts by social media users in Malaysia, HCPs could focus on initiating positive discussions around this topic to raise awareness about the harmful effects of alcohol consumption. In addition, most of the alcohol-related posts made by social media users were not aligned with recommended health practices. There is an increased need for HCPs to emphasize on limiting and stopping alcohol consumption, while also acknowledging that the users' attitudes towards alcohol consumption may still vary among different religions in Malaysia. Health advocacy for positive lifestyle behaviors on social media should continue for the other three lifestyle behaviors.

Further research could be proposed to explore the opinions of social media users toward lifestyle behaviors in Malaysia. First, despite the statistical significance observed in the associations between sentiment classification and lifestyle behaviors, the percentage difference between both sentiment classes in tobacco-related posts was small. Therefore, it would be interesting to investigate whether tobacco sentiments would vary over time. We may want to further track sentiments by time series analysis to explore changes in users' emotions towards tobacco across a time period. The tracking of real-time sentiments across a time period was previously conducted in a review examining public health data on X that included posts on alcohol consumption [62]. In addition, since posts are scraped based on location metadata, future studies could leverage on this data to explore the relationship between the prevalence of specific lifestyle behaviors in certain locations (eg, urban areas in Malaysia) and the intensity of lifestyle behavior-related discussions on social media. A similar study has previously been conducted in the United States; therefore, conducting such studies in the Malaysian context would be beneficial [63].

Second, the majority of posts involved content related to self-narratives of lifestyle behaviors. These self-narratives outlined X's roles as a microblog for users to freely express the behaviors they practice from a first-person perspective. As self-narratives encompass a broad and generalized category, it may be beneficial to conduct a more detailed examination of posts that only described users' self-narratives. This in-depth analysis would provide insights into the specific themes commonly discussed by users from a first-person perspective. In addition, the examination of posts could be extended to other lifestyle behaviors such as sleep patterns, which is particularly relevant as active social media users are mainly adolescents and young adults who are commonly affected by sleep-related issues [64].

Third, this study was conducted on the microblogging platform X. It is also important to examine social media posts made by users on other platforms, such as Facebook. Future research is proposed to analyze the sentiments and content of posts on these platforms. Audience demographics can vary across these platforms. For instance, younger millennials may be more active on X, whereas Facebook often attracts a slightly older audience [65,66]. Comparing our study findings with those obtained from Facebook could help HCPs to deliver health messages that suit the audiences of different social media platforms.

Fourth, our study emphasizes accessibility and simplicity in data visualization and reporting to effectively communicate findings to a diverse audience, including non-technical stakeholders such as HCPs, public health practitioners and policymakers. To achieve this, we employed techniques like word clouds, which provide a visually appealing representation of frequently mentioned terms in the dataset, and lexicon-based sentiment analysis, which is straightforward to implement as it does not require additional labeled data or extensive training. We recognize the potential value of more advanced methods and suggest exploring these techniques in future studies related to the conduct of in-depth text analysis. These may include approaches like topic modeling or keyword co-occurrence analysis to summarize text data through word groups, as well as training machine learning models such as support vector machines or Naïve Bayes to classify sentiments. Furthermore, hybrid methods of sentiment analysis could be explored by integrating machine learning models with lexicon-based approaches. These combined models can then be assessed for accuracy and robustness through comparative analysis. Similar studies have been conducted previously in both health and non-health posts [67,68].

XSL•FO
RenderX

Conclusion

In conclusion, the incorporation of lexicon-based sentiment analysis holds significance as it enabled the use of large amounts of data to capture users' emotions whilst posting on lifestyle behaviors. Positive sentiments were significantly expressed in posts for all lifestyle behaviors. Nevertheless, there was a small percentage difference observed in tobacco-related posts, indicating a more varied sentiment among users. Most of the posts showed users' own narratives and planned actions towards the conduct of a behavior. As the majority of alcohol-related discussions were not aligned with recommended health practices, this reflects the need for individual HCPs and health organizations to increase their delivery of health information pertaining to alcohol consumption on social media platforms. It is also equally important for HCPs to continue providing health information on other lifestyle behaviors to social media users, while monitoring ongoing discussions by users on social media.

Acknowledgments

The authors would like to thank the director-general of Health Malaysia for his permission to publish this article. The work was supported by the Ministry of Higher Education of Malaysia's Fundamental Research Grant Scheme under grant FRGS/1/2020/SS0/UKM/02/11. The funders played no role in study design, collection, analysis, interpretation of data, or writing of the report.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

YYY, MRY, and WWC contributed to the conception or design of the study. YYY and MIAL contributed to data collection of the study. YYY, MIAL, and WWC contributed to data analysis of the study. All authors (YYY, MRY, MM-B, MIAL, and WWC) contributed to data interpretation and provided scientific inputs and technical improvement. YYY drafted the manuscript while MRY and WWC guided the revisions. All authors (YYY, MRY, MM-B, MIAL, and WWC) read and approved the final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of keywords for data scraping of posts.
[DOCX File, 18 KB - infodemiology_v5i1e65835_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File, 27 KB - infodemiology_v5i1e65835_app2.docx]

Multimedia Appendix 3
Word cloud representation of overall X dataset (n=3320).
[PNG File, 2512 KB - infodemiology_v5i1e65835_app3.png]

Multimedia Appendix 4
Examples of posts with positive and negative sentiments according to each lifestyle behavior.
[DOCX File, 20 KB - infodemiology_v5i1e65835_app4.docx]

Multimedia Appendix 5
Example of posts selected for manual content analysis.
[DOCX File, 23 KB - infodemiology_v5i1e65835_app5.docx]

References

1. Noncommunicable diseases. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2022-12-06]
2. Malaysian Burden of Disease and Injury Study 2009-2014: Institute for Public Health; 2017. URL: https://iku.moh.gov.my/images/IKU/Document/REPORT/BOD/BOD2009-2014.pdf [accessed 2022-12-06]

XSL•FO
RenderX

3. National Health and Morbidity Survey (NHMS) 2023: Non-Communicable Diseases and Healthcare Demand: Institute for Public Health; 2024. URL: https://iku.nih.gov.my/images/nhms2023/key-findings-nhms-2023.pdf [accessed 2024-05-24]
4. Reducing modifiable risk factors for noncommunicable diseases. World Health Organization. 2022. URL: https://www.who.int/westernpacific/activities/reducing-modifiable-risk-factors-for-noncommunicable-diseases [accessed 2022-12-06]
5. Miyamoto SW, Henderson S, Young HM, Pande A, Han JJ. Tracking health data is not enough: a qualitative exploration of the role of healthcare partnerships and mhealth technology to promote physical activity and to sustain behavior change. JMIR Mhealth Uhealth 2016 Jan 20;4(1):e5. [doi: 10.2196/mhealth.4814] [Medline: 26792225]
6. Islam SMS, Tabassum R, Liu Y, et al. The role of social media in preventing and managing non-communicable diseases in low-and-middle income countries: Hope or hype? Health Policy Technol 2019 Mar;8(1):96-101. [doi: 10.1016/j.hlpt.2019.01.001]
7. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520. [Medline: 25083128]
8. Balicer RD, Luengo-Oroz M, Cohen-Stavi C, et al. Using big data for non-communicable disease surveillance. Lancet Diabetes Endocrinol 2018 Aug;6(8):595-598. [doi: 10.1016/S2213-8587(17)30372-8] [Medline: 29146206]
9. Digital 2024: Malaysia. Data Reportal. 2024. URL: https://datareportal.com/reports/digital-2024-malaysia [accessed 2024-04-24]
10. Tonkin EL. In: Tonkin EL, Tourte GJL, editors. A Day at Work (with Text): A Brief Introduction, 1st edition: Chandos Publishing; 2016:23-60.
11. van Atteveldt W, van der Velden M, Boukes M. The validity of sentiment analysis: comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms. Commun Methods Meas 2021 Apr 3;15(2):121-140. [doi: 10.1080/19312458.2020.1869198]
12. Abualigah L, Alfar HE, Shehab M, Hussein AMA. Sentiment analysis in healthcare: a brief review. In: Abd Elaziz M, Al-qaness MAA, Ewees AA, Dahou A, editors. Recent Advances in NLP: The Case of Arabic Language: Springer International Publishing; 2020:29-141.
13. Lu X, Sun L, Xie Z, Li D. Perception of the food and drug administration electronic cigarette flavor enforcement policy on twitter: observational study. JMIR Public Health Surveill 2022 Mar 29;8(3):e25697. [doi: 10.2196/25697] [Medline: 35348461]
14. Shamoi E, Turdybay A, Shamoi P, Akhmetov I, Jaxylykova A, Pak A. Sentiment analysis of vegan related tweets using mutual information for feature selection. PeerJ Comput Sci 2022;8:e1149. [doi: 10.7717/peerj-cs.1149] [Medline: 36532810]
15. Rintyarna B. Mapping acceptance of indonesian organic food consumption under COVID-19 pandemic using sentiment analysis of Twitter dataset. J Theor Appl Inf Technol 2021;99:1009-1019 [FREE Full text]
16. Kaity M, Balakrishnan V. An integrated semi-automated framework for domain-based polarity words extraction from an unannotated non-English corpus. J Supercomput 2020 Dec;76(12):9772-9799. [doi: 10.1007/s11227-020-03222-0]
17. Mogaji E, Balakrishnan J, Kieu TA. Examining consumer behaviour in the UK Energy sector through the sentimental and thematic analysis of tweets. J of Consumer Behaviour 2021 Mar;20(2):218-230. [doi: 10.1002/cb.1820]
18. Karmegam D, Mappillairaju B. Social media analytics and reachability evaluation - #Diabetes. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 2022 Jan;16(1):102359. [doi: 10.1016/j.dsx.2021.102359]
19. Najafizada M, Rahman A, Donnan J, Dong Z, Bishop L. Analyzing sentiments and themes on cannabis in Canada using 2018 to 2020 Twitter data. J Cannabis Res 2022 Apr 13;4(1):22. [doi: 10.1186/s42238-022-00132-1] [Medline: 35418169]
20. Kasson E, Singh AK, Huang M, Wu D, Cavazos-Rehg P. Using a mixed methods approach to identify public perception of vaping risks and overall health outcomes on Twitter during the 2019 EVALI outbreak. Int J Med Inform 2021 Nov;155:104574. [doi: 10.1016/j.ijmedinf.2021.104574] [Medline: 34592539]
21. Chatterjee A, Prinz A, Gerdes M, Martinez S. Digital Interventions on healthy lifestyle management: systematic review. J Med Internet Res 2021 Nov 17;23(11):e26931. [doi: 10.2196/26931] [Medline: 34787575]
22. Kohl LFM, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013 Jul 16;15(7):e146. [doi: 10.2196/jmir.2665] [Medline: 23859884]
23. Ong SQ, Pauzi MBM, Gan KH. Text mining and determinants of sentiments towards the COVID-19 vaccine booster of Twitter users in Malaysia. Healthcare (Basel) 2022 May 27;10(6):994. [doi: 10.3390/healthcare10060994] [Medline: 35742045]
24. Kent EE, Prestin A, Gaysynsky A, et al. “Obesity is the New Major Cause of Cancer”: connections between obesity and cancer on Facebook and Twitter. J Cancer Educ 2016 Sep;31(3):453-459. [doi: 10.1007/s13187-015-0824-1] [Medline: 25865399]
25. Hutto C, Gilbert E, editors. VADER: a parsimonious rule-based model for sentiment analysis of social media text. Presented at: Proceedings of the International AAAI Conference on Web and Social Media; Jun 1-4, 2014; Ann Arbor, Michigan. [doi: 10.1609/icwsm.v8i1.14550]
26. Elbagir S, Yang J, editors. Twitter sentiment analysis using natural language toolkit and VADER sentiment. Presented at: Proceedings of the International Multiconference of Engineers and Computer Scientists; Mar 13-15, 2019; Hong Kong. [doi: 10.1142/9789811215094_0005]

XSL•FO
RenderX

27. Wan Min WNS, Zulkarnain NZ. Comparative evaluation of lexicons in performing sentiment analysis. J Adv Res Comput Tech Software Appl 2020 Jun 18;2(1):14-20 [FREE Full text]
28. Agarwal A, Xie B, Vovsha I, Rambow O. Sentiment analysis of Twitter data. Presented at: Proceedings of the Workshop on Languages in Social Media; Jun 23, 2011; Portland, Oregon.
29. Emblem H. Are you scared, VADER? Understanding how NLP pre-processing impacts VADER scoring. Medium. 2021. URL: https://medium.com/data-science/are-you-scared-vader-understanding-how-nlp-pre-processing-impacts-vader-scoring-4f4edadbc91d [accessed 2023-07-19]
30. Mathieson S, O'Keeffe M, Traeger AC, Ferreira GE, Abdel Shaheed C. Content and sentiment analysis of gabapentinoid-related tweets: an infodemiology study. Drug Alcohol Rev 2024 Jan;43(1):45-55. [doi: 10.1111/dar.13590] [Medline: 36539307]
31. Miller CA, Jung Kim S, Schwartz-Bloom RD, Bloom PN, Murphy SK, Fuemmeler BF. Informing women about the risks of exposing babies to tobacco smoke: outreach and education efforts using Facebook "boost posts". Transl Behav Med 2022 May 26;12(5):714-720. [doi: 10.1093/tbm/ibab158] [Medline: 35532323]
32. Healthy living: what is a healthy lifestyle? WHO Regional Office for Europe. 1999. URL: https://iris.who.int/bitstream/handle/10665/108180/EUR_ICP_;jsessionid=A62C6B339873362BC61C7C8BC6088C38?sequence=1 [accessed 2022-12-06]
33. Norimah AK, Safiah M, Jamal K, Haslinda S, Zuhaida H, Rohida S, et al. Food consumption patterns: Findings from the Malaysian Adult Nutrition Survey (MANS). Malays J Med Sci 2008 Mar;14(1) [FREE Full text]
34. Kortteinen T. Alcohol in Malaysia: the impact of social transformation. Contemp Drug Probl 1999 Sep;26(3):391-411. [doi: 10.1177/009145099902600303]
35. Robert Lourdes TG, Abd Hamid HA, Riyadzi MR, et al. Findings from a nationwide study on alcohol consumption patterns in an upper middle-income country. Int J Environ Res Public Health 2022 Jul 21;19(14):35886700. [doi: 10.3390/ijerph19148851] [Medline: 35886700]
36. Shaw G, Zimmerman M, Vasquez-Huot L, Karami A. Deciphering latent health Information in social media using a mixed-methods design. Healthcare (Basel) 2022 Nov 19;10(11):2320. [doi: 10.3390/healthcare10112320] [Medline: 36421644]
37. Hagger MS, Hardcastle SJ, Chater A, Mallett C, Pal S, Chatzisarantis NLD. Autonomous and controlled motivational regulations for multiple health-related behaviors: between- and within-participants analyses. Health Psychol Behav Med 2014 Jan 1;2(1):565-601. [doi: 10.1080/21642850.2014.912945] [Medline: 25750803]
38. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
39. Rutten GM, Meis JJM, Hendriks MRC, Hamers FJM, Veenhof C, Kremers SPJ. The contribution of lifestyle coaching of overweight patients in primary care to more autonomous motivation for physical activity and healthy dietary behaviour: results of a longitudinal study. Int J Behav Nutr Phys Act 2014 Dec;11(1):86. [doi: 10.1186/s12966-014-0086-z]
40. Teixeira PJ, Patrick H, Mata J. Why we eat what we eat: the role of autonomous motivation in eating behaviour regulation. Nutr Bull 2011 Mar;36(1):102-107. [doi: 10.1111/j.1467-3010.2010.01876.x]
41. Silva MN, Markland D, Carraça EV, et al. Exercise autonomous motivation predicts 3-yr weight loss in women. Med Sci Sports Exerc 2011 Apr;43(4):728-737. [doi: 10.1249/MSS.0b013e3181f3818f] [Medline: 20689448]
42. Hartmann-Boyce J. Why can't scientists agree on e-cigarettes. The Guardian. 2016. URL: https://www.theguardian.com/science/sifting-the-evidence/2016/sep/14/why-cant-scientists-agree-on-e-cigarettes-vaping [accessed 2023-05-06]
43. Kwon M, Park E. Perceptions and sentiments about electronic cigarettes on social media platforms: systematic review. JMIR Public Health Surveill 2020 Jan 15;6(1):e13673. [doi: 10.2196/13673] [Medline: 31939747]
44. Health minster: MOH to review tobacco generational endgame policy. Malay Mail. 2022. URL: https://www.malaymail.com/news/malaysia/2022/12/08/health-minster-moh-to-review-tobacco-generational-endgame-policy/44238 [accessed 2023-05-06]
45. Mixed sentiments towards smoking 'end game. New Straits Times. 2022. URL: https://www.nst.com.my/news/nation/2022/07/811829/mixed-sentiments-towards-smoking-end-game [accessed 2023-05-06]
46. National strategic plan for cancer control programe, 2021-2025. International Cancer Control Partnership. 2021. URL: https://www.iccp-portal.org/system/files/plans/National_Strategic_Plan_for_Cancer_Control_Programme_2021-2025.pdf [accessed 2024-08-20]
47. Jusoh S, Naing NN, Wan-Arfah N, et al. Prevalence and factors influencing smoking behavior among female inmates in Malaysia. Healthcare (Basel) 2023 Jan 9;11(2):203. [doi: 10.3390/healthcare11020203] [Medline: 36673571]
48. Evans-Polce RJ, Castaldelli-Maia JM, Schomerus G, Evans-Lacko SE. The downside of tobacco control? Smoking and self-stigma: a systematic review. Soc Sci Med 2015 Nov;145:26-34. [doi: 10.1016/j.socscimed.2015.09.026] [Medline: 26439764]
49. Castaldelli-Maia JM, Ventriglio A, Bhugra D. Tobacco smoking: From "glamour" to "stigma". A comprehensive review. Psychiatry Clin Neurosci 2016 Jan;70(1):24-33. [doi: 10.1111/pcn.12365] [Medline: 26449875]
50. Flashback #star50: when smoking was banned at eateries. Star. 2021. URL: https://www.thestar.com.my/news/nation/2021/10/20/flashback-star50-when-smoking-was-banned-at-eateries#:~:text=Malaysia%20enforced%20a%20no%2Dsmoking,rooms%20and%20ashtrays%20were%20prohibited [accessed 2023-05-06]

XSL•FO
RenderX

51. Hawkins LK, Farrow C, Thomas JM. Do perceived norms of social media users' eating habits and preferences predict our own food consumption and BMI? Appetite 2020 Jun 1;149:104611. [doi: 10.1016/j.appet.2020.104611] [Medline: 31958481]
52. Durau J, Diehl S, Terlutter R. Motivate me to exercise with you: the effects of social media fitness influencers on users' intentions to engage in physical activity and the role of user gender. Digit Health 2022;8:20552076221102769. [doi: 10.1177/20552076221102769] [Medline: 35615268]
53. Johnston C, Davis WE. Motivating exercise through social media: is a picture always worth a thousand words? Psychol Sport Exerc 2019 Mar;41:119-126. [doi: 10.1016/j.psychsport.2018.12.012]
54. Mohamad Saleh MS, Mehellou A, Huang M, Briandana R. Social media impact on sustainable intention and behaviour: a comparative study between university students in Malaysia and Indonesia. JARHE . [doi: 10.1108/JARHE-10-2023-0479]
55. Bauman A, Bull F, Chey T, et al. The International Prevalence Study on Physical Activity: results from 20 countries. Int J Behav Nutr Phys Act 2009 Mar 31;6(1):21. [doi: 10.1186/1479-5868-6-21] [Medline: 19335883]
56. Nik-Nasir NM, Md-Yasin M, Ariffin F, et al. Physical activity in Malaysia: are we doing enough? Findings from the REDISCOVER Study. Int J Environ Res Public Health 2022 Dec 15;19(24):16888. [doi: 10.3390/ijerph192416888] [Medline: 36554769]
57. Assanangkornchai S, Sam-Angsri N, Rerngpongpan S, Lertnakorn A. Patterns of alcohol consumption in the Thai population: results of the National Household Survey of 2007. Alcohol Alcohol 2010;45(3):278-285. [doi: 10.1093/alcalc/agq018] [Medline: 20348437]
58. Reporting about alcohol: a guide for journalists. World Health Organization. 2023. URL: https://iris.who.int/bitstream/handle/10665/366715/9789240071490-eng.pdf?sequence=1 [accessed 2023-09-06]
59. Abu Samah KAF, Nor Azharludin NM, Riza LS, Hasrol Jono MNH, Moketar NA. Classification and visualization: Twitter sentiment analysis of Malaysia's private hospitals. IJ-AI 2023 Mar 10;12(4):1793. [doi: 10.11591/ijai.v12.i4.pp1793-1802]
60. Mohd Yuswardi P, Ahmad NA. Sentiment analysis of Malaysians citizen's emotion towards cyberbullying in Twitter. IJARBSS 2023 Apr;13(4):769-780. [doi: 10.6007/IJARBSS/v13-i4/16777]
61. Malik M, Lamba H, Nakos C. Population bias in geotagged tweets. Presented at: Proceedings of the International AAAI Conference on Web and Social Media; May 26-29, 2015; University of Oxford, Oxford, UK. [doi: 10.1609/icwsm.v9i4.14688]
62. Lane JM, Habib D, Curtis B. Linguistic methodologies to surveil the leading causes of mortality: scoping review of twitter for public health data. J Med Internet Res 2023 Jun 12;25:e39484. [doi: 10.2196/39484] [Medline: 37307062]
63. Gore RJ, Diallo S, Padilla J. You are what you tweet: connecting the geographic variation in America's obesity rate to Twitter content. PLoS ONE 2015;10(9):e0133505. [doi: 10.1371/journal.pone.0133505] [Medline: 26332588]
64. Owens J, Adolescent Sleep Working Group, Committee on Adolescence. Insufficient sleep in adolescents and young adults: an update on causes and consequences. Pediatrics 2014 Sep;134(3):e921-e932. [doi: 10.1542/peds.2014-1696] [Medline: 25157012]
65. How to determine the best social media platforms for your business. Emphatic. 2024. URL: https://emphatic.co/how-to-determine-the-best-social-media-platforms-for-your-business [accessed 2024-01-08]
66. Riserbato R. What a social media target audience is and how to find it. Hubspot. 2024. URL: https://blog.hubspot.com/marketing/social-media-target-audience [accessed 2024-01-08]
67. Huang M, Rasool A, Jiang Q, Qu Q, Kamyab M, editors. HSMC: hybrid sentiment method for correlation to analyze COVID-19 tweets. In: Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery: Cham: Springer International Publishing; 2021. [doi: 10.1007/978-3-030-89698-0_101]
68. Rasool A, Tao R, Marjan K, Naveed T. Twitter sentiment analysis: a case study for apparel brands. J Phys: Conf Ser 2019 Mar 1;1176(2):022015. [doi: 10.1088/1742-6596/1176/2/022015]

Abbreviations

HCP: health care professional
NCD: noncommunicable disease
NHMS: National Health and Morbidity Survey
SDT: self-determination theory
VADER: Valence Aware Dictionary and Sentiment Reasoner
WHO: World Health Organization

XSL•FO
RenderX

Edited by M Haupt; submitted 24.09.24; peer-reviewed by A Rasool, R Gore; revised version received 26.01.25; accepted 13.02.25; published 25.06.25.

Please cite as:
Yip YY, Yaakub MR, Makmor-Bakry M, Abu Latiffi MI, Chong WW
Exploring Social Media Posts on Lifestyle Behaviors: Sentiment and Content Analysis
JMIR Infodemiology 2025;5:e65835
URL: https://infodemiology.jmir.org/2025/1/e65835
doi:10.2196/65835

© Yan Yee Yip, Mohd Ridzwan Yaakub, Mohd Makmor-Bakry, Muhammad Iqbal Abu Latiffi, Wei Wen Chong. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 25.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Reliability and Validity of Principles for Health-Related Information on Social Media (PRHISM) for Evaluating Breast Cancer Treatment Videos on YouTube: Instrument Validation Study

Hiroki Kusama[1], MD, PhD; Yoshimitsu Takahashi[2], DrPH; Shunichiro Orihara[3], PhD; Kayo Adachi[1], MD; Yumiko Ishizuka[4], MD, PhD; Ryoko Semba[4], MD, PhD; Hidetaka Shima[5], MD, PhD; Yoshiya Horimoto[1,4], MD, PhD; Hiroshi Kaise[6], MD, PhD; Masataka Taguri[3], PhD; Sho Inoue[7], MD, PhD; Takeo Nakayama[8], MD, PhD; Takashi Ishikawa[1], MD, PhD

[1]Department of Breast Surgical Oncology, Tokyo Medical University Hospital, 6-7-1, Nishishinjuku, Shinjuku-ku, Tokyo, Japan
[2]Department of Implementation Science in Public Health, School of Public Health, Kyoto University, Kyoto, Japan
[3]Department of Health Data Science, Tokyo Medical University Hospital, Tokyo, Japan
[4]Department of Breast Oncology, Faculty of Medicine, Juntendo University, Tokyo, Japan
[5]Department of Breast Surgery, Douaikai Ozawa Hospital, Kanagawa, Japan
[6]Department of Breast Surgical Oncology, Tokyo Medical University Ibaraki Medical Center, Ibaraki, Japan
[7]YCU Co-Creation Innovation Center, Yokohama City University, Kanagawa, Japan
[8]Department of Health Informatics, School of Public Health, Kyoto University, Kyoto, Japan

Corresponding Author:
Hiroki Kusama, MD, PhD
Department of Breast Surgical Oncology, Tokyo Medical University Hospital, 6-7-1, Nishishinjuku, Shinjuku-ku, Tokyo, Japan

Abstract

Background: There is breast cancer–related medical information on social media, but there is no established method for objectively evaluating the quality of this information. Principles for Health-Related Information on Social Media (PRHISM) is a newly developed tool for objectively assessing the quality of health-related information on social media; however, there have been no reports evaluating its reliability and validity.

Objective: The purpose of this study was to statistically examine the reliability and validity of PRHISM using videos about breast cancer treatment on YouTube (Google).

Methods: In total, 60 YouTube videos were selected on January 5, 2024, with the Japanese words for "breast cancer," "treatment," and "chemotherapy," and assessed by 6 Japanese physicians with expertise in breast cancer. These evaluators independently evaluated the videos using PRHISM and an established tool for assessing the quality of health-related information, DISCERN, as well as through subjective assessments. We calculated interrater and intrarater agreement among evaluators with CIs, measuring agreement using weighted Cohen kappa.

Results: The interrater agreement for PRHISM overall quality was κ=0.52 (90% CI 0.49-0.55), indicating that the expected level of agreement, statistically defined by the lower limit of the 90% CI exceeding 0.53, was not achieved. However, PRHISM demonstrated higher agreement compared with DISCERN overall quality, which had a κ=0.45 (90% CI 0.41-0.48). In terms of validity, the intrarater agreement between PRHISM and subjective assessments by breast experts was κ=0.37 (95% CI 0.14-0.60), while DISCERN showed an agreement of κ=0.27 (95% CI 0.07-0.48), indicating fair agreement and no significant difference in validity.

Conclusions: PRHISM has demonstrated sufficient reliability and validity for evaluating the quality of health-related information on YouTube, making it a promising new metric. To further enhance objectivity, it is necessary to explore the use of artificial intelligence and other approaches.

(JMIR Infodemiology 2025;5:e66416) doi:10.2196/66416

KEYWORDS

information quality; social media; YouTube; PRHISM; breast cancer treatment; videos; reliability; validity; instrument validation study; medical information; online health information; cancer treatment; Japan; Principles for Health-Related Information on Social Media

XSL•FO
RenderX

Introduction

In recent years, advances in medical technology and diagnostic methods have made health care increasingly complex, leading to a growing tendency for patients to seek information about their disease and treatments [1]. Many patients feel anxious during this process, and internet use for collecting medical information is increasing [2]. In Japan, about 50% of people use the internet as a method for collecting medical information [3]. The sources of information include those officially provided by national cancer centers, as well as those from social media. On social media, anyone can post information regardless of their expertise or qualifications, making it very difficult for patients to judge the quality of that information, which is often a mix of reliable and unreliable sources [4]. The spread of misinformation about vaccines during the COVID-19 pandemic highlighted the problem of inaccurate medical information on social media [5,6]. In response to this issue, the National Academy of Medicine (NAM) published a guide in 2021 to help identify reliable sources of health information on social media [4].

Methods for objectively evaluating the quality of medical information available on the internet have included Diagnostic Information Support Communication Evaluation Report Network (DISCERN) [7]. DISCERN was developed as a tool to assess the quality of written information about treatment choices. It is a metric that evaluates the quality of information using 16 questions, consisting of 15 basic assessments and 1 overall assessment, each scored on a 1 (No) to 5 (Yes) Likert scale. This metric has long been considered reliable and is currently used to assess the quality of breast cancer treatment information available on the internet [8,9]. However, since they were developed before the year 2000 and were not designed with social media in mind, there are concerns that they may be inadequate for evaluating social media [10].

Given the increasing reliance on the internet for health information, it is crucial to ensure that the information available to patients is accurate and reliable. Misinformation can lead to incorrect self-diagnosis, inappropriate treatment choices, and increased anxiety, ultimately affecting patient outcomes and public health. Therefore, there is a pressing need to develop and validate tools specifically designed to evaluate the quality of medical information on social media platforms.

Denniss and colleagues [10] developed the Principles for Health-Related Information on Social Media (PRHISM) tool to evaluate health-related information on social media using a modified Delphi method. This tool assesses 13 principles on a 0 (completely unmet)-4 (completely met) Likert scale and allows nonexperts to evaluate the quality of information, potentially reducing reviewer bias. Compared with existing metrics for evaluating information, PRHISM was specifically designed for social media, offering greater logical validity. Its questions are tailored for social media platforms, with additional considerations for readability and accommodations for vision and hearing impairments. Although PRHISM has not yet been widely used for social media evaluation, its development through appropriate methods and its adaptability suggest it could be valuable for this purpose.

To address these challenges, we adapted a Japanese version of PRHISM and evaluated its reliability and validity. We also conducted assessments using DISCERN, widely used for evaluating medical information quality, and compared the results with PRHISM. Our goal was to evaluate whether PRHISM could adequately assess information quality on Japanese social media and, using this tool, create an environment where patients can access accurate medical information.

Methods

Translation of the PRHISM

To use PRHISM in Japan, we obtained permission from the developers of PRHISM. First, PRHISM was translated from English to Japanese by a native Japanese speaker (H Kusama). Then, the Japanese version was back-translated into English by a native English speaker. Any discrepancies between the back-translated English version and the original text were identified, and the Japanese version appropriately adjusted (Multimedia Appendix 1).

Social Media Platform

The evaluation of PRHISM was conducted using YouTube (Google). YouTube is the second most used platform in Japan after LINE (LY Corporation) and is used across all age groups [11]. Among various platforms, we determined that YouTube is suitable for this evaluation because it allows the posting of longer videos, enabling experts to adequately assess the medical information provided.

How to Search

Since YouTube uses an algorithm that analyzes viewing history through artificial intelligence (AI) to prioritize related videos, searches were conducted using the internet browser Google Chrome in incognito mode, with no login session active. The search terms used were the Japanese words "乳癌" (にゅうがん, Nyugan, breast cancer), "治療" (ちりょう, Tiryou, treatment), and "抗癌剤" (こうがんざい, Kouganzai, chemotherapy). The search was conducted on January 5, 2024, and the results were sorted by relevance, with the videos listed in order from the top. The exclusion criteria were (1) languages other than Japanese, (2) fewer than 3000 views, (3) shorter than 60 seconds, (4) irrelevant videos, (5) without audio, (6) YouTube shorts, (7) duplicates, and (8) advertisements.

The sources of the videos were categorized into eight categories: (1) health profession schools and other educational institutions (schools of medicine, pharmacy, etc); (2) health care facilities (hospitals, clinics, etc); (3) nonprofit health plans; (4) public health departments (national statement, regional statement, etc); (5) individual health care professionals (doctors, nurses, occupational therapists, etc); (6) entertainment, media, news; (7) personal blogs; and (8) other.

These categories were created with reference to the reliable sources identified by the National Academy of Medicine [4]. Categories 1 - 4 were defined as content from reliable sources, while categories 5 - 8 were defined as content from other

XSL•FO
RenderX

sources. In addition, general information, such as the uploader, number of channel subscribers, number of views, video length, upload date, number of likes, and time since posting, was also collected.

Evaluators

In total, 6 physicians with expertise in breast cancer conducted the evaluations. All of them were surgical oncologists, and the authors recruited participants from their own institution and affiliated hospitals for this study. : KA (8 y as a physician, 1 y as a breast expert), YI (9 y as a physician, 2 y as a breast expert), RS (10 y as a physician, 2 y as a breast expert), HS (16 y as a physician, 7 y as a breast expert), YH (24 y as a physician, 15 y as a breast expert), and H Kaise (37 y as a physician, 20 y as a breast expert).

Evaluation Method

In total, 6 physicians with expertise in breast cancer evaluated a common set of 60 videos. For each video, they used PRHISM and DISCERN for evaluation. In addition, the accuracy and potential harm of the information was assessed.

PRHISM

PRHISM evaluates the quality of health-related information on social media using 13 principles, which are scored on a 0 - 4 Likert scale [9].

Since some items may not be applicable depending on the content of the video, the score is calculated based on the applicable questions and converted to a score out of 100 (PRHISM score). Scores of 100 - 76 are rated as excellent, 75 - 51 as good, 50 - 26 as mediocre, and 25 - 0 as poor (Textbox 1).

XSL•FO
RenderX

Textbox 1. Summary of the evaluation tools used in this study.

Principles for Health-Related Information on Social Media (PRHISM)

PRHISM is comprised of 13 principles. Each principle is scored on a 5-point Likert scale (0-4).

Principles:

1. Authorship
2. Authoritative
3. Action-oriented
4. Financial disclosure
5. Attribution
6. Balance and justifiability
7. Risks and benefits
8. Privacy
9. Complementary information
10. Referrals and support
11. Readability and comprehensibility
12. Accessibility
13. Images

DISCERN

DISCERN is comprised of 8 reliability assessments, 7 information quality assessments, and one overall quality assessment. Each principle is scored on a 5-point Likert scale (1-5).

Reliability:

1. Are the aims clear?
2. Does it achieve its aims?
3. Is it relevant?
4. Is it clear what sources of information were used to compile the publication (other than the author or producer)?
5. Is it clear when the information used or reported in the publication was produced?
6. How good is the quality of information treatment choices?
7. Is it balanced and unbiased?
8. Does it provide details of additional sources of support and information?

Information quality:

1. Does it refer to areas of uncertainty?
2. Does it describe how each treatment works?
3. Does it describe the benefits of each treatment?
4. Does it describe the risks of each treatment?
5. Does it describe what would happen if no treatment is used?
6. Does it describe how the treatment choices affect overall quality of life?
7. Is it clear that there may be more than one possible treatment choice? Does it provide support for shared decision-making?

Overall evaluation:

1. Overall rating of the publications.

Cancer Expert Assessment Tool

Cancer expert assessment tool is comprised of 4 question assessments. Two assessments consist of whether the information is true or false and harmful or not harmful, and 2 review reasons why the evaluation was chosen.

Expert Panel Member Assessment:

XSL•FO
RenderX

1. In your opinion, are the primary medical claims within the article accurate?
   - 5: True, 4: Mostly true, 3: Mixture both True/False, 2: Mostly False, 1: False.
2. If you answered, “Mixture both True/False,” “Mostly False” or “False,” why did you answer this way?
3. In your opinion, could the primary medical claims within the article cause harm?
   - 5: Certainly NOT Harmful, 4: Probably NOT Harmful, 3: Uncertain, 2: Probably Harmful, 1: Certainly Harmful.
4. If you answered, “Uncertain” or “Probably Harmful” or “Certainly Harmful,” why did you answer this way?

DISCERN

DISCERN evaluates the scientific reliability of medical information related to treatment descriptions and assigns a score [8]. The total score is out of 80 points, with each of the 16 assessments contributing up to 5 points (DISCERN score) (Textbox 1).

Accuracy and Potential Harm of the Information

The subjective assessment was conducted using an assessment tool developed by Johnson et al [12]. This tool assesses whether the provided medical information is accurate or inaccurate, and if considered inaccurate, the reason is marked in a checkbox. In addition, it evaluates whether the information is harmful or not using the same tool. Accurate information is rated as 1, and inaccurate information as 5, on a Likert scale. Similarly, nonharmful information is rated as 1, and harmful information as 5. To align the scoring methods of PRHISM and DISCERN, in this study, accurate or nonharmful information was rated as 5, and inaccurate or harmful information was rated as 1 (Textbox 1).

Training and Protocol for Evaluators

Each evaluator received a lecture once on how to use the metrics during a preliminary meeting. The principal investigator, who is also the first author of the study, conducted the lecture in an internet-based group session in a group session for approximately 60 minutes. Subsequently, evaluation sheets, along with Japanese translations of the PRHISM and DISCERN guides [10,13], were provided. The evaluators conducted their assessments based on these guides. No further training were conducted thereafter.

Discussion among evaluators regarding each evaluation was not permitted.

Reliability

Reliability was assessed by examining the agreement of scores between evaluators. In the previous literature on DISCERN, evaluations were based on the agreement of the overall quality score. However, PRHISM does not have a corresponding criterion. For the purpose of statistical evaluation and comparison in this study, we added a “PRHISM overall quality” component, similar to that in DISCERN, which assessed the entire video after evaluating the 13 principles. This component was also scored on a 0 - 4 Likert scale, and its agreement was evaluated.

Validity

There is no gold standard for evaluating the validity of the quality of medical information. For validity, the subjective assessment of the quality of medical information by physicians with expertise in breast cancer was considered an appropriate assessment, and this was used as the standard for validity. Validity was assessed by examining the agreement between the PRHISM overall quality and the experts’ subjective assessments. The validity of DISCERN was also examined in the same way.

Sample Size Determination

The number of videos to be evaluated was determined based on the interrater agreement for PRHISM overall quality. As there are no reports examining the level of interrater agreement for PRHISM, DISCERN was used as a reference. Currently, DISCERN is the primary tool used to evaluate the quality of medical information on social media. DISCERN assesses 15 criteria and then evaluates the overall quality of the information. In previous studies, the agreement on the overall quality has been evaluated. Cohen kappa was used to evaluate the degree of agreement. The degree of agreement varies depending on the expertise of the evaluators, with a kappa of 0.23 reported for self-help group members, 0.40 for information providers, and 0.53 for an expert [7].

The expected agreement for PRHISM overall quality was assumed to have a threshold of κ=0.53 and an expected value of κ=0.61. The threshold of 0.53 was determined based on the assumption that PRHISM, being a tool specifically designed for evaluating social media, would achieve a higher level of agreement than DISCERN for an expert. The expected value of 0.61 is generally considered to indicate “sufficient agreement” in terms of the kappa coefficient [14]. For the number of videos to be evaluated in this study, it was necessary to assume the distribution of the overall quality scores in PRHISM. Therefore, as a pilot test, the investigator (HK [Hiroki Kusama]) evaluated 50 videos to establish this distribution in Multimedia Appendix 2. Based on the above settings, the number of videos was determined through a simulation experiment conducted 10,000 times. Assuming an alternative hypothesis kappa coefficient of 0.61, the number of videos required to reject the null hypothesis of 0.53 with over 80% power at a 2-sided significance level of 10% was calculated to be 55 for 6 evaluators. The z test, approximated by a normal distribution, was used to calculate the test statistics [15]. Anticipating that some videos might be difficult to evaluate, we decided to have 6 evaluators assess 60 videos. The 6 evaluators were distributed, with 3 physicians having more than 10 years of experience and 3 physicians having less than 10 years of experience.

XSL•FO
RenderX

Statistical Analysis

The primary analysis focused on examining the interrater agreement for PRHISM overall quality (reliability of PRHISM) and DISCERN overall quality (reliability of DISCERN).

As secondary analyses, we examined the following:

- The intrarater agreement between PRHISM overall quality and DISCERN overall quality.
- The intrarater agreement between PRHISM overall quality and expert evaluations (validity of PRHISM).
- The intrarater agreement between DISCERN overall quality and expert evaluations (validity of DISCERN).
- The interrater agreement for PRHISM score and its categories (reliability of PRHISM).
- The interrater agreement for DISCERN score and its categories (reliability of DISCERN).

To compare the DISCERN score with the PRHISM score, we calculated the modified DISCERN score by subtracting the DISCERN overall quality score from the total DISCERN score, dividing the result by the maximum possible score based on the number of applicable questions, and then converting it to a score out of 100 (modified DISCERN score).

The agreement of evaluations was assessed using the kappa coefficient. The interpretation of agreement levels is as follows: <0.00 (no agreement); 0.00 - 0.20 (slight); 0.21 - 0.40 (fair); 0.41 - 0.60 (moderate); 0.61 - 0.80 (substantial); and 0.81 - 1.00 (almost perfect) [14]. The interobserver agreement was calculated using the kappa coefficient and a 90% CI, and an agreement level of 0.61 or higher was interpreted as sufficient.

For the interrater agreement of the PRHISM overall quality, 2 evaluators were selected from a group of 6, and the agreement for 15 different patterns was calculated. The average and the 90% CI were calculated [15]. The 90% CI was calculated using normal approximation. If the lower limit of the 90% CI exceeded 0.53, the primary analysis (PRHISM reliability) was considered to have been achieved.

For the secondary analyses, the kappa coefficient and its 95% CI were calculated for each pair of evaluators to assess intrarater agreement. Secondary analyses for the interrater agreement were performed using the same approach as the primary analysis.

In addition, since the PRHISM score and the modified DISCERN score are continuous variables, we also performed an analysis using intraclass correlation coefficient. Statistical analyses were performed using R software (version 4.2.3; R Core Team).

Ethical Considerations

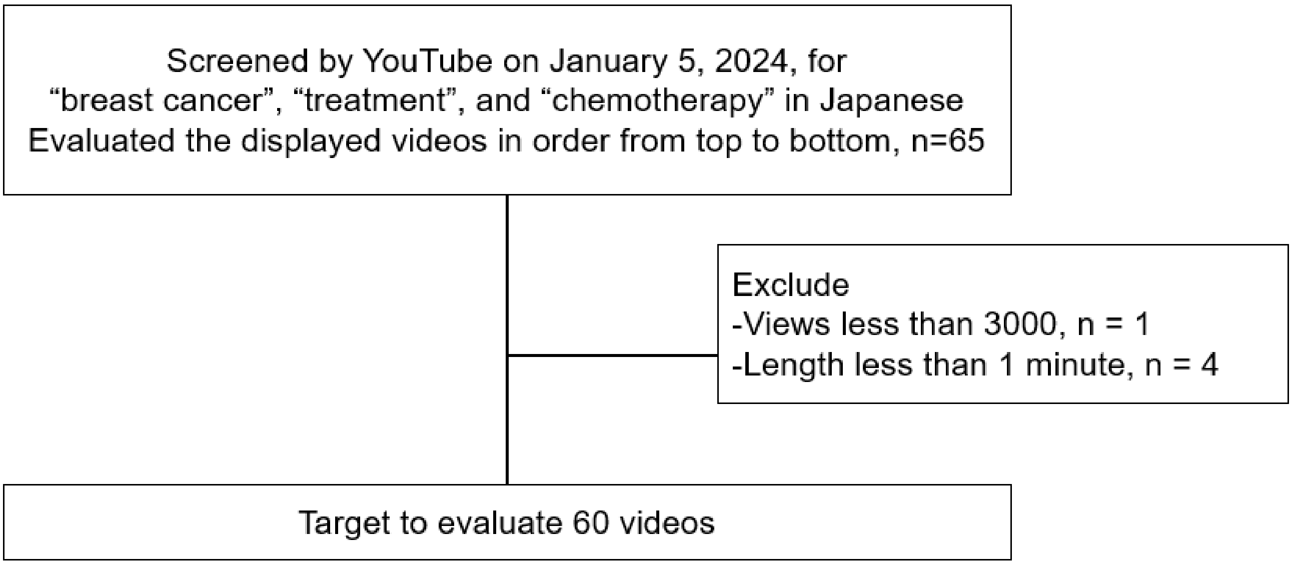
The study was approved by the Institutional Review Board of Tokyo Medical University (T2024-0034). The study involved an analysis of publicly available YouTube data, which does not require individual consent from participants. However, ethical approval was obtained to ensure that the research adhered to institutional guidelines for research involving public data. No compensation was provided to participants as this study involved an analysis of publicly available data. Should any concerns or complaints be raised by video contributors or their families regarding ethical or social issues, the principal investigator will respond sincerely and appropriately in line with institutional procedures.

Results

Overview

Using the predefined search method, we excluded 5 videos in total (1 video with fewer than 3000 views and 4 videos with a duration of less than 1 minute), resulting in a list of 60 videos. A CONSORT (Consolidated Standards of Reporting Trials) diagram is shown in Figure 1.

Figure 1. Flowchart of video selection in this study. In total, 60 videos were selected after excluding those with fewer views or short duration.



The median video length was 8 (range: 1 - 126) minutes, the median number of views was 30,542.5 (range: 3921 - 978,676), and the median time since posting was 29 (range: 7 - 123) months. The sources of the videos were as follows: 5 individual health care professionals accounted for the largest category with 15 videos (25%); followed by 6 entertainment, media, and news

XSL•FO
RenderX

with 13 videos (22%); 7 personal blogs with 12 videos (20%); 3 nonprofit health plans with 9 videos (15%); 1 health professions schools and other educational institutions with 6 videos (10%); 2 health care organizations with 4 videos (7%); and 8 other sources with 1 video (2%). No videos were posted from the public health departments (Table 1).

Table . The characteristics of selected videos on YouTube.

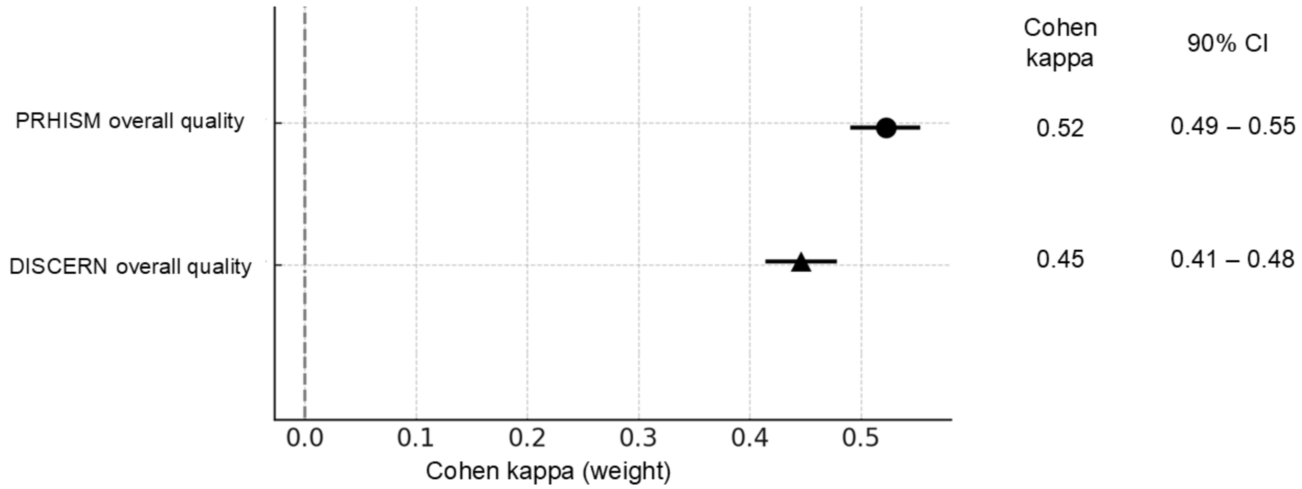
Characteristics		Statistical value (n=60)
Video length (mins), median (range)		8 (1-126)
Views, n (range)		30,542.5 (3921-978,676)
Time since posting (months), median (range)		29 (7-123)
Sources of videos, n (%)		
	Health professions schools and other educational institutions (eg, medical schools and pharmacy schools)	6 (10)
	Health care organizations (eg, academic medical centers and specialty hospitals)	4 (7)
	Nonprofit health plans	9 (15)
	Public health departments	0 (0)
	Individual health care professionals	15 (25)
	Entertainment and media news	13 (22)
	Breast cancer survivor's blog	12 (20)
	Others	1 (2)

Primary Analysis

Reliability

The interrater agreement for PRHISM overall quality was κ=0.52 (90% CI 0.49-0.55), indicating moderate agreement. Since the lower limit of the 90% CI was below 0.53, the primary analysis was not achieved. However, the interrater agreement for DISCERN overall quality was κ=0.45 (90% CI 0.41-0.48), also indicating moderate agreement. The 90% CI did not overlap, suggesting that PRHISM may be a superior measure in terms of interrater agreement compared with DISCERN (Figures 2 and 3).

Figure 2. Interrater agreement on PRHISM and DISCERN overall quality. The circle and triangle represent the mean kappa and CIs (90% CI) are represented by horizontal error bars. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.



XSL•FO
RenderX

Figure 3. Results of reliability and validity agreement. The circle represents the mean kappa and CIs (90 or 95% CI) are represented by horizontal error bars. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; ICC: intraclass correlation coefficients; NaN: not a number; PRHISM: Principles for Health-Related Information on Social Media.

XSL•FO
RenderX

Subgroup Analysis in Primary Analysis

For videos originating from reliable sources (categories 1 - 4), the PRHISM overall quality was κ=0.45 (90% CI 0.38-0.51), and the DISCERN overall quality was κ=0.28 (90% CI 0.22-0.33). For videos from other sources (categories 5 - 8), the PRHISM overall quality was κ=0.40 (90% CI 0.36-0.44), and the DISCERN overall quality was κ=0.34 (90% CI 0.30-0.38; Figure 3, Multimedia Appendix 3). The 90% CI for PRHISM from reliable sources did not overlap with the 90% CI for DISCERN.

For those with over 10 years of experience, the PRHISM overall quality was κ=0.57 (90% CI 0.51-0.63), and the DISCERN overall quality was κ=0.60 (90% CI 0.54-0.65). For those with less than 10 years of experience, the PRHISM overall quality was κ=0.30 (90% CI 0.20-0.40), and the DISCERN overall quality was κ=0.21 (90% CI: 0.12-0.29). When the years of experience were 10 or more, the agreement on overall quality was higher for both PRHISM and DISCERN compared with those with less than 10 years of experience (Figure 3, Multimedia Appendix 4).

Secondary Analysis

Intrarater agreement for PRHISM overall quality and DISCERN overall quality was κ=0.63 (95% CI 0.55-0.73), indicating substantial agreement. In evaluating medical information on social media, there was no significant difference between the assessments of PRHISM overall quality and DISCERN overall quality (Figure 3).

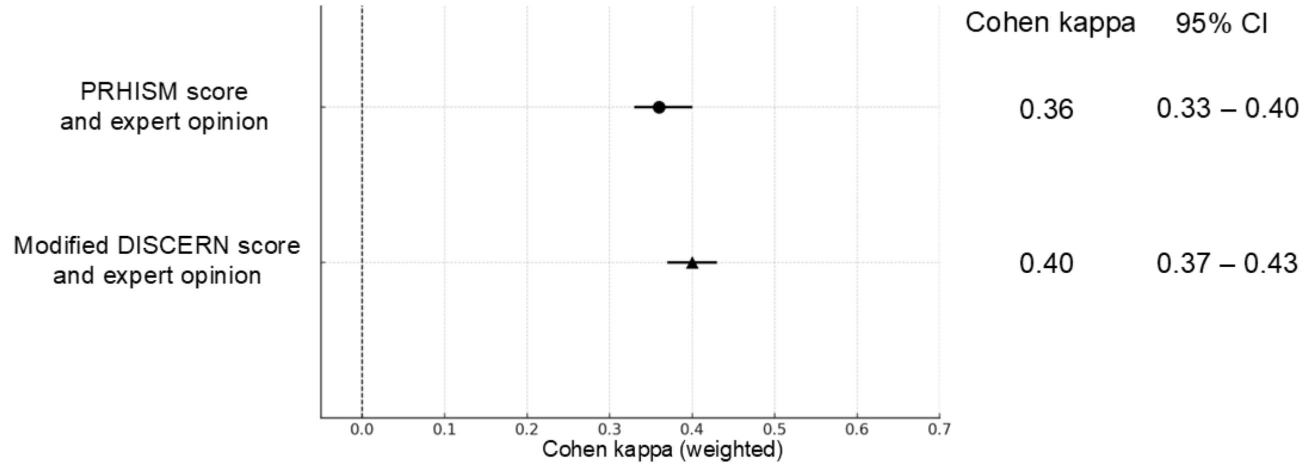
The intrarater agreement among experts was κ=0.54 (95% CI 0.50-0.57), indicating moderate agreement (Figure 3).

Validity

We evaluated the agreement between the PRHISM overall quality and DISCERN overall quality scores with the quality of information as subjectively assessed by experts. The agreement between PRHISM overall quality and the experts' subjective assessment was κ=0.37 (95% CI 0.14-0.60), indicating fair agreement. The agreement between DISCERN overall quality and the experts' subjective assessment was κ=0.27 (95% CI 0.07-0.48), indicating fair agreement. The 95% CIs overlapped, suggesting that the validity was considered equivalent (Figures 3 and 4). The 95% CIs for the agreement of each of the 6 evaluators all overlapped, but the level of agreement varied among specialists for both PRHISM and DISCERN, ranging from κ=0.07 to 0.65 and κ=0.01 to 0.51, respectively. (Multimedia Appendix 5).

The circle represents the mean kappa for PRHISM, and triangle represent the mean kappa for DISCERN. The 95% CIs are represented by horizontal error bars.

Figure 4. Intrarater agreement between PRHISM or DISCERN overall quality and expert opinion, the circle and triangle represent the mean kappa and CIs (95% CI) are represented by horizontal error bars. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.



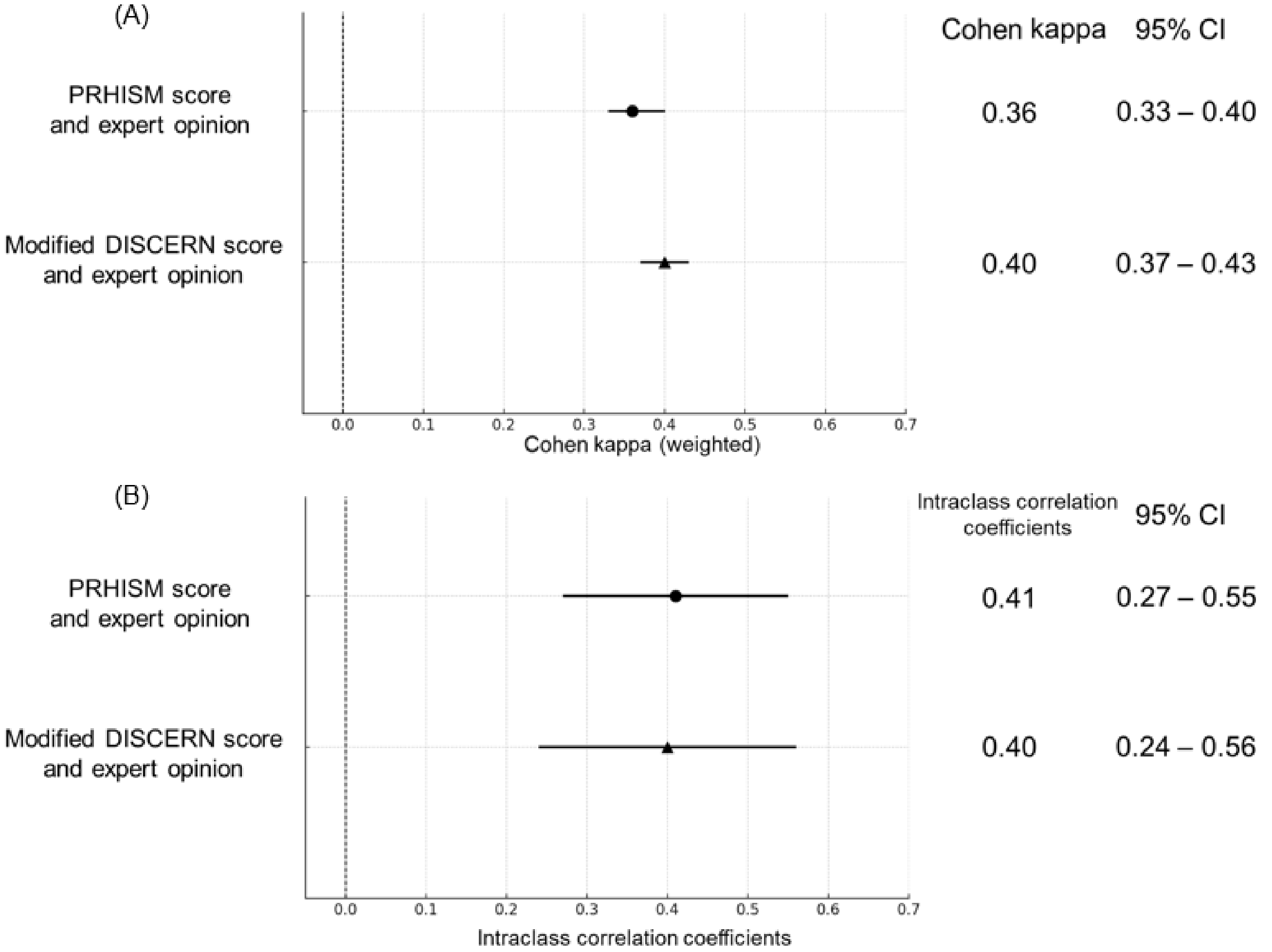
Agreement Between PRHISM Score and DISCERN Score

We evaluated the interrater agreement among 6 evaluators for the PRHISM score and the modified DISCERN score, both rated on a score out of 100. The interrater agreement for the PRHISM score and the modified DISCERN score was κ=0.36 (95% CI 0.33-0.40) and κ=0.40 (95% CI 0.37-0.43), respectively (Figure 5A). Using the intraclass correlation coefficient, the PRHISM score was 0.41 (95% CI 0.27-0.55) and the DISCERN score was 0.40 (95% CI 0.24-0.56) (Figure 5B).

The circle represents the mean kappa for PRHISM, and the triangle represent the mean kappa for DISCERN. The 95% CIs are represented by horizontal error bars.

XSL•FO
RenderX

Figure 5. Intra-rater agreement between PRHISM score and modified DISCERN score and expert opinion. The circle and triangle represents the mean kappa and 95% CIs are represented by horizontal error bars. (**A**) Cohen kappa. (**B**) Intraclass correlation coefficient. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.



Agreement for Each Evaluation Question of PRHISM and DISCERN

For PRHISM, when a question was judged as "not applicable," it was excluded from the score calculation. Therefore, some were assessed using unweighted Cohen kappa. The highest agreement was for question 2, "authoritative," with κ=0.73 (95% CI 0.70-0.76). For DISCERN, the highest agreement was for item 12, "benefits of treatment," with κ=0.60 (95% CI 0.56-0.63) (Figure 3).

In addition, PRHISM is a metric that classifies video quality as poor, mediocre, good, or excellent based on the PRHISM score. We evaluated the agreement between the subjective assessments by experts and the PRHISM scoring classification. When assigning 1 to poor and 4 to excellent, the agreement was κ=0.54 (95% CI 0.45-0.64), indicating moderate agreement (Multimedia Appendix 6).

Evaluation of the Quality of Breast Cancer Treatment Information on YouTube in Japan Using PRHISM and DISCERN Score

Although this study primarily examined the utility of PRHISM, we also evaluated the quality of breast cancer treatment information on YouTube in Japan using PRHISM and DISCERN scores. The mean PRHISM and DISCERN scores for medical information related to breast cancer treatment on Japanese YouTube were 60.6 (SD 11.5) and 58.9 (SD 11.5), respectively (Multimedia Appendix 7).

Discussion

Principal Findings

In evaluating the quality of medical information on YouTube, the interrater agreement for the overall quality score of PRHISM was κ=0.52 (90% CI 0.49-0.55), and the primary end point was not achieved. However, the 90% CI for interrater agreement of PRHISM was superior to that of DISCERN, indicating that PRHISM is a more reliable metric when evaluating the quality of medical information on YouTube. In terms of validity, the agreement between the experts' subjective assessment and PRHISM overall quality was κ=0.37 (95% CI 0.14-0.60), indicating fair agreement. It was found that PRHISM has validity equivalent to that of DISCERN.

Although the primary end point was not achieved, 1 possible reason was the difficulty in setting the threshold and expected values for interrater agreement of PRHISM. In setting the threshold and expected values, we used DISCERN as a reference because no studies have examined the interrater agreement of PRHISM [7,10]. While the number of evaluators and videos was determined statistically, it is possible that a larger number

XSL•FO
RenderX

of both was necessary to adequately evaluate the quality of information. However, DISCERN was designed for books available in public libraries and bookstores, and leaflets produced by professional organizations and national self-help groups. We referred to studies that used DISCERN to evaluate medical information on social media, but many of them used a modified DISCERN with fewer evaluation items, making them difficult to reference [16-22]. DISCERN was not developed for social media, so the agreement may differ from previous studies. In fact, the agreement for DISCERN in this study, which focused on social media, was 0.45, lower than the 0.53 reported in previous studies [7] , suggesting that the threshold and expected values might have been better set slightly lower. Therefore, given the results obtained with this threshold, no definitive conclusion can be drawn about the robustness of PRHISM.

In previous studies, the agreement for DISCERN decreased depending on the profession of the evaluators [7]. In this study as well, a difference in agreement was observed depending on whether the years of experience were 10 or more, or less than 10. Although PRHISM is designed to allow for evaluation by nonexperts, these results suggest that evaluations conducted by experts may be more accurate.

We also examined validity. However, there is no gold standard for evaluating the quality of medical information. Therefore, it is necessary to establish a consensus among experts. During the development of PRHISM, the modified Delphi method, a consensus-building technique, was used. In addition, we assessed validity by setting alternative criteria. In previous reports, some studies have assessed validity based on guidelines and evidence [23] , while others have used expert evaluations as the standard [12,24] . In addition, there are studies that have used DISCERN as an alternative criterion. In this study, using expert evaluations as the standard, the agreement among experts was κ=0.54 (95% CI 0.50-0.57), showing a moderate level of consistency. However, the agreement between the experts' evaluations and PRHISM overall quality was κ=0.37 (95% CI 0.14-0.60), showing only fair agreement. Nevertheless, since the agreement was comparable with that of DISCERN, it suggests that PRHISM is also sufficiently valid for evaluating the quality of information. The lack of strong agreement may be due to the inevitable subjectivity of the assessments, leading to variations in judgments based on each expert's preferences and perspectives. In fact, the level of agreement varied among experts (Multimedia Appendix 5). It can be considered to have at least comparable validity to DISCERN, but there may be a need to consider how to use this tool regarding its validity.

Future Prospects

We aim to explore the use of AI to make the evaluation of medical information more objective, efficient, and with higher validity, allowing for the assessment of a larger volume of information in a shorter time. In fact, there are reports investigating whether AI can be used to evaluate online health information [25] . This study demonstrated that PRHISM is a suitable tool for evaluating the quality of medical information on social media. Therefore, this research serves as an important first step toward further investigations using PRHISM to assess the quality of medical information in various social media contexts. If low-quality information could automatically trigger warnings, it would help ensure that patients receive higher-quality medical information. Future research will explore the extent to which AI can be integrated into the evaluation process.

Limitations

This study has some limitations. First, the PRHISM overall quality is a metric independently established by the authors and was created specifically for statistical analysis. Since PRHISM is a tool that evaluates information using the PRHISM score or PRHISM scoring classification [10], the study may not directly evaluate the tool itself.

Second, there was a small sample size and a limited number of evaluators. Although a statistically valid number was considered, there was variability in the evaluations among the experts, potentially influenced by differences in their years of clinical experience. This indicates that a larger sample size and a more diverse group of evaluators with varying levels of expertise might be needed. To address this issue, future studies could include standardized training modules to improve consistency. In addition, integrating AI could automate certain aspects of scoring, reducing human bias and increasing efficiency.

Third, this study presents results solely from Japan. There may be an influence on the results due to biases in breast cancer treatment practices and expertise in Japan, as well as differences in medical environments. In addition, this study was limited to information about breast cancer treatment on YouTube, and the findings may not be applicable to other diseases or health-related information on different social media platforms. Similar studies need to be conducted in other countries and on different social media platforms. Although YouTube has regulations on posted videos, some reports indicate that the quality of health-related videos on YouTube varies widely, from low to high. Therefore, further research is needed to determine whether PRHISM is effective for evaluating content in other countries or on different social media platforms, as this will enhance its utility and relevance globally. We plan to conduct evaluations on other social media platforms.

Conclusions

PRHISM has greater reliability than DISCERN in evaluating the quality of medical information on social media, with comparable validity. It has the potential to become a standard metric for assessing the quality of medical information on social media.

XSL•FO
RenderX

Acknowledgments

We would like to express our sincere gratitude to Denniss for granting us permission to use the PRHISM (Principles for Health-Related Information on Social Media) in this study. We thank Crimson Interactive Pvt Ltd (Ulatus) for their assistance in manuscript translation and editing.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

H Kusama contributed to conceptualization, data curation, formal analysis, investigation, methodology, visualization, and writing – original draft. YT contributed to conceptualization, methodology, and writing – review and editing. SO contributed to formal analysis. KA, YI, RS, HS, and H Kaise contributed to resources. YH contributed to resources and writing – review and editing. MT contributed to formal analysis. SI and TN contributed to conceptualization. TI handled the project administration and writing – review and editing.

Conflicts of Interest

YT is employed through a joint research fund between Kyoto University and HealthTech Laboratory Inc. The authors declare that this relationship had no influence on the study design, data collection, analysis, interpretation, or manuscript preparation.

Multimedia Appendix 1
PRHISM English translation. PRHISM: Principles for Health-Related Information on Social Media.
[DOCX File, 28 KB - infodemiology_v5i1e66416_app1.docx]

Multimedia Appendix 2
Result of pilot study. The results of an evaluation of the top 50 YouTube videos searched by the representative (H Kusama). The search terms, search method, and exclusion criteria are the same as those used in this study. The y-axis indicates the number of videos, while the x-axis represents the PRHISM overall quality score. PRHISM: Principles for Health-Related Information on Social Media.
[PPTX File, 41 KB - infodemiology_v5i1e66416_app2.pptx]

Multimedia Appendix 3
Intrarater agreement between PRHISM/DISCERN overall quality and each expert opinion. The y-axis displays the initials of the breast experts who made the assessments. The circle represents the mean kappa for PRHISM, and the triangle represent the mean kappa for DISCERN. The 95% CIs are represented by horizontal error bars. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.
[PPTX File, 118 KB - infodemiology_v5i1e66416_app3.pptx]

Multimedia Appendix 4
Subgroup analysis of interrater agreement on PRHISM and DISCERN. The y-axis displays the initials of the breast experts, divided by experience level (over 10 years and 10 years or less). The circle represents the mean kappa for PRHISM, and the triangle represent the mean kappa for DISCERN. The 95% CIs are represented by horizontal error bars. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.
[PPTX File, 248 KB - infodemiology_v5i1e66416_app4.pptx]

Multimedia Appendix 5
Subgroup analysis of interrater agreement on PRHISM and DISCERN. The y-axis displays the initials of the breast experts, divided by video categories (1–4 and 5–8). The circle represents the mean kappa for PRHISM, and the triangle represent the mean kappa for DISCERN. The 95% CIs are represented by horizontal error bars. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.
[PPTX File, 195 KB - infodemiology_v5i1e66416_app5.pptx]

Multimedia Appendix 6
Intrarater agreement between PRHISM score classification and expert opinion. The y-axis displays the initials of the breast experts who made the assessments. The 95% CIs are represented by horizontal error bars. PRHISM: Principles for Health-Related Information on Social Media.

XSL•FO
RenderX

[PPTX File, 107 KB - infodemiology_v5i1e66416_app6.pptx]

Multimedia Appendix 7

PRHISM and DISCERN Scores for evaluating medical information on Japanese YouTube. Box plots showing the distribution of PRHISM and DISCERN scores for medical information related to breast cancer treatment on Japanese YouTube. The boxes represent the IQR, with the horizontal line indicating the median. The vertical extending lines shows the minimum and maximum values within 1.5 times the IQR. DISCERN: Diagnostic Information Support Communication Evaluation Report Network; PRHISM: Principles for Health-Related Information on Social Media.

[PPTX File, 48 KB - infodemiology_v5i1e66416_app7.pptx]

References

1. Zeraatkar D, Obeda M, Ginsberg JS, Hirsh J. The development and validation of an instrument to measure the quality of health research reports in the lay media. BMC Public Health 2017 Apr 20;17(1):343. [doi: 10.1186/s12889-017-4259-y] [Medline: 28427426]
2. McMullan M. Patients using the internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
3. Telecommunication usage trend survey 2023. Information & Communications Statistics Database. 2023. URL: https://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html [accessed 2024-02-13]
4. Kington RS, Arnesen S, Chou WYS, Curry SJ, Lazer D, Villarruel AM. Identifying credible sources of health information in social media: principles and attributes. NAM Perspect 2021;2021:34611600. [doi: 10.31478/202107a] [Medline: 34611600]
5. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187. [doi: 10.2196/17187] [Medline: 33470931]
6. Lee JJ, Kang KA, Wang MP, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. J Med Internet Res 2020 Nov 13;22(11):e22205. [doi: 10.2196/22205] [Medline: 33048825]
7. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111. [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
8. Blackwood J, Wright FC, Hong NJL, Gagliardi AR. Quality of DCIS information on the internet: a content analysis. Breast Cancer Res Treat 2019 Sep;177(2):295-305. [doi: 10.1007/s10549-019-05315-8] [Medline: 31214858]
9. Iesato A, Fushimi A, Tahara R, et al. A novel system to provide information via online YouTube videos and an evaluation of current online information about hereditary breast cancer. Breast Cancer 2024 Jan;31(1):63-74. [doi: 10.1007/s12282-023-01512-y] [Medline: 37995024]
10. Denniss E, Lindberg R, McNaughton SA. Development of principles for health-related information on social media: Delphi study. J Med Internet Res 2022 Sep 8;24(9):e37337. [doi: 10.2196/37337] [Medline: 36074544]
11. Social media stats Japan. Statcounter Global Stats. 2024. URL: https://gs.statcounter.com/social-media-stats/all/japan [accessed 2025-05-13]
12. Johnson SB, Parsons M, Dorff T, et al. Cancer misinformation and harmful information on Facebook and other social media: a brief report. J Natl Cancer Inst 2022 Jul 11;114(7):1036-1039. [doi: 10.1093/jnci/djab141] [Medline: 34291289]
13. Charnock D. The DISCERN Handbook: Quality Criteria for Consumer Health Information on Treatment Choices University of Oxford and The British Library: Radcliffe Medical Press; 1998.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310] [Medline: 843571]
15. Fleiss JL, Levin B, Paik MC. Statistical Methods for Rates and Proportions, 3rd edition: John Wiley & Sons, Inc; 2003:xxvii-xx760.
16. Aktas BK, Demirel D, Celikkaleli F, et al. YouTube as a source of information on prostatitis: a quality and reliability analysis. Int J Impot Res 2024 May;36(3):242-247. [doi: 10.1038/s41443-023-00666-9] [Medline: 36646834]
17. Çintesun F, Çintesun E, Seçilmiş Ö. YouTube as a source of information on gonadotropin self-injections. Eur J Obstet Gynecol Reprod Biol 2021 Sep;264(135-40):135-140. [doi: 10.1016/j.ejogrb.2021.07.015] [Medline: 34303073]
18. Gul M, Diri MA. YouTube as a source of information about premature ejaculation treatment. J Sex Med 2019 Nov;16(11):1734-1740. [doi: 10.1016/j.jsxm.2019.08.008] [Medline: 31522984]
19. Khatri P, Singh SR, Belani NK, et al. YouTube as source of information on 2019 novel coronavirus outbreak: a cross sectional study of English and Mandarin content. Travel Med Infect Dis 2020 May;35(101636):101636. [doi: 10.1016/j.tmaid.2020.101636]
20. Kidy S, McGoldrick DM, Stockton P. YouTube as a source of information on extraction of third molars. Oral Maxillofac Surg 2021 Dec;25(4):519-524. [doi: 10.1007/s10006-021-00941-3] [Medline: 33608797]

XSL•FO
RenderX

21. Martinez VH, Ojo D, Gutierrez-Naranjo JM, Proffitt M, Hartzler RU. The most popular YouTube videos about shoulder replacement are of poor quality for patient education. Arthrosc Sports Med Rehabil 2023 Jun;5(3):e623-e628. [doi: 10.1016/j.asmr.2023.03.001] [Medline: 37388878]
22. Reddy R, Cheng H, Jufas N, Patel N. Assessing the quality of patient information for cholesteatoma on the video sharing platform YouTube. Otol Neurotol 2023 Apr 1;44(4):e230-e234. [doi: 10.1097/MAO.0000000000003824] [Medline: 36706445]
23. Robillard JM, Jun JH, Lai JA, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. BMC Med Inform Decis Mak 2018 Oct 19;18(1):87. [doi: 10.1186/s12911-018-0668-9] [Medline: 30340488]
24. Kureyama N, Terada M, Kusudo M, et al. Fact-checking cancer information on social media in Japan: retrospective study using Twitter. JMIR Form Res 2023 Sep 6;7:e49452. [doi: 10.2196/49452] [Medline: 37672310]
25. Kinkead L, Allam A, Krauthammer M. AutoDiscern: rating the quality of online health information with hierarchical encoder attention-based neural networks. BMC Med Inform Decis Mak 2020 Jun 9;20(1):104. [doi: 10.1186/s12911-020-01131-z] [Medline: 32517759]

Abbreviations

AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials
DISCERN: Diagnostic Information Support Communication Evaluation Report Network
PRHISM: Principles for Health-Related Information on Social Media

Edited by T Mackey; submitted 19.09.24; peer-reviewed by M Agbede, Z Ehtesham; revised version received 14.03.25; accepted 19.03.25; published 11.06.25.

Please cite as:
Kusama H, Takahashi Y, Orihara S, Adachi K, Ishizuka Y, Semba R, Shima H, Horimoto Y, Kaise H, Taguri M, Inoue S, Nakayama T, Ishikawa T
Assessing the Reliability and Validity of Principles for Health-Related Information on Social Media (PRHISM) for Evaluating Breast Cancer Treatment Videos on YouTube: Instrument Validation Study
JMIR Infodemiology 2025;5:e66416
URL: https://infodemiology.jmir.org/2025/1/e66416
doi:10.2196/66416

© Hiroki Kusama, Yoshimitsu Takahashi, Shunichiro Orihara, Kayo Adachi, Yumiko Ishizuka, Ryoko Semba, Hidetaka Shima, Yoshiya Horimoto, Hiroshi Kaise, Masataka Taguri, Sho Inoue, Takeo Nakayama, Takashi Ishikawa. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 11.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Natural Language Processing Methods to Build the Hypersexuality in Bipolar Reddit Corpus: Infodemiology Study of Reddit

Daisy Harvey[1], PhD; Paul Rayson[2], PhD; Fiona Lobban[1], PhD; Jasper Palmier-Claus[1,3], PhD; Clare Dolman[4], PhD; Anne Chataigné[5], MSc; Steven Jones[1], PhD

[1]Spectrum Centre for Mental Health Research, Division of Health Research, Lancaster University, Lancaster, United Kingdom

[2]School of Computing and Communications, Lancaster University, Lancaster, United Kingdom

[3]Lancashire & South Cumbria NHS Foundation Trust, Lancashire, United Kingdom

[4]School of Mental Health & Psychological Sciences, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[5]Lived Experience Researcher, London, United Kingdom

Corresponding Author:
Daisy Harvey, PhD
Spectrum Centre for Mental Health Research
Division of Health Research
Lancaster University
Bailrigg
Lancaster, LA1 4YW
United Kingdom
Phone: 44 1524 65201
Email: d.harvey4@lancaster.ac.uk

Abstract

Background: Bipolar is a severe mental health condition affecting at least 2% of the global population, with clinical observations suggesting that individuals experiencing elevated mood states, such as mania or hypomania, may have an increased propensity for engaging in risk-taking behaviors, including hypersexuality. Hypersexuality has historically been stigmatized in society and in health care provision, which makes it more difficult for service users to talk about their behaviors. There is a need for greater understanding of hypersexuality to develop better evidence-based treatment, support, and training for health professionals.

Objective: This study aimed to develop and assess effective methodologies for identifying posts on Reddit related to hypersexuality posted by people with a self-reported bipolar diagnosis. Using natural language processing techniques, this research presents a specialized dataset, the Talking About Bipolar on Reddit Corpus (TABoRC). We used various computational tools to filter and categorize posts that mentioned hypersexuality, forming the Hypersexuality in Bipolar Reddit Corpus (HiB-RC). This paper introduces a novel methodology for detecting hypersexuality-related conversations on Reddit and offers both methodological insights and preliminary findings, laying the groundwork for further research in this emerging field.

Methods: A toolbox of computational linguistic methods was used to create the corpora and infer demographic variables for the Redditors in the dataset. The key psychological domains in the corpus were measured using Linguistic Inquiry and Word Count, and a topic model was built using BERTopic to identify salient language clusters. This paper also discusses ethical considerations associated with this type of analysis.

Results: The TABoRC is a corpus of 6,679,485 posts from 5177 Redditors, and the HiB-RC is a corpus totaling 2146 posts from 816 Redditors. The results demonstrate that, between 2012 and 2021, there was a 91.65% average yearly increase in posts in the HiB-RC (SD 119.6%) compared to 48.14% in the TABoRC (SD 51.2%) and an 86.97% average yearly increase in users (SD 93.8%) compared to 27.17% in the TABoRC (SD 38.7%). These statistics suggest that there was an increase in posting activity related to hypersexuality that exceeded the increase in general Reddit use over the same period. Several key psychological domains were identified as significant in the HiB-RC ($P<.001$), including more negative tone, more discussion of sex, and less discussion of wellness compared to the TABoRC. Finally, BERTopic was used to identify 9 key topics from the dataset.

Conclusions: Hypersexuality is an important symptom that is discussed by people with bipolar on Reddit and needs to be systematically recognized as a symptom of this illness. This research demonstrates the utility of a computational linguistic

XSL•FO
RenderX

framework and offers a high-level overview of hypersexuality in bipolar, providing empirical evidence that paves the way for a deeper understanding of hypersexuality from a lived experience perspective.

(JMIR Infodemiology 2025;5:e65632) doi:10.2196/65632

KEYWORDS

bipolar; hypersexuality; natural language processing; Linguistic Inquiry and Word Count; LIWC; BERTopic; topic modeling; computational linguistics

Introduction

Background

Bipolar is a severe mental health condition characterized by recurring episodes of high mood and low mood that is thought to affect at least 2% of the global population [1]. Clinical observations suggest that individuals with bipolar face difficulties regulating emotions and impairments to their cognitive processing, which can contribute to an association with high-risk behaviors [2], and research has demonstrated that these behaviors are often associated with a period of elevated mood [3-5]. Most of the existing research in this area has focused on trying to isolate the biological and behavioral mechanisms that drive risky behavior in people living with bipolar [2,6-14], whereas how these behaviors are exhibited in reality has been comparatively underresearched. Existing research presents a preliminary classification system for the types of risk-taking behavior that people living with bipolar may engage in [3], and through this study, we hope to contribute a more nuanced understanding of one facet of risk-taking behavior, the presentation of hypersexuality, based on large-scale social media data.

This research approaches hypersexuality through the lens of risk-taking behavior and as a symptom of bipolar, focusing on its potential to harm personal safety. However, hypersexuality is a complex concept lacking a universal definition and is shaped by cultural, individual, and situational factors. Perrotta [15] describes it as "a psychological and behavioural alteration as a result of which sexually motivated stimuli are sought in inappropriate ways and often experienced in a way that is not completely satisfactory" and further highlights that hypersexuality is challenging to diagnose due to the lack of established criteria and the impracticality of rigid diagnostic standards in addressing the subjective emotional universe of individuals. Walton et al [16] emphasize that diagnosing hypersexuality requires observable symptoms, subjective perceptions, adverse consequences, and distress. While it is included in the *International Classification of Diseases, 11th Revision*, as compulsive sexual behavior, the rejection of hypersexuality as a distinct diagnosis from the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, underscores ongoing debates about its classification, reflecting concerns about stigmatization and definitional challenges. The term *hypersexuality* may be used by some individuals to articulate personal experiences without negative consequences, and while these self-descriptions may not align with the definition adopted in this paper, they represent meaningful aspects of lived experience.

There are only a limited number of studies that have focused on the topic of hypersexuality and sexual risk taking in bipolar, and the literature on hypersexuality is sparse and not systematically defined [4,17-19]. Krantz et al [5] found that hypomania often precedes risky sexual behavior, with two-thirds of sexually active youth with bipolar engaging in behaviors categorized as above minimal risk and one-third reporting pregnancy, and Mazza et al [19] observed increased sexual interest in women with bipolar type I compared to bipolar type II. Raja and Azzoni [20] noted high awareness of sexually transmitted infection risks but prevalent risky sexual behaviors among individuals with bipolar, schizophrenia, or schizoaffective disorder, and Marengo et al [21,22] found a link between unplanned pregnancies and hypersexuality in manic episodes, also finding higher rates of casual and nonmonogamous sex among women with bipolar, including during euthymia. Krogh et al [4] explored the impact of mood swings on sexuality in bipolar through qualitative interviews, identifying 5 key themes: sexual drive, behavior, thoughts, intimate relationships, and identity. Their results suggest that elevated mood states increased sexual drive and interactions and that mood-related shifts had significant relational impacts. Observing the existing literature critically, a number of studies that have investigated hypersexuality in bipolar are >30 years old [23-25], making it "subject to the biases of sexual and gender norms" of those times [17]. There is also evidence of stigma attached to hypersexuality and the discussion of sexual experiences from health care professionals [26], as well as a lack of qualitative research into the sexual behaviors of people living with bipolar [4,27].

In this paper, we present a toolbox of computational linguistic techniques, including pretrained machine learning models for demographic inference, the extraction of key psychological domains using the 2022 version of Linguistic Inquiry and Word Count (LIWC-22; Pennebaker Conglomerates, Inc) [28], and unsupervised topic modeling using BERTopic [29], to provide an understanding of what kinds of topics are talked about in discussions regarding hypersexuality. This is the first study to use such methods on data that relate to hypersexuality in general and specifically to bipolar and demonstrates the utility of large-scale language analysis in health research. We acknowledge that there are serious ethical implications associated with the collection of such sensitive information but believe that the benefit of improved understanding and awareness that can be obtained using Reddit (Reddit, Inc) posts is of significant value to people who experience the symptom of hypersexuality as part of their diagnosis of bipolar. We provide a comprehensive outline of our ethical considerations, including consultation with lived experience experts, in the Methods section.

XSL•FO
RenderX

This research aimed to form the foundation for future work in the area by developing a dataset of qualitative information, addressing a significant gap in the field, and presenting key themes. The objective was not to provide an exhaustive analysis of all posts in the dataset as this lies beyond the scope of this study. Instead, the focus of this study was on the methodology used to construct the corpus and on foregrounding this topic as a critical area of scientific interest. We hope that this study supports calls for novel research to "address sexual symptomatology in bipolar within the context of current sexual, cultural, and gender norms" [27]. Our research questions are defined in the following section.

Research Questions

The research questions for this study were as follows:

- Is hypersexuality talked about on Reddit?
    - How can we recognize Redditors who post about hypersexuality on Reddit?
    - What are these Redditors' posting behaviors?
- How can computational linguistic methods be used for exploratory analysis of the Hypersexuality in Bipolar Reddit Corpus (HiB-RC)? This includes the following:
    - Psychological domains
    - Topic modeling

Methods

The Talking About Bipolar on Reddit Corpus

Application Programming Interface Data Collection

The posts in this dataset were collected using the Pushshift and PRAW application programming interfaces (APIs) in July 2022 through adaptation of existing code [30]. The 2 subreddits related to bipolar with the highest number of followers—r/bipolar and r/BipolarReddit (approximately 300,000 users)—were scraped to include data posted between July 2017 and July 2022. Applying a similar framework to those in the studies by Coppersmith et al [31], Sekulic et al [32], Cohan et al [33], and Jagfeld et al [34], we then used pattern-matching methods on this corpus to detect Redditors who self-reported a clinical diagnosis of bipolar using a framework implemented by Jagfeld et al [34,35]. We adapted this framework to identify self-reported diagnosis patterns from Reddit posts and comments that (1) contained at least one condition term for bipolar, (2) matched at least one inclusion pattern (ie, bipolar diagnosis of any type by a professional), and (3) did not match any exclusion pattern (eg, self-diagnosis).

After identifying posts from Redditors who had self-reported a diagnosis, we then collected the entire posting history for these users across all subreddits using a custom Python script (Python Software Foundation). This script collected the following information for each comment or submission made by a user: (1) post ID, (2) text body, (3) username, (4) subreddit, (5) post title (for main submissions and not for comments), and (6) time stamp.

We note that there are limitations to using self-reported diagnoses as these have not been clinically verified within the dataset.

Demographic Inference

Overview

To develop a more comprehensive understanding of the Redditors in our dataset, we used a number of methods for demographic inference (age, gender, and location) presented originally in the work by Jagfeld et al [34], Tigunova et al [36], and Harrigian [37]. While we acknowledge that these methods do not necessarily implement state-of-the-art technologies such as large language models, they are to date the only publicly available models for this type of demographic inference within the Reddit domain. Ethical considerations associated with using inference models are presented in the Ethical Considerations section.

Age and Gender

First, we manually identified self-reported instances of age and gender using the pattern-matching code provided in the work by Jagfeld et al [34]. These patterns identify self-reported instances of age and gender from submission titles, which are captured between square brackets as is typical notation on Reddit, for example, "I {28f} am posting here for the first time." Age was calculated using a function that estimates date of birth based on the age provided in the submission title compared to the submission posting date. Labels for gender were assigned using manual extraction for 675 users, and labels for age were assigned using manual extraction for 643 users. We then used pretrained models to determine age and gender for the remaining users in the dataset for whom a self-reported age or gender could not be determined. The pretrained models used for automated age and gender inference were developed by Tigunova et al [36], who presented a hidden attribute model using a convolutional neural network with attention mechanism architecture to develop representations of demographic information based on language use. The models were trained on similar domain data using the posts from >350,000 Redditors included in the RedDust dataset [36]. The reported accuracies for the age and gender algorithms are an area under the receiver operating characteristic curve of 0.88 for age and an area under the receiver operating characteristic curve of 0.91 and accuracy of 0.86 for gender [36]. Using a subset of gold truth labels that were manually extracted from the dataset for age and gender (675 users for gender and 643 users for age), we manually calculated a weighted F_1 accuracy of 0.8 for gender and 0.6 for age for our dataset. The text used as input to the models was preprocessed before being used as input, which involved cleaning the data to remove hyperlinks and non–English-language words and converting the text to the vector representation format expected by the model (adapting the scripts provided by Tigunova et al [36,38]). Both submissions and comments were used as input to the model provided that the content was between 10 and 100 words in length and that users had at least 10 posts that matched these criteria and using only the most recent 100 posts for each Redditor as input. The inference methods for gender that were used in this study were designed only to detect binary genders

XSL•FO
RenderX

(man and woman), the implications of which are discussed further in the Discussion section.

Geolocation

We used a pretrained model presented by Harrigian [37] to infer location identifiers for each user in the dataset at the country level. This model was trained using the distribution of words, posts per subreddit, and posts per hour of the day for Reddit users. When applying this model to our data, we included only users with >50 posts and up to 250 posts as specified in the documentation for the package to improve the accuracy of predictions [39]. The global model provided by Harrigian [37] was used, which achieves 35.6% accuracy, and as reported by Jagfeld et al [34], the accuracy is generally higher for users with more training data (95.1% for the United States, 65.1% for Canada, 82.8% for the United Kingdom, 44.1% for Australia, and 41.1% for Germany).

Developing the HiB-RC

After implementing the inference models, any users whose posting history did not satisfy the criteria for the pretrained models were removed from the dataset. This resulted in a snapshot corpus that contains data that span 13 years, with the earliest post dating back to June 2009 and the latest submission date in August 2022.

To detect posts with content related to hypersexuality, we created an initial set of seed terms to generate a subcorpus (the HiB-RC) of users with a self-reported history of hypersexuality. To develop this vocabulary of seed terms, we identified the keywords and phrases related to hypersexuality from a previous study that used lived experience interview data [3] and trained both word2vec (Google AI) [40] and fastText (Facebook's Artificial Intelligence Research laboratory) [41] embedding models on the Talking About Bipolar on Reddit Corpus (TABoRC) to find synonyms (words and phrases) and misspellings of these keywords and phrases. The fastText algorithm produces character-level embeddings that find numeric representations of words by looking at their character-level compositions, thus enabling us to detect common typographical errors for the hypersexuality seed terms. Traditional word- and character-level embeddings were deemed to be sufficient for this task as the embeddings were not being used as part of a predictive algorithm and, thus, there was a cost benefit in terms of lower computational and environmental cost for training these simpler models versus fine-tuning a contextual large language model. The final list of seed terms used to collect posts related to hypersexuality is presented in Textbox 1.

Textbox 1. Hypersexuality keywords used to create the Hypersexuality in Bipolar Reddit Corpus. These keywords were generated by finding the most similar terms to the input keywords using word2vec (Google AI) and fastText (Facebook's Artificial Intelligence Research laboratory) embeddings trained on the Talking About Bipolar on Reddit Corpus.

Input keyword to the word2vec and fastText models

- "Hypersexual"
- "Hypersexuality"
- "Hyper-sexual"
- "Hyper_sexual"

Output—*most similar* keywords

- "Hypersexual"
- "Hyper sexual"
- "Hypersexuallity"
- "Hypersex"
- "Hyper sexualised"
- "Hyper sexuality"
- "Oversexual"
- "Hyposexual"
- "Hyper sexualized"
- "Hypersexualized"
- "Overly sexual"
- "Hyper sexualization"
- "Hypersexualization"
- "Hyposexuality"
- "Hypersexuality"

XSL•FO
RenderX

At the early stages of data collection, we used a much longer list of seed terms to search for posts related to hypersexuality, including phrases such as “hook up with strangers,” “high sex drive,” and “threesomes.” This list of vocabulary was generated using the same word embedding methodology but included a more diverse set of keywords as input when using the models to search for similar words and phrases. This resulted in a much noisier dataset where it was apparent after manual inspection that a large number of the posts were not written in the context of experiencing hypersexuality as a symptom but rather in the context of people sharing and discussing sexual experiences. Due to the infancy of this field of work and to avoid compounding the stigma regarding sex or incorrectly categorizing diverse sexual experiences as hypersexuality, we chose to refine the keyword list used as input to the word embedding models to words and phrases that directly related to the notion of “hypersexuality.” We considered it more ethical to collect data from instances in which individuals self-reported the symptom of hypersexuality rather than inferring hypersexuality through more nuanced descriptions of sexual behavior. The result was that there was less ambiguity and greater reliability in the dataset of posts, with the disadvantage that we filtered out an unknown amount of data related to hypersexuality that talked about the topic in more nuanced ways. We refer in this paper to the concept of a corpus being “acceptably representative,” whereby “we have to make do with studying merely a sample of the language use, or variety, as a whole” due to restrictions on time and resources and, in this case, ethical considerations [42].

After we had generated the final seed list of hypersexuality terms, we created a filter and applied this to the TABoRC. After preprocessing the returned posts to remove duplicates and only include posts that were >30 words in length, we manually annotated this dataset using the doccano tool to verify a post’s inclusion in the corpus, with the posts annotated as confirming a hypersexuality report forming the HiB-RC. The corpus was annotated in full by DH, and circa 10% of the corpus (300 posts) was annotated by second and third annotators (SJ and PR). Interannotator agreement achieved a Krippendorff α score of 0.77 [43], and majority voting was used to solve annotator disagreements. Disagreements primarily occurred in cases in which an experience of hypersexuality was described but there was ambiguity on whether the author of the post was the one who had experienced the symptom. The annotation guidelines are presented in Multimedia Appendix 1.

Analysis Methodology

Interpreting the HiB-RC

To begin the exploratory analysis of our dataset, we produced descriptive statistics to detail the user and posting characteristics of the corpus. These analyses were conducted using Python, and the results are presented in the Results section to show demographic characteristics, the number of new users posting in the HiB-RC each year and the number of new posts referencing hypersexuality each year (using the TABoRC as a comparison dataset), and the top subreddits to which posts about hypersexuality were posted.

Linguistic Inquiry and Word Count

After exploring the Redditor characteristics of our dataset, we used LIWC-22 [28] to understand the key psychological domains within the HiB-RC.

LIWC-22 is a text analysis application that maps psychosocial constructs to words, phrases, and linguistic constructions [28]. Linguistic Inquiry and Word Count (LIWC) processes text using software and a dictionary, where the dictionary contains groups of words that relate to a particular domain (eg, positive or negative tone). Documents of interest (the input text) are analyzed by the software to map the domains to the text, calculating the percentage of each document that comprises words in these dictionary domains. LIWC was designed on the premise that the words that people use tell us about “their psychological states: their beliefs, emotions, thinking habits, lived experiences, social relationships, and personalities” [28]. The LIWC-22 dictionary is based on >12,000 words, phrases, and emoticons, and the authors describe that “in the advent of more powerful analytic methods and more diverse language samples, we have been able to build more internally consistent language dictionaries with enhanced psychometric properties” in this latest release of the software [28]. Modern text analysis has been influenced by >100 years of psychological research [44], and previous research has demonstrated how language analysis can provide insights into cognitive mechanisms, with “an increasing number of studies [which] demonstrate, [that] the ways in which people use words is reliable over time” [45].

LIWC domains have been used in various existing studies that explore how language is used by people living with bipolar, including as input for prediction and classification models [31-33,46-53] and exploratory analysis of mental health datasets [54,55]. In this research, we used LIWC to identify psychological domains that appear significantly more or less by comparing the HiB-RC to a control corpus formed of the same users’ entire posting history across Reddit.

Modeling Hypersexuality

Egger and Yu [56] describe that social media data have opened up new pathways for scientific research but that the short and unstructured nature of the documents within social media datasets can cause methodological issues for analysis. The authors describe that topic modeling has increasingly been applied to the topic of social science, where topic models are defined as “probabilistic models for uncovering the underlying semantic structure of a document collection” [57].

Topic models seek to identify patterns between similar documents to add structure to an otherwise unstructured collection of text to facilitate exploration and understanding. Latent Dirichlet allocation (LDA) is one of the most widely used traditional methods for topic modeling and is a generative statistical model introduced by Blei et al [58]. Despite the popularity of LDA, the reliability and validity of the results have been criticized because there is no definitive method of model evaluation and there is a lack of guidance related to fine-tuning. The efficacy of LDA for analyzing social media data has been further criticized because the noisy and sparse

XSL•FO
RenderX

datasets generated in social science research often do not contain enough features for statistical learning [56].

More recent topic-modeling algorithms that have been implemented as an alternative to LDA [56] include embedding models [29,59] that rely on the vectorization of text data to locate semantically similar words and documents. BERTopic [29] is an algorithm that uses pretrained embedding models to create word and document embeddings so that documents that occupy similar vector space can be grouped together to form topics. By default, BERTopic incorporates Bidirectional Encoder Representations From Transformers embeddings and a term frequency–inverse document frequency algorithm, which compares the importance of terms within a cluster and creates term representation based on this [60]. This means that the higher the value is for a term, the more representative it is of its topic. Due to the sparse nature of social media data, BERTopic also includes a default module for dimension reduction using uniform manifold approximation and projection, which enables these dimensions to be reduced to the extent that hierarchical density-based spatial clustering of applications with noise can be used to identify dense regions in the documents [56,59].

On the basis of the comparison of topic-modeling methods presented in the work by Egger and Yu [56], BERTopic presents a number of advantages that influenced our decision to use this method in our research. These include its ability to perform well across multiple domains due to the use of pretrained embeddings and the fact that little to no preprocessing of text is required before training. There still remain limitations, which are described in the Discussion section.

BERTopic Setup

BERTopic was adapted for this study from the code provided by Grootendorst [61]. The parameters that had a significant impact on the topic output included the following:

First, KeyBERTInspired as the main representation input to the model. KeyBERTInspired [62] extracts representative keywords for topics using word embeddings, ensuring more context-aware representations. First, document embeddings are generated to capture the overall meaning of a document. Word embeddings are then created for N-gram words and phrases. Finally, cosine similarity is used to identify the words and phrases that are most similar to the document embedding. Textbox 2 shows the difference in representations produced using the default term frequency–inverse document frequency and KeyBERTInspired representation models.

Second, the use of *mxbai-embed-large-v1* sentence embeddings [63] as the pretrained embeddings for the model, which demonstrate very high performance for low memory use (ranked 13 in the Massive Text Embedding Benchmark leaderboard at the time of writing). We also tested topic generation using MentalBERT embeddings that have been trained on Reddit data within the mental health domain, but the resulting topic representations were less defined and noisier [64].

Third, a custom list of stop words were provided to the CountVectorizer module and, thus, excluded from clusters after training. This list included generic English stop words (eg, “and,” “or,” “this,” and “was”) as well as frequently occurring words such as “hypersex*” and “bipolar”—keywords that appeared in nearly every post due to the seed list of vocabulary used to generate the corpus or the topic domain.

Textbox 2. Default versus KeyBERTInspired representation of the example topics generated by BERTopic.

Default representation

- “Ve,” “manic,” “feel,” “really,” “don,” “mania,” “time,” “people,” “sleep,” and “know”
- “Age,” “years,” “sexual,” “older,” “csa,” “remember,” “trauma,” “know,” “young,” and “happened”

KeyBERTInspired representation

- “Hypomanic,” “manic,” “mania,” “depressed,” “depressive,” “depression,” “disorder,” “psychiatrist,” and “mood”
- “Abuser,” “abused,” “abuse,” “sexual,” “trauma,” “memories,” “rape,” “therapy,” “touched,” and “older”

After our model setup had been finalized, we manually merged similar topics after inspecting the posts included within each topic using the *merge_topic()* method of the model. Finally, we manually assigned topic labels for our topics to be used in visualizations and saved the model as a pickle file for future analysis. As noted when describing the limitations of BERTopic, the topics produced by the model may change each time the model is run. After altering the parameters of the model, implementing *mxbai-embed-large-v1* as the sentence embedding model, and using KeyBERTInspired as the main representation model, we found the generation of topics to be relatively stable with each iteration.

Ethical Considerations

We recognize the importance of developing an ethical framework when working with sensitive data that describe personal lived experience, especially when collecting data from a public site such as Reddit. We outline in this section our considerations regarding consent, anonymization, the right to be forgotten, and dataset retention. Our framework was informed by multiple sources, including institutional resources from the British Psychological Society, the British Sociological Association, and the UK government [65-67] as well as sources from academic research and guidelines [34,68-72]. This study was conducted as part of a PhD thesis on the topic of risk-taking behaviors in bipolar, and we consulted a panel of lived experience advisors through Lancaster University Spectrum Connect at the early stages of design. We also engaged with

XSL•FO
RenderX

Bipolar UK on a webinar on hypersexuality in 2024 [73] and sought invaluable guidance from lived experience researchers who coauthored this paper. Ethics approval was granted for the project by Lancaster University in December 2021 (FHMREC21042).

Reddit is colloquially known as "the front page of the internet," with >50 million daily users and 100,000 *active* subreddits in 2024 [74,75], and research has shown that the anonymity afforded by social media sites enables users to self-disclose on sensitive topics that they may otherwise find difficult to talk about [76]. As researchers, we wholly acknowledge that the Reddit posts used in our study contain sensitive information and that the forum users were not aware that their discussions would be used for research. We did not seek informed consent from the Redditors whose posts we collected due to the impractical nature of this task considering that the posts of >5000 Redditors were included in the TABoRC, but we note that Reddit users are made aware that their posts are publicly accessible through Reddit's terms and conditions. From a legal perspective, although Reddit is by nature an anonymous platform, we cannot know that Redditors do not use the same username across other social media sites or platforms, and therefore, we treat the information collected from the site as personal data. In accordance with the Data Protection Act 2018 and General Data Protection Regulation, an exemption for conducting research for "special purposes" would be relevant for nonconsent as we intend to publish our research and are confident that the publication of any research associated with the collection of these Reddit data "would be in the public interest" [67]. Further to the legal grounding of work conducted in the public interest, the motivation for this study was to learn more about experiences of a typically stigmatized symptom to identify how people experiencing hypersexuality could be better supported. There is existing evidence from lived experience suggesting that data on this topic can be difficult to access within a health care setting, so we acknowledge the limitations of using data sourced from the web but also recognize the unique insights that the analysis of such data can provide [3,27,77].

Following previous guidance [65,68,69], as we did not rely on consent for this study, we masked the usernames in this dataset (created alternative alphanumeric usernames for each Redditor in the dataset) and have only included paraphrased and depersonalized quotes in research outputs. We have also minimized the amount of qualitative data reported by using computational methods such as topic modeling and LIWC, which enable us to present key themes and insights from the data in an aggregate format without needing to rely heavily on quotes. Where we presented paraphrased quotes, we verified that Redditors could not be reidentified based on an internet search of the reworded quotes. Using these methods, we strived to maintain the privacy of the Redditors included in our corpus as much as possible.

We would also like to draw attention to the demographic inference methods that we used. Performing inference of such data enables us to offer predicted demographic information about the study population, which may allow for comparison to other domains, for example, clinical populations. Reporting on aspects such as gender also contributes toward more ethical natural language processing data collection as these predictions can suggest how experimental results might be generalized and also highlights where the data include bias [78]. However, inferring demographic information adds an extra level of personal data to the corpus, and we acknowledge that this comes with its own risks. The demographic data that we inferred are not intended to be used for identification or targeting of users in any way, and we understand that these inferred statistics are not 100% accurate, nor have they been used as features in any predictive models. The demographic data were only reported in aggregate format and will not be publicly released, although the code used is available open source. We would also like to strongly emphasize that any analysis reported using the demographic data indicates correlation and not causality.

Using Reddit as a primary data source is not "wholly problematic or must be ceased," but "careful handling and anonymization of such materials is of paramount importance for maximising ethical research practice going forward" [71]. We have decided to only publish redacted versions of both the TABoRC and HiB-RC with the UK Data Service, as requested by the funder of this research (the Economic and Social Research Council). The redacted versions of the datasets will include only the IDs for the posts that form the corpora. The corpora will be disseminated upon request on a case-by-case basis to researchers with an institutional email address, and future researchers will be required to access the content of the posts using an API. This complies with Article 17 of the UK General Data Protection Regulation and an individual's rights to data erasure because any content that has been removed since the creation of our datasets will appear as "[removed]" upon retrieving the post ID using an API.

Results

Posting Characteristics on Reddit

The TABoRC comprises 6,679,485 posts from 5177 users, and the HiB-RC comprises 2146 posts from 816 users. The demographic statistics for the TABoRC and HiB-RC corpora are presented in Table 1. The data suggest that >15% (816/5177, 15.76%) of the users in the TABoRC reported experiences of hypersexuality.

XSL•FO
RenderX

Table 1. Demographic information for the Hypersexuality in Bipolar Reddit Corpus (HiB-RC), the Talking About Bipolar on Reddit Corpus (TABoRC), and the benchmarking dataset [34].

	Proportion of users
TABoRC (n=5177), n (%)	
Age (y)[a]	
14-23 (teenagers and young adults)	1385 (26.8)
24-45 (adults)	3371 (65.1)
46-65 (middle-aged adults)	389 (7.5)
66-100 (older adults)	32 (0.6)
Gender	
Female	3668 (70.8)
Male	1509 (29.1)
Country	
United States	3970 (76.7)
United Kingdom	366 (7.1)
Canada	337 (6.5)
Germany	108 (2.1)
Australia	100 (1.9)
Sweden	58 (1.1)
Other countries	238 (4.6)
HiB-RC (n=816), n (%)	
Age (y)[a]	
14-23 (teenagers and young adults)	207 (25.4)
24-45 (adults)	531 (65.1)
46-65 (middle-aged adults)	74 (9.1)
66-100 (older adults)	4 (0.5)
Gender	
Female	626 (76.7)
Male	190 (23.3)
Country	
United States	600 (73.5)
United Kingdom	62 (7.6)
Canada	61 (7.5)
Germany	21 (2.6)
Australia	24 (2.9)
Sweden	12 (1.5)
Other countries	36 (4.4)
Benchmarking dataset [1][b], %	
Mean age (y)	
13-17	16.1
18-29	29.8
30-49	47.5
50-64	6.6
≥65	0

XSL•FO
RenderX

	Proportion of users
Gender	
Female	52.2
Male	47.8
Country	
United States	81.9
United Kingdom	5.6
Canada	4.9
Germany	1.4
Australia	1.7
Sweden	—[c]
Other countries	4.5

[a]The pretrained model [2] included an additional age category of 0 to 13 years (child). For any users who were manually or automatically included within this age group, we removed their data from the dataset as Reddit requires a minimum sign-up age of 13 years.

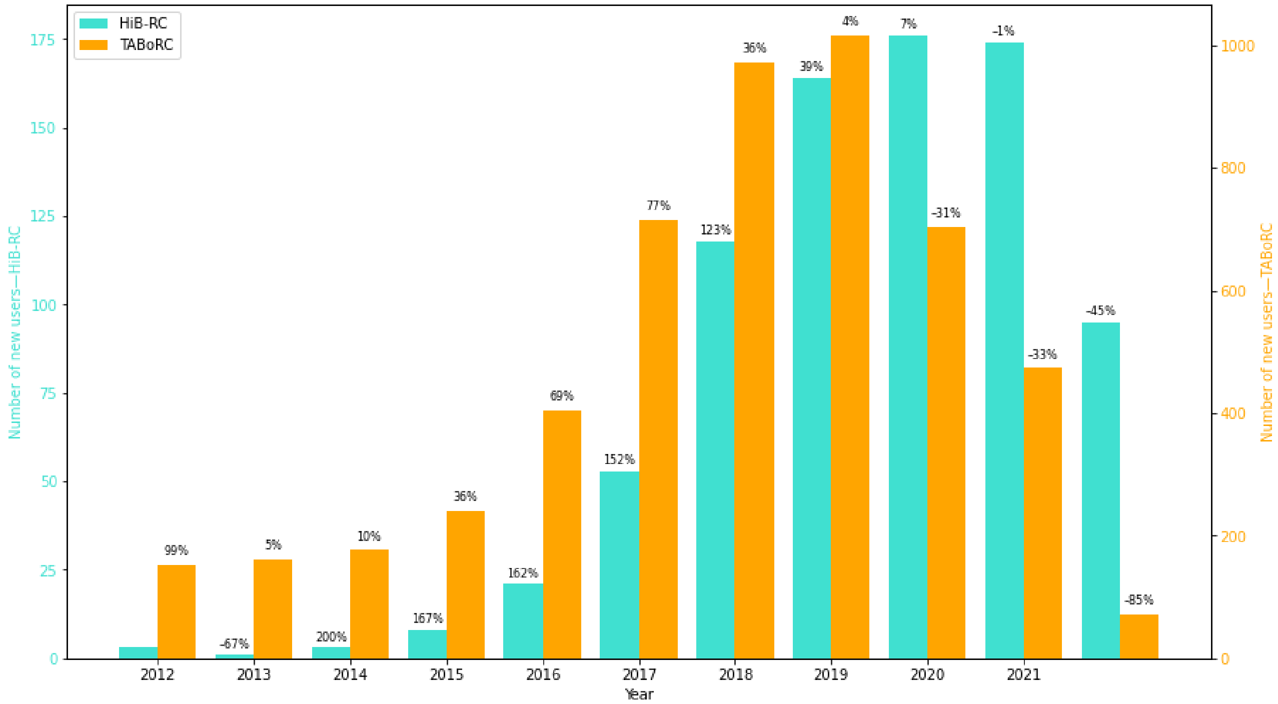
[b]Original data values were not provided with the dataset, so we have only presented percentages in this section.

[c]Not available.

Figure 1 compares the number of new users between 2012 and 2021 in the HiB-RC and the TABoRC, with an average yearly increase of 86.97% (SD 93.8%) and 27.17% (SD 38.7%), respectively. Figure 2 compares the number of new posts between 2012 and 2021 in the HiB-RC and the TABoRC, with an average yearly increase of 91.65% (SD 119.6%) and 48.14% (SD 51.2%), respectively. The bars represent the raw number of posts and the labels demonstrate the yearly percentage increase compared to the previous year. Table 2 shows how many posts that reference hypersexuality are made by each user.

Table 3 shows the top subreddits where posts related to hypersexuality were made within the HiB-RC (where >5 posts were made to the same subreddit), with the most visited subreddits including r/bipolar, r/BipolarReddit, r/bipolar2, r/AskReddit, and r/BipolarSOs.

Figure 1. Comparing the number of new users each year in the Hypersexuality in Bipolar Reddit Corpus (HiB-RC) and the Talking About Bipolar on Reddit Corpus (TABoRC). There is no percentage increase reported for the HiB-RC in 2012 because the first post in the HiB-RC was reported in 2012. Data collection ended in July 2022, so the observed trend in user growth may not fully reflect subsequent changes.



XSL•FO
RenderX

Figure 2. Comparing the number of new posts each year in the Hypersexuality in Bipolar Reddit Corpus (HiB-RC) and the Talking About Bipolar on Reddit Corpus (TABoRC). There is no percentage increase reported for the HiB-RC in 2012 because the first post in the HiB-RC was reported in 2012. Data collection ended in July 2022, so the observed trend in post growth may not fully reflect subsequent changes.

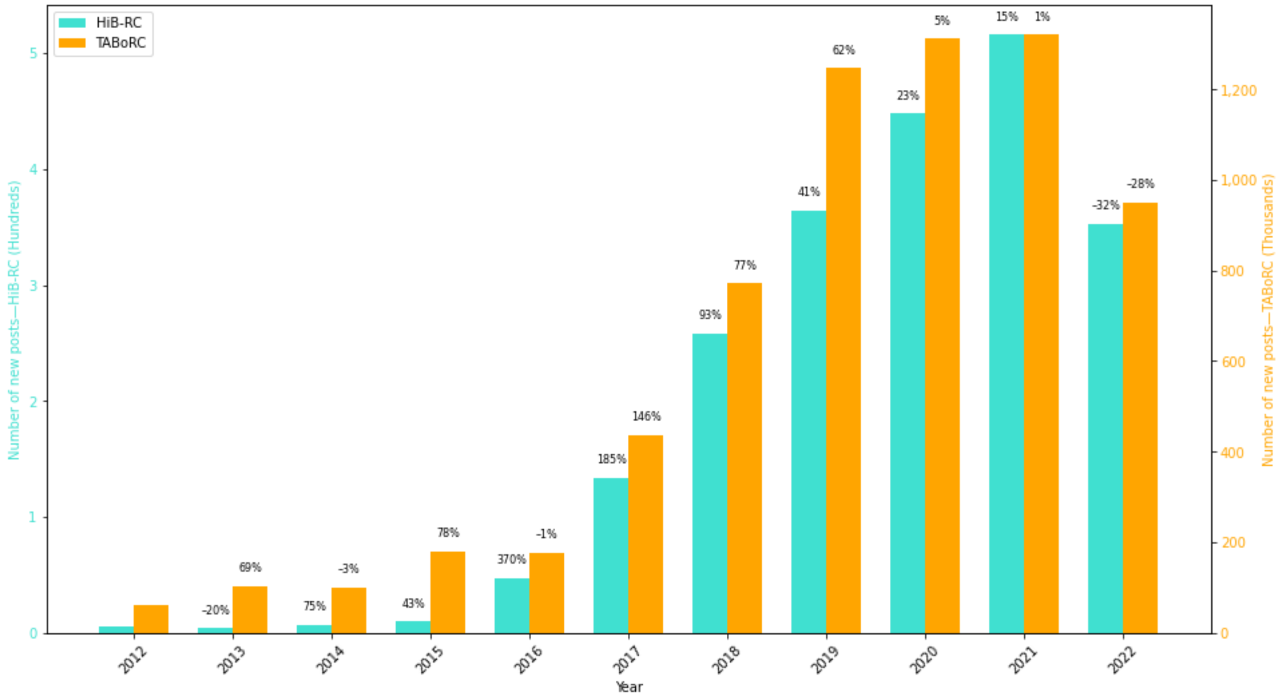


Table 2. Number of posts per user referencing hypersexuality (N=816).

Number of posts per user referencing hypersexuality	Users, n (%)
1	453 (55.5)
≥1 to <5	270 (33.1)
≥5 to ≤10	65 (8)
>10	28 (3.4)

XSL•FO
RenderX

Table 3. Top subreddits for posts related to hypersexuality (where >5 posts were made to the same subreddit; N=2146).

Subreddit	Posts, n (%)
r/bipolar	1027 (47.86)
r/BipolarReddit	421 (19.62)
r/bipolar2	169 (7.88)
r/AskReddit	53 (2.47)
r/BipolarSOs	43 (2)
r/polyamory	28 (1.3)
r/BPD	28 (1.3)
r/hypersexuality	26 (1.21)
r/sex	16 (0.75)
r/adultsurvivors	13 (0.61)
r/ADHD	11 (0.51)
r/BDSMAdvice	10 (0.47)
r/CPTSD	10 (0.47)
r/relationship_advice	9 (0.42)
r/AskRedditAfterDark	9 (0.42)
r/demisexuality	8 (0.37)
r/relationships	7 (0.33)
r/AskMen	6 (0.28)
r/BorderlinePDisorder	6 (0.28)
r/depression	6 (0.28)
r/mentalillness	6 (0.28)

LIWC Results

Table 4 presents a selection of LIWC domains that were statistically significant when comparing the HiB-RC to a control corpus from the same users. The control corpus contains all posting history from each user in the HiB-RC across Reddit after removing the posts that are included in the HiB-RC. The total word count of the HiB-RC is 344,786, and the total word count of the control corpus is 69,495,570. We built the control corpus based on the hypothesis that these data would be representative of more general language use across Reddit by the same group of users based on manual inspection of a sample of the data. After identifying a nonnormal distribution in most LIWC domains based on paired scores using the Shapiro-Wilk test [79], we determined statistical significance using a paired Wilcoxon signed rank test [80] to identify significant differences in domain scores between the control and hypersexuality corpora. All domains included in Table 4 are significant at a *P* value of <.001. The table presents the Wilcoxon score and associated *P* value together with the effect size (Cohen *d*, with directionality represented by the minus sign [–]), which ranges between small (0.01 to 0.2) and huge (≥2) [81]. The methodology for the LIWC analysis was adapted from the work by Cohan et al [33].

XSL•FO
RenderX

Table 4. Significant Linguistic Inquiry and Word Count domains in the Hypersexuality in Bipolar Reddit Corpus (HiB-RC) compared to a control corpus of Reddit posts from the same set of users.

Domain	Description or most frequently used exemplars (from LIWC-22[a] dictionary)	Direction of significance[b]	Wilcoxon signed rank score	*P* value	Cohen *d*
Linguistic dimensions					
First person singular	"I," "me," "my," and "myself"	Positive	34,402.0	<.001	0.37
First person plural	"We," "our," "us," and "lets"	Negative	66,744.0	<.001	−1.14
Second person	"You," "your," "u," and "yourself"	Negative	52,244.5	<.001	−0.55
Third person singular	"He," "she," "her," and "his"	Negative	71,742.5	<.001	−0.55
Third person plural	"They," "their," "them," and "themsel*"	Negative	49,597.5	<.001	−1.57
Psychological processes					
Achievement	"Work," "better," "best," and "working"	Negative	91,466.0	<.001	−0.61
Power	"Own," "order," "allow," and "power"	Negative	111,908.5	<.001	−0.34
Cognition	"Is," "was," "but," and "are	Positive	126,646.5	<.001	0.09
Cognitive processes	"But," "not," "if," "or," and "know"	Positive	126,921.0	<.001	0.09
Insight	"Know," "how," "think," and "feel"	Positive	138,386.0	<.001	0.17
Positive tone	"Good," "well," "new," and "love"	Negative	95,852.5	<.001	−0.36
Negative tone	"Bad," "wrong," "too much," and "hate"	Positive	119,137.5	<.001	0.27
Emotion	"Good," "love," "happy," and "hope"	Positive	132,424.5	<.001	0.24
Positive emotion	"Good," "love," "happy," and "hope"	Negative	131,386.0	<.001	−0.12
Negative emotion	"Bad," "hate," "hurt," and "tired"	Positive	30,310.0	<.001	0.34
Social behavior	"Said," "love," "say," and "care"	Negative	121,529.5	<.001	−0.16
Prosocial behavior	"Care," "help," "thank," and "please"	Negative	107,645.5	<.001	−0.22
Politeness	"Thank," "please," "thanks," and "good morning"	Negative	64,811.0	<.001	−1.63
Communication	"Said," "say," "tell," and "thank*"	Negative	105,069.0	<.001	−0.42
Social referents	"You," "we," "he," and "she"	Negative	46,417.5	<.001	−0.39
Family	"Parent*," "mother*," "father*," and "baby"	Negative	98,628.5	<.001	−0.31
Female references	"She," "her," "girl," and "woman"	Negative	84,008.5	<.001	−0.37
Male references	"He," "his," "him," and "man"	Negative	96,669.5	<.001	−0.29
Expanded LIWC-22 dictionary					
Lifestyle	"Work," "home," "school," and "working"	Negative	53,011.0	<.001	−0.69
Leisure	"Game*," "fun," "play," and "party*"	Negative	82,334.0	<.001	−0.74
Home	"Home," "house," "room," and "bed"	Negative	66,942.5	<.001	−1.52
Work	"Work," "school," "working," and "class"	Negative	57,181.0	<.001	−0.96
Money	"Business*," "pay*," "price*," and "market*"	Negative	94,900.5	<.001	−0.51
Religion	"God," "hell," "christmas*," and "church"	Negative	78,149.5	<.001	−0.47
Physical	"Medic*," "food*," "patients," and "eye*"	Positive	64,808.5	<.001	0.38
Health	"Medic*," "patients," "physician*," and "health"	Positive	97,079.0	<.001	0.31
Wellness	"Healthy," "gym*," "supported," and "diet"	Negative	50,662.5	<.001	−2.35
Mental health	"Mental health," "depressed," "suicid*," and "trauma*"	Positive	73,266.5	<.001	0.58
Substances	"Beer*," "wine," "drunk," and "cigar*"	Negative	73,783.0	<.001	−0.29

XSL•FO
RenderX

Domain	Description or most frequently used exemplars (from LIWC-22[a] dictionary)	Direction of significance[b]	Wilcoxon signed rank score	*P* value	Cohen *d*
Sexual	"Sex," "gay," "pregnan*," and "dick"	Positive	40,559.5	<.001	0.78
Reward	"Opportun*," "win," "gain*," and "benefit*"	Negative	52,059.0	<.001	−2.45
Time	"When," "now," "then," and "day"	Positive	106,340.5	<.001	0.22
Past focus	"Was," "had," "were," and "been"	Positive	125,182.5	<.001	0.14
Future focus	"Will," "going to," "have to," and "may"	Negative	72,929.0	<.001	−0.90

[a]LIWC-22: 2022 version of Linguistic Inquiry and Word Count

[b]Positive direction indicates that the domain is more prevalent in the HiB-RC than the control corpus. Negative direction indicates that the domain is less prevalent in the HiB-RC than the control corpus.

BERTopic Results

Our implementation of BERTopic initially yielded 14 topics and 1 outlier class (which contained posts that were determined to be too noisy to accurately cluster into one of the topics by the algorithm). After manual analysis of these topics, we merged a number of similar clusters using the inbuilt function in BERTopic to produce 9 final topics (shown in Table 5).

Figure 3 shows how the representation of hypersexuality topics has changed over time, with all topics showing an increase in representation since the inception of the dataset.

XSL•FO
RenderX

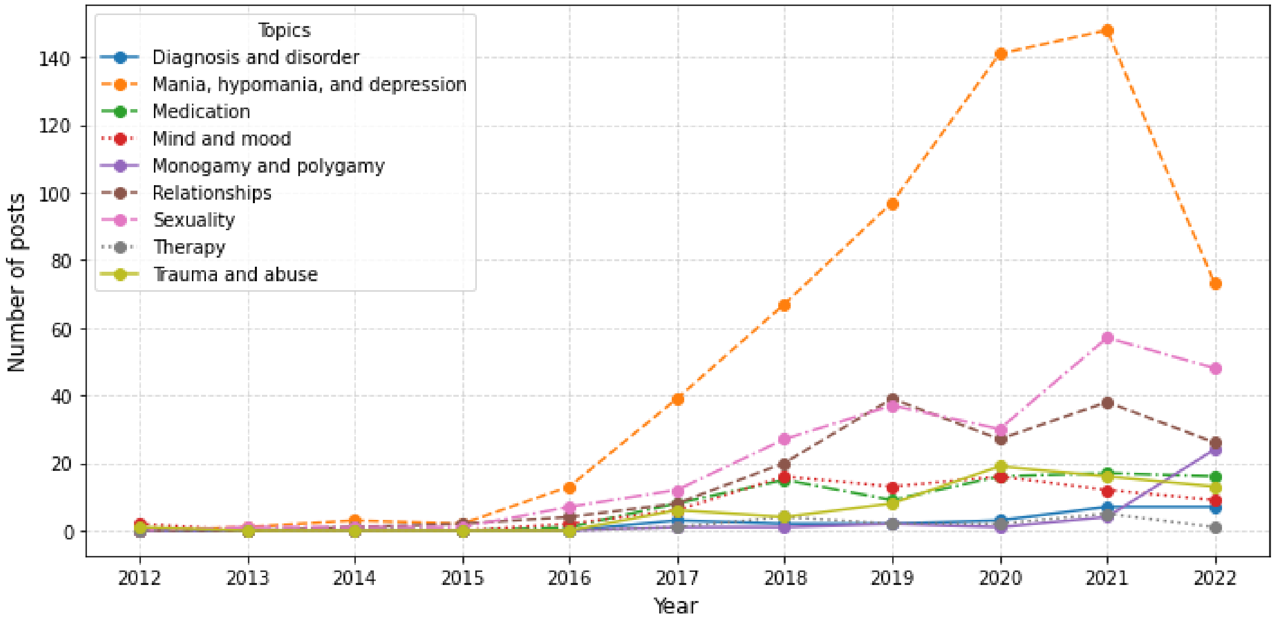
Table 5. Topics produced by BERTopic (with the manually inferred topic name), the top 10 keywords for each cluster, and paraphrased excerpts from the most representative post for each topic. Additional examples for each topic are provided in Multimedia Appendix 1 (n=2146).

Topic name (inferred)	Posts, n (%)	Top 10 keywords in the cluster	Extract from the most representative post for each topic (paraphrased)
—[a]	878 (40.91)	Outliers	—
Mania, hypomania, and depression	584 (27.21)	"Hypomanic," "hypomania," "manic," "mania," "disorder," "depressive," "depressed," "depression," "diagnosed," and "psychiatrist"	"Over 3-4 months, I left home, almost divorced, and indulged in reckless sexual encounters due to hypersexuality, hurting my family and behaving poorly. Reflecting on my manic episode, I now see the embarrassment and realize it's a common experience for many. As I came down, I recognized my strange behavior."
Sexuality	221 (10.3)	"Sexuality," "sexually," "sexual," "relationship," "feelings," "manic," "bisexual," "aroused," "feeling," and "boyfriend"	"I define myself as demisexual because I only experience attraction towards those I'm emotionally connected to, none of whom share the sentiment. Despite this, I have a strong sexual drive, feeling intense arousal monthly, and occasionally endure extended periods of hypersexuality lasting days or weeks."
Relationships	165 (7.69)	"Relationship," "relationships," "manic," "boyfriend," "disorder," "sexuality," "mania," "dating," "mental," and "diagnosed"	"I'm a challenging partner due to my manic episodes, leading to outbursts, bouts of hypersexuality (increasing the temptation to cheat), excessive drinking, and impulsive life-altering choices. Also, I believe I haven't completely healed from my previous abusive relationship."
Medication	83 (3.87)	"Hypomanic," "hypomania," "lamictal," "manic," "wellbutrin," "seroquel," "antipsychotic," "lithium," "zoloft," and "psychiatrist"	"In the last two months of taking it, there's been no improvement. Even after a week on 200mg, I'm still stuck in a severe mixed episode. I'm overwhelmed with hypersexuality, impulsivity, late nights, and a complete lack of motivation. My mood appears to be cycling rapidly, possibly even faster than before."
Mind and mood	76 (3.54)	"Hypomanic," "manic," "mood," "mania," "lithium," "feeling," "anxiety," "days," "thoughts," and "mind"	"I'm beginning to understand that although I experience cycling, my episodes often extend beyond a few days. Recent weeks of mood tracking reveal durations of a week or even two, with my current mood episode already lasting four days. In this most recent episode I've been feeling hypersexual, and like my head is full of thoughts. I'm also anxious and I've been focusing a lot on work."
Trauma and abuse	67 (3.12)	"Abuser," "abused," "abuse," "sexual," "raped," "trauma," "feelings," "memories," "therapy," and "touched"	"I started having cyber-sex with men in their 20s when I was 13, I would have online sex with anyone who was there, I wasn't thinking about their age. After this hypersexuality, I became very anxious and scared of men, and now I become very triggered when the topic of sexual abuse comes up."
Monogamy and polygamy	33 (1.54)	"Polyamory," "polyamorous," "monogamy," "monogamous," "relationship," "relationships," "poly," "married," "spouse," and "boyfriend"	"Following almost two decades of monogamous marriage, I divorced due to manic hypersexuality from bipolar, finding monogamy challenging. For five years, I explored different non-monogamous arrangements, aiming to find a new partner for monogamy. However, after another failed attempt, I encountered a married polyamorous man and chose to explore that avenue instead."
Diagnosis and disorder	24 (1.12)	"Disorder," "sexually," "sexual," "addiction," "manic," "adolescence," "mania," "psychological," "addicts," and "diagnosed"	"At 32, I was diagnosed with BP2, prompting reflection on missed signs in my childhood and adolescence. Back then, mental health wasn't a focus in my large family, and I concealed much of my struggles. With a BPD diagnosis too, distinguishing between disorders complicates understanding my experiences and symptoms. I completely relate to the hypersexuality. I have been very sexual since my early teens with a boyfriend who was years older than me."
Therapy	15 (0.7)	"Therapist," "therapy," "therapists," "counseling," "psychologist," "relationship," "intimacy," "psych," "helped," and "talking"	"I always remember them saying to never underestimate libido although that may not be the best advice for someone who's hypersexual."

[a]This is the outlier category that is automatically created by BERTopic to filter posts that are ambiguous and cannot be clustered into one of the topics.

XSL•FO
RenderX

Figure 3. Graph representing the dynamic topic modeling over time. Data collection ended in July 2022, so the observed trends may not fully reflect subsequent changes.



Discussion

Posting Behaviors

The results demonstrate that natural language processing methods were successfully used to create a corpus of Reddit posts from users who had self-reported a diagnosis of bipolar and who created content that relates to hypersexuality. There were 816 users in the dataset who posted to Reddit about hypersexuality, forming a corpus of >2000 posts. While most of the users (453/816, 55.5%) in the HiB-RC had only posted about hypersexuality once (within the data that we collected), 44.5% (363/816) of the Redditors did post repeatedly about hypersexuality—which could indicate repeat episodes of hypersexuality or sharing the same experience across multiple threads. The data demonstrate that there has been a substantial increase in the discussion of hypersexuality in terms of both the number of posts and the number of users when comparing the HiB-RC posts to the TABoRC, suggesting that this is a salient topic being discussed on Reddit.

The data suggest that the HiB-RC encompasses approximately 15% of the Redditors from the TABoRC (816/5177, 15.76%), although the number of Reddit users who talk about hypersexuality more widely on Reddit is likely to be much higher than this. We make this assumption based on the fact that we used a restrictive set of keywords and phrases to retrieve posts related to hypersexuality, as discussed in the Methods section, and based on reports that 63% of women in a recent survey on experiences of bipolar reported hypersexuality as a symptom of bipolar [27,82]. Our dataset relied on Redditors who had self-reported a diagnosis and were already aware of the terminology of "hypersexuality," but we recognize that there is a large number of people who may be sharing their hypersexual experiences on the web before receiving a diagnosis using nonclinical terminology without knowing that this is a symptom of bipolar [27,77,82]. This is an important area of exploration for future research.

When comparing the demographic inference of the HiB-RC to data from a study that profiled Reddit users with a self-reported diagnosis of bipolar [34], our statistics for age and geolocation correlate. Most Redditors in the HiB-RC were based in the United States, the United Kingdom, Canada, Germany, and Australia (768/816, 94.1%) and were between the ages of 24 and 45 years (531/816, 65.1%). However, the inferred gender data for the TABoRC suggest that most Redditors were women (3668/5177, 70.85%), which is an interesting observation compared to findings that most Reddit users in general are men [83] and previous research on bipolar that identified a more equitable distribution of Redditors who present as men and women [34]. One interpretation could stem from different methodologies of data collection; we initially sourced our Redditors from subreddits that were specific to bipolar, whereas Jagfeld et al [34] sourced Redditors across Reddit from the outset. This notion correlates with research that Redditors who present as women are 33% more likely to post in mental health–related subreddits than Redditors who present as men [55] and, thus, we would assume are also more likely to self-report a diagnosis of bipolar in these subreddits. This gender inequality is further conflated in the HiB-RC (626/816, 76.7% of the dataset presented as women). While the interpretation of this statistic requires consideration of a number of sociological perspectives and a full understanding of this topic is beyond the scope of this study, existing research reports on the "sexual double standard" [84,85]. It is well documented that "behaviours associated with high sexual activity [are] expected more and evaluated more positively" [84] in men than in women, and therefore, it is conceivable that women could feel more stigmatized about hypersexual experiences and may be more likely to post in an online "safer" space [76]: "women must strike the right balance between what society deems to be too

XSL•FO
RenderX

much sex or not enough; men suffer from the pressure of performance” [77].

Finally, when considering where Redditors in the HiB-RC posted, we can observe that 77.82% (1670/2146) of the content was posted in subreddits associated with bipolar (r/bipolar, r/BipolarReddit, r/bipolar2, and r/BipolarSOs), suggesting that most of the Redditors in the dataset were aware that this is a symptom that is linked to bipolar. As described previously, this corpus is unlikely to be fully representative of the multiple and nuanced ways in which hypersexuality could be described on the web, and therefore, we should not misrepresent this statistic and assume that the wider population of people with a diagnosis of bipolar are aware of hypersexuality as a symptom. We also note that 7.88% (169/2146) of the posts appeared in the r/bipolar2 subreddit, which has typically been ignored in academic literature related to hypersexuality in bipolar [27,86].

LIWC Analysis

The significant LIWC domains presented in the HiB-RC yielded a number of interesting insights, of which we will only discuss the most salient in this section.

With reference to the *cognition* domains, posts in the HiB-RC were more likely to demonstrate *negative tone* and *negative emotion* and less likely to present *positive tone* and *positive emotion*. This is logical when we consider the potential impact that the symptom of hypersexuality can have on a person’s life and correlates with the significantly higher presence of the *mental health* domain, which matches words such as *depressed*, *suicide*, and *trauma*. It is also logical that the *sexual* domain was significantly more frequent in the HiB-RC, where Redditors focused on sharing sexual experiences. For the domains of *reward* and *wellness*, we observed huge effect sizes of >−2, indicating that words such as *healthy*, *supported*, *gain*, and *benefit* (from the LIWC-22 dictionary) were significantly less prevalent in the HiB-RC, suggesting that Redditors do not view hypersexuality as a rewarding behavior. Finally, the domain of *past focus* was significantly more prevalent in the HiB-RC, whereby manual analysis of posts suggests that Redditors were primarily recounting histories and past experiences of hypersexuality. The significantly lower presence of the *future focus* domain correlates with this finding, as well as signifying the impulsive nature of hypersexuality that has been documented in the literature [77,86].

BERTopic Analysis

The clusters produced by BERTopic included 9 topics and 1 outlier class, and each topic was presented alongside a text excerpt from the most representative post (determined by BERTopic). Holistically, the model provided what we consider to be fairly distinct and identifiable topics, which is impressive considering the relatively small corpus and the niche domain of the dataset. Although topic modeling is not capable of capturing every nuance of the data, the model output provides a good starting point for understanding the data without needing to train a supervised model. The number of posts that were clustered into each topic by the model does not mean that these were the only posts that referenced a specific topic as some posts talked about more than one topic, and it is also likely that insightful data may have inadvertently been clustered into the outlier category. We can see that there was an increasing trend for all identified topics since 2017, which was especially pronounced for the topics of *sexuality* and *monogamy and polygamy* since 2020.

Evidence from the existing literature correlates with some of the topics identified by the automated model, including the onset of hypersexuality during an elevated mood [4,5,86], sexuality and sexual orientation [4,87], managing hypersexuality within a relationship [4,17], hypersexuality and medication [88-90], the role of child sexual abuse in hypersexuality [91-93], and vulnerability to sexual assault due to hypersexuality [27,77,82].

The Utility of a Computational Linguistic Framework

Current evidence from lived experience underscores the severe and multifaceted consequences of hypersexuality. These include risks such as sexual assault, unplanned pregnancies, vulnerability to sexually transmitted infections, traumatic abortions, and significant disruptions in personal relationships [82]. Findings from a Bipolar Commission survey involving >1500 individuals reveal that 88% of respondents experienced hypersexual behaviors, highlighting the symptom’s prevalence and potential to impact thousands of people across the United Kingdom [27,94]. Over half of the participants reported experiencing ≥8 episodes of hypersexuality during their lifetime. Furthermore, 54% reported putting themselves in dangerous situations, 54% experienced relationship breakdowns, and 22% reported being raped during a period of hypersexuality. In total, 1 in 5 respondents attempted suicide due to hypersexual behavior or its consequences, aligning with previous findings that link hypersexuality in bipolar to increased suicidal ideation [95]. The data reveal a troubling gap in clinical practice, with 60% of respondents reporting that health care professionals had not addressed hypersexuality as part of their care [82]. This disconnect between the prevalence of hypersexuality and its clinical recognition underscores an urgent need for a more comprehensive understanding of hypersexual behaviors, particularly from the perspective of those with lived experience. The development of the HiB-RC and exploratory analysis using computational linguistic methods highlights the potential of this framework in advancing our understanding of hypersexuality as a symptom experienced by individuals with bipolar. The HiB-RC represents a significant resource for future research, enabling deeper exploration of the complex relationship between hypersexuality and bipolar to help bridge the gap between clinical knowledge and practice. The use of Reddit as a data source provides unique advantages, offering insights from real-time, user-generated narratives that are free from the constraints of predefined categories typically observed in self-report questionnaires or controlled laboratory settings [76]. This approach captures an authentic and dynamic perspective, reflecting the lived experiences of individuals as they occur. Future research using this dataset will use a corpus-assisted discourse analysis to explore key thematic concepts discussed by Reddit users and describe how these findings can inform and improve clinical practice for people with bipolar.

Additional avenues for future research could build on the exploratory nature of this study using alternative methodologies

XSL•FO
RenderX

to verify the findings and deepen insights. For instance, ethnographic or participatory studies could provide a more immersive understanding, whereas large-scale qualitative studies using interviews could triangulate the results. Applying the same computational methods to clinical datasets would offer valuable cross-validation. Collecting more detailed demographic information, such as relationship status, could also shed light on how hypersexuality manifests across different life contexts, enriching our understanding of this complex symptom.

Strengths and Limitations

This study offered a unique insight into the presentation of hypersexuality within a Reddit population who self-reported a professional diagnosis of bipolar. This is the first study to observe hypersexuality in such a population, and we endeavored to not only contribute to the literature on hypersexuality but also provide a rigorous and ethical framework for doing this. We used novel computational methods to identify salient patterns in the language used by Redditors, which signpost to common experiences shared by people who experience the symptom of hypersexuality. It is also important to consider the limitations of research conducted using social media data and predictive models, and these are outlined in this section.

First, as referenced in the Methods section, we relied on self-reported diagnoses of bipolar. As is the risk with any analysis conducted using social media data, we are assuming that the posts within our corpus are truthful. As described by Coppersmith et al [49], due to "the stigma often associated with mental illness," it seems unlikely that Redditors would post about symptoms of a mental health condition that they do not have. We also tried to reduce false-positive reports of a bipolar diagnosis in the dataset by using pattern matching to capture self-reported diagnoses by Redditors.

Second, we also acknowledge limitations associated with demographic inference. The first limitation is that the gender inference model was restricted to the binary prediction of men and women as there is no tool currently available that predicts beyond these two genders, and this is a limitation of the demographic predictions. A tangential avenue for further research could involve the development of a multiclass predictive model to avoid binary classification. Future research that involves the collection of primary lived experience data (eg, through interviews) should also focus on inclusive data collection to encompass a broader set of gender identities. The second demographic limitation that we would like to address is that most of the inferred geolocations were based in America, and although the data that we report are consistent with existing literature on hypersexuality and bipolar, we cannot assume that these findings will be fully representative of international experiences. For example, Redditors worldwide are likely to be affected differently by varying health care provisions, which could have an impact on experiences with access to psychosocial support and medication costs.

Third, there are a number of limitations associated with using an unsupervised topic model, including the generation of a large number of outliers and a lack of objective evaluation metrics (which is consistent across topic-modeling methodologies). The interpretation of the topic models generated by BERTopic also still relies on human interpretation and domain knowledge, but BERTopic does provide an option to use an "auto" parameter in the setup of the model, which reduces the number of topics by merging similar clusters after the model has been trained to produce the "optimum" number of topics (as opposed to defining *k* number of topics in LDA). Finally, due to the stochastic nature of uniform manifold approximation and projection (the dimension reduction algorithm used by BERTopic), the resulting topics produced by the BERTopic model may differ when running the same code multiple times [29].

Finally, as we have acknowledged throughout this paper, we used a restrictive set of keywords to search for posts that contained references to hypersexuality, and therefore, the data presented in this paper are not definitively representative of all experiences and understandings of hypersexuality in bipolar across Reddit. Future research could use word embeddings on the HiB-RC to identify words and phrases that appear in a similar context to variants of the lemma *hypersexual* and then search for these words in the TABoRC to return a large corpus of posts that potentially describe hypersexuality. To avoid confusing hypersexuality with experiences of increased sex drive or discussion of nonnormophilic sexuality [16], these posts would need to be manually verified for inclusion, and strict coding guidelines would need to be developed.

Conclusions

This paper has presented a novel methodology for generating a corpus of data related to experiences of hypersexuality in bipolar—inferring demographic information for these data—and 2 computational linguistic methods for exploratory analysis. We demonstrated that hypersexuality is an important symptom that is discussed by people living with bipolar, with significant associated factors suggested by the topic model, including the impact on relationships, discussion of medication, sexual assault, and correlation with an elevated mood. Our LIWC analysis demonstrated that posts describing hypersexuality were significantly more likely to include language that denoted mental illness and negative emotions, and we signposted to areas of further research that could be informative in guiding future clinical interventions. This study not only fills a critical gap by providing a dataset of experiences of hypersexuality within the context of bipolar but also highlights the potential of computational linguistic methods in mental health research. The findings underscore the importance of using innovative methodologies to bridge the gap between anecdotal experiences and empirical evidence, providing data that can help develop more informed and impactful psychosocial interventions in the future.

XSL•FO
RenderX

Acknowledgments

This study was completed as part of an Economic and Social Research Council Collaborative Studentship Competition PhD studentship (grant ES/P000665/1). The funder had no role in the study design; collection, analysis, or interpretation of the data; writing of the manuscript; or the decision to submit the paper for publication.

Data Availability

A redacted version of both the Talking About Bipolar on Reddit Corpus and Hypersexuality in Bipolar Reddit Corpus are available from the UK Data Service.

Authors' Contributions

DH designed this study, collected the Reddit data, and conducted the analysis. SJ, PR, FL, JP-C, CD, and AC provided comments and guidance throughout this study and provided valuable insights for the manuscript draft. PR and SJ performed second annotations for 10% of the Hypersexuality in Bipolar Reddit Corpus, and all the authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Annotation guidelines.
[DOCX File , 17 KB - infodemiology_v5i1e65632_app1.docx]

References

1. Zhong Y, Chen Y, Su X, Wang M, Li Q, Shao Z, et al. Global, regional and national burdens of bipolar disorders in adolescents and young adults: a trend analysis from 1990 to 2019. Gen Psychiatr 2024 Feb 21;37(1):e101255 [FREE Full text] [doi: 10.1136/gpsych-2023-101255] [Medline: 38390238]
2. Lasagna CA, Pleskac TJ, Burton CZ, McInnis MG, Taylor SF, Tso IF. Mathematical modeling of risk-taking in bipolar disorder: evidence of reduced behavioral consistency, with altered loss aversion specific to those with history of substance use disorder. Comput Psychiatr 2022;6(1):96-116 [FREE Full text] [doi: 10.5334/cpsy.61] [Medline: 36743406]
3. Harvey D, Rayson P, Lobban F, Palmier-Claus J, Jones S. Lived experience at the core: a classification system for risk-taking behaviours in bipolar. Digit Health 2024 Aug 05;10:20552076241269580 [FREE Full text] [doi: 10.1177/20552076241269580] [Medline: 39108254]
4. Krogh HB, Vinberg M, Mortensen GL, Skakke D, Nielsen D, Giraldi A. Bipolar disorder and sexuality: a preliminary qualitative pilot study. Int J Bipolar Disord 2023 Feb 03;11(1):5 [FREE Full text] [doi: 10.1186/s40345-023-00285-9] [Medline: 36735118]
5. Krantz M, Goldstein T, Rooks B, Merranko J, Liao F, Gill MK, et al. Sexual risk behavior among youth with bipolar disorder: identifying demographic and clinical risk factors. J Am Acad Child Adolesc Psychiatry 2018 Feb;57(2):118-124 [FREE Full text] [doi: 10.1016/j.jaac.2017.11.015] [Medline: 29413144]
6. Ramírez-Martín A, Ramos-Martín J, Mayoral-Cleries F, Moreno-Küstner B, Guzman-Parra J. Impulsivity, decision-making and risk-taking behaviour in bipolar disorder: a systematic review and meta-analysis. Psychol Med 2020 Oct 03;50(13):2141-2153. [doi: 10.1017/S0033291720003086] [Medline: 32878660]
7. Sicilia AC, Lukacs JN, Jones S, Perez Algorta G. Decision-making and risk in bipolar disorder: a quantitative study using fuzzy trace theory. Psychol Psychother 2020 Mar 24;93(1):105-121. [doi: 10.1111/papt.12215] [Medline: 30677209]
8. Reyna VF, Mills BA. Theoretically motivated interventions for reducing sexual risk taking in adolescence: a randomized controlled experiment applying fuzzy-trace theory. J Exp Psychol Gen 2014 Aug;143(4):1627-1648 [FREE Full text] [doi: 10.1037/a0036717] [Medline: 24773191]
9. Bauer IE, Meyer TD, Sanches M, Spiker D, Zunta-Soares G, Soares JC. Are self-rated and behavioural measures of impulsivity in bipolar disorder mainly related to comorbid substance use problems? Cogn Neuropsychiatry 2017 Jul 11;22(4):298-314. [doi: 10.1080/13546805.2017.1324951] [Medline: 28490234]
10. Hıdıroğlu C, Demirci Esen Ö, Tunca Z, Neslihan Gűrz Yalçìn S, Lombardo L, Glahn DC, et al. Can risk-taking be an endophenotype for bipolar disorder? A study on patients with bipolar disorder type I and their first-degree relatives. J Int Neuropsychol Soc 2013 Apr 14;19(4):474-482 [FREE Full text] [doi: 10.1017/S1355617713000015] [Medline: 23410848]
11. Kathleen Holmes M, Bearden CE, Barguil M, Fonseca M, Serap Monkul E, Nery FG, et al. Conceptualizing impulsivity and risk taking in bipolar disorder: importance of history of alcohol abuse. Bipolar Disord 2009 Feb 09;11(1):33-40 [FREE Full text] [doi: 10.1111/j.1399-5618.2008.00657.x] [Medline: 19133964]
12. Linke J, King AV, Poupon C, Hennerici MG, Gass A, Wessa M. Impaired anatomical connectivity and related executive functions: differentiating vulnerability and disease marker in bipolar disorder. Biol Psychiatry 2013 Dec 15;74(12):908-916. [doi: 10.1016/j.biopsych.2013.04.010] [Medline: 23684382]

XSL•FO
RenderX

13. Reddy LF, Lee J, Davis MC, Altshuler L, Glahn DC, Miklowitz DJ, et al. Impulsivity and risk taking in bipolar disorder and schizophrenia. Neuropsychopharmacology 2014 Jan 21;39(2):456-463 [FREE Full text] [doi: 10.1038/npp.2013.218] [Medline: 23963117]
14. Scholz V, Houenou J, Kollmann B, Duclap D, Poupon C, Wessa M. Dysfunctional decision-making related to white matter alterations in bipolar I disorder. J Affect Disord 2016 Apr;194:72-79. [doi: 10.1016/j.jad.2015.12.019] [Medline: 26803778]
15. Perrotta G. The concept of "hypersexuality" in the boundary between physiological and pathological sexuality. Int J Environ Res Public Health 2023 May 17;20(10):5844 [FREE Full text] [doi: 10.3390/ijerph20105844] [Medline: 37239570]
16. Walton MT, Cantor JM, Bhullar N, Lykins AD. Hypersexuality: a critical review and introduction to the "sexhavior cycle". Arch Sex Behav 2017 Nov 07;46(8):2231-2251. [doi: 10.1007/s10508-017-0991-8] [Medline: 28687897]
17. Kopeykina I, Kim HJ, Khatun T, Boland J, Haeri S, Cohen LJ, et al. Hypersexuality and couple relationships in bipolar disorder: a review. J Affect Disord 2016 May;195:1-14. [doi: 10.1016/j.jad.2016.01.035] [Medline: 26851616]
18. Yaghmaei S, Babamohamadi H, Shamsaei F, Ghods AA. Unheard voices: the sexual challenges of families of patients with bipolar disorder: a qualitative content analysis. Bipolar Disord 2024 Mar 31;26(2):148-159. [doi: 10.1111/bdi.13369] [Medline: 37524374]
19. Mazza M, Harnic D, Catalano V, Di Nicola M, Bruschi A, Bria P, et al. Sexual behavior in women with bipolar disorder. J Affect Disord 2011 Jun;131(1-3):364-367. [doi: 10.1016/j.jad.2010.11.010] [Medline: 21130498]
20. Raja M, Azzoni A. Sexual behavior and sexual problems among patients with severe chronic psychoses. Eur Psychiatry 2003 Mar 16;18(2):70-76. [doi: 10.1016/s0924-9338(03)00009-9] [Medline: 12711402]
21. Marengo E, Martino DJ, Igoa A, Scápola M, Fassi G, Baamonde MU, et al. Unplanned pregnancies and reproductive health among women with bipolar disorder. J Affect Disord 2015 Jun 01;178:201-205. [doi: 10.1016/j.jad.2015.02.033] [Medline: 25827504]
22. Marengo E, Martino DJ, Igoa A, Fassi G, Scápola M, Urtueta Baamonde M, et al. Sexual risk behaviors among women with bipolar disorder. Psychiatry Res 2015 Dec 30;230(3):835-838. [doi: 10.1016/j.psychres.2015.10.021] [Medline: 26564549]
23. Clayton PJ, Pitts Jr F. Affective disorder: IV. Mania. Compr Psychiatry 1965 Oct;6(5):313-322. [doi: 10.1016/s0010-440x(65)80025-6] [Medline: 5825998]
24. Allison JB, Wilson WP. Sexual behavior of manic patients: a preliminary report. South Med J 1960;53(7):870-874. [doi: 10.1097/00007611-196007000-00009]
25. Carlson GA, Goodwin FK. The stages of mania. A longitudinal analysis of the manic episode. Arch Gen Psychiatry 1973 Feb 01;28(2):221-228. [doi: 10.1001/archpsyc.1973.01750320053009] [Medline: 4684288]
26. Gilliland R, South M, Carpenter BN, Hardy SA. The roles of shame and guilt in hypersexual behavior. Sex Addict Compulsivity 2011 Mar 08;18(1):12-29. [doi: 10.1080/10720162.2011.551182]
27. Dolman C, Howard L, Goodwin GM, Young A, Jones IR, Bergink V, et al. Hypersexuality in bipolar disorder: an overlooked symptom. Lancet Psychiatry 2024 Jun;11(6):405-406. [doi: 10.1016/S2215-0366(24)00093-2] [Medline: 38760105]
28. Boyd RL, Ashokkumar A, Seraj S, Pennebaker JW. The development and psychometric properties of LIWC-22. University of Texas at Austin. 2021. URL: https://www.liwc.app/static/documents/LIWC-22%20Manual%20-%20Development%20and%20Psychometrics.pdf [accessed 2024-04-29]
29. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online March 11, 2022 [FREE Full text]
30. pistocop - subreddit-comments-dl. GitHub. URL: https://github.com/pistocop/subreddit-comments-dl [accessed 2024-06-16]
31. Coppersmith G, Dredze M, Harman C, Hollingshead K. From ADHD to SAD: analyzing the language of mental health on Twitter through self-reported diagnoses. In: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2015 Presented at: CLPsych '15; June 5, 2015; Denver, Colorado p. 1-10 URL: https://aclanthology.org/W15-1201.pdf [doi: 10.3115/v1/w15-1201]
32. Sekulic I, Gjurković M, Šnajder J. Not just depressed: bipolar disorder prediction on Reddit. In: Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. 2018 Presented at: WASSA '18; October 31, 2018; Brussels, Belgium p. 72-78 URL: https://aclanthology.org/W18-6211.pdf [doi: 10.18653/v1/w18-6211]
33. Cohan A, Desmet B, Yates A, Soldaini L, MacAvaney S, Goharian N. SMHD: a large-scale resource for exploring online language usage for multiple mental health conditions. In: Proceedings of the 27th International Conference on Computational Linguistics. 2018 Presented at: COLING '18; August 20-26, 2018; Santa Fe, NM p. 20-26 URL: https://aclanthology.org/C18-1126.pdf [doi: 10.18653/v1/w18-0618]
34. Jagfeld G, Lobban F, Rayson P, Jones SH. Understanding who uses Reddit: profiling individuals with a self-reported bipolar disorder diagnosis. The Seventh Workshop on Computational Linguistics and Clinical Psychology 2021 [FREE Full text] [doi: 10.18653/v1/2021.clpsych-1.1]
35. glorisonne - reddit_bd_user_characteristics. GitHub. 2023. URL: https://github.com/glorisonne/reddit_bd_user_characteristics [accessed 2024-06-16]
36. Tigunova A, Mirza P, Yates A, Weikum G. RedDust: a large reusable dataset of reddit user traits. In: Proceedings of the 12th Conference on Language Resources and Evaluation. 2020 Presented at: LREC '20; May 11-16, 2020; Marseille, France p. 6118-6126 URL: https://aclanthology.org/2020.lrec-1.751.pdf

XSL•FO
RenderX

37. Harrigian K. Geocoding without geotags: a text-based approach for Reddit. In: Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-generated Text. 2018 Presented at: WNUT '18; November 1, 2018; Brussels, Belgium p. 1-50 URL: https://aclanthology.org/W18-6100.pdf [doi: 10.18653/v1/w18-6103]
38. Anna146 / HiddenAttributeModels. GitHub. URL: https://github.com/Anna146/HiddenAttributeModels/tree/master [accessed 2024-06-16]
39. khariggian - smgeo. GitHub. 2023. URL: https://github.com/kharrigian/smgeo [accessed 2024-07-05]
40. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv Preprint posted online January 16, 2013 [FREE Full text] [doi: 10.5860/choice.45-0602]
41. Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching word vectors with subword information. Trans Assoc Comput Linguist 2017 Dec;5:135-146. [doi: 10.1162/tacl_a_00051]
42. Ädel A. Corpus compilation. In: Paquot M, Gries S, editors. A Practical Handbook of Corpus Linguistics. Cham, Switzerland: Springer; 2020:3-24.
43. Krippendorff K. Computing Krippendorff's alpha-reliability. Annenberg School for Communication. URL: https://www.asc.upenn.edu/sites/default/files/2021-03/Computing%20Krippendorff%27s%20Alpha-Reliability.pdf [accessed 2024-04-29]
44. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927X09351676]
45. Boyd RL, Pennebaker JW. Language-based personality: a new approach to personality in a digital world. Curr Opin Behav Sci 2017 Dec;18:63-68. [doi: 10.1016/j.cobeha.2017.07.017]
46. Nguyen T, O'Dea B, Larsen M, Phung D, Venkatesh S, Christensen H. Differentiating sub-groups of online depression-related communities using textual cues. In: Proceedings of the 16th International Conference on Web Information Systems Engineering. 2015 Presented at: WISE '15; November 1-3, 2015; Miami, FL p. 216-224 URL: https://www.springerprofessional.de/differentiating-sub-groups-of-online-depression-related-communit/6876920 [doi: 10.1007/978-3-319-26187-4_17]
47. Chen X, Sykora MD, Jackson TW, Elayan S, Munir F. Tweeting your mental health: exploration of different classifiers and features with emotional signals in identifying mental health conditions. In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Presented at: HICSS '18; January 2-6, 2018; Honolulu, HI p. 5225-5233 URL: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/f3ae6e6f-f96a-41a9-ac75-fda2972ea5ca/content [doi: 10.24251/hicss.2018.421]
48. Chang CH, Saravia E, Chen YS. Subconscious crowdsourcing: a feasible data collection mechanism for mental disorder detection on social media. In: Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2016 Presented at: ASONAM '16; August 18-21, 2016; San Francisco, CA p. 374-379 URL: https://ieeexplore.ieee.org/document/7752261 [doi: 10.1109/asonam.2016.7752261]
49. Coppersmith G, Dredze M, Harman C. Quantifying mental health signals in Twitter. In: Proceedings of the 2014 Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2014 Presented at: CLPsych '14; June 27, 2014; Baltimore, ML p. 51-60 URL: https://aclanthology.org/W14-3207.pdf [doi: 10.3115/v1/w14-3207]
50. Huang YH, Chen YH, Alvarado FH, Lee SR, Wu SI, Lai Y, et al. Leveraging linguistic characteristics for bipolar disorder recognition with gender differences. arXiv Preprint posted online July 17, 2019 [FREE Full text]
51. Jiang Z, Levitan SI, Zomick J, Hirschberg J. Detection of mental health from Reddit via deep contextualized representations. In: Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis. 2020 Presented at: LOUHI 2020; November 20, 2020; Virtual Event p. 147-156 URL: https://aclanthology.org/2020.louhi-1.16.pdf [doi: 10.18653/v1/2020.louhi-1.16]
52. Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: observational study. J Med Internet Res 2020 Oct 12;22(10):e22635 [FREE Full text] [doi: 10.2196/22635] [Medline: 32936777]
53. Saha B, Nguyen T, Phung D, Venkatesh S. A framework for classifying online mental health-related communities with an interest in depression. IEEE J Biomed Health Inform 2016 Jul;20(4):1008-1015. [doi: 10.1109/JBHI.2016.2543741] [Medline: 27008680]
54. Olsson V, Lindow M. How does bipolar and depressive diagnoses reflect in linguistic usage on Twitter: a study using LIWC and other tools. KTH Royal Institute of Technology. URL: http://kth.diva-portal.org/smash/get/diva2:1212987/FULLTEXT01.pdf [accessed 2024-04-29]
55. Jagfeld G, Lobban F, Davies R, Boyd RL, Rayson P, Jones S. Posting patterns in peer online support forums and their associations with emotions and mood in bipolar disorder: exploratory analysis. PLoS One 2023 Sep 25;18(9):e0291369 [FREE Full text] [doi: 10.1371/journal.pone.0291369] [Medline: 37747891]
56. Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Front Sociol 2022 May 6;7:886498 [FREE Full text] [doi: 10.3389/fsoc.2022.886498] [Medline: 35602001]
57. Blei DM, La'erty JD. Topic models. In: Srivastava AN, Sahami M, editors. Text Mining: Classification, Clustering, and Applications. New York, NY: Chapman and Hall/CRC; 2009:101-124.

XSL•FO
RenderX

58. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022. [doi: 10.7551/mitpress/1120.003.0082]
59. Angelov D. Top2Vec: distributed representations of topics. arXiv Preprint posted online August 19,2020.
60. Sánchez‐Franco MJ, Rey‐Moreno M. Do travelers' reviews depend on the destination? An analysis in coastal and urban peer‐to‐peer lodgings. Psychol Mark 2021 Oct 16;39(2):441-459. [doi: 10.1002/mar.21608]
61. Grootendorst M. Topic modeling with BERTopic. Google Drive. URL: https://colab.research.google.com/drive/1FieRA9fLdkQEGDIMYl0I3MCjSUKVF8C-?usp=sharing [accessed 2024-07-16]
62. Grootendorst M. Representation models - BERTopic. GitHub. URL: https://maartengr.github.io/BERTopic/getting_started/representation/representation.html [accessed 2024-07-16]
63. Lee S, Shakir A, Koenig D, Lipp J. Open source strikes bread - new fluffy embeddings model. mixedbread ai inc. URL: https://www.mixedbread.ai/blog/mxbai-embed-large-v1 [accessed 2024-06-16]
64. Ji S, Zhang T, Ansari L, Fu J, Tiwari P, Cambria E. MentalBERT: publicly available pretrained language models for mental healthcare. GitHub. URL: https://github.com/Inusette/ [accessed 2024-04-29]
65. BPS code of human research ethics. The British Psychological Society. URL: https://www.bps.org.uk/guideline/bps-code-human-research-ethics [accessed 2024-04-29]
66. Researching online forums. British Sociological Association. 2016. URL: https://www.britsoc.co.uk/media/24834/j000208_researching_online_forums_-cs1-_v3.pdf [accessed 2024-07-05]
67. Data protection act. Government of UK. URL: https://www.gov.uk/data-protection [accessed 2024-07-05]
68. Benton A, Coppersmith G, Dredze M. Ethical research protocols for social media health research. In: Proceedings of the 1st ACL Workshop on Ethics in Natural Language Processing. 2017 Presented at: EthNLP '17; April 4th, 2017; Valencia, Spain p. 94-102. [doi: 10.18653/v1/w17-1612]
69. Williams ML, Burnap P, Sloan L. Towards an ethical framework for publishing twitter data in social research: taking into account users' views, online context and algorithmic estimation. Sociology 2017 Dec 26;51(6):1149-1168 [FREE Full text] [doi: 10.1177/0038038517708140] [Medline: 29276313]
70. Smedley RM, Coulson NS. A practical guide to analysing online support forums. Qual Res Psychol 2018 May 30;18(1):76-103. [doi: 10.1080/14780887.2018.1475532]
71. Norman Adams N. ‘Scraping’ Reddit posts for academic research? Addressing some blurred lines of consent in growing internet-based research trend during the time of Covid-19. Int J Soc Res Methodol 2022 Aug 18;27(1):47-62. [doi: 10.1080/13645579.2022.2111816]
72. Di Cristofaro M. Corpus Approaches to Language in Social Media. New York, NY: Taylor & Francis; 2023.
73. Hypersexual behaviour webinar 2024. Bipolar UK. URL: https://www.youtube.com/watch?v=ES514cUouyU [accessed 2024-07-16]
74. Lin Y. 10 Reddit statistics. Oberlo. URL: https://www.oberlo.com/blog/reddit-statistics [accessed 2024-07-05]
75. New subreddits by month. Reddit Metrics. URL: https://frontpagemetrics.com/month/ [accessed 2024-07-05]
76. Miller B. Investigating Reddit self-disclosure and confessions in relation to connectedness, social support, and life satisfaction. J Soc Media Stud 2020;9(1):39-62 [FREE Full text]
77. Chataigné A. Naming and coming to terms with bipolar hypersexuality. Lancet Psychiatry 2024 Jun;11(6):415-416. [doi: 10.1016/S2215-0366(24)00146-9] [Medline: 38760111]
78. Bender EM, Friedman B. Data statements for natural language processing: toward mitigating system bias and enabling better science. Trans Assoc Comput Linguist 2018 Dec;6:587-604. [doi: 10.1162/tacl_a_00041]
79. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965 Dec 01;52(3-4):591-611. [doi: 10.1093/biomet/52.3-4.591]
80. Rey D, Neuhäuser M. Wilcoxon-signed-rank test. In: Lovric M, editor. International Encyclopedia of Statistical Science. Cham, Switzerland: Springer; 2011:1658-1659.
81. Sawilowsky SS. New effect size rules of thumb. J Mod App Stat Meth 2009 Nov 01;8(2):597-599. [doi: 10.22237/jmasm/1257035100]
82. Bipolar minds matter. The Bipolar Commission. URL: https://www.bipolaruk.org/ [accessed 2024-04-29]
83. Barthel M, Stocking G, Holcomb J, Mitchell A. Nearly eight-in-ten Reddit users get news on the site. Pew Research Center. 2016. URL: https://www.pewresearch.org/wp-content/uploads/sites/8/2016/02/PJ_2016.02.25_Reddit_FINAL.pdf [accessed 2024-04-29]
84. Endendijk JJ, van Baar AL, Deković M. He is a stud, she is a slut!: a meta-analysis on the continued existence of sexual double standards. Pers Soc Psychol Rev 2020 May 27;24(2):163-190 [FREE Full text] [doi: 10.1177/1088868319891310] [Medline: 31880971]
85. Marks MJ, Young TM, Zaikman Y. The sexual double standard in the real world. Soc Psychol 2019 Mar;50(2):67-79. [doi: 10.1027/1864-9335/a000362]
86. Fletcher K, Parker G, Paterson A, Synnott H. High-risk behaviour in hypomanic states. J Affect Disord 2013 Aug 15;150(1):50-56 [FREE Full text] [doi: 10.1016/j.jad.2013.02.018] [Medline: 23489397]

XSL•FO
RenderX

87. Abé C, Rahman Q, Långström N, Rydén E, Ingvar M, Landén M. Cortical brain structure and sexual orientation in adult females with bipolar disorder or attention deficit hyperactivity disorder. Brain Behav 2018 Jul 29;8(7):e00998 [FREE Full text] [doi: 10.1002/brb3.998] [Medline: 29845776]
88. Yuan S, Deban CE. SSRI-induced hypersexuality. Am J Psychiatry Resid J 2021 Mar 12;16(3):9-12 [FREE Full text] [doi: 10.1176/appi.ajp-rj.2021.160305] [Medline: 33796812]
89. Bulbena-Cabré A, Bulbena A. Aripiprazole-induced hypersexuality. Prim Care Companion CNS Disord 2016 Dec 29;18(6):26650. [doi: 10.4088/pcc.16l01983]
90. Thomson SR, Patil N, Ommurugan B, Bhandary RK. A case of hyper sexuality probably associated with clozapine. Psychopharmacol Bull 2018;48(4):20. [doi: 10.1254/jpssuppl.wcp2018.0_po3-12-17]
91. Fontanesi L, Marchetti D, Limoncin E, Rossi R, Nimbi F, Mollaioli D, et al. Hypersexuality and Trauma: a mediation and moderation model from psychopathology to problematic sexual behavior. J Affect Disord 2021 Feb 15;281:631-637. [doi: 10.1016/j.jad.2020.11.100] [Medline: 33229025]
92. Vibhakar V, Allen LR, Gee B, Meiser-Stedman R. A systematic review and meta-analysis on the prevalence of depression in children and adolescents after exposure to trauma. J Affect Disord 2019 Aug 01;255:77-89. [doi: 10.1016/j.jad.2019.05.005] [Medline: 31203106]
93. Leverich GS, McElroy SL, Suppes T, Keck PE, Denicoff KD, Nolen WA, et al. Early physical and sexual abuse associated with an adverse course of bipolar illness. Biological Psychiatry 2002 Feb;51(4):288-297. [doi: 10.1016/S0006-3223(01)01239-2]
94. Hypersexuality as a symptom of bipolar. Bipolar UK. URL: https://www.bipolaruk.org/hypersexuality-as-a-symptom-of-bipolar#:~:text=?Most%20respondents%20said%20that%20they,dangerous%20situations%20as%20a%20result [accessed 2024-10-15]
95. Dell'Osso L, Carmassi C, Carlini M, Rucci P, Torri P, Cesari D, et al. Sexual dysfunctions and suicidality in patients with bipolar disorder and unipolar depression. J Sex Med 2009 Nov;6(11):3063-3070 [FREE Full text] [doi: 10.1111/j.1743-6109.2009.01455.x] [Medline: 19732307]

Abbreviations

API: application programming interface
HiB-RC: Hypersexuality in Bipolar Reddit Corpus
LDA: latent Dirichlet allocation
LIWC: Linguistic Inquiry and Word Count
LIWC-22: 2022 version of Linguistic Inquiry and Word Count
TABoRC: Talking About Bipolar on Reddit Corpus

Edited by T Mackey; submitted 21.08.24; peer-reviewed by E Morton, B Najand; comments to author 09.11.24; revised version received 05.12.24; accepted 25.01.25; published 06.03.25.

Please cite as:
Harvey D, Rayson P, Lobban F, Palmier-Claus J, Dolman C, Chataigné A, Jones S
Using Natural Language Processing Methods to Build the Hypersexuality in Bipolar Reddit Corpus: Infodemiology Study of Reddit
JMIR Infodemiology 2025;5:e65632
URL: https://infodemiology.jmir.org/2025/1/e65632
doi:10.2196/65632
PMID:40053804

©Daisy Harvey, Paul Rayson, Fiona Lobban, Jasper Palmier-Claus, Clare Dolman, Anne Chataigné, Steven Jones. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 06.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Versus Academic Discourse on ChatGPT in Health Care: Mixed Methods Study

Patrick Baxter, PhD; Meng-Hao Li, PhD; Jiaxin Wei, MA; Naoru Koizumi, PhD

Schar School of Policy and Government, George Mason University, 3351 Fairfax Dr, Arlington, VA, United States

Corresponding Author:
Meng-Hao Li, PhD
Schar School of Policy and Government, George Mason University, 3351 Fairfax Dr, Arlington, VA, United States

Abstract

Background: The rapid emergence of artificial intelligence–based large language models (LLMs) in 2022 has initiated extensive discussions within the academic community. While proponents highlight LLMs' potential to improve writing and analytical tasks, critics caution against the ethical and cultural implications of widespread reliance on these models. Existing literature has explored various aspects of LLMs, including their integration, performance, and utility, yet there is a gap in understanding the nature of these discussions and how public perception contrasts with expert opinion in the field of public health.

Objective: This study sought to explore how the general public's views and sentiments regarding LLMs, using OpenAI's ChatGPT as an example, differ from those of academic researchers and experts in the field, with the goal of gaining a more comprehensive understanding of the future role of LLMs in health care.

Methods: We used a hybrid sentiment analysis approach, integrating the Syuzhet package in R (R Core Team) with GPT-3.5, achieving an 84% accuracy rate in sentiment classification. Also, structural topic modeling was applied to identify and analyze 8 key discussion topics, capturing both optimistic and critical perspectives on LLMs.

Results: Findings revealed a predominantly positive sentiment toward LLM integration in health care, particularly in areas such as patient care and clinical decision-making. However, concerns were raised regarding their suitability for mental health support and patient communication, highlighting potential limitations and ethical challenges.

Conclusions: This study underscores the transformative potential of LLMs in public health while emphasizing the need to address ethical and practical concerns. By comparing public discourse with academic perspectives, our findings contribute to the ongoing scholarly debate on the opportunities and risks associated with LLM adoption in health care.

(JMIR Infodemiology 2025;5:e64509) doi:10.2196/64509

KEYWORDS

large language models; sentiment analysis; natural language processing; structural topic modeling; social media discourse; ethics, medical; health knowledge, attitudes, practice

Introduction

Artificial Intelligence (AI)-based large language models (LLMs) have sparked extensive discussions within the academic community since their 2022 emergence. The rhetoric is multifaceted, with rapt users touting the highly sophisticated chatbots' potential to assist in writing tasks. Critics caution, however, that the cultural and ethical ramifications associated with such reliance on LLMs may be a burden too costly to bear. Thus far, literature assessing ramifications of LLMs spans multiple fields, including finance [1], education [2], software programming [3], public health [4], and environmental studies [5]. These studies often focus on overlapping themes, that is, appropriate integration of LLMs, their analytical performance, and practical benefits for users. What seems to be missing is an examination of the nature of such deliberations. LLMs' societal impact ultimately relies on these users' verdicts, as warring technophile and luddite factions set the stage for successful technological adoption. To this end, our study assessed public opinion and perception regarding the most popular LLM widely available: OpenAI's ChatGPT.

This study specifically focused on health care, analyzing tweets that discussed the impact of ChatGPT on the health care sector since its inception and examined how ChatGPT interacts with various public health domains, including health care management, public health, digital health, clinical medicine, and nursing science [6-8]. We reviewed the opinions and sentiment shared on X (formerly known as Twitter) in order to answer the key question: Where does public and expert sentiment lie on ChatGPT's use in health care? Health care as LLMs technology has drawn recent accolade for the potential use, in part due to ChatGPT's recent accomplishment of passing the US Medical Licensing Exam [9]. Key concerns raised by the scientific community thus far include consequences of potential biases introduced in the LLM algorithm training

XSL•FO
RenderX

process, which may exacerbate the existing health disparities [10], spreading misinformation [11], and medical record breaches and cyber security [12]. In our current work, we aimed to understand how public opinions and sentiment about LLMs contrast with the opinions shared among academic researchers and other field experts to gain a broader view for the future direction of LLMs in health care. Through opinion mining, we identified 8 topics that represent general concerns and optimisms toward ChatGPT in health care. For the analysis and classification of twitters sharing positive and negative sentiments, we implemented 4 algorithms, determining their accuracy based upon a manual review of the tweets themselves conducted by our research team. We find our novel enhanced method that combines Syuzhet and GPT 3.5 had an 84% accuracy rate, 12 percentage points better than other classification algorithms used in our analysis.

The remaining article is ordered as follows: first, we state our research goal and key question underlying our analysis. Next, we state out our statistical methods. We then present our findings and discuss the implications for future research.

Methods

Ethical Considerations

To protect the privacy and confidentiality of study data, all IDs, usernames, and tweets have been deidentified. Additionally, any full tweets cited in the paper have been paraphrased to prevent them from being traced back to the original user.

Data Source

We used the academic Twitter API to retrieve tweets with search terms "ChatGPT AND (health OR healthcare OR hospital OR physician OR nurse OR nursing OR patient)" [13,14]. This data collection process was executed for the period between December 1, 2022, the day after ChatGPT became publicly available, and March 20, 2023. After removing duplicates using the Jaccard Similarity score [15], there were 6138 unique tweets authored by 4837 distinct accounts. The Jaccard Similarity is expressed as:

$$J(A,B)=\frac{|A\cap B|}{|A\cup B|}$$

where A and B are 2 sets, $|A\cap B|$ represents the number of common words between them, and $|A\cup B|$ represents the total number of unique words across both sets. The Jaccard score ranges from 0 to 1, where 0 indicates no similarity and 1 indicates identical sets.

It is worth noting that the academic Twitter API can be used to collect all tweets instead of just sampled tweets. Academic researchers are granted special access to the Twitter V2 API, which provides access to X's real-time and all historical public data (unbiased tweets). This API is no longer supported by X Corporation. However, users can still obtain full access to X through a paid subscription.

Sentiment Classification and Analysis Procedure

Our analysis consisted of three phases: (1) human-labeled sentiment tweet classification, (2) algorithm-based sentiment tweet classification; and (3) structural topic modeling (STM) to distinctly group tweet content. Each phase is detailed below.

Human-Labeled Sentiment Classification

A team of 2 public health faculty members and two PhD students first reviewed each tweet and classified them into 3 mutually exclusive sentiment categories: positive, negative, and neutral. The team categorized tweet sentiment based on lexical content, context, emojis (eg, for positive, for negative), and tone. Positive tweets typically use words like "happy" or "love," while negative tweets include "terrible" or "hate"; neutral tweets lack strong emotions, informative contexts or advertisements. Each tweet was reviewed by 2 reviewers. If the 2 reviewers disagreed, the team discussed how to label the tweet in order to reach a consensus. Of the 6138 tweets, the majority of tweets were classified as neutral sentiment (4359/6138, 71%), while only a small percentage of tweets were classified as positive (1350/6138, 22%) and negative sentiments (460/6138, 7.5%). However, sentiment analysis models typically struggle with neutral context. Most of the collected tweets with neutral sentiment were those that included both positive and negative sentiments, potentially leading to misinterpretation of the results. Due to the inherent challenges associated with accurately interpreting neutral tweets, we excluded them from the analysis [16]. After removing neutral sentiment tweets, our final dataset contained 1806 tweets authored by 1586 distinct accounts.

Algorithm-Based Sentiment Classification

Next, we compared sentiment classification between 3 ML algorithms and those derived from the research team's manual review. These ML algorithms consisted of 2 of OpenAI's API models, Gpt-3.5-Turbo-0301 and Gpt-4.0, and the conventional dictionary-based Syuzhet method [17]. Both GPTs and Syuzhet can be used for sentiment classification, although they serve distinct purposes within this domain. GPT, generating text based on extensive pretraining text data, has a better understanding of context and is suitable for various tasks. In contrast, Syuzhet is specifically designed for sentiment analysis tasks and provides predicted labels based on its training on sentiment-labeled data. We used the Syuzhet package in R to evaluate the textual content extracted from the X data, tokenize the input text into words, then mapping them to predefined sentiment scores based on the chosen lexicon. For example, Syuzhet assigns a positive word (eg, love or happy) a+1and a negative word (eg, bad or terrible) a −1 and then adds them together for the sentence.

We explored the potential for enhancing the performance of algorithms by combining predictions from the Syuzhet and OpenAI GPT algorithms. Here, we identified the GPT algorithm with the superior performance, either GPT 3.5 or GPT 4.0, and then compared the GPT predictions with the Syuzhet predictions. If both the GPT and Syuzhet predictions exhibit the same direction, the predictions are retained and compared with the ground truth values. Each true match (ie, a tweet with the same algorithm and human labeling result) was considered a success, while contradictory results were considered an inaccurate algorithm classification. The accuracy rate was calculated as:

$$Accuracy=\frac{TruePositive+TrueNegative}{TruePositive+TrueNegative+FalsePositive+FalseNegative}$$

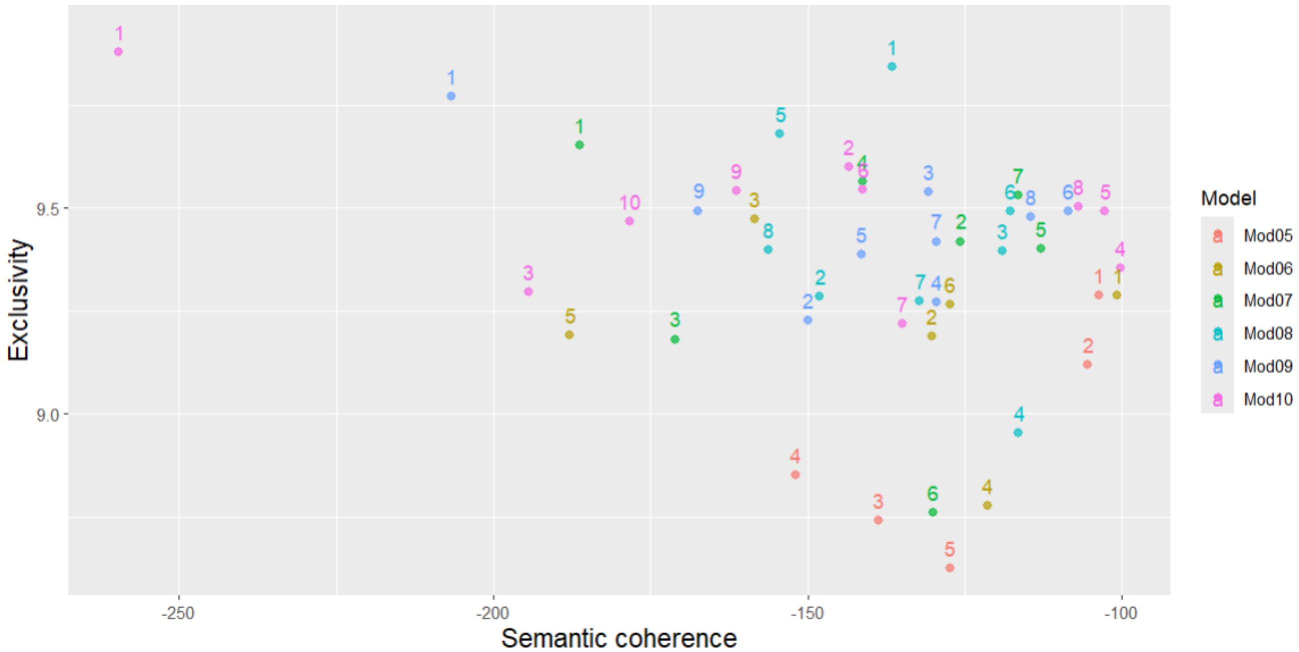
XSL•FO
RenderX

Application of STM

We used a STM to classify the tweets into distinct content topics discussing ChatGPT [18]. STM is a natural language processing and text analysis statistical technique that builds upon traditional topic modeling approaches, such as Latent Dirichlet Allocation [19]. Rather than simply identifying content topics, STM analyzes both the overall structure of the documents (eg, the arrangement of sentences, paragraphs, and other textual components) as well as the associations with other metadata (covariates). The approach allows for a more nuanced understanding of the topics, as well as how the topics relate to each other and to other contextual covariates. In this study, we incorporated 2 specific covariates, namely our human-labeled sentiment (positive and negative) and opinion leader status. The opinion leader covariate is defined by the number of followers associated with a given X account [20,21]. We categorize an "opinion leader" as an account in the top 10% (159/1586) of followed accounts, requiring at least 10,048 followers.

The STM method itself was an elaborative process. Initially, researchers arbitrarily determined the appropriate number of topics to batch the corpus content. We experimented with 5 to 10 topics (Figure 1). Once the number of topics was established, the STM algorithm proceeded to randomly assign topics to each word in every document within the corpus. The algorithm then refined these initial assignments iteratively by analyzing both the frequency of each topic's appearance in a document and the distribution of words within each topic. This dynamic process resulted in ongoing adjustments to the topic assignments for each tweet, while taking the probabilities of topic occurrences and word distributions into consideration. Two key measures determined the optimal number of STM topics: semantic coherence and exclusivity. Semantic coherence assesses how frequently words co-occur within a topic, reflecting the strength of their thematic connection. Topics with high semantic coherence contained words that often appear together, indicating a strong thematic link. Exclusivity quantified the distinctiveness of words within a topic, indicating how unique they are to that topic compared to others. Topics with high exclusivity contained words that are specific to that topic and rarely appear in other topics [18].

Figure 1. Comparison between exclusivity and semantic coherence.



After finalizing the topics, we evaluated each topic closely in order to assign a specific label to the topic. Further, we descriptively compared the opinion leaders' and nonopinion leaders' tweets.

Results

Comparison of Sentiment Classification Algorithms

Table 1 presents the results of sentiment analyses performed using: (1) human labeling; (2) Syuzhet; (3) OpenAI GPT 3.5; and (4) GPT 4.0. True positive and true negative cases are presented in bold in the table. As seen in the table, the model performed far better in predicting positive cases with the accuracy rates of 47% (Syuzhet) and 55% (OpenAI GPT 3.5 and 4.0). For the negative cases, OpenAI GPT 3.5 performed far better than the other algorithms with the accuracy rates of 9% (Syuzhet), 18% (OpenAI GPT 3.5) and 10% (OpenAI GPT 4.0). Consequently, for the overall accuracy, OpenAI GPT 3.5 outperformed GPT 4.0 and Syuzhet with the accuracy rates of 72.04%, 65.23%, and 55.82% respectively.

XSL•FO
RenderX

Table . Human labelling versus sentiment classification algorithms.

	Syuzhet			GPT3.5			GPT4.0		
	Negative	Positive	Neutral	Negative	Positive	Neutral	Negative	Positive	Neutral
Ground Truth									
Negative, n (%)	163 (9.03)[a]	165 (9.14)[b]	135 (7.48)[c]	316 (17.50)[a]	17 (0.94)[b]	130 (7.20)[c]	178 (9.86)[a]	15 (0.83)[b]	270 (14.95)[c]
Positive, n (%)	73 (4.04)[c]	845 (46.79)[d]	425 (23.53)[c]	43 (2.38)[c]	985 (54.54)[d]	315 (17.44)[c]	20 (1.11)[c]	1000 (55.37)[d]	323 (17.88)[c]

[a]True negative.
[b]False positive.
[c]False negative.
[d]True positive.

Table 2 presents the result of the algorithm that combined Syuzhet and OpenAI GPT 3.5. The enhanced ML approach yielded a superior accuracy rate of 84.04%, surpassing the accuracy of any single algorithm mentioned above by at least a 12% improvement rate.

Table . Human labelling versus. (Syuzhet+ OpenAI GPT 3.5)[a].

	Syuzhet + OpenAI GPT 3.5		
	Negative	Positive	Neutral
Ground truth			
Negative, n (%)	111 (12.74)[a]	6 (0.69)[b]	34 (3.9)[c]
Positive, n (%)	11 (1.26)[c]	621 (71.3)[d]	88 (10.1)[c]

[a]True negative.
[b] False positive.
[c]False negative.
[d]True positive.

Evaluation of Structural Topic Model Results

Figure 1 demonstrates how these topics were assessed using semantic coherence (x-axis) and exclusivity (y-axis). This evaluation guided our decision in selecting the most appropriate models for clustering the tweets. Topics positioned near the upper right corner represent higher semantic coherence and exclusivity. The evaluation indicates the “Mod08” model, which consists of 8 topics, demonstrates both high semantic coherence and high exclusivity. In other words, Mod08 best embodies a strong thematic grouping with distinctive and closely related words compared to other STMs. The corpus was therefore divided into 8 topics for the subsequent analysis.

Table 3 lists the top 20 words with the highest probabilities that appear in the tweets classified under each of the eight topics. While most tweets focused on ChatGPT’s utility in the health arena, some tweets referred to the utility of ChatGPT in multiple sectors or industries, including the health sector. These tweets were classified under Topic 1. The tweets under Topic 2 predominantly focused on the potential and plausible future roles of ChatGPT in our daily lives with significant optimism. Under Topic 3, there were a number of tweets referring to the limitations of ChatGPT due to its sole reliance on data in making decisions. The tweets under Topic 4 referred to the utility of ChatGPT as an existing mental health service provider. Tweets classified under Topics 5 and 6 focused on plausible near future transformation in health care system (Topic 5) and services (Topic 6) triggered by ChatGPT. Topic 7 tweets also commented on future roles of ChatGPT in healthcare services, but the discussions were more on the concerns stemming from biases and inaccurate information identified in ChatGPT’s responses. Finally, Topic 8 contained tweets expressing positive sentiment about ChatGPT’s capability in improving clinical documentation’s effectiveness and precision as well as its availability to respond to medical questions around the clock, many of which referred to the possibility of physicians and other healthcare professionals being replaced in the near future.

XSL•FO
RenderX

Table . Top 20 words with the highest probabilities of each topic.

Topic	Top 20 words
Topic 1	health care, industries, industry, revolutionize, finance, applications, customer, revolutionizing, exciting, impact, possibilities, various, healthtech, development, efficiency, efficient, forbes, save, lets, delivery
Topic 2	potential, technology, improve, education, future, like, care, openai, artificialintelligence, new, see, ways, way, health, medicine, innovation, service, outcomes, world, work
Topic 3	time, tool, data, one, like, human, better, see, much, hospital, able, read, responses, day, school, never, imagine, dont, far, next
Topic 4	health, mental, can, support, using, advice, issues, used, people, chatbot, mentalhealth, users, public, provide, app, openai, chat, gpt, company, ethical
Topic 5	will, use, just, think, great, going, well, writing, already, take, amazing, change, cases, example, interesting, things, tech, field, many, cant
Topic 6	can, medical, patient, used, patients, language, treatment, tools, diagnosis, professionals, clinical, chatbots, intelligence, provide, help, artificial, models, accurate, data, large
Topic 7	can, help, even, care, get, people, may, information, doctor/s, patient, better, system, made, make, replace, way, want, wrong, like
Topic 8	asked, write, google, questions, patient, good, physician, answer, ask, still, using, lot, work, now, medical, best, doctors, say, need, asking

We reviewed all of the tweets under each of the topics and used these words as a supplement to label the eight topics as mentioned in Textbox 1:

Textbox 1. Eight topics.

- Topic 1: ChatGPT’s potential in advancing various industries
- Topic 2: ChatGPT’s potential in improving our daily lives
- Topic 3: Concerns related to ChatGPT’s reliance on data
- Topic 4: ChatGPT in mental health services
- Topic 5: ChatGPT as text generator
- Topic 6: ChatGPT as an analytical tool
- Topic 7: Fairness in ChatGPT responses
- Topic 8: ChatGPT’s potential in replacing healthcare professionals

Representative Tweets

We summarized each of the 8 topics identified via opinion mining and provided several representative tweets under each topic to demonstrate key points and discussions surrounding each topic.

Topic 1: ChatGPT’s Potential in Advancing Various Industries

The 49 tweets under Topic 1 highlighted ChatGPT’s potential to trigger significant transformations across various sectors, such as retail, health care, and entertainment, through real-time applications. Approximately 90% (44/49) of these tweets expressed a positive outlook. These tweets provided concrete examples of how these industries could benefit from advancements in AI and state-of-the-art technologies.

> *Across sectors like retail, healthcare, and entertainment, visual ChatGPT offers a range of real-time applications poised to transform entire industries.*

> *With advances in AI and other cutting-edge technologies, ChatGPT is steadily gaining traction in the market. Its integration into the healthcare industry? Absolutely possible.*

> *The impact of #ChatGPT across industries is already unfolding—from healthcare to finance, its potential is immense. As conversational AI continues to evolve, it’ll be exciting to see what the future brings.*

Topic 2: ChatGPT’s Potential in Improving Our Daily Lives

Topic 2 contained 355 tweets, with about 97% (344/355) expressing a positive sentiment. These tweets highlighted the

XSL•FO
RenderX

role of innovative technologies, such as ChatGPT, in improving various aspects of our lives, including health care and transportation. Well known figures, such as Bill Gates, have emphasized ChatGPT 's significance in revolutionizing office operations, health care, and education for better outcomes and efficiency. ChatGPT is seen as a pivotal innovation with the potential to reshape diverse domains and create numerous opportunities for innovation ecosystems. In summary, these tweets emphasize the transformative impact of AI technologies, particularly ChatGPT, in enhancing our daily lives and addressing pressing challenges, underscoring the opportunities they offer for innovation ecosystems.

> *It's exciting to see how #AI is enhancing our lives across the board—from healthcare to transportation. With technology on our side, the future is looking brighter than ever.*
>
> *Bill Gates believes AI—especially tools like #ChatGPT—is currently the "most important" innovation. AI technology offers powerful opportunities to boost efficiency and outcomes in workplaces, healthcare systems, and educational settings.*
>
> *In conclusion, ChatGPT is as a versatile AI tool with the potential to significantly improve many areas of our lives—from personal productivity and education to health, career growth, financial planning, customer service, and even virtual events.*

Topic 3: Concerns Related to ChatGPT's Reliance on Data

The 136 tweets under Topic 3 placed an emphasis on concerns related to the utilization of ChatGPT in health care applications. Approximately 43% (58/136) of these tweets conveyed a negative sentiment. A primary concern pertained to the feasibility of integrating ChatGPT into health care scenarios with limited patient data and the fiscal constraints imposed by insurance companies, thus underscoring the financial considerations associated with data acquisition. Another worry centered around the security of sensitive medical data, and risks associated with disclosing such information on public domain. In addition, doubts were expressed regarding ChatGPT's ability to offer appropriate medical advice. In summary, these tweets collectively highlighted concerns about the adoption of ChatGPT in health care services and advocated for a more cautious approach in handling sensitive medical data.

> *Feeding ChatGPT data equations, flowcharts, and calculations? That's the easy part. Now put it in a patient room—with a poor historian and limited information—and expect a clear answer? Good luck. Oh, and by the way—insurance doesn't always approve more tests. Data isn't free.*
>
> *Y'all, what are you doing? Treat any data you give to OpenAI like it's going on your public Facebook feed. Would you post patient info on FB? Then don't put it in ChatGPT. Would you share your company's financials on FB? Then don't feed them to ChatGPT either. Come on—this is rookie league stuff.*
>
> *I've been steering clear of (chat)gpt's hype—and turns out, I was right. Don't rely on ChatGPT for health advice. Tools like https://t.co/z0nixzpowi are far better for lit reviews and won't mislead you with false hope or questionable treatments.*

Topic 4: ChatGPT in Mental Health Services

Topic 4 encompassed 387 tweets, with approximately 52% (201/387) of them conveying a negative sentiment. These tweets shed light on the debates surrounding the utilization of ChatGPT in mental health applications. Supporters argued that ChatGPT could offer a valuable alternative to traditional therapy for individuals who face financial constraints or prefer remote interactions, thereby enhancing the accessibility of mental health services. Furthermore, health apps using ChatGPT as an AI health coach can provide personalized and round-the-clock assistance, potentially revolutionizing the field of health coaching. However, ethical concerns were raised, particularly in terms of informed consent. Some view experiments like Koko's use of ChatGPT for mental health support as ethically questionable. In conclusion, although ChatGPT shows promise in addressing mental health needs, it is imperative to carefully navigate ethical considerations and consent issues to ensure its responsible implementation in this domain.

> *ChatGPT has the potential to serve as a digital therapist for those who can't afford counseling or prefer to avoid in-person sessions. It could help expand access to mental health support for people who need it the most.*
>
> *A health app now uses ChatGPT to take the place of human health coaches—giving users round-the-clock access to an AI coach that offers clear, helpful support and advice based on their needs.*
>
> *I honestly don't see how an experiment like this could be exempt from informed consent requirements. It's flat-out #unethical. A company using #ChatGPT for mental health support without proper safeguards brings serious ethical concerns to the table.*
>
> *Here we go—Koko, a nonprofit focused on peer mental health support, ran a test using ChatGPT on its users without getting their consent. That's a serious breach of trust.*

Topic 5: ChatGPT as Text Generator

Topic 5 contained a total of 169 tweets, with the majority (132/161, 82%) reflecting a positive sentiment. These tweets highlight the growing recognition of ChatGPT's remarkable applications in various domains, particularly for remote health care delivery and summarizing virtual meetings. Its automation capabilities and ability to provide insightful summaries generated positive feedback from users, suggesting its potential to disrupt conventional industries such as consulting and academia. In short, ChatGPT is demonstrating its transformative potential across multiple domains, reshaping our approaches to writing, summarizing, and decision-making.

> *At our organization, we're currently experimenting with ChatGPT in the healthcare space—especially in delivering remote care globally. It's been helpful for*

XSL•FO
RenderX

generating automated summaries and transcripts of virtual sessions. At this point, its usefulness is hard to deny.

The tools coming out of OpenAI are already shaking up the overpriced consulting world—and honestly, good! Just yesterday, a friend told me their buddy, who's on the hunt for a nursing job, was amazed at how much ChatGPT helped them craft a strong resume.

Truly impressive—I asked ChatGPT how the #healthcare system might evolve after COVID-19, and it delivered a thoughtful response in just 3–5 seconds. It's clear that #academia needs to start engaging with this tool in a smart, thoughtful way.

Topic 6: ChatGPT as an Analytical Tool

Topic 6 contained a total of 288 tweets, with the vast majority (268/288, 93%) conveying a positive sentiment. These tweets emphasized the crucial role of ChatGPT in health care decision-making and highlighted several key aspects. ChatGPT assists physicians with analyzing patient symptoms and medical history, facilitating diagnoses, and personalized treatment recommendations, and helping physicians make more informed decisions. Finally, ChatGPT serves as a valuable resource for lay people to understand medical conditions, drug interactions and treatment options, supporting decisions based on individual needs.

One way ChatGPT can support the medical field is by helping professionals consider possible diagnoses and treatment paths. By reviewing patient symptoms and medical history, it can suggest options that aid doctors in making more informed choices.

ChatGPT can support healthcare by delivering fast, accurate responses to medical questions, aiding in clinical decision-making, and offering suggestions that reflect each patient's individual needs and situation.

ChatGPT can assist doctors by offering information on medical conditions, potential drug interactions, and available treatment options—helping support their work in diagnosing and treating patients.

Conversational AI can play a role in telemedicine by helping patients reach healthcare professionals and offering useful information about their health along the way.

ChatGPT could serve as a source of reliable, up-to-date information on a wide variety of health topics—from common illnesses to available treatments and therapies.

Topic 7: Fairness in ChatGPT Responses

There were 176 tweets categorized under Topic 7. Approximately 35% (62/176) of the tweets expressed negative sentiments. The negative tweets highlighted concerns about biases seen in ChatGPT's responses in health-related conversations. When requesting stories involving doctors and nurses, the AI often portrays nurses as women and doctors as men. Similarly, some professions are likely to be portrayed with a specific sex and performance reviews generated by ChatGPT tend to be more critical for female employees, exhibiting possible gender as well as role biases. In addition, some tweets acknowledged that small errors generated by ChatGPT could potentially cause serious harm to patients.

I asked ChatGPT for ten stories involving a doctor and a nurse. Only one featured a female doctor and a male nurse. Every story had a heterosexual pairing, and the names were overwhelmingly Anglophone—doctors named Alex, Jack, and Rachel; nurses with similar naming patterns. AI is still echoing bias, not representing reality.

AI often mirrors the biases in its training data. Ask for a picture of a nurse, and you'll probably see a woman. Ask for a doctor, and chances are you'll get a man. This isn't just coincidence—it's a reflection of long-standing stereotypes baked into the data.

ChatGPT tends to write longer and more critical performance reviews when it assumes the employee is a woman. It associates roles like nurse, receptionist, and kindergarten teacher with women, while seeing mechanic as male, and banker or engineer as male or neutral. These patterns show how gender bias can still surface in AI-generated content.

The issue with ChatGPT is that it can still make mistakes—even in basic essays, including historical or factual ones. Sure, it might get 95% of the information right, but that remaining 5%? In medicine, that margin of error could cost a life. We're still years away from trusting chatbots in the operating room.

Topic 8: ChatGPT's Potential in Replacing Healthcare Professionals

Topic 8 consisted of 266 tweets, and 29% (77/266) of them expressed negative sentiments. These tweets offered a glimpse into the various perspectives on ChatGPT's potential in substituting health care professionals. Some individuals believed that ChatGPT has demonstrated promising prospects in the health care arena and has the ability to generate text of human-like quality. For example, ChatGPT could transcribe patient audio and produce medical correspondence, possibly improving clinical documentation's effectiveness and precision. However, it is critical to acknowledge that ChatGPT is not a substitute for human expertise at this point. While ChatGPT can offer valuable assistance, particularly in generating high-quality human-like text, its performance in certain areas, such as health care real estate, is still inadequate. This implies that more refinement and development is required before it can completely replace human expertise in every aspect of health care.

Using OpenAI's open-source Whisper to transcribe patient audio, then running it through ChatGPT for responses—it's a setup that could be built at low cost and, frankly, might outperform your average BetterHelp therapist.

XSL•FO
RenderX

I just fed some brief (fictional) patient notes into ChatGPT and asked it to draft a medical letter—the outcome was surprisingly solid. Honestly, it's getting close to dictation-level quality.

ChatGPT is out here solving problems we didn't really have. What I actually need is a system that makes healthcare affordable. Or a robot that can clean my place. Not a faster version of Google.

I ran some detailed healthcare real estate questions by ChatGPT—stuff I already knew the answers to—and it came back with a vague, off-the-mark response. Safe to say, my job's not going anywhere anytime soon.

Comparison Between Public and Opinion Leaders

Table 4 presents the number and percentage of tweets by opinion leader status. Overall, both groups exhibited similar tweeting patterns. Opinion leaders were, however, more likely to discuss: (1) ChatGPT's potential in improving our daily lives (Topic 2: 22.48% vs 18.01%); and (2) the possibility of ChatGPT replacing health care professionals (Topic 8: 20.18% vs 13.98%). In contrast, nonopinion leaders expressed more concern about the use of ChatGPT as an analytical tool in their daily lives (Topic 6:16.75% vs 10.09%). These findings illustrated the diverse perspectives and priorities within the ChatGPT discourse, underscoring the importance of considering multiple viewpoints when evaluating its impact on industries and daily lives.

Table . Comparison between opinion leaders' and nonopinion leaders' tweets.

	Opinion leader, n (%)	Nonopinion leader, n (%)
Topic 1: ChatGPT's potential in advancing various industries	2 (0.92)	47 (2.96)
Topic 2: ChatGPT's potential in improving our daily lives	49 (22.48)	286 (18.01)
Topic 3: Concerns related to ChatGPT's reliance on data	19 (8.72)	117 (7.37)
Topic 4: ChatGPT in mental health services	46 (21.1)	341 (21.47)
Topic 5: ChatGPT as text generator	16 (7.34)	153 (9.63)
Topic 6: ChatGPT as an analytical tool	22 (10.09)	266 (16.75)
Topic 7: Fairness in ChatGPT responses	20 (9.17)	156 (9.82)
Topic 8: ChatGPT's potential in replacing healthcare professionals	44 (20.18)	222 (13.98)

Discussion

Principal Findings

This study performed public sentiment analysis regarding ChatGPT's impact on health care. The findings provide several implications for both the deployment of LLMs in healthcare and the need for broader understanding of public opinion towards AI technologies in medical contexts.

The predominance of positive sentiment toward ChatGPT indicates a general optimism amongst X or Twitter users about the integration of AI into the field. This optimism was notably strong in discussions surrounding ChatGPT's potential to enhance patient care and health care decision-making, perhaps an acknowledgment of AI's capacity to process and synthesize large amounts of medical information quickly and accurately, which can support medical professionals in diagnosing and treating patients more effectively. In addition, users were excited about ChatGPT's availability to respond to questions.

Conversely, the areas of concern highlighted by the negative sentiments—primarily around mental health support and patient communication—point to critical ethical and practical challenges. Concerns of the reliability of AI-generated advice, the management of patient data, and the potential for perpetuating biases within AI algorithms are prevalent across topics. This suggests while there is readiness to embrace AI for certain technical tasks within health care, cautious public concern remains regarding the limits of AI's role in sensitive aspects of health care. Empathetic human interaction can and should play a crucial role in these areas.

The added sentiment classification accuracy of the enhanced ML approach is also worthy of note. Combining Syuzhet and OpenAI GPT 3.5 algorithm predictions resulted in an 84% accuracy rate, by far the best performing classification strategy we tested for our analysis. We theorized that by integrating the Syuzhet approach, which focuses on extracting the underlying emotional trajectory of a narrative with the predictive power of OpenAI's GPT 3.5. The model's synergy allowed for a more robust interpretation of data. Future researchers should iterate our approach with more sophisticated LLM models, like ChatGPT 4.0 or ChatGPT 4.5, to further enhance accurate sentiment analysis. The high accuracy rate implies that using both the GPT models and the Syuzhet package, researchers or policy makers can efficiently monitor public sentiment on emerging health crises and quickly analyze public health sentiment and meaningful insights on Twitter (or other social media) without extensive expertise in natural language processing. Also, recent studies on sentiment analysis of ChatGPT tweets typically use machine learning approaches for classification and require researchers to manually label tweets to create a training set [22-24]. The human labelling process is labor-intensive and time-consuming. In our study, we

XSL•FO
RenderX

demonstrated that pretrained LLMs are a potential tool to classify tweet sentiment without the need for manual labeling, significantly reducing the time and effort required by researchers.

Overall, public sentiments towards AI adoption in health care mirrored opinions found in academic literature. In particular, an abundance of literature has been published on ChatGPT's utility as a text generator (Topic 5). Here, a number of academic publications focus on the use of ChatGPT in scientific publications, which has led many journals and publishers to set new restrictions on the use of ChatGPT in generating manuscripts [25-28]. The main rationales for such restrictions are ChatGPT's "artificial hallucination", particularly in generating references that do not exist [29] as well as potential plagiarism for which non-human authors cannot be held accountable [30]. On the positive side, ChatGPT's usefulness as a generator of patient clinical notes and discharge summaries has also been explored and discussed by academics and health professionals [31-33] which corresponds to many tweets found under Topic 5. Racial or ethnic and gender biases as well as misinformation in ChatGPT's responses (Topic 7) are also noted in academic literature [34,35], and there is a general consensus in the literature that the challenge is likely to persist despite the use of more training data and novel algorithms [36].

Both academic literature and public tweets commented on the overconfidence of ChatGPT's responses in providing misinformation. Topic 6 (ChatGPT as an Analytical Tool) has also been widely discussed in academic and health professional communities as a potential diagnostic tool and a recommender and a potential decision maker of treatment regimens [37-39]. These studies conclude that ChatGPT's responses are mostly accurate on common, nonspecialized, topics, while, for specialized topics, the accuracy remains subpar. This aspect was not discussed in the tweets from the general public. Discussions about the negative consequences of limited training data (Topic 3) were seen ubiquitously as the major limitation of the currently available LLM in academic literature [36]. And the literature often perceived this issue as a long-term challenge. The literature unanimously states that health care workers cannot be replaced by AI (Topic 8), highlighting the importance of human-AI collaboration [40-42], while the general public was more likely to emphasize that human replacement is likely in the near future.

Finally, there is abundant literature on ChatGPT and mental health (Topic 4). Most of the academic literature on this topic focuses on the risks involved in mental health patients relying on ChatGPT. This was somewhat in contrast with the many tweets found on Topic 4 that highlighted the availability of ChatGPT as a provider of human-like interactions and personalized advice around the clock. The negative consequences discussed by the academicians and other professionals included escalation of self-isolation, which is known to lead to suicide or self-harm [43,44], risk of exposing sensitive personal information about themselves and their caregivers, which could lead to privacy violations [44,45], and ChatGPT's failure in capturing nonverbal cues and subtle human signals and its tendency to underestimate the suicide risk [44,46]. The literature suggests that ChatGPT is particularly not equipped to serve younger generation and children who are more likely to rely on the LLM applications [44,47]. Also, ChatGPT in childcare can easily provide false information. The collection of data from children raises significant concerns regarding privacy, security and the risk of potential data misuse [48]. Furthermore, informed consent is a fundamental ethical requirement for AI-driven mental health apps, as outlined in WHO's Regulatory Considerations on Artificial Intelligence for Health [49], which emphasizes the need for privacy and data protection, such as the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA).

In summary, we found that the general public is clearly privy to the main opinions raised by the academic community and health professionals, while the discrepancy in opinions was more notable in ChatGPT's capability as a mental health service provider as well as its potential as an analytical tool. For both topics, the general public was somewhat more optimistic or less specific in providing negative opinions. However, it is important to note that academic literature, which provides a more authoritative and evidence-based perspective, holds greater significance and value than public opinion in evaluating these aspects. Public opinion, while informative, often lacks the depth and rigor that scholarly analysis offers in these domains.

This study acknowledges several limitations. First, tweets often express a mix of positive and negative sentiment and may also contain advertisements. This complexity can challenge both LLMs and human analysts in accurately classifying them. Future research could address this by developing methods to categorize each tweet based on percentages of positive and negative sentiment, and by training LLMs to predict the likelihood of advertisements. Second, the vast number of LLMs (over 10,000) makes it impractical to test them all within a single study. Future work could involve building a centralized platform that stores LLM parameters and facilitates the replication of research findings. Third, the previously free academic Twitter API for collecting census data is no longer available. This necessitates exploring paid alternatives for future studies involving census tweets. Fourth, besides X or Twitter, other social media platforms such as Facebook, Instagram, and Threads can also serve as valuable sources of public opinion. However, accessing data from these platforms presents significant challenges due to strict privacy policies and data access restrictions. Unlike X or Twitter, where public data is more accessible, these platforms impose restrictions that limit large-scale data collection and analysis. More resources are required to collect data and analyze such variations in pattern. Finally, studying X or Twitter users may not fully reflect general public opinion, as the characteristics of X or Twitter users may not be representative of the general population. X or Twitter users from different backgrounds may exhibit varying behaviors. Also, our data do not allow us to assess differences in sentiment and opinion between health care professionals and patients. A comparative study focusing on these differences would be valuable to capture a full spectrum of perspectives. Thus, the generalization of our results to the general population should be approached with caution.

XSL•FO
RenderX

Conclusion

The public's cautious optimism serves as a call to action for both technological developers and regulatory bodies to prioritize transparency, ethical standards, and the safeguarding of patient data as integral components of AI development in health care. Ensuring these measures should not only build public confidence but also enhance the efficacy and acceptance of AI in the health care sector overall. Further, divergent opinions across different healthcare AI topics indicate further research is warranted to better understand where AI can best add value without compromising ethical standards. Future studies should continue to track such public sentiment discussion and its correlation with real-world AI integration outcomes in healthcare. Over time, deeper insights into how public perceptions will evolve, effectively guiding successful LLM adoption in the health care space.

Conflicts of Interest

None declared.

References

1. Oehler A, Horn M. Does ChatGPT provide better advice than robo-advisors? Finance Research Letters 2024 Feb;60:104898. [doi: 10.1016/j.frl.2023.104898]
2. Rawas S. ChatGPT: Empowering lifelong learning in the digital age of higher education. Educ Inf Technol 2024 Apr;29(6):6895-6908. [doi: 10.1007/s10639-023-12114-8]
3. Ságodi Z, Siket I, Ferenc R. Methodology for code synthesis evaluation of LLMs presented by a case study of ChatGPT and copilot. IEEE Access 2024;12:72303-72316. [doi: 10.1109/ACCESS.2024.3403858]
4. Ong JCL, Seng BJJ, Law JZF, et al. Artificial intelligence, ChatGPT, and other large language models for social determinants of health: Current state and future directions. Cell Rep Med 2024 Jan 16;5(1):101356. [doi: 10.1016/j.xcrm.2023.101356] [Medline: 38232690]
5. Kim J, Lee J, Jang KM, Lourentzou I. Exploring the limitations in how ChatGPT introduces environmental justice issues in the United States: A case study of 3,108 counties. Telematics and Informatics 2024 Feb;86:102085. [doi: 10.1016/j.tele.2023.102085]
6. Frieden TR. A framework for public health action: the health impact pyramid. Am J Public Health 2010 Apr;100(4):590-595. [doi: 10.2105/AJPH.2009.185652] [Medline: 20167880]
7. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century: National Academies Press; 2001.
8. Wang Y, Kung L, Byrd TA. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol Forecast Soc Change 2018 Jan;126:3-13. [doi: 10.1016/j.techfore.2015.12.019]
9. Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Med Educ 2023 Feb 8;9(1):e45312. [doi: 10.2196/45312] [Medline: 36753318]
10. Andreadis K, Newman DR, Twan C, Shunk A, Mann DM, Stevens ER. Mixed methods assessment of the influence of demographics on medical advice of ChatGPT. J Am Med Inform Assoc 2024 Sep 1;31(9):2002-2009. [doi: 10.1093/jamia/ocae086] [Medline: 38679900]
11. Shah SB, Thapa S, Acharya A, et al. Navigating the web of disinformation and misinformation: large language models as double-edged swords. IEEE Access 2024:1-1. [doi: 10.1109/ACCESS.2024.3406644]
12. Lehman K, Aroney E, Wu I. ChatGPT in private practice: the opportunities and pitfalls of novel technology. Australas Psychiatry 2024 Jun;32(3):214-219. [doi: 10.1177/10398562241241473] [Medline: 38545872]
13. Gohil S, Vuik S, Darzi A. Sentiment analysis of health care tweets: review of the methods used. JMIR Public Health Surveill 2018 Apr 23;4(2):e43. [doi: 10.2196/publichealth.5789] [Medline: 29685871]
14. Morita PP, Zakir Hussain I, Kaur J, Lotto M, Butt ZA. Tweeting for health using real-time mining and artificial intelligence-based analytics: design and development of a big data ecosystem for detecting and analyzing misinformation on Twitter. J Med Internet Res 2023 Jun 9;25(1):e44356. [doi: 10.2196/44356] [Medline: 37294603]
15. Bag S, Kumar SK, Tiwari MK. An efficient recommendation generation using relevant Jaccard similarity. Inf Sci (Ny) 2019 May;483:53-64. [doi: 10.1016/j.ins.2019.01.023]
16. Jain PK, Saravanan V, Pamula R. A hybrid CNN-LSTM: a deep learning approach for consumer sentiment analysis using qualitative user-generated contents. ACM Trans Asian Low-Resour Lang Inf Process 2021 Sep 30;20(5):1-15. [doi: 10.1145/3457206]
17. Jockers M. Syuzhet: extracts sentiment and sentiment-derived plot arcs from text. The Comprehensive R Archive Network. 2023. URL: https://cran.r-project.org/web/packages/syuzhet/syuzhet.pdf [accessed 2025-06-11]
18. Roberts ME, Stewart BM, Tingley D. stm: an R Package for Structural Topic Models. J Stat Softw 2019;91:1-40. [doi: 10.18637/jss.v091.i02]
19. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text]

XSL•FO
RenderX

20. Mohammadi S, Moradi P, Trufanov A, Jafari GR. A structural approach to detecting opinion leaders in Twitter by random matrix theory. Sci Rep 2023 Dec 8;13(1):21788. [doi: 10.1038/s41598-023-48682-5] [Medline: 38066201]
21. Rehman AU, Jiang A, Rehman A, Paul A, din S, Sadiq MT. Identification and role of opinion leaders in information diffusion for online discussion network. J Ambient Intell Human Comput 2023 Nov;14(11):15301-15313. [doi: 10.1007/s12652-019-01623-5]
22. Hussain I, Zaidi SMH, Khan U, Ahmed A. Sentiment analysis classification of ChatGPT tweets using machine learning and deep learning algorithms. Journal of Computing & Biomedical Informatics 2024;8(1).
23. Kumar R, Ayyasamy RK, Sangodiah A, Krishnan K, Jebna AK, Theam LJ. Sentiment analysis of ChatGPT healthcare discourse: insights from Twitter data. Presented at: 2023 15th International Conference on Software, Knowledge, Information Management and Applications (SKIMA; Dec 8-10, 2023; Kuala Lumpur, Malaysia. [doi: 10.1109/SKIMA59232.2023.10387306]
24. Sabir A, Ali HA, Aljabery MA. ChatGPT tweets sentiment analysis using machine learning and data classification. IJCAI (U S) 2024;48(7). [doi: 10.31449/inf.v48i7.5535]
25. Brainard J. Journals take up arms against AI-written text. Science 2023 Feb 24;379(6634):740-741. [doi: 10.1126/science.adh2762] [Medline: 36821673]
26. Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove. Nature New Biol 2023 Jan;613(7945):620-621. [doi: 10.1038/d41586-023-00107-z] [Medline: 36653617]
27. Thorp HH. ChatGPT is fun, but not an author. Science 2023 Jan 27;379(6630):313-313. [doi: 10.1126/science.adg7879]
28. Zielinski C, Winker MA, Aggarwal R, et al. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. Colomb Med (Cali) 2023;54(3):e1015868. [doi: 10.25100/cm.v54i3.5868] [Medline: 38089825]
29. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus 2023 Feb;15(2):e35179. [doi: 10.7759/cureus.35179] [Medline: 36811129]
30. van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. Nature New Biol 2023 Feb;614(7947):224-226. [doi: 10.1038/d41586-023-00288-7] [Medline: 36737653]
31. Ali SR, Dobbs TD, Hutchings HA, Whitaker IS. Using ChatGPT to write patient clinic letters. Lancet Digit Health 2023 Apr;5(4):e179-e181. [doi: 10.1016/S2589-7500(23)00048-1] [Medline: 36894409]
32. Lawson McLean A. Artificial intelligence in surgical documentation: a critical review of the role of large language models. Ann Biomed Eng 2023 Dec;51(12):2641-2642. [doi: 10.1007/s10439-023-03282-2] [Medline: 37338711]
33. Patel SB, Lam K. ChatGPT: the future of discharge summaries? Lancet Digit Health 2023 Mar;5(3):e107-e108. [doi: 10.1016/S2589-7500(23)00021-3] [Medline: 36754724]
34. Hosseini M, Horbach S. Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. Res Integr Peer Rev 2023 May 18;8(1):4. [doi: 10.1186/s41073-023-00133-5] [Medline: 37198671]
35. Ito N, Kadomatsu S, Fujisawa M, et al. The accuracy and potential racial and ethnic biases of GPT-4 in the diagnosis and triage of health conditions: evaluation study. JMIR Med Educ 2023 Nov 2;9:e47532. [doi: 10.2196/47532] [Medline: 37917120]
36. Li J, Dada A, Puladi B, Kleesiek J, Egger J. ChatGPT in healthcare: a taxonomy and systematic review. Comput Methods Programs Biomed 2024 Mar;245:108013. [doi: 10.1016/j.cmpb.2024.108013] [Medline: 38262126]
37. Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T. Diagnostic accuracy of differential-diagnosis lists generated by generative pretrained transformer 3 chatbot for clinical vignettes with common chief complaints: a pilot study. Int J Environ Res Public Health 2023 Feb 15;20(4):3378. [doi: 10.3390/ijerph20043378] [Medline: 36834073]
38. Kuroiwa T, Sarcon A, Ibara T, et al. The potential of ChatGPT as a self-diagnostic tool in common orthopedic diseases: exploratory study. J Med Internet Res 2023 Sep 15;25:e47621. [doi: 10.2196/47621] [Medline: 37713254]
39. Lecler A, Duron L, Soyer P. Revolutionizing radiology with GPT-based models: current applications, future possibilities and limitations of ChatGPT. Diagn Interv Imaging 2023 Jun;104(6):269-274. [doi: 10.1016/j.diii.2023.02.003]
40. Gupta R, Pande P, Herzog I, et al. Application of ChatGPT in cosmetic plastic surgery: ally or antagonist? Aesthet Surg J 2023 Jun 14;43(7):NP587-NP590. [doi: 10.1093/asj/sjad042] [Medline: 36840479]
41. Javaid M, Haleem A, Singh RP. ChatGPT for healthcare services: an emerging stage for an innovative perspective. BenchCouncil Transactions on Benchmarks, Standards and Evaluations 2023 Feb;3(1):100105. [doi: 10.1016/j.tbench.2023.100105]
42. Lahat A, Shachar E, Avidan B, Shatz Z, Glicksberg BS, Klang E. Evaluating the use of large language model in identifying top research questions in gastroenterology. Sci Rep 2023 Mar 13;13(1):4164. [doi: 10.1038/s41598-023-31412-2] [Medline: 36914821]
43. Guo D, Chen H, Wu R, Wang Y. AIGC challenges and opportunities related to public safety: a case study of ChatGPT. Journal of Safety Science and Resilience 2023 Dec;4(4):329-339. [doi: 10.1016/j.jnlssr.2023.08.001]
44. Kalam KT, Rahman JM, Islam M, Dewan SMR. ChatGPT and mental health: friends or foes? Health Sci Rep 2024 Feb;7(2):e1912. [doi: 10.1002/hsr2.1912] [Medline: 38361805]

XSL•FO
RenderX

45. Singh OP. Artificial intelligence in the era of ChatGPT - opportunities and challenges in mental health care. Indian J Psychiatry 2023 Mar;65(3):297-298. [doi: 10.4103/indianjpsychiatry.indianjpsychiatry_112_23] [Medline: 37204980]
46. Elyoseph Z, Levkovich I. Beyond human expertise: the promise and limitations of ChatGPT in suicide risk assessment. Front Psychiatry 2023;14:1213141. [doi: 10.3389/fpsyt.2023.1213141] [Medline: 37593450]
47. Imran N, Hashmi A, Imran A. Chat-GPT: opportunities and challenges in child mental healthcare. Pak J Med Sci 2023;39(4):1191-1193. [doi: 10.12669/pjms.39.4.8118] [Medline: 37492313]
48. Ashraf M. A systematic review on the potential of AI and ChatGPT for parental support and child well-being. arXiv. Preprint posted online on May 23, 2024. [doi: 10.48550/arXiv.2407.09492]
49. Regulatory Considerations on Artificial Intelligence for Health: World Health Organization; 2023. URL: https://www.who.int/publications/i/item/9789240078871 [accessed 2025-06-11]

Abbreviations

AI: artificial intelligence
GDPR: General Data Protection Regulation
HIPAA: Health Insurance Portability and Accountability Act
LLM: large language model
STM: structural topic modeling

Edited by M Haupt; submitted 18.07.24; peer-reviewed by A Rasool, C Zielinski, L Raymond Guo, LH Yao, M Chatzimina; revised version received 23.03.25; accepted 24.03.25; published 23.06.25.

Please cite as:
Baxter P, Li MH, Wei J, Koizumi N
Public Versus Academic Discourse on ChatGPT in Health Care: Mixed Methods Study
JMIR Infodemiology 2025;5:e64509
URL: https://infodemiology.jmir.org/2025/1/e64509
doi:10.2196/64509

© Patrick Baxter, Meng-Hao Li, Jiaxin Wei, Naoru Koizumi. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 23.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Infodemic Versus Viral Information Spread: Key Differences and Open Challenges

Matteo Cinelli[1], PhD; Francesco Gesualdo[2], MD

[1]Department of Computer Science, Sapienza University of Rome, Via Regina Elena 295, Rome, Italy

[2]Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy

Corresponding Author:
Matteo Cinelli, PhD
Department of Computer Science, Sapienza University of Rome, Via Regina Elena 295, Rome, Italy

Abstract

As we move beyond the COVID-19 pandemic, the risk of future infodemics remains significant, driven by emerging health crises and the increasing influence of artificial intelligence in the information ecosystem. During periods of apparent stability, proactive efforts to advance infodemiology are essential for enhancing preparedness and improving public health outcomes. This requires a thorough examination of the foundations of this evolving discipline, particularly in understanding how to accurately identify an infodemic at the appropriate time and scale, and how to distinguish it from other processes of viral information spread, both within and outside the realm of public health. In this paper, we integrate expertise from data science and public health to examine the key differences between information production during an infodemic and viral information spread. We explore both clear and subtle distinctions, including context and contingency (ie, the association of an infodemic and viral information spread with a health crisis); information dynamics in terms of volume, spread, and predictability; the role of misinformation and information voids; societal impact; and mitigation strategies. By analyzing these differences, we highlight challenges and open questions. These include whether an infodemic is solely associated with pandemics or whether it could arise from other health emergencies; if infodemics are limited to health-related issues or if they could emerge from crises initially unrelated to health (like climate events); and whether infodemics are exclusively global phenomena or if they can occur on national or local scales. Finally, we propose directions for future quantitative research to help the scientific community more robustly differentiate between these phenomena and develop tailored management strategies.

(JMIR Infodemiology 2025;5:e57455) doi:10.2196/57455

KEYWORDS

infodemic; information spreading; infodemiology; misinformation; artificial intelligence; information virality; public health; multidisciplinary; data science; AI; difference; challenge

Introduction

The definition of infodemic has evolved over the years. It started from being an "epidemic of information," as defined by Rothkopf in 2003 [1] in the context of the severe acute respiratory syndrome (SARS) outbreak, to then include the element of misinformation [2-4], especially when the concept gained momentum during the COVID-19 pandemic. Since then, thanks to the efforts of the World Health Organization (WHO) and collaboration with academics, infodemiology, that is, "the science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy" [2,5], has evolved from a primarily descriptive discipline into a more comprehensive field [6]. This transformation has integrated insights from behavioral, medical, and complex systems sciences, leading to a more comprehensive understanding that encompasses a broader range of phenomena and challenges.

From the earliest WHO reports on the COVID-19 infodemic and the global technical consultation of April 2020 [7], the infodemic was defined as "an overabundance of information—some accurate and some not—that occurs during an epidemic." Later definitions introduced the concept of information overflow [8], emphasizing the difficulty individuals face in identifying trustworthy information during crises. This broader conceptualization acknowledges that an infodemic extends beyond misinformation, as it is shaped by public concerns, information voids, and digital transformation processes that amplify content circulation.

Recent discussions, including those from the Fifth WHO Infodemic Management Conference [9], have further refined this perspective, by highlighting the layered and multifaceted nature of infodemics. The conference underscored how infodemics emerge from a confluence of structural, behavioral, and technological factors, including the characteristics of the information ecosystem, the role of digital platforms in amplifying or mitigating information spread, the psychological

XSL•FO
RenderX

and social drivers of user engagement, and the dimension of trust in public institutions.

As digital platforms continue to evolve, new technological advancements further complicate the landscape: after the COVID-19 pandemic, the release of large language models to the general public raised discussions on the potential role of generative artificial intelligence in fueling future infodemics [10], adding another dimension to the challenge of managing information flows during crises.

While acknowledging this complexity, in this paper, we focus on one specific dimension of the infodemic phenomenon: the quantification of information production. By narrowing our analysis to this aspect, we aim to clarify how excess information during an infodemic differs from a general viral information spread. A more precise characterization of the quantitative aspects of information overabundance can help differentiate infodemics from other forms of information dissemination, guiding efforts to improve detection methods and to ensure that public health strategies can respond to crises in an even more informed manner.

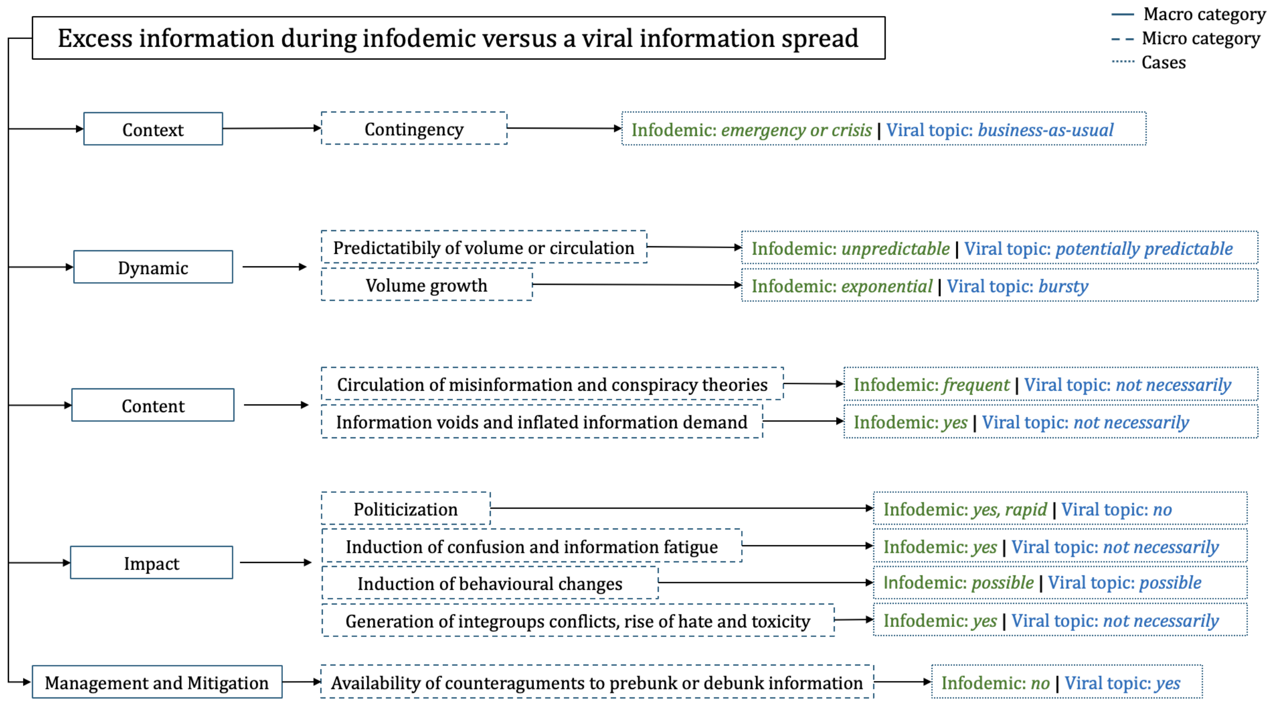
The concept of an overabundance or excess of information, which underpins the current definition of infodemic, could be better described quantitatively by focusing on different aspects of information production and circulation. First, to our knowledge, there is broad recognition that individuals are continuously exposed to vast amounts of information, raising the question of how to benchmark overabundance. In other words, should it be measured with respect to content production regarding different topics of discourse or with respect to the same topic (assuming it was present in the debate) in a previous time window? Additionally, are different time windows comparable on social media, given the high turnover of users within digital communities? Second, should a spike in interest or a heated public discourse about a health-related issue happening on the web automatically be labeled as an infodemic or can it "just" be considered a noteworthy episode? And, further, can we provide estimates of the minimum lasting time for information overabundance to become an infodemic? In this scenario, we believe that understanding the differences between excess information during an infodemic and information virality can be useful to advance our understanding of infodemic processes. This distinction can inform the development of more precise indicators for enhanced infoveillance systems, enabling better monitoring and intervention strategies.

For these reasons, in what follows, we will list a set of key points that, building upon recent literature, differentiate excess information during an infodemic from a viral process of information diffusion and that could be turned, in certain instances, into viable measurements, with the aim of pointing towards new research directions. Instrumental to our purpose is to clarify what we mean by a viral process of information spread: a viral information spread or virality refers to the rapid and widespread dissemination of information or content, often through social networks and web-based platforms [11]. It follows that a viral process also generates an "overabundance" of information that, in theory, should be distinguishable from that generated by an infodemic in many respects.

Differences Between an Infodemic and a Viral Information Spread

In this section, we list a set of differences between the excess information observed during an infodemic and a viral information spread, considering up to 5 characterizing dimensions used for classification. Categories and differences are presented in a conceptual map in Figure 1.

Figure 1. Summary of differences between excess information during an infodemic and a viral information spread.



XSL•FO
RenderX

Context and Contingency

An infodemic is associated with a (health) emergency or crisis, occurring in contexts where there is a heightened need for timely information and leading to an increased need for information dissemination to the public. In contrast, a viral topic may emerge during periods of "business-as-usual," meaning it does not necessarily coincide with any crisis or emergency. The timing and context of these events are crucial for differentiating an infodemic from a viral topic, with infodemics being crisis-driven and viral topics being more routine. However, in some cases, the initial magnitude of a health crisis may not be immediately clear, making it difficult to distinguish between a transient viral circulation of information and an evolving infodemic that could escalate in magnitude and impact.

Dynamic: Volume Growth and Predictability

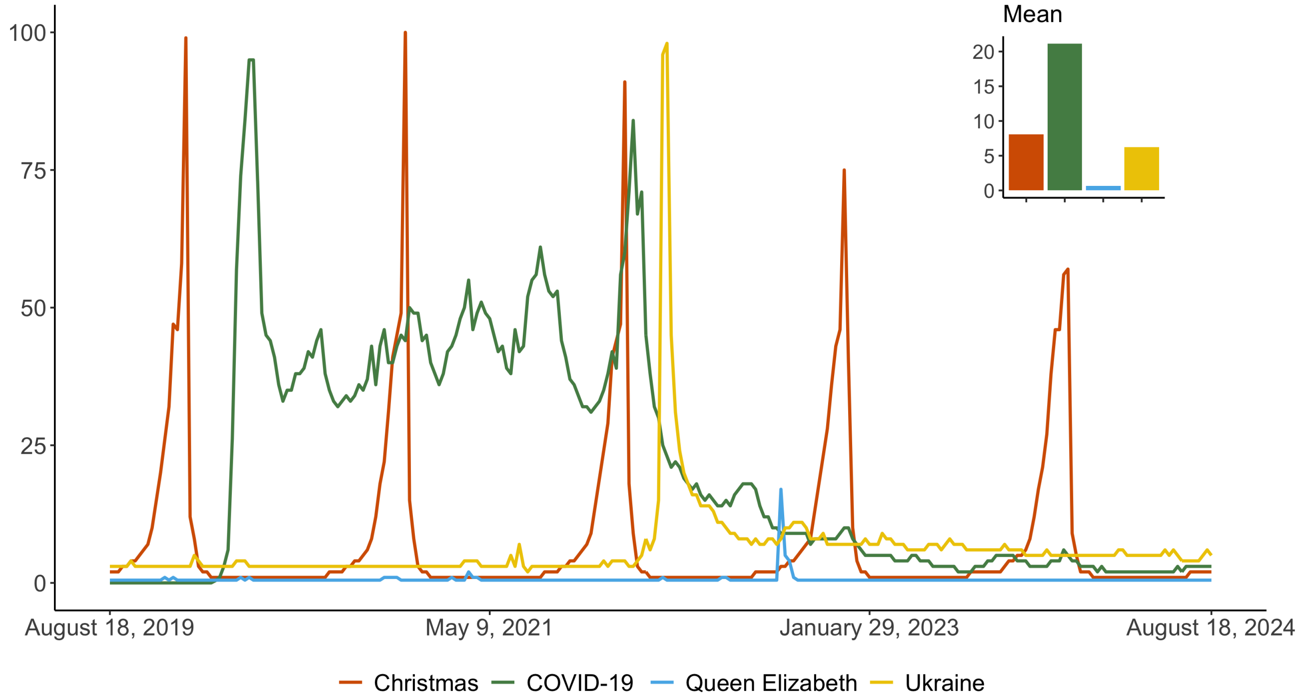
The dynamics of how information spreads during infodemics differ significantly from those of viral topics. The production of information during infodemics is characterized by its unpredictable volume and spread. Content volume often grows exponentially during an infodemic and remains steady over extended periods. On the other hand, specific topics characterized by a certain seasonality or scheduled events of great resonance (eg, elections), that fall into the definition of virality, tend to have more predictable patterns. A similar reasoning holds for topics displaying sublinear, superlinear, or wavy behaviors in their volume growth, whose evolution, unlike in the case of infodemics, can be inferred by precise mathematical models [12,13]. However, any output by a forecasting model would require some initial input data, meaning that no predictions can be made until some pieces of content are released. It is also worth reminding that any expectation or prediction related to events in the social sphere carries a degree of uncertainty that depends on the contingency of exogenous events. Such uncertainty can lead to significant fluctuations in the phenomenon being studied, potentially rendering model predictions unreliable. This is a fundamental characteristic of complex systems and a key challenge in studying a constantly evolving society.

An example of the differences between viral topics and infodemics can be observed by comparing 4 cases or events that have gained substantial collective attention, namely the COVID-19 pandemic, the Russo-Ukrainian conflict, the death of Queen Elizabeth II, and Christmas. To illustrate this, and under the hypothesis that search interest in a topic correlates with content production [14], we used Google Trends data [15], which provide weekly time series of worldwide searches on Google over a 5-year period. We analyzed the search volumes for the following topics: "COVID-19," "Queen Elizabeth," "Ukraine," and "Christmas," though similar results were obtained when substituting "Russia" for "Ukraine." The weekly search volumes for these 4 terms were normalized on a scale of 0 to 100, where 100 represents the highest recorded search volume across all 4 terms, and 50 indicates half of that peak. A score of 0 indicates insufficient data.

From Figure 2, we can observe a clear difference between the 4 terms. While the death of Queen Elizabeth II gathered significant attention and became a viral event, its trajectory is markedly different from the other cases. COVID-19 shows a trend that takes about 6 weeks to peak, after which it continues to receive a high volume of searches over an extended period. The Russo-Ukrainian conflict presents a different pattern, falling somewhere in between the other 2 cases. Search volume peaks within approximately 3 weeks, followed by a decline to much lower, though still notable, levels—perhaps suggestive of an infodemic (note also that the situation in Ukraine is listed among the WHO's health emergencies [16]). Lastly, the case of Christmas serves as an example of a seasonal and thus predictable topic, which still falls within the definition of virality.

XSL•FO
RenderX

Figure 2. Time series of relevant topics including 1-time virality (Queen Elizabeth), seasonal virality (Christmas), infodemic (COVID-19), and infodemic-like (Ukraine). Data were obtained by searching keywords on Google Trends. The inset shows the average value of normalized searches for each keyword.



Content: Circulation of Misinformation and Information Voids

Content-wise, infodemics are frequently associated with a flood of mis- or disinformation and conspiracy theories, which is not necessarily the case with viral topics. During infodemics, the increased production of content co-occurs with the presence of information voids and inflated information demand [17,18], where gaps in credible information contribute to the uncontrolled spread of both accurate and misleading content. In contrast, viral topics do not necessarily involve such voids or demands, and the spread of misinformation may not be as prevalent.

Impact: Politicization, Confusion, Behavioral Changes, Social Cohesion, and Toxicity

The impact of an infodemic can be far-reaching. Infodemics tend to become rapidly politicized [19,20], with misinformation being shared and consumed as it aligns with the views of specific social and political groups. This rapid politicization is less likely with viral topics. Furthermore, infodemics can induce confusion and information fatigue [21-23] (sometimes referred to as information overload or infoxication), contributing to significant behavioral changes among individuals [24]. Viral topics, while they may also influence behavior [25], do not necessarily lead to such widespread confusion or fatigue.

An infodemic can erode social cohesion [26] by generating intergroup conflicts and increasing levels of toxicity in the public discourse [27]. This is a significant consequence of infodemics, fueled by the spread of divisive misinformation. Viral topics, while they may stir debate, do not typically lead to the same level of social disruption or rise in toxicity. At present, no systematic research has examined the correlation between topic virality and toxic speech usage. However, in the case of infodemics, researchers have observed a consistent association with toxic speech, starting with Sinophobia at the onset of the COVID-19 pandemic [27] and continuing with heated debates over vaccines in later stages [28], whereas in the case of viral topics, such occurrences have been limited to only a few instances [29,30].

Management and Mitigation: Availability of Counterarguments

When it comes to managing and mitigating the effects of infodemics and viral topics, a stark contrast emerges. During an infodemic, there is often a lack of established counterarguments, making it challenging to pre- or debunk misinformation [20]. In such situations, the scientific community may need to rely heavily on the best evidence available at the time due to the absence of consensus statements or previously established facts. On the other hand, for viral topics related to public health, such as vaccines, there is often an abundance of counterarguments from past debates that can be employed to address mis- or disinformation. These resources can also be effective in crafting prebunking strategies.

For instance, in recent years, a decline in coverage rates for the human papillomavirus (HPV) vaccine has been documented in Ireland [31], Denmark [32,33] and Colombia [34], primarily due to the viral circulation of web-based and social media content questioning the vaccine's safety profile. While one could argue that the spread of misinformation in these cases can be described as an infodemic, this characterization remains debatable due to insufficient data for measuring information overabundance in digital and physical environments. Instead, it can be more accurately hypothesized that these incidents represent a viral spread of misinformation. In some cases, the impact of this misinformation has been long-lasting. However, in most instances, the availability of robust safety data on the HPV vaccine has enabled public health agencies to conduct

XSL•FO
RenderX

communication campaigns that effectively restored public confidence in the HPV vaccine. By contrast, during the COVID-19 pandemic, public health agencies faced a tougher challenge in countering vaccine hesitancy, as they initially lacked real-world safety data to support their efforts.

Discussion

Open Points

As a concluding step, moved by the reasoning on the difference between excess information in an infodemic and a viral information spread, we discuss certain open points regarding the infodemic phenomenon.

Is an infodemic solely associated with a pandemic? Not necessarily; other health emergencies, as in the case of abortion [35], may generate an infodemic. Furthermore, health emergencies can arise from various factors beyond the spread of communicable diseases. For instance, the WHO includes events such as conflicts in Sudan and the neighboring countries [36] and other health and humanitarian crises [37] in the list of present and past health emergencies.

Is an infodemic limited to health-related issues? It is probable that future infodemics will arise during crises initially unrelated to health. For example, the increasing likelihood of extreme climate events over time may trigger prolonged (infodemic-like) debates fueled by various factors such as disasters or government interventions. In such cases, as crises are likely to touch every societal layer and health emergencies are deeply interconnected with socioeconomic and environmental factors, the health community would be anyway involved. Recent examples include the Israeli–Palestinian and the Russo-Ukrainian conflicts. On a related note, an infodemic might have also occurred following the launch of ChatGPT, an event that engaged a broad audience in the artificial intelligence debate, raising concerns and causing confusion about the role of this new technology in our future.

Is the infodemic exclusively a global phenomenon? Given that health emergencies and topics of public concern are not necessarily international (ie, they are not public health emergencies of international concern), an infodemic may occur at a national level or even at a more local scale. For instance, not all countries may be subject to an overabundance of information at the same time, switching in turn from an infodemic regime to a noninfodemic one. An example regarding the limited geographical reach of an infodemic may be represented by the previously mentioned case of the contagion of psychogenic reactions and the consequent HPV vaccine drop in Colombia due to the spread of videos, published on social media platforms, documenting false adverse reactions post vaccination [34].

Conclusion

In this paper, we have established a list of key differences that distinguish the excess information production during an infodemic from viral episodes of information diffusion. One key point is that the primary distinctions often become apparent in the medium to long term, including the overabundance of information, the protracted virality over time, the confusion of individuals, and the clear consequences for society. It follows that, in the shorter period, the two phenomena may appear somewhat indistinguishable, most likely as in the case of the transition from an epidemic to a pandemic. However, considering the vast amount of data and early signals potentially detectable on web-based and social media platforms, a significant phenomenon such as an infodemic cannot be regarded as something one realizes is happening as it occurs. This conceptual work aims to contribute to a better understanding of infodemics and information virality and may serve as a baseline for developing improved measures of information overabundance to support the multilayered process of infoveillance. In practice, possible efforts to measure the circulation of information during an infodemic could include developing causal models of infodemic evolution based on a series of independent variables, both epidemiological and nonepidemiological, to understand how non–health-related factors contribute to its progression. Establishing and utilizing appropriate baseline models to measure information overabundance across topics of similar and different natures would also be crucial. Additionally, leveraging the wide availability of viral web-based phenomena to create a set of case studies could help to compare the growth rates characterizing the initial waves of information diffusion of a new potential infodemic, thus providing reliable early warnings. While this may not be a primary concern for public health, it is valuable to highlight how different information-spreading processes may share certain characteristics (eg, overabundance) with the infodemic. Recognizing these commonalities could allow research to incorporate insights and practices developed in the study of infodemics into other contexts and vice versa, thus enhancing an always relevant multidisciplinary debate around this domain of knowledge. Future research efforts could shift the focus on how the dynamics of information consumption differ in infodemic versus viral contexts, as health-related content tends to have even less impact on behavior when individuals are not experiencing a personal, family, or social event that alters their need for information.

Finally, beyond serving as a checklist for infodemic recognition, the listed differences account for observations in the data regarding the COVID-19 infodemic and other topics, providing a challenging perspective that the public health community could employ to enhance existing infodemic management frameworks.

Acknowledgments

MC acknowledges support from Project SEED (SP122184858BEDB3).

XSL•FO
RenderX

Data Availability

The data generated and analyzed in this paper are publicly available on Google Trends [15].

Authors' Contributions

MC conceptualized and drafted the paper. FG contributed to the conception of the paper and revised the manuscript.

Conflicts of Interest

FG received an honorarium for participating in meetings by Merck Sharp & Dohme (MSD) and Moderna.

References

1. Rothkopf DJ. When the buzz bites back. The Washington Post. 2003. URL: https://www.washingtonpost.com/archive/opinions/2003/05/11/when-the-buzz-bites-back/bc8cd84f-cab6-4648-bf58-0277261af6cd/ [accessed 2025-04-22]
2. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med 2002 Dec 15;113(9):763-765. [doi: 10.1016/s0002-9343(02)01473-0] [Medline: 12517369]
3. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb 29;395(10225):676. [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
4. Cinelli M, Quattrociocchi W, Galeazzi A, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 6;10(1):16598. [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
5. Mavragani A. Infodemiology and infoveillance: scoping review. J Med Internet Res 2020 Apr 28;22(4):e16206. [doi: 10.2196/16206] [Medline: 32310818]
6. Briand SC, Cinelli M, Nguyen T, et al. Infodemics: a new challenge for public health. Cell 2021 Dec 9;184(25):6010-6014. [doi: 10.1016/j.cell.2021.10.031] [Medline: 34890548]
7. Tangcharoensathien V, Calleja N, Nguyen T, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced who technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659. [doi: 10.2196/19659] [Medline: 32558655]
8. Calleja N, AbdAllah A, Abad N, et al. A public health research agenda for managing infodemics: methods and results of the first WHO infodemiology conference. JMIR Infodemiology 2021;1(1):e30979. [doi: 10.2196/30979] [Medline: 34604708]
9. Wilhelm E, Ballalai I, Belanger ME, et al. Measuring the burden of infodemics: summary of the methods and results of the Fifth WHO Infodemic Management Conference. JMIR Infodemiology 2023;3(1):e44207. [doi: 10.2196/44207] [Medline: 37012998]
10. De Angelis L, Baglivo F, Arzilli G, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front Public Health 2023;11:1166120. [doi: 10.3389/fpubh.2023.1166120] [Medline: 37181697]
11. Berger J, Milkman KL. What makes online content viral? Journal of Marketing Research 2012 Apr;49(2):192-205. [doi: 10.1509/jmr.10.0353]
12. Etta G, Sangiorgio E, Di Marco N, et al. Characterizing engagement dynamics across topics on Facebook. PLoS ONE 2023;18(6):e0286150. [doi: 10.1371/journal.pone.0286150] [Medline: 37379268]
13. Barabási AL. The origin of bursts and heavy tails in human dynamics. Nature New Biol 2005 May 12;435(7039):207-211. [doi: 10.1038/nature03459] [Medline: 15889093]
14. Gravino P, Prevedello G, Galletti M, Loreto V. The supply and demand of news during COVID-19 and assessment of questionable sources production. Nat Hum Behav 2022 Aug;6(8):1069-1078. [doi: 10.1038/s41562-022-01353-3] [Medline: 35606514]
15. Google Trends. URL: https://www.google.com/trends [accessed 2024-08-22]
16. Ukraine emergency. World Health Organization. URL: https://www.who.int/emergencies/situations/ukraine-emergency [accessed 2024-09-04]
17. Purnat TD, Vacca P, Czerniak C, et al. Infodemic signal detection during the COVID-19 pandemic: development of a methodology for identifying potential information voids in online conversations. JMIR Infodemiology 2021;1(1):e30971. [doi: 10.2196/30971] [Medline: 34447926]
18. Piltch-Loeb R, James R, Albrecht SS, et al. What were the information voids? A qualitative analysis of questions asked by Dear Pandemic readers between August 2020-August 2021. J Health Commun 2023 Apr 7;28(sup1):25-33. [doi: 10.1080/10810730.2023.2214986] [Medline: 37390014]
19. Gagliardone I, Diepeveen S, Findlay K, Olaniran S, Pohjonen M, Tallam E. Demystifying the COVID-19 infodemic: conspiracies, context, and the agency of users. Social Media + Society 2021 Jul;7(3):20563051211044233. [doi: 10.1177/20563051211044233]
20. van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. Nat Med 2022 Mar;28(3):460-467. [doi: 10.1038/s41591-022-01713-6] [Medline: 35273402]
21. Lilleholt L, Zettler I, Betsch C, Böhm R. Development and validation of the pandemic fatigue scale. Nat Commun 2023 Oct 10;14(1):6352. [doi: 10.1038/s41467-023-42063-2] [Medline: 37816702]

XSL•FO
RenderX

22. Kim S, So J. How message fatigue toward health messages leads to ineffective persuasive outcomes: examining the mediating roles of reactance and inattention. J Health Commun 2018;23(1):109-116. [doi: 10.1080/10810730.2017.1414900] [Medline: 29272208]
23. Shi CF, So MC, Stelmach S, Earn A, Earn DJD, Dushoff J. From science to politics: COVID-19 information fatigue on YouTube. BMC Public Health 2022 Apr 23;22(1):816. [doi: 10.1186/s12889-022-13151-7] [Medline: 35461254]
24. Loomba S, de Figueiredo A, Piatek SJ, de Graaf K, Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat Hum Behav 2021 Mar;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
25. Centola D. The spread of behavior in an online social network experiment. Science 2010 Sep 3;329(5996):1194-1197. [doi: 10.1126/science.1185231] [Medline: 20813952]
26. Jewett RL, Mah SM, Howell N, Larsen MM. Social cohesion and community resilience during COVID-19 and pandemics: a rapid scoping review to inform the United Nations Research Roadmap for COVID-19 Recovery. Int J Health Serv 2021 Jul;51(3):325-336. [doi: 10.1177/0020731421997092] [Medline: 33827308]
27. Tahmasbi F, Schild L, Ling C, et al. “Go eat a bat, Chang!”: on the emergence of Sinophobic behavior on web communities in the face of COVID-19. Presented at: WWW ’21; Apr 19-23, 2021; Ljubljana, Slovenia. [doi: 10.1145/3442381.3450024]
28. Stevens H, Rasul ME, Oh YJ. Emotions and incivility in vaccine mandate discourse: natural language processing insights. JMIR Infodemiology 2022;2(2):e37635. [doi: 10.2196/37635] [Medline: 36188420]
29. Lupu Y, Sear R, Velásquez N, et al. Offline events and online hate. PLoS ONE 2023;18(1):e0278511. [doi: 10.1371/journal.pone.0278511] [Medline: 36696388]
30. Avalle M, Di Marco N, Etta G, et al. Persistent interaction patterns across social media platforms and over time. Nature New Biol 2024 Apr;628(8008):582-589. [doi: 10.1038/s41586-024-07229-y] [Medline: 38509370]
4. Corcoran B, Clarke A, Barrett T. Rapid response to HPV vaccination crisis in Ireland. The Lancet 2018 May;391(10135):2103. [doi: 10.1016/S0140-6736(18)30854-7]
32. Humlum MK, Skipper N, Thingholm PR. Vaccine hesitancy and differential susceptibility to media coverage: a critical documentary led to substantial reductions in human papillomavirus vaccine uptake in Denmark. Med Decis Making 2021 Jul;41(5):550-558. [doi: 10.1177/0272989X211003589] [Medline: 33899553]
5. Hansen PR, Schmidtblaicher M, Brewer NT. Resilience of HPV vaccine uptake in Denmark: decline and recovery. Vaccine (Auckl) 2020 Feb;38(7):1842-1848. [doi: 10.1016/j.vaccine.2019.12.019]
34. Simas C, Munoz N, Arregoces L, Larson HJ. HPV vaccine confidence and cases of mass psychogenic illness following immunization in Carmen de Bolivar, Colombia. Hum Vaccin Immunother 2019;15(1):163-166. [doi: 10.1080/21645515.2018.1511667] [Medline: 30118381]
35. Pagoto SL, Palmer L, Horwitz-Willis N. The next infodemic: abortion misinformation. J Med Internet Res 2023 May 4;25:e42582. [doi: 10.2196/42582] [Medline: 37140975]
36. Sudan emergency. World Health Organization. URL: https://www.who.int/emergencies/situations/sudan-emergency [accessed 2024-02-12]
37. List of emergencies. World Health Organization. URL: https://www.who.int/emergencies/situations [accessed 2024-02-12]

Abbreviations

HPV: human papillomavirus
SARS: severe acute respiratory syndrome
WHO: World Health Organization

Edited by T Purnat; submitted 17.02.24; peer-reviewed by F Medina, I Ballalai, JF Fuertes-Bucheli, P Lopalco; revised version received 14.02.25; accepted 15.02.25; published 07.05.25.

Please cite as:
Cinelli M, Gesualdo F
Infodemic Versus Viral Information Spread: Key Differences and Open Challenges
JMIR Infodemiology 2025;5:e57455
URL: https://infodemiology.jmir.org/2025/1/e57455
doi:10.2196/57455

© Matteo Cinelli, Francesco Gesualdo. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 7.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a

XSL•FO
RenderX

link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX