JMIR Infodemiology

Journal Impact Factor (JIF) (2023): 3.5
Volume 4 (2024) ISSN 2564-1891 Editor-in-Chief: Tim Ken Mackey, MAS, PhD

Contents

Viewpoints

Corrigenda and Addendas

Research Letter

Review

Ethical Considerations in Infodemic Management: Systematic Scoping Review

Federico Germani[1*], PhD; Giovanni Spitale[1*], PhD; Sandra Varaidzo Machiri[2], MPH; Calvin Wai Loon Ho[3], PhD; Isabella Ballalai[4], MD; Nikola Biller-Andorno[1], MD, MHBA, PhD; Andreas Alois Reis[5], MD, MSc

[1]Institute of Biomedical Ethics and History of Medicine, University of Zurich, Zurich, Switzerland

[2]Unit for High Impact Events Preparedness, Department of Epidemic and Pandemic Preparedness and Prevention, World Health Organization, Genève, Switzerland

[3]Monash Law School, Monash University, Melbourne, Australia

[4]Brazilian Immunization Society, Sao Paolo, Brazil

[5]Health Ethics and Governance Unit, Department of Research for Health, World Health Organization, Genève, Switzerland

[*]these authors contributed equally

Corresponding Author:
Andreas Alois Reis, MD, MSc
Health Ethics and Governance Unit
Department of Research for Health
World Health Organization
Av. Appia 20
Genève, 1202
Switzerland
Phone: 41 794673574
Email: reisa@who.int

Abstract

Background: During health emergencies, effective infodemic management has become a paramount challenge. A new era marked by a rapidly changing information ecosystem, combined with the widespread dissemination of misinformation and disinformation, has magnified the complexity of the issue. For infodemic management measures to be effective, acceptable, and trustworthy, a robust framework of ethical considerations is needed.

Objective: This systematic scoping review aims to identify and analyze ethical considerations and procedural principles relevant to infodemic management, ultimately enhancing the effectiveness of these practices and increasing trust in stakeholders performing infodemic management practices with the goal of safeguarding public health.

Methods: The review involved a comprehensive examination of the literature related to ethical considerations in infodemic management from 2002 to 2022, drawing from publications in PubMed, Scopus, and Web of Science. Policy documents and relevant material were included in the search strategy. Papers were screened against inclusion and exclusion criteria, and core thematic areas were systematically identified and categorized following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. We analyzed the literature to identify *substantive ethical principles* that were crucial for guiding actions in the realms of infodemic management and social listening, as well as related *procedural ethical principles*. In this review, we consider ethical principles that are extensively deliberated upon in the literature, such as equity, justice, or respect for autonomy. However, we acknowledge the existence and relevance of procedural practices, which we also consider as ethical principles or practices that, when implemented, enhance the efficacy of infodemic management while ensuring the respect of substantive ethical principles.

Results: Drawing from 103 publications, the review yielded several key findings related to ethical principles, approaches, and guidelines for practice in the context of infodemic management. Community engagement, empowerment through education, and inclusivity emerged as procedural principles and practices that enhance the quality and effectiveness of communication and social listening efforts, fostering trust, a key emerging theme and crucial ethical principle. The review also emphasized the significance of transparency, privacy, and cybersecurity in data collection.

Conclusions: This review underscores the pivotal role of ethics in bolstering the efficacy of infodemic management. From the analyzed body of literature, it becomes evident that ethical considerations serve as essential instruments for cultivating trust and credibility while also facilitating the medium-term and long-term viability of infodemic management approaches.

(JMIR Infodemiology 2024;4:e56307) doi:10.2196/56307

KEYWORDS

World Health Organization; ethics; infodemic management; social listening; review; infodemic; health emergency; health emergencies; misinformation; disinformation; scoping review; ethical principles; community engagement; empowerment; data privacy; effectiveness

Introduction

Background

In an age dominated by the digital dissemination of information and in an information ecosystem [1] in which the digital divide continues to be a global challenge, a new term has recently emerged, one that reflects the profound impact of the digital age on our information landscape: "infodemics," derived from the fusion of the terms *information* and *epidemic* [2,3]. The World Health Organization (WHO) defines an infodemic as the surge of information, both accurate and false, that inundates the public during acute health events such as outbreaks and epidemics. These infodemics hold far-reaching consequences, affecting public health, shaping societal decision-making, and influencing individual behaviors [4].

Infodemics can have a detrimental effect on public health efforts by raising questions, concerns, and doubts, which, if unresolved, may lead to information voids, alongside an overabundance of information, accurate or not, that can incite panic and confusion while hindering the dissemination of vital information [5-7]. In the digital era, distinguishing fact from fiction is a difficult task for the public [8,9], and the successful recognition of the accuracy of information requires information literacy and critical thinking [10-12]. Equally complex is the role of public health institutions and infodemic managers, as they navigate a polarized society that often rejects well-intentioned, safety-focused information [13].

Infodemic management refers to the processes and strategies put in place to monitor and improve the information ecosystem, including handling and controlling the spread of misinformation and excessive information during health crises [14]. Social listening involves monitoring and analyzing online and offline conversations (eg, on social media) to gain insights into public sentiment, concerns, and behaviors during such crises [15]. Social listening is an integral component of infodemic management, as it is an important source of insights to improve public health interventions [15,16].

The importance of ethics in infodemic management has been underscored amid the COVID-19 pandemic [17-19], as the lack of the integration of ethical approaches in the management of infodemics has diminished the potential for long-term effectiveness of these strategies, leading to a decline in trust for those same institutions safeguarding public health [20].

Ethical considerations encompass a broad scope, including determining when and under what circumstances intervention in the public information space is justifiable. Privacy, autonomy, trust, and the potential for censorship are all salient issues [21-24]. Neglecting these ethical dimensions can have profound consequences, eroding public trust and inadvertently causing harm [25]. Therefore, it is crucial to explore how ethical awareness may contribute to improving the effectiveness of these practices and how ethics can offer practical tools to solve the problems posed by infodemics. Infodemic management is a discipline driven by a moral imperative to improve the quality of the information ecosystem, thereby ensuring better public health outcomes and saving lives. In this review, ethical considerations are therefore defined as reasonings on morally significant principles intended to shape and guide the actions of stakeholders involved in and executing infodemic management practices [26]. We consider as "ethical considerations" not only those concerning the implementation of the infodemic management moral imperative but also those that enhance the short-term and long-term effectiveness of infodemic management.

The concept of infodemics gained prominence during the COVID-19 pandemic [2,7,14] and is expected to remain a pressing concern even as COVID-19 is no longer a major public health emergency [7,27,28]. Given the recent focus on integrating ethics into infodemic management and social listening [25,29], coupled with the expanding scope of infodemic-related challenges and the widespread adoption of social listening techniques for monitoring public health concerns and behaviors, it is imperative to investigate how existing literature addresses the ethical dimensions of infodemic management and social listening. This exploration can provide valuable insights to guide the integration of ethics into infodemic management and social listening practices while ensuring their effectiveness. The aim of this systematic scoping review is to pinpoint ethical considerations that have proven beneficial in the past, thereby informing and guiding future advancements in the field.

Recognizing the global urgency of integrating ethics into infodemic management and social listening practices, the WHO established the expert group (EG) on ethical considerations in infodemic management and social listening in 2023 [29]. This EG, coordinated by the Unit for High Impact Events Preparedness and the Health Ethics and Governance Unit at WHO, highlights the need for a comprehensive and ethically sound response to infodemics, with an eye to improve infodemic management practices as devised during the acute phases of the COVID-19 pandemic. The EG is developing practical tools to guide infodemic managers to ethically monitor infodemic trends and guide interventions. The EG discusses emerging dilemmas, practical applications of guidance in the field, and areas requiring deeper exploration. Importantly, the EG is working toward the goal of creating WHO ethics guidance and a practical implementation framework, along with accompanying tools, for infodemic managers and public health institutions involved in shaping or conducting infodemic management. The systematic scoping review described in this paper is essential for the work of the WHO's EG on ethical considerations in infodemic management and social listening by providing literature-driven insights and grounding their discussions and guidance in empirical evidence. In addition to detailing the practical initiatives already implemented to integrate ethics into the practice of infodemic management and social listening, this systematic scoping review also aims to describe the global health research community's perspective and understanding of the ethical dimensions crucial to infodemic management, thus blending practical, theoretical, and experimental perspectives to advance the field.

Objectives

The primary objective of this review is to identify, categorize, and analyze the ethical challenges and issues related to infodemics and their management. Our scope primarily covers literature published between 2002 and 2022, extracted from PubMed, Web of Science, and Scopus and enriched by a substantial amount of gray literature and policy documents contributed by distinguished experts in the field.

Methods

A Methodological Note on Our Approach to Ethics

Traditionally, ethical inquiry in applied ethics has been centered on establishing, defining, and elucidating substantive ethical principles [30,31]. These principles guide ethical decision-making and conduct. Commonly considered substantive principles include equity, justice, beneficence, and respect for autonomy. Within the context of our investigation of the literature concerning ethics and infodemic management, we fundamentally consider these principles, which provide a theoretical foundation for ethical considerations in infodemic management. However, our approach diverges from traditional frameworks by incorporating not only substantive ethical principles but also procedural principles that we consider as "proethical" [32]. While substantive ethical principles offer overarching moral guidance, procedural principles operationalize substantive principles into actionable steps for implementation in practice. In this systematic scoping review, we recognize the significance of both types of ethical considerations in the domains of infodemic management. By systematically examining the literature through the lens of both substantive and procedural principles, we aimed to provide a comprehensive understanding of ethics in the context of infodemic management during emergency health crises. Our analysis goes beyond defining ethics and identifying ethical principles that need to be respected in infodemic management; rather, it seeks to identify practical strategies and methodologies that can enhance the ethical conduct of infodemic management while augmenting the effectiveness of these practices. We used this approach as we acknowledge the dynamic and rapidly evolving nature of ethical challenges in public health and infodemic management specifically.

Approach

This review intertwines elements from both scoping and systematic reviews. Traditionally, a scoping review maps an expansive and varied body of literature to provide an overview of a broad subject area, identifying knowledge gaps [33]. A systematic review consolidates empirical evidence from a set of studies centered around a precise research question [33]. Our review has the goal of a scoping review and adopts the methodology of a systematic review [34]. We adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [35]. The PRISMA checklist is available in Multimedia Appendix 1 [6,7,16,17,27,28,36-78].

Query Definition

To initiate the search and extraction process, we used TopicTracker for text mining within PubMed records [79]. Using TopicTracker, we executed our initial query (query v0). All versions of our queries are available in Multimedia Appendix 1. This preliminary attempt harvested 34 papers. Interestingly, only 3 (9%) of these 34 papers were published prior to 2019, focusing on infodemics or social listening aspects that were not tethered to the COVID-19 narrative. A discernible challenge emerged from this exercise: the term "infodemic" was yet to be recognized as a standard keyword to describe research in the field of infodemic management and social listening before 2020. In response, we opted for an expansive approach [80], mining the initial results for synonyms, medical subject heading terms, and lemmas that encapsulate the essence of "infodemic." Underpinning "infodemic" as the overabundance of information, encompassing misinformation, particularly visible in digital spheres during significant health crises [81], we refined our initial query, leading to the formulation of query v1. This revised query retrieved 151 papers. The earliest papers trace back to 2003, with 54 (36%) of the 151 papers published before the onset of the COVID-19 pandemic. Following a deliberation session with the WHO EG panel on ethical considerations in infodemic management and social listening, we decided to further amplify our search strategy, mining for an expanded array of keywords resonating with "social listening" and incorporating synonymous descriptors for "infodemics," namely, "information overload," "information pollution," "information quality," "health information," "information voids," and "information deficits," gleaned from a related research project with the goal of mapping infodemic management interventions during health emergencies [82]. The above expansions led to the formulation of query v2. Despite a noticeable expansion in OR keywords, this query fetched only 1 additional paper compared to its predecessor (n=152), suggesting a saturated and robust query formulation. Finally, after another collaborative session with the EG, where query v2 was discussed, we added a few keywords to ensure that the query was not overly restrictive to a few fields of application in the realm of infodemic management and social listening. This iteration, termed query v3, yielded a list of 225 papers. The structure of this query is comprehensively detailed, alongside all previous versions of the query, in Multimedia Appendix 1 and in the study's protocol stored on our Open Science Framework (OSF) repository [83].

Query Translation

Once the query was validated by the EG, we translated query v3 for compatibility with Scopus and Web of Science. Detailed representations of these translations are presented in Multimedia Appendix 1 and in the study's OSF repository [83]. Further enriching our corpus, the EG provided additional material to be added to the list of retrieved items: this comprised not only traditional literature but also gray literature, references from United Nations and WHO work, and recently published work (beyond the cutoff imposed by our study design at the end of 2022). All these elements were integrated into the corpus of the literature retrieved with query v3 and translated queries for Scopus and Web of Science.

Data Retrieval and Screening of Records

Records extracted from the various sources were collated and stored in a publicly accessible Zotero (Corporation for Digital Scholarship) project [84]. Initial sifting was based on the exclusion criteria applied to the record titles and abstracts. The defined exclusion criteria were as follows: record does not mention social listening or infodemic management (directly or indirectly; see "infodemics (expanded)" in query definition and query v3; or record does not mention outbreak, epidemic, or pandemic; or record does not mention public health, risk to public health, public health emergency, and related concepts [both "acute" and "chronic"]); record does not mention ethics or ethical aspects; and record is not in English

Screening of Full Texts

We made use of Zotero's automatic download feature to retrieve the full texts of the previously shortlisted records. For papers that were not amenable to automatic downloading, we conducted a manual search to obtain them. For a paper to be incorporated into our main corpus, it had to meet the following inclusion criteria: full text is available; full text mentions social listening or infodemic management (directly or indirectly; see "infodemics (expanded)" in query definition or query v3; or full text mentions outbreak, epidemic, or pandemic; or full text mentions public health, risk to public health, public health emergency, and related concepts [both "acute" and "chronic"]); full text mentions ethics or ethical aspects; and full text is in English

Paper Assessment

To streamline the process of assessing all retrieved items, we designed a specialized web app leveraging Python (Python Software Foundation) and its Streamlit framework [85]. A comprehensive assessment of the corpus is available for scrutiny via the aforementioned web app or in the study's OSF repository [83]. This custom-built platform facilitates multiuser access and interaction and is securely hosted on a Firebase (Google LLC) database. We recorded a wide array of details related to each paper, including adherence to the established inclusion criteria; country of origin or focus of the research; year of the study (which might differ from the year of publication); specific health emergency or health-related issue tackled; type of study, whether theoretical, empirical, a literature review, and so on; methodological approach used; in-depth understanding and definition of infodemic management and social listening strategies presented; exploration of ethical considerations concerning infodemics; ethical challenges in infodemic management; aims of infodemic management and social listening strategies; and concluding insights and recommendations.

Analysis

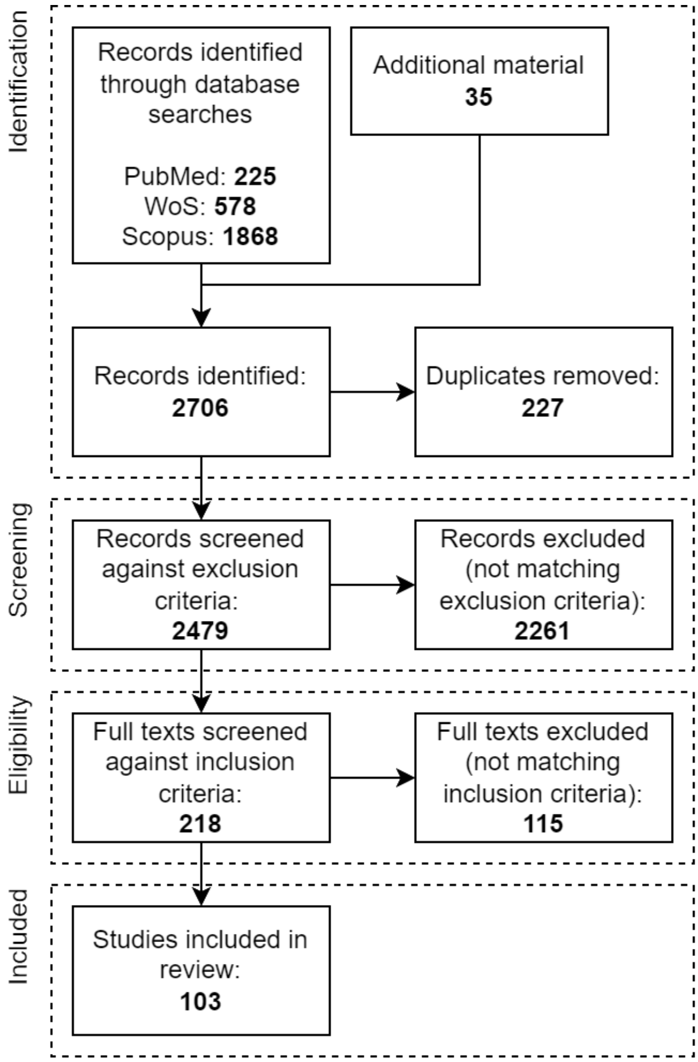
We conducted an analysis of papers that met our inclusion criteria to evaluate the fundamental themes that emerged from the literature. Our thematic analysis entailed identifying recurring patterns, key concepts, and trends related to ethical considerations in the context of infodemic management and social listening during health emergencies. It is important to note that not all these concepts fit the standard definition of "principle"; some encompass processes and conceptual frameworks that were not previously categorized as principles (see discussion in the *A Methodological Note on our Approach to Ethics* section). In addition to the substantive principles essential for addressing ethical concerns in infodemic management, we have also identified proethical procedural principles [32]. As mentioned earlier, these may not conform strictly to the traditional definition of principles in ethics; they encompass a mix of principles, concepts, and processes. However, when implemented, they ensure adherence to substantive principles and effectively address ethical tensions and issues in infodemic management. To structure our analysis, we created a comprehensive analytical framework that revolved around 2 core areas: the ethical issues pertinent to infodemic management and the ethical aims or values to be pursued in this context. To ensure a systematic approach, we associated these themes with their respective source papers. This phase laid the groundwork for the subsequent synthesis of findings and the development of a profound understanding of the ethical dimensions of infodemic management, as elaborated in the *Results* section. Our coding procedure was conducted in a blinded manner. Initially, 2 independent researchers took detailed notes on ethical issues in infodemic management using a specialized web application built with Python and the Streamlit framework [85]. In the second step, they categorized these issues into thematic clusters based on the output of the application, which presented data in a tabular data set without direct reference to the source papers. Finally, 2 independent researchers collaboratively coded and mapped the different ethical issues and objectives in infodemic management.

Results

Characterization of Included Papers

Through the query described in the *Methods* section of the paper, we identified 225 records through PubMed, 578 through Web of Science, and 1868 through Scopus, as described schematically in the PRISMA diagram in Figure 1. We also manually included 35 additional items, which were considered to be relevant by the WHO EG on the ethics of infodemic management and social listening [29], most of which were published recently and would have otherwise been excluded by our inclusion criteria (ie, papers until the end of 2022). We identified 2706 records and removed 227 (8.4%) duplicates. We screened the remaining 2479 (91.6%) records against the exclusion criteria and excluded 2261 (91.2%) items. Of the remaining 218 full texts, we excluded those that did not match the inclusion criteria; we removed 115 (52.8%) full texts for a total of 103 (47.2%) studies included in the review (Figure 1). Among the studies included in our systematic scoping review, we encountered a diverse array of publication types, including 88 journal papers (85.4%), 9 documents (8.7%), 2 preprints (1.9%), 2 reports (1.9%), 1 presentation (1%), and 1 book (1%). These studies encompassed a spectrum of research types, most of which were theoretical studies (n=45, 43.7%), followed by empirical (n=35, 33.9%), viewpoints or commentaries (n=14, 13.6%), literature reviews (n=4, 3.9%), or other type of studies (n=5, 4.9%). Focusing on empirical research items, most studies were either observational (n=9, 25.7%) or cross-sectional research (n=8, 22.9%), while experimental studies were the least frequent (n=5, 14.3%). The included literature discussed different types of health emergencies. Of the 103 papers, 49 (47.6%) papers explored ethical aspects related to infodemic management and social listening during pandemics and epidemics, including COVID-19, H1N1 influenza, HIV, measles, H5N1, and dengue. An additional 24 (23.3%) of the 103 papers focused on infodemics related to the COVID-19 pandemic or vaccine hesitancy. A few (n=6, 5.8%) papers addressed chronic health emergencies, including those related to smoking, alcohol, obesity, nutrition, and food risk, while others discussed environmental hazards such as radioactivity, floods, disasters, water pollution, and contamination. Of note, most research items (n=70, 68%) were published during the COVID-19 pandemic, with a significant surge from 2020 onward, peaking in 2021 (n=37, 35.9%) and remaining substantial in 2022 (n=18, 17.5%). We also looked at the geographical origin of the papers included in the review. Breaking down the analysis per continent, we found that Europe (n=58, 56.3%) is the most represented continent in our review database, followed by North America (n=38, 36.9%) and Asia (n=27, 26.2%). Only 1 (9.7%) item published by researchers from Central and South America was included (Figure S1 in Multimedia Appendix 1). Since the cutoff was at the end of 2022, we initially made reference to the Open Science Framework repository of 1 paper [82]; however, during review, this paper was published and we decided to link to the published version instead.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram.



Ethical Issues and Ethical Aims in Infodemic Management and Social Listening

We analyzed the 103 papers included in our review. Noteworthy themes emerged, in particular the relevance of adopting nuanced strategies for effective communication, outreach efforts, and the importance of disseminating truthful information. Trust and the consequences of mistrust in infodemic management contexts were also significant areas of discussion. Furthermore, ethical practices concerning surveillance and social listening were also key points of discussion and, in particular, the ethical implications of social listening practices on privacy rights. In addition, several papers evaluated the importance of addressing vulnerability and equity in the design of infodemic management and social listening strategies, as well as navigating the balance between safeguarding free speech and combating misinformation. Other topics included the lack of community engagement in infodemic management and social listening, the ethical dilemma of informing versus manipulating the public toward desired health behaviors, conflicts of interest, and the importance of honesty, as well as the critical need for education and fostering critical thinking skills to build autonomy. Table 1 provides a comprehensive overview of the identified ethical issues and references.

In addition to exploring ethical issues, we examined the overarching ethical aims of infodemic management and social listening discussed in the corpus of papers. Our analysis revealed primarily the ethical importance of disseminating truthful information. Notably, social listening emerged not only as an issue but also as an ethical goal in infodemic management, that is, listening to and understanding concerns and reacting to them in a timely manner. Furthermore, crucial aims included community engagement, trust-building initiatives, transparency, and educational strategies. In addition, ensuring inclusivity and equity, leveraging fact-checking mechanisms to counter misinformation, and prioritizing privacy and anonymization in social listening practices were among the top aims identified. While many of these ethical aspects were already recognized as crucial ethical issues requiring resolution, their prominence as aims in infodemic management, and not only as issues to solve, highlights the primary importance placed on resolving ethical tensions and integrating ethics into these practices. Table 2 shows a comprehensive list of the identified ethical aims and references.

To visualize the data, we graphically represented the relevance of the different ethical issues and aims in infodemic management by creating a bubble graph (Figure 2A), in which we defined, on the x-axis, the relevance of each ethical issue based on how often it was discussed in the corpus of publications retrieved with our query. The value on the y-axis is determined by how often each ethical aim was represented in our corpus of publications. The entire list of ethical issues and aims is presented in Multimedia Appendix 2. The size of the bubbles, determined by the sum of the x and y values, visually represents the relevance of the ethical concepts and principles. This serves as a starting point to define which aspects have been considered by the literature when integrating ethical approaches in infodemic management. We also split the different ethical concepts and principles into 6 different categories, as highlighted by the different colors of the bubbles (Figure 2A). The main categories we identified were linked to “communication, media, and information”; “privacy, surveillance, and data ethics”; “ethics, responsibility, and governance”; “social equity and inclusivity”; and “public engagement and education.” We identified 2 relevant clusters housing the most prevalent ethical concepts and principles within our corpus of publications. These clusters will be subject to our analysis in this review, with cluster 1 principles being discussed in the main manuscript and cluster 2 ethical considerations being discussed in Multimedia Appendix 1 (Figure 2B).

Table 1. List of ethical issues in infodemic management and social listening and the frequency with which the issue has been reported in the literature analyzed in the review (n=103).

Rank number	Ethical issue in infodemic management and social listening	Papers, n (%)	References
1	Right to be informed truthfully, communication, and outreach	30 (29.1)	[17,18,36-40,45,53,61-64,86-102]
2	Trust and mistrust	28 (27.2)	[16,18,38,40,51,52,57,58,65,66,88,90,91,93,94,96,99-101,103-111]
3	Surveillance and social listening	28 (27.2)	[6,7,16,17,27,36,40,44-50,54,55,60,67-69,86,98,108,112-116]
4	Vulnerability and inequity	25 (24.3)	[36-38,50-53,55,59,63,68,82,87-89,98,101,103,110,113,117-121]
5	Free speech versus regulation	16 (15.5)	[2,7,17,40,42,59,62,64,70,71,82,90,107,110,122,123]
6	Right to privacy	14 (13.6)	[7,47,51,52,54,55,65,69,108,114-117,124]
7	Lack of community engagement	11 (10.7)	[17,18,46,54,57,63,86,98,112,125,126]
8	Informing versus manipulating	10 (9.7)	[17,36,40,61,90,97,104,110,127,128]
9	Honesty and conflicts of interest	10 (9.7)	[36,38,39,51,63,71,105,109,126,129]
10	Lack of education	9 (8.7)	[17,36,46,62,88,96,110,123,130]
11	Necessity	9 (8.7)	[17,49,61,63,68,86,92,105,126]
12	Cybersecurity	9 (8.7)	[7,36,52,54,65,82,98,108,124]
13	Lack of transparency	7 (6.8)	[36,38,54,63,86,92,93]
14	Individual versus collective health	7 (6.8)	[51,58,59,61,94,109,122]
15	Good governance	6 (5.8)	[7,39,60,63,71,129]
16	Epistemic underdetermination	6 (5.8)	[39,61,63,92,96,99]
17	Lack of autonomy	6 (5.8)	[52,59,61,62,104,121]
18	Power imbalances	6 (5.8)	[38,58,59,97,100,103]
19	Translation of evidence into public health practice	6 (5.8)	[63,71,96,98,99,110]
20	Responsibility	5 (4.8)	[58,62,105,109,131]
21	Different cultural perspectives	5 (4.8)	[41,51,87,96,126]
22	Stigma	4 (3.9)	[109,119,120,123]
23	Definition of truth	4 (3.9)	[2,62,71,99]
24	Alignment with human rights framework	4 (3.9)	[86,89,121,122]
25	Legality	4 (3.9)	[69,112,123,131]
26	Proportionality	4 (3.9)	[61,93,112,126]
27	Social media practices	3 (2.9)	[42,129,130]
28	Control of citizens	3 (2.9)	[54,60,64]
29	Selection bias and information bias	3 (2.9)	[2,17,66]
30	Fairness	3 (2.9)	[53,109,126]
31	Appeal to fear	3 (2.9)	[102,109,127]
32	Data and representation inclusiveness	3 (2.9)	[44,98,113]
33	Lack of research	2 (1.9)	[28,98]
34	Beneficence	2 (1.9)	[42,89]
35	Solidarity	2 (1.9)	[52,63]
36	Lack of openness	1 (1)	[48]
37	Criminalization	1 (1)	[70]
38	Lack of independent oversight	1 (1)	[112]
39	Reciprocity	1 (1)	[63]

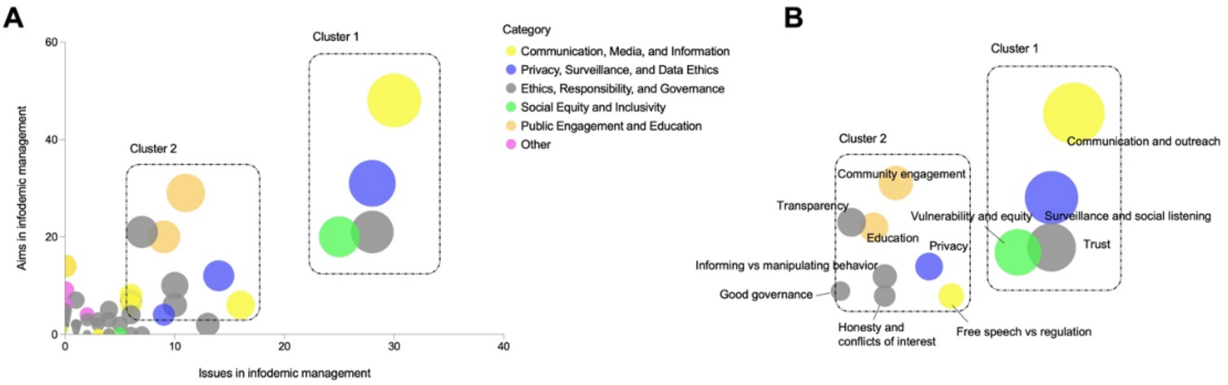
Rank number	Ethical issue in infodemic management and social listening	Papers, n (%)	References
40	Absence of an ethical framework	1 (1)	[86]

Table 2. List of ethical aims of infodemic management and social listening and the frequency with which the aims have been reported in the literature analyzed in the review (n=103).

Rank number	Ethical aims in infodemic management and social listening	Papers, n (%)	References
1	Truthful communication and outreach	48 (46.6)	[6, 16-18, 27, 28, 36, 39, 43-45, 47, 48, 50, 54, 55, 57, 60-64, 69, 71, 82, 86, 87, 88, 92, 95, 100, 104, 107, 117-119, 123-125, 127, 132-139]
2	Surveillance and social listening	31 (30.1)	[6, 7, 16, 17, 27, 36, 37, 43, 44, 47, 50, 51, 53, 54, 59, 60, 82, 86, 87, 96, 98, 108, 113, 117, 123, 124, 128, 130, 131, 134, 136]
3	Community engagement	29 (28.2)	[16,17,41,43-47,50,53,55,60-63,67,82,86,96,112,114,117,125,133-136,140,141]
4	Trust	21 (20.4)	[7,17,28,36,43-45,57,61,66,70,82,106,108,117,132,133,135,139,141,142]
5	Transparency	21 (20.4)	[7,17,41,43-45,48,50,61-63,72,112-114,119,124,134,139-141]
6	Empowerment through education and educational strategies	20 (19.4)	[16,27,28,42,47,55,57,59,62-64,70-72,82,96,106,118,128,138]
7	Inclusivity and equity	20 (19.4)	[6,17,18,28,37,44,50,53,55,61-63,67,72,82,92,98,117,138,139]
8	Effectiveness of targeted interventions	14 (13.6)	[17,36,39,45,50,60,61,63,112,119,127,132,135,141]
9	Fact checking and labeling misinformation and debunking misinformation	14 (13.6)	[6,27,36,37,62,70,104,105,110,118,121,123,124,128]
10	Guarantee privacy and anonymization	12 (11.6)	[7,17,48-50,60,67,68,112,114,116,124]
11	Influencing health behavior and improving health	10 (9.7)	[52,59,61,64,82,92,96,97,100,117]
12	Foster cooperation between institutions	9 (8.7)	[41,43,45,50,55,72,117,124,136]
13	Transform evidence in communication and policies	8 (7.8)	[16,17,47,55,61,112,134,137]
14	Openness	7 (6.8)	[17,50,54,60,86,134,141]
15	Good governance	7 (6.8)	[6,7,39,44,50,63,64]
16	Honesty and integrity	6 (5.8)	[27,39,45,61,63,134]
17	Acknowledgment of failure and evaluation of impact	6 (5.8)	[43-45,51,69,143]
18	Acknowledging uncertainty (epistemic underdetermination)	6 (5.8)	[37,43,44,63,134,143]
19	Respect for human rights	5 (4.8)	[45,82,112,114,134]
20	Respect for dignity and persons	5 (4.8)	[45,49,50,109,112]
21	Research and generation of new ideas	4 (3.9)	[6,45,57,125]
22	Cybersecurity	4 (3.9)	[7,49,50,124]
23	Justice	4 (3.9)	[17,59,107,143]
24	Autonomy	4 (3.9)	[17,45,49,50]
25	Blocking and removing misinformation and conspiracy theories	4 (3.9)	[36,37,40,93]
26	Accountability	3 (2.9)	[45,72,112]
27	Nondiscrimination and stigma	3 (2.9)	[61,109,112]
28	Solidarity	3 (2.9)	[63,139,143]
29	Population tracing and control	3 (2.9)	[54,98,136]
30	Proportionality	2 (1.9)	[63,139]
31	Using ethical approaches (ethics for ethics)	2 (1.9)	[36,140]
32	Guaranteeing free speech	2 (1.9)	[71,118]
33	Fairness	2 (1.9)	[17,45]
34	Independent oversight	2 (1.9)	[67,72]

Rank number	Ethical aims in infodemic management and social listening	Papers, n (%)	References
35	Improve society, social cohesion, reduce polarization	2 (1.9)	[66,82]
36	Responsibility	2 (1.9)	[45,95]
37	Stewardship	2 (1.9)	[61,139]
38	Enforcement of recommendations and restrictions	2 (1.9)	[52,94]
39	No harm	1 (1)	[112]
40	Acknowledge the limitations of social listening practices	1 (1)	[44]
41	Reciprocity	1 (1)	[63]
42	Protect health care professionals	1 (1)	[42]
43	Maintenance of peace	1 (1)	[107]
44	Maintenance of democracy	1 (1)	[107]
45	Collect data into a single and accessible platform	1 (1)	[69]

Figure 2. Ethical issues and aims in infodemic management and social listening. The x-axis illustrates the frequency of specific ethical issues discussed in the literature. (A) The y-axis measures the frequency of ethical aims in infodemic management, which is also based on the number of papers discussing them in the analyzed literature. (B) The size of the bubbles represents the sum of the x and y values and serves as a graphical representation of the overall relevance of the ethical aspect or principle under consideration.



Analysis and Application for Each Principle

In this second part of the *Results* section, we will introduce each concept and principle identified in cluster 1 (Figure 2B) and define, for each, which specific ethical issues emerged from the literature, which specific associated aims and goals should be achieved, and why these aims should be achieved in the context of infodemic management. For each of these concepts and principles, we will identify procedural principles, which, if applied to infodemic management, can provide practical guidance and recommendations on how to ensure that substantive underlying ethical principles are respected while safeguarding the effectiveness of infodemic management practices.

Truthful Communication and Outreach

The systematic scoping review encompassed a comprehensive examination of literature pertaining to communication and outreach in the context of infodemic management and the practice of social listening. Several critical findings emerged, and they were categorized into distinct thematic aspects, each of which holds significant implications for ethical and effective communication and outreach in the context of infodemic management.

The first pertains to inclusivity. The reviewed literature emphasizes the importance of crafting communication strategies that take into account the needs of vulnerable groups [36,86,87]. This includes individuals with limited or no access to the internet or with restricted use of social media platforms, such as those who rely solely on services such as WhatsApp for information. Recognizing the fragility of communication technologies is paramount, as technical disruptions can impede information dissemination efforts [37]. In particular, the literature highlights the vulnerability of individuals with low information and media literacy, who are at a heightened risk of falling victim to misinformation [88]. Consequently, the ethical principle of vulnerability intersects with the imperative for education and literacy. It is therefore fundamental to address these vulnerabilities in communication strategies [38,89].

The second thematic aspect underscores the significance of maintaining consistency and reliability in information dissemination to foster public trust [88]. Addressing information gaps and uncertainties is crucial to mitigate the spread of misinformation [90]. However, it is essential to exercise caution when providing information in situations characterized by epistemic underdetermination [39,91,92]. Incorrect or imprecise information can have detrimental effects, contributing to information overload and confusion among recipients [40]. This can erode institutional trust and hinder the receptivity of future public health advice [18]. Therefore, a foundation of evidence-based and epistemically truthful communication is advocated to guide the development of public health messages and strategies, which ultimately serve to bolster public trust [93].

Furthermore, the literature emphasizes the risk of information overload, even when the information is accurate. To mitigate this, communication from reliable sources should strike a balance between countering disinformation and avoiding overwhelming the intended audience. Information should be timely, accurate, disseminated through appropriate channels, and designed for the specific target population [94]. Ensuring clarity and timeliness is considered fundamental, and the use of plain language and suitable metaphors is recommended to enhance public comprehension [16,144]. The employment of personnel experienced in scientific communication can be instrumental in conveying complex scientific information to the public. Tailoring messages to specific audiences significantly improves understanding and engagement, especially in risk and crisis communication situations, where reassurance and panic mitigation are integral strategies [17].

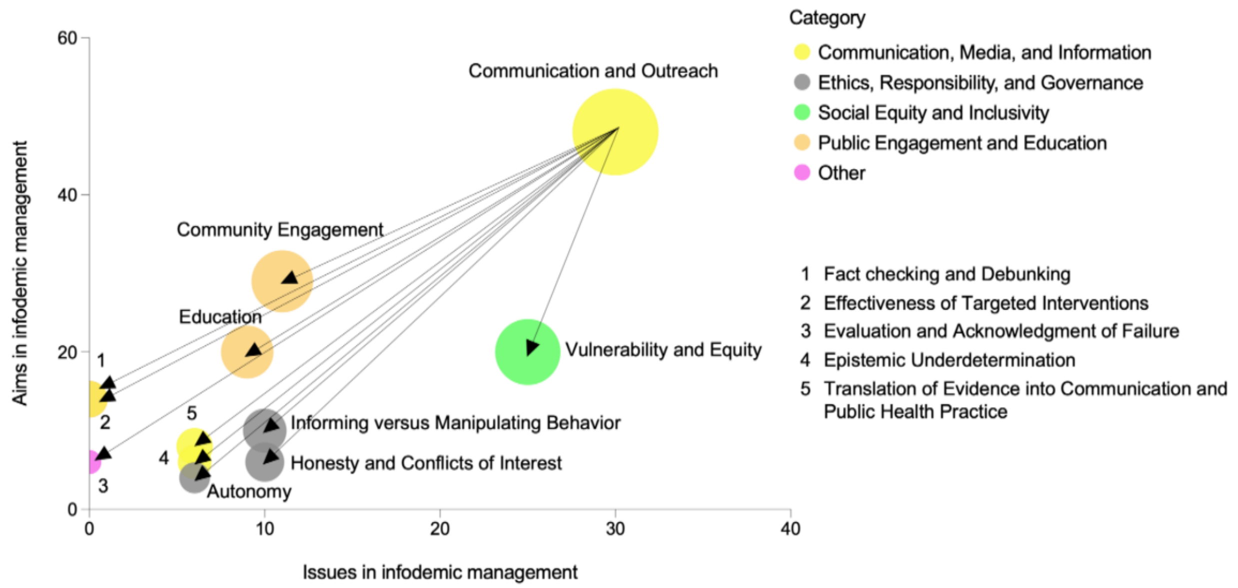
A third aspect identified in the literature centers on the dynamics of social media communication as a key component in infodemic management [145,146]. Social media platforms are recognized as significant channels for information dissemination, making social media literacy an essential skill for effective communication [147]. The choice of communication channels should align with the specific message being conveyed and the intended target public [17]. Engaging with influencers, key opinion leaders, and religious figures is seen as a means to aid in the dissemination of accurate information [17,41,145].

Importantly, the reviewed literature underscores the importance of improving the information ecosystem as a fundamental goal in infodemic management. This involves enhancing access and exposure to credible health information and encouraging positive changes in information-seeking behavior [18,92,95,96]. Promoting information literacy among the public empowers individuals to critically evaluate and understand the information they encounter, thereby reducing the impact of misinformation [17,42].

Finally, the literature stresses the necessity of continuously assessing the effectiveness of communication strategies. Feedback should inform the development of new and improved approaches. In this context, advocacy and community engagement are recognized as pivotal in ensuring effective communication and outreach [43,44,97]. Involving representatives of the target public in the design of communication strategies enhances effectiveness and ensures that community voices are heard [45,46]. Small-scale social listening approaches, in which community members provide real-time feedback, offer valuable insights into the effectiveness of communication and outreach strategies [43,44]. These elements collectively contribute to the efficacy of public health measures and ensure that communication is conducted in an ethical and responsible manner.

In light of the insights and considerations drawn from the reviewed literature, we have articulated a set of ethical procedural principles. When applied to communication and outreach within the framework of infodemic management practices, these principles enhance their effectiveness while upholding ethical standards, as visually represented in Figure 3.

Figure 3. Procedural principles to ensure ethical and effective communication and outreach for infodemic management. The x-axis illustrates the frequency of specific ethical issues being discussed in the literature. The y-axis measures the frequency of ethical aims in infodemic management, which is also based on the number of papers discussing them in the analyzed literature. The size of the bubbles represents the sum of the x and y values and serves as a graphical representation of the overall relevance of the ethical aspect or principle under consideration.



The first principle, encapsulated by the terms *vulnerability and equity*, underscores the imperative of ensuring equitable access to information, particularly for vulnerable groups, thus promoting inclusivity. Another critical facet is *community engagement*, emphasizing the active involvement of the public in infodemic management and social listening practices. This involvement, notably in the design of communication strategies and the generation of feedback, facilitates the integration of public concerns into public health initiatives, whether related to prevention or risk and crisis communication [17]. Empowerment through *education* serves as a substantive ethical requirement, emphasizing the significance of fostering information literacy to nurture a healthier information ecosystem and mitigate the adverse effects of misinformation. The concept of *epistemic underdetermination* acknowledges the need to address information gaps transparently, especially when evidence-based information is unavailable. In Figure 3, the issue of *informing versus manipulating behavior* intends to highlight that a highly literate health information ecosystem should provide adequate information to enhance individual and public health without resorting to manipulative tactics, which could erode institutional trust over time [18,88,127]. The principles of honesty and avoiding conflicts of interest, represented as *honesty and conflicts of interest* in Figure 3, emphasize the pivotal role of honesty and the avoidance of conflicts of interest in shaping communication and outreach within infodemic management, thereby safeguarding trust. The principle of *autonomy* relates to a combination of building information literacy through educational approaches and ensuring inclusivity and equity. The principle of *evaluating and acknowledging failure*, presented in Figure 3, encourages the development of more effective information campaigns through iterative processes involving communities. *Effectiveness of targeted interventions* underscores that ethically sound communication campaigns are a prerequisite for their effectiveness. *Fact checking and debunking*, presented in Figure 3, further underscores the importance of these activities as primary objectives in communication and outreach efforts to combat misinformation. Finally, *translating evidence into public health practices* emphasizes the critical task of translating evidence into tangible health benefits for individuals and communities, even in the face of challenges such as epistemic underdetermination or a polarized and misinformation-rich information ecosystem.

These ethical procedural principles together constitute a robust foundation for the development and implementation of communication and outreach strategies within the context of infodemic management. Their application not only bolsters the effectiveness of these practices but also ensures ethical integrity and adherence to ethical standards.

Monitoring and Social Listening

We also examined the landscape of ethical considerations regarding monitoring and social listening in infodemic management. This comprehensive exploration has revealed several pivotal findings and strategic approaches, each bearing profound implications for the ethical conduct of monitoring and social listening as they relate to the management of information epidemics.

In the ethically complex arena of surveillance and data collection, under specific circumstances requiring comprehensive data sets and rigorous data protection, individuals may have an ethical obligation to contribute to monitoring even without explicit consent [47-50]. That said, the principle of autonomy, implemented through the practice of obtaining informed consent, should stand as the foremost pillar whenever possible. It advocates for integrating informed consent into social listening practices, ensuring that individuals are aware of and consent to data collection [47]. The literature widely advocates for informed consent as a key ethical practice in data collection, emphasizing that it is essential, not a hindrance, in sustaining institutional trust and respecting privacy rights [47,48,50]. This approach enhances public confidence in data-driven methods. Balancing data and privacy, raising privacy awareness, and maintaining confidentiality and anonymity are crucial for upholding the ethical standards of data collection and therefore mitigating the erosion of trust [51,52].

Transparency is the bedrock upon which trust is constructed. To foster public trust and ethical conduct, articulating data collection practices to the public and offering a clear overview of the social listening strategy is suggested to be fundamental in the analyzed corpus of literature [38,86,93]. This transparency assures the public that their data are managed responsibly. A cardinal rule in ethical data collection is to prioritize *active* social listening practices over *passive* social listening approaches [38,48]. Engaging with the public actively, considering their concerns, and respecting their autonomy are in line with key ethical principles. Instead, passively extracting data, for example, to monitor public concerns and rumors, may impact public trust in the medium and long term [17,38,48]. The corpus of literature thus underlines that ethical decision-making demands that all public concerns are taken into account when converting social listening insights into infodemic management strategies [14,112]. Community engagement is not just an ethical necessity but a source of valuable insights [53,148,149].

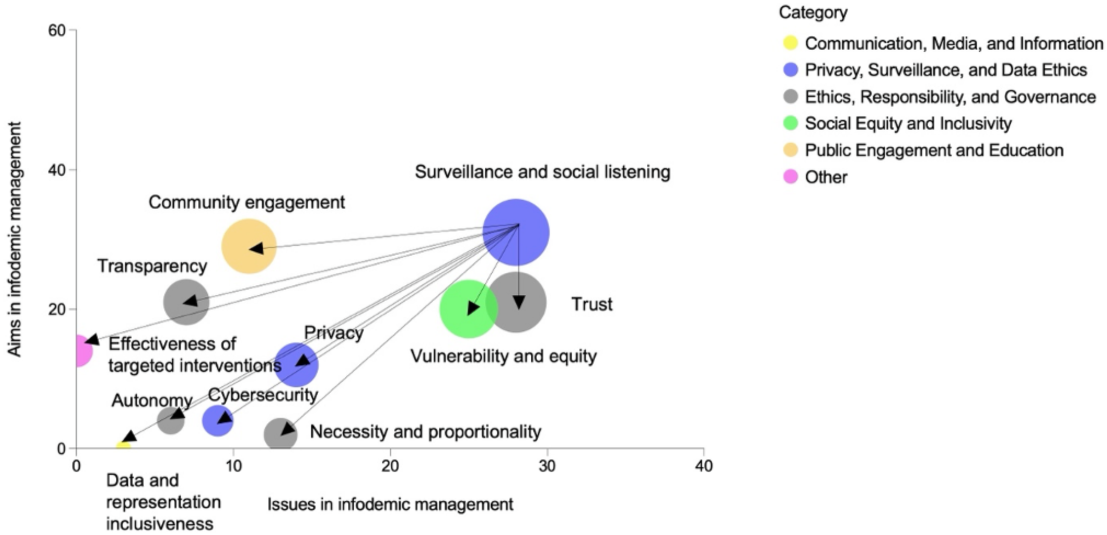
Furthermore, the literature underlines that it is crucial to emphasize that social listening must not be used for tracking dissent, population control, or governmental monitoring [48,54]. This ethical use safeguards against potential misuse and violations of privacy [50,54,55].

A second important aspect to ensure ethical data collection is that representativeness in data is the standard. This inclusivity encompasses demographic diversity, language considerations, and the types of data captured [44,87,98,113]. By avoiding research biases, ethical data collection becomes a more powerful tool for understanding the information landscape [17,44]. If the design of social listening is not inclusive, the conclusions drawn from the data may have limited or even negative repercussions for vulnerable groups or groups that were not considered or integrated into the social listening design [14,44].

To ensure effective social listening, monitoring information should happen in real time, including both online and offline sources. This is instrumental in detecting narratives, questions, concerns, and misinformation within the information ecosystem. Despite the focus on effectiveness, data security remains paramount, ensuring the protection of sensitive data and preventing unauthorized access or use [53,54,148,149]. Evidence suggests that a breach in security would lead to a reduction of trust, and a reduction of trust would lead to decreased effectiveness of social listening [54]. Thus, also in this case, ethical social listening is necessary to guarantee the effectiveness of infodemic management practices in the short and long term. In addition to cybersecurity, respecting individuals' right to privacy is considered fundamental [50,54,55]. This includes data anonymization and the implementation of robust data security measures to safeguard sensitive information. When anonymization and data security are guaranteed, in situations of necessity, social listening practices can shift more toward passive approaches [17,48,54].

In line with the methodology outlined in our communication and outreach framework and guided by the recommendations and insights gleaned from the reviewed literature, we have formulated a set of ethical procedural principles applied to the domains of surveillance and social listening; these principles are instrumental in bolstering the efficacy of social listening practices while upholding the highest ethical standards, as visually depicted in Figure 4.

Figure 4. Procedural principles to ensure ethical and effective social listening practices. The x-axis illustrates the frequency of specific ethical issues discussed in the literature. The y-axis measures the frequency of ethical aims in infodemic management, which is also based on the number of papers discussing them in the analyzed literature. The size of the bubbles represents the sum of the x and y values and serves as a graphical representation of the overall relevance of the ethical aspect or principle under consideration.



The most important procedural principle is *trust*. Prioritizing trust fosters the perception of surveillance and social listening endeavors as constructive measures dedicated to ensuring public health. The second principle emphasizes inclusivity and equity (Figure 4; *vulnerability and equity*), advocating for the development of social listening practices that are free from biases and sensitive to the needs of vulnerable populations. It is imperative that the design and implementation of these practices aim for the utmost representation, extending to the insights derived from social listening (*data and representation inclusiveness* as another procedural principle). *Community engagement* is the third aspect to consider, suggesting the preference for the adoption of *active* social listening strategies that directly involve the target audience, thereby cultivating trust in the processes. *Privacy* and *cybersecurity* constitute 2 fundamental principles; these principles demand that privacy and anonymity be guaranteed, especially when *passive* social listening methods are used. In addition, *transparency* is another procedural principle in this context, closely linked to trust and requiring clear communication of the purpose behind any social listening action to dispel any negative perceptions held by the public. *Necessity and proportionality* are key operational principles in this context. To avoid invasiveness in social listening, it is essential to ensure that such practices are only used when absolutely necessary and in proportion to the specific circumstances [50]. *Autonomy* is also considered a principle to guarantee ethically sound and effective surveillance and social listening. It underscores the importance of ensuring that the target audience comprehends the practices they are subject to and that they have the power to assert control over their own privacy, cybersecurity, and personal information.

Trust and Mistrust

At the core of trust building lies the imperative to involve the public actively [17,38], steering clear of top-down approaches; this is especially valid for health departments and governmental entities [17,38,150]. Collaborative decision-making and involving the public in shaping policies engender a sense of inclusivity and shared responsibility [18]. Furthermore, discouraging the pursuit of profit-driven objectives in public health is crucial. Rather, decisions and actions should be grounded in a commitment to the well-being of the public, prioritizing their welfare above all else [37,56].

The literature further highlights that a critical building block for trust building is the elevation of trust in research and researchers [57,93,103]. This involves ensuring that research is conducted ethically, findings are communicated transparently, and public health initiatives are rooted in sound scientific evidence. Ethical communication principles are paramount, emphasizing the importance of not resorting to manipulative tactics, even when pursuing noble causes [90,104]. Informed and transparent communication with the target public is essential, promoting honesty and avoiding any perception of manipulation (for a detailed explanation, refer to the *Truthful Communication and Outreach* section) [38,93].

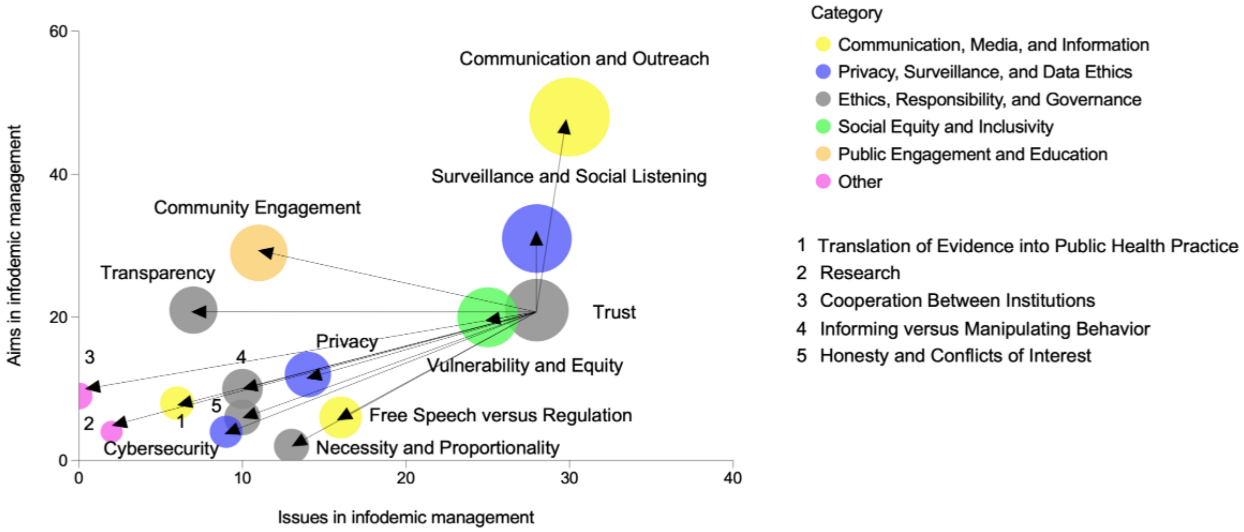
Furthermore, another important aspect of trust is its connection to credibility. Trust is bolstered by credibility and the use of expertise. Establishing oneself (this, for example, includes researchers, public health institutions, and health departments within governments) as a reliable and knowledgeable source of information is pivotal in instilling trust in public health communications [58,114,117].

Another important aspect in building trust is the engagement of various public groups as coactors in the planning and execution of infodemic management and social listening initiatives; this is not merely an ethical requirement but also provides a wealth of valuable insights that ensure the effectiveness of infodemic management and social listening actions. A pluralistic approach ensures that all voices are heard and considered in the decision-making process [17,18,51,57].

Finally, the literature suggests that even public health institutions, which have no financial interests, can benefit from adopting branding and advertising strategies similar to those used by business-oriented organizations. This approach helps to effectively highlight the services they offer to the public. This strategic approach can contribute to a stronger and more recognizable public health identity. These strategic solutions collectively form a robust framework for enhancing trust and countering mistrust within the domain of public health, and that indirectly reflect on trust for infodemic management and social listening practices performed by such institutions [16,17].

Trust appears to be a key core principle to ensure the effectiveness of infodemic management and social listening in the medium- and long-term horizons, closely linked to several procedural principles as shown in Figure 5. By integrating these principles into practice, thus strengthening trust, public health efforts through infodemic management can be not only more deeply rooted in ethical conduct but also more effective.

Figure 5. Procedural principles to build trust and reduce mistrust. The x-axis illustrates the frequency of specific ethical issues discussed in the literature. The y-axis measures the frequency of ethical aims in infodemic management, which is also based on the number of papers discussing them in the analyzed literature. The size of the bubbles represents the sum of the x and y values and serves as a graphical representation of the overall relevance of the ethical aspect or principle under consideration.



As highlighted by the procedural principles for trust in Figure 5, at its core, trust is nurtured through transparent, informed communication, free from manipulation [38,90,93,104]. This substantive principle ensures that the public receives accurate information. *Privacy* and *cybersecurity* principles, as highlighted in Figure 5, are nonnegotiable, safeguarding data protection and personal privacy within surveillance and social listening. Upholding these ethical standards reduces the likelihood of developing public mistrust in the institution carrying on infodemic management and social listening practices. Furthermore, *community engagement* ensures that the voices of the community are actively incorporated into the decision-making process. In fact, building trust involves embracing inclusive approaches, and similarly, it involves addressing the needs of vulnerable populations. Left out and marginalized voices that are not considered by the institutions leading infodemic management and social listening efforts would lead to increased mistrust. Similarly, adhering to the principles of *necessity and proportionality* is key to justifying the invasive nature of these practices and maintaining trust. The literature also identified *transparency* as a key procedural principle in this context: by explaining the purpose of social listening practices and dispelling doubts and misconceptions, transparency helps mitigate mistrust [38,93]. Another important aspect considered in the literature is that striking the right balance between free speech and necessary regulation is crucial, preventing overreaching censorship that could erode trust [17,18,59] (*free speech vs regulation* in Figure 5). Furthermore, aligning policies and communications with the existing evidence is essential, as is building trust in research and fostering collaboration between researchers and public institutions [57,93,103] (*translation of evidence into communication and public health practice* and *cooperation between institutions* in Figure 5). It is vital to clarify that communication and policies are designed to inform rather than manipulate, emphasizing the ethical intent behind these practices [90,104] (*informing vs manipulating behavior* in Figure 5). Finally, honesty and the disclosure of any conflicts of interest at every step of infodemic management help guarantee transparency and ensure that the decision-making process remains free from bias or manipulation, enhancing trust and overall effectiveness [38,93,105] (*honesty and conflicts of interest* in Figure 5).

Vulnerability, Equity, and Inclusivity

The first thematic aspect emphasizes the need to strengthen vulnerable media information ecosystems. To do so, the body of literature suggests that it is crucial to empower individual members of the public and communities to be autonomous and resilient against manipulation tactics [16,122]. This entails strategies to enhance critical thinking; media literacy; and the ability to discern reliable sources, especially among those who are most susceptible to manipulation [17,88,103].

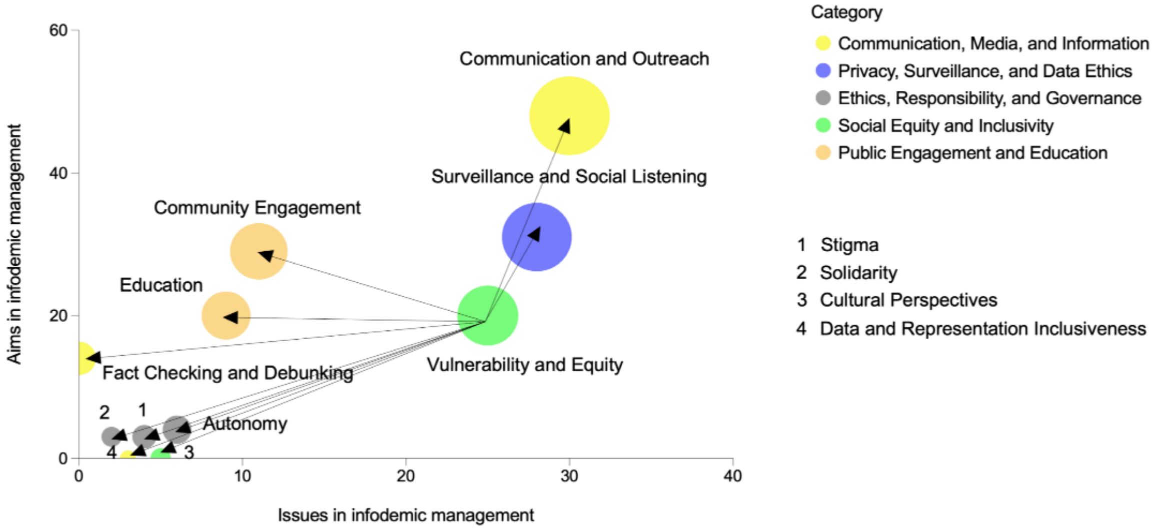
The second thematic aspect underscores the importance of combating polarization within the information ecosystem. The detrimental effects of epistemic echo chambers and bubbles must be mitigated to ensure that all individuals, regardless of their background or beliefs, have access to a balanced and diverse information landscape [16]. Beyond information, the literature advocates for holistic improvements in socioeconomic, cultural, environmental, and "infospherical" conditions [39]. This includes addressing disparities in living and working conditions, access to water and sanitation, housing, education, health care services, and food production. It extends to fostering inclusive social, community, and web-based networks, considering individual lifestyle factors, age, sex, and genetics [17,39]. Equity in these areas is essential to reduce vulnerabilities and is a primary goal of infodemic management. Similarly, infodemic management efforts should focus on improving access to information. Ensuring an equal distribution of such resources among all segments of the population, regardless of socioeconomic status, is paramount in promoting equity and reducing vulnerabilities [53,55,60,151].

The third aspect highlights the monitoring of ethnically targeted disinformation and misinformation that exploits the fears of vulnerable groups, including older adults, leading to mental health burdens [36,59]. Accurate and inclusive social listening designs, alongside proactive measures, are needed to prevent and counteract such disinformation, protecting the most vulnerable groups. Similarly, the literature highlights the importance of preventing information-related discrimination based on ethnicity, religion, or political beliefs. Equitable access to information must be safeguarded for all without discrimination.

The fourth and final thematic aspect that we identified underscores the fundamental right to receive accurate health information. Equitable access to accurate health information is essential for everyone, regardless of their background or circumstances. This requires the establishment of adequate and fair communication channels that cater to different segments of the public, with special attention to minority and vulnerable groups. These channels should ensure that information is accessible and comprehensible [36,38,86,87,89].

On the basis of the recommendations and insights from the reviewed literature, we have formulated a set of ethical procedural principles applied to infodemic management; these principles are instrumental in upholding the highest ethical standards during infodemic management actions, as represented in Figure 6.

Figure 6. Procedural principles to ensure inclusivity and equity and address vulnerability. The x-axis illustrates the frequency of specific ethical issues discussed in the literature. The y-axis measures the frequency of ethical aims in infodemic management, which is also based on the number of papers discussing them in the analyzed literature. The size of the bubbles represents the sum of the x and y values and serves as a graphical representation of the overall relevance of the ethical aspect or principle under consideration.



The principles of inclusivity and equity are interwoven with various procedural principles that help ensure their application. A holistic approach to inclusivity and equity is pivotal. Inclusive *communication and outreach* strategies must ensure that all voices, with particular emphasis on vulnerable groups, are not only heard but also actively integrated into the decision-making process. Inclusive social listening design is paramount [16,44] (*surveillance and social listening* in Figure 6), demanding that the concerns and perspectives of vulnerable populations be central to the process. *Community engagement* becomes an important bridge, actively incorporating diverse voices in social listening design, decision-making, and communication efforts.

Equally important is the drive to reduce vulnerabilities through *education* and literacy, especially among the marginalized groups [17,88,103]. The battle against misinformation and disinformation targeting vulnerable groups within polarized information ecosystems thus becomes an ethical imperative. In this context, *autonomy* should be granted, empowering individuals to protect themselves from the dangers of disinformation. Furthermore, addressing *stigma* is critical, ensuring that no one is excluded from social listening initiatives. In this context, *solidarity* serves as the moral backbone, affirming that all individuals are equally valued [16]. Importantly, adapting to various *cultural perspectives* is important; that is, infodemic management actions need to be tailored to different cultural perspectives, recognizing that infodemic management is not a one-size-fits-all practice [60]. Finally, *data and representation inclusiveness* guarantees that the concerns of vulnerable groups are acknowledged and responded to, fostering equity and inclusivity [17]. This entails ensuring that insights from infodemic management reports incorporate the concerns of vulnerable groups and propose tailored solutions.

Principles in Cluster 2

In this cluster, we identified and discussed the most represented substantive principles in cluster 1 (Figure 2B) in the analyzed body of literature, discussed the solutions proposed by the literature to implement and follow them in infodemic management practices, and defined procedural principles that help to ensure that such substantive principles are integrated into infodemic management practices. In Multimedia Appendix 1, we delve into the detailed analysis of principles highlighted in cluster 2 (Figure 2B). These ethical principles are *community engagement*, empowerment through *education*, *transparency*, *free speech versus regulation*, *informing versus manipulating behavior*, *honesty and conflicts of interest*, and *good governance*.

Discussion

The Implications of the Findings

The review has illuminated various crucial ethical considerations that can be instrumental in enhancing the effectiveness of infodemic management. The literature surveyed predominantly emanates from the backdrop of the COVID-19 pandemic. This temporal context is both a strength and a weakness: on the positive side, it signifies a wealth of learnings from a recent global crisis that has driven substantial advances in infodemic management strategies; however, it implies that the ethical readiness for infodemics lacked a solid foundation in evidence. Scientific support for ethics within infodemic management was not yet accessible based on the lessons learned before the COVID-19 pandemic and existing literature. It is promising to see a few studies extending their gaze beyond acute health events to investigate chronic health issues. Still, the field must continuously adapt and evolve as new challenges emerge. In general, the literature on infodemic management still lacks a robust foundation in ethics and ethical considerations. This observation highlights the importance of advocating for expanded research efforts in this domain.

The literature indicates a limited number of experimental empirical approaches within ethics in infodemic management. Only a few studies have taken this approach [51,58,87,91,97], underlining the need for more work of this kind. This limitation is particularly concerning, as empirical research is vital for the improvement of infodemic management strategies, serving as the foundation for prevention and preparedness in the face of future infodemics [152]. Another evident limitation of the existing literature is the dominance of Western approaches to ethics in infodemic management. To ensure a comprehensive understanding and inclusivity in ethical considerations, this Western-centric bias should be addressed by incorporating diverse global ethical perspectives [16].

As highlighted by the literature reviewed in this study, it is paramount to emphasize that ethics is not a hindrance but a tool for enhancing the effectiveness of infodemic management and social listening. The review underscored that ethical considerations are instrumental for achieving medium- and long-term effectiveness in these practices.

A few key ethical aspects have emerged as fundamental for different practices linked to infodemic management and social listening. The first is community engagement, which emerged as a central procedural principle, enhancing the quality and effectiveness of communication, surveillance, and social listening efforts. It not only fosters trust in the institutions carrying out these activities but also contributes to improving the strategies themselves through feedback mechanisms [17].

Second, empowering individuals through educational approaches was identified as a fundamental procedural principle. Education, that is, information and media literacy, equips them with the ability to discern between accurate and inaccurate information [18,88,127]. When facing educated and literate publics, institutions that conduct infodemic management activities need to rely less on censorship or manipulative communication strategies, which are detrimental in the medium and long run [110]. Ethical strategies involving empowerment through education, resilience building, and autonomy are thus vital for the efficacy of infodemic management at all stages [62,110].

Third, the importance of inclusivity and equity extends beyond beneficence; it is integral to effectiveness. This is directly connected to community engagement, ensuring that vulnerable individuals and communities have a voice in the design of infodemic management and social listening strategies improves their effectiveness and helps prevent the formation of pockets of polarized resistance to public health communication [17,45,46]. This inclusivity also applies to minority groups holding opinions that do not reflect science-based evidence, such as antivaxxers; engaging with these groups and listening to their concerns are essential for ethical and effective infodemic management and social listening strategies [48]. Of note, engaging and listening to the concerns of these groups does not imply embracing, endorsing, or justifying their opinions [17,43,44,51,57].

Finally, central to all the principles discussed above is trust. Trust is fundamental in ensuring the publics' receptivity to public health communication and the willingness to share data for social listening purposes. The literature emphasizes that trust plays a crucial role in minimizing the negative effects of information received by the public when such information, albeit being accurate and designed to promote individual and public health, is regarded as manipulative, conspiratorial, and biased toward the interest of the institution that is performing infodemic management activities [90,104,110]. Trustworthy institutions disseminating public health messages encounter less resistance and can leverage the publics' sense of responsibility [18,62,105]. This concept has implications that could extend beyond infodemic management and social listening, possibly impacting democracy and peace, since trustworthy institutions are thriving in nonpolarized information ecosystems [9]. While these aspects are currently underexplored in the literature, we advocate for further exploration of the potential far-reaching effects of maintaining a healthy information ecosystem with trusted actors and educated, autonomous publics.

Some themes and ethical principles remain underrepresented in the literature included in this review. These aspects should not be necessarily considered as less relevant per se in the context of infodemic management since the literature included in this study only represents the views of the research community limited to the period and data taken into consideration by this systematic scoping review; this underrepresentation should rather highlight opportunities for further research and reflection in the field. For example, independent oversight of ethical infodemic management and social listening practices ensures that none of these practices are conducted in unnecessary situations, without considerations about their proportionality [112], valuing transparency and preventing conflicts of interest [36,38]; all these aspects ensure the maintenance or enhancement of institutional trust. A second example is the integration of different cultural perspectives in infodemic management, circling back to the importance of inclusivity and equity since some of the highlighted procedural principles that enhance the effectiveness of infodemic management may not hold the same value and importance in different cultural contexts [46,51]. These underexplored areas should not be underestimated in terms of their ethical significance and potential impact on infodemic management and social listening effectiveness.

In sum, this systematic scoping review provides a comprehensive understanding of the ethical dimensions of infodemic management. It highlights the critical role of ethics in enhancing the effectiveness of these practices and underscores the need for an ethically and empirically informed approach to infodemics. The findings and principles identified in this review are integral to the continuous improvement and adaptation of strategies for tackling infodemics and safeguarding public health. These findings serve as a foundational element for structuring a WHO ethics guidance and a practical implementation framework on the ethics of infodemic management and social listening, which aims to combine learned lessons from the literature and know-how and expert opinions from a WHO EG on ethical considerations in infodemic management and social listening [29].

Limitations

While our systematic scoping review offers valuable insights into the ethical dimensions of infodemic management, it is essential to acknowledge certain limitations that shape the scope and generalizability of our findings. First, our review is based on the literature published between 2002 and 2022 (although it includes a few papers and documents contributed by the WHO EG on infodemic management and social listening published after 2022), thereby excluding potentially relevant studies published before or after our cutoff date. Of note, since the cutoff of our inclusion criteria was at the end of 2022, we initially made reference to the Open Science Framework repository of 1 paper [82]; however, during review in 2023, this paper was published and we decided to link to the published version instead. Furthermore, our research is constrained to the literature published in English. Second, while our search strategy encompassed prominent databases such as PubMed, Scopus, and Web of Science and included additional material with a focus on policy documents, there may be relevant literature not included in our search, including gray literature, news articles, and blog posts. It must also be recognized that the assessment and categorization of papers, as well as the identification of core thematic areas, involve an element of subjectivity. Despite rigorous methodology and intersubjective blinded coding, interpretational variations may exist. Moreover, it is worth noting that the review offers interpretations and recommendations for considering and applying ethics within the field of infodemic management solely based on the ethical considerations and approaches identified within the analyzed body of literature. Given the rapid evolution of this field, it is essential to acknowledge that many pertinent aspects related to the ethics of infodemic management have not yet been thoroughly discussed. Therefore, we strongly encourage the research community, as well as infodemic managers and policy makers, to deepen our understanding of ethics within the context of infodemic management. This commitment to knowledge enhancement is essential for maintaining ethical standards and promoting responsible practices in an ever-changing information landscape. Furthermore, as previously discussed, most studies in our corpus originate from Western contexts, potentially limiting the generalizability of our findings to diverse global settings with distinct cultural and societal norms. Finally, published studies may not fully represent the spectrum of research or of the practice conducted in the field of infodemic management.

Conclusions

Infodemic management presents relevant ethical challenges. The insights derived from our systematic scoping review highlight that ethical approaches in infodemic management and social listening are necessary for the medium- and long-term effectiveness of infodemic management practices. In this review, several fundamental and procedural ethical principles have been identified, including community engagement, education, inclusivity, equity, and trust, among others, all of which enhance the quality and efficacy of these crucial public health activities. Our review provides a foundational understanding of the ethical issues arising in infodemic management. It will hopefully contribute to improving ethical guidance in this field and help to adequately address these issues in future infodemic management programs. To fully realize the potential of ethical infodemic management, future research should strive for empirical studies and incorporate diverse global perspectives to further advance the field and protect public health during acute or chronic health events.

Acknowledgments

The authors would like to acknowledge Tina D Purnat and Atsuyoshi Ishizumi (Unit for High Impact Events Preparedness, Department of Epidemic and Pandemic Preparedness and Prevention, and Health Ethics and Governance Unit, Department of Research for Health, and World Health Organization).

ChatGPT was used for editorial assistance during the editing of this paper to improve the manuscript's readability. The authors have thoroughly checked all the text and take full responsibility for its content. This research, funded by the World Health Organization, was conducted impartially and without any external influences that could have biased the interpretation or presentation of the findings.

Data Availability

All data are accessible within the manuscript, in the Open Science Framework repository, or in the accompanying multimedia appendix.

Conflicts of Interest

The authors alone are responsible for the views expressed in this article and they do not necessarily represent the views, decisions or policies of the institutions with which they are affiliated.

Multimedia Appendix 1
Queries, PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist, supplementary figures, and other emerging themes.
[DOCX File , 253 KB - infodemiology_v4i1e56307_app1.docx]

Multimedia Appendix 2
Issues and aims.
[XLSX File (Microsoft Excel File), 12 KB - infodemiology_v4i1e56307_app2.xlsx]

References

1. Kuehn EF. The information ecosystem concept in information literacy: a theoretical approach and definition. Assoc Inf Sci Technol 2022 Dec 23;74(4):434-443. [doi: 10.1002/asi.24733]
2. 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]
3. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic#tab=tab_1 [accessed 2023-09-11]
4. Ranjit YS, Shin H, First JM, Houston JB. COVID-19 protective model: the role of threat perceptions and informational cues in influencing behavior. J Risk Res 2021 Feb 18;24(3-4):449-465. [doi: 10.1080/13669877.2021.1887328]
5. Aiyer I, Shaik L, Kashyap R, Surani S. COVID-19 misinformation: a potent co-factor in the COVID-19 pandemic. Cureus 2022 Oct;14(10):e30026 [FREE Full text] [doi: 10.7759/cureus.30026] [Medline: 36348909]
6. An ad hoc WHO technical consultation managing the COVID-19 infodemic: call for action. World Health Organization. 2020 Sep 15. URL: https://www.who.int/publications/i/item/9789240010314 [accessed 2023-01-30]
7. WHO public health research agenda for managing infodemics. World Health Organization. 2021 Feb 3. URL: https://www.who.int/publications/i/item/9789240019508 [accessed 2023-01-30]
8. Prainsack B, do Céu Patrão Neves M, Sahlin NE, Biller-Andorno N, Laukyte M, Łuków P, et al. Values in challenging times: strategic crisis management in the EU. Lancet Reg Health Eur 2023 Jan;24:100553 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100553] [Medline: 36643663]
9. European Commission, Directorate-General for Research and Innovation, European Group on Ethics in Science and New Technologies. Opinion on democracy in the digital age. Publications Office of the European Union. 2023. URL: https://data.europa.eu/doi/10.2777/773647 [accessed 2023-11-01]
10. Spitale G, Germani F, Biller-Andorno N, Redaelli S, Glöckler S, Brown J. Can critical thinking skills help us recognize misinformation better? OSF Home. 2022. URL: https://osf.io/f3467/ [accessed 2023-11-02]
11. Khan ML, Idris IK. Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. Behav Inf Technol 2019 Feb 11;38(12):1194-1212 [FREE Full text] [doi: 10.1080/0144929x.2019.1578828]

12. Orhan A. Fake news detection on social media: the predictive role of university students' critical thinking dispositions and new media literacy. Smart Learn Environ 2023 Apr 26;10:29. [doi: 10.1186/s40561-023-00248-8]
13. Kerr J, Panagopoulos C, van der Linden S. Political polarization on COVID-19 pandemic response in the United States. Pers Individ Dif 2021 Sep;179:110892 [FREE Full text] [doi: 10.1016/j.paid.2021.110892] [Medline: 34866723]
14. Briand S, Hess S, Nguyen T, Purnat TD. Infodemic management in the twenty-first century. In: Purnat TD, Nguyen T, Briand S, editors. Managing Infodemics in the 21st Century. Cham, Switzerland: Springer; 2023.
15. Purnat TD, Vacca P, Czerniak C, Ball S, Burzo S, Zecchin T, 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 Jul 28;1(1):e30971 [FREE Full text] [doi: 10.2196/30971] [Medline: 34447926]
16. Purnat TD, Nguyen T, Briand S. Managing Infodemics in the 21st Century: Addressing New Public Health Challenges in the Information Ecosystem. Cham, Switzerland: Springer; 2023.
17. Spitale G, Germani F, Biller-Andorno N. The PHERCC matrix. An ethical framework for planning, governing, and evaluating risk and crisis communication in the context of public health emergencies. Am J Bioeth 2024 Apr;24(4):67-82 [FREE Full text] [doi: 10.1080/15265161.2023.2201191] [Medline: 37114888]
18. Biller-Andorno N, Spitale G. Addressing volatile ethical issues of Covid-19 with the core five enduring values list for health care professionals. NEJM Catalyst. 2022 Aug 4. URL: https://catalyst.nejm.org/doi/full/10.1056/CAT.22.0108 [accessed 2022-08-05]
19. Lotto M, Hanjahanja-Phiri T, Padalko H, Oetomo A, Butt ZA, Boger J, et al. Ethical principles for infodemiology and infoveillance studies concerning infodemic management on social media. Front Public Health 2023;11:1130079 [FREE Full text] [doi: 10.3389/fpubh.2023.1130079] [Medline: 37033062]
20. Bayram AB, Shields T. Who trusts the WHO? Heuristics and Americans' trust in the World Health Organization during the COVID-19 pandemic. Soc Sci Q 2021 Sep;102(5):2312-2330 [FREE Full text] [doi: 10.1111/ssqu.12977] [Medline: 34226772]
21. Benoit WL. Image repair discourse and crisis communication. Public Relat Rev 1997 Jun;23(2):177-186. [doi: 10.1016/s0363-8111(97)90023-0]
22. Coombs WT. Attribution Theory as a guide for post-crisis communication research. Public Relat Rev 2007 Jun;33(2):135-139. [doi: 10.1016/j.pubrev.2006.11.016]
23. Xu K, Li W. An ethical stakeholder approach to crisis communication: a case study of Foxconn's 2010 employee suicide crisis. J Bus Ethics 2012 Oct 26;117(2):371-386. [doi: 10.1007/s10551-012-1522-0]
24. Contreras-Pacheco OE. Care ethics and crisis communication: examining two experiences in South America. Cuad Adm 2018 Oct 16;34(62):20-32. [doi: 10.25100/10.25100/cdea.2018v34n62.6866]
25. Machiri S, Purnat T, Nguyen T, Ho C, Ballalai I, Biller-Andorno N, et al. An ethics framework for social listening and infodemic management. Eur J Public Health 2023 Oct;33(Suppl 2):ckad160.661. [doi: 10.1093/eurpub/ckad160.661]
26. Global health ethics: key issues. World Health Organization. 2015 Jul 16. URL: https://www.who.int/publications/i/item/global-health-ethics-key-issues [accessed 2024-08-15]
27. WHO competency framework: building a response workforce to manage infodemics. World Health Organization. 2021 Sep 15. URL: https://www.who.int/publications/i/item/9789240035287 [accessed 2023-01-30]
28. Wilhelm E, Ballalai I, Belanger ME, Benjamin P, Bertrand-Ferrandis C, Bezbaruah S, et al. Measuring the burden of infodemics: summary of the methods and results of the fifth WHO Infodemic Management Conference. JMIR Infodemiology 2023 Feb 20;3:e44207 [FREE Full text] [doi: 10.2196/44207] [Medline: 37012998]
29. WHO kicks off deliberations on ethical framework and tools for social listening and infodemic management. World Health Organization. 2023 Feb 10. URL: https://www.who.int/news/item/10-02-2023-who-kicks-off-deliberations-on-ethical-framework-and-tools-for-social-listening-and-infodemic-management [accessed 2023-11-02]
30. Singer P. Practical Ethics, Third Edition. Cambridge, MA: Cambridge University Press; 2011.
31. Harris J. The Value of Life: An Introduction to Medical Ethics. Milton Park, UK: Routledge; 1985.
32. Floridi L. Tolerant paternalism: pro-ethical design as a resolution of the dilemma of toleration. Sci Eng Ethics 2016 Dec 9;22(6):1669-1688. [doi: 10.1007/s11948-015-9733-2] [Medline: 26649432]
33. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J 2009 Jun 27;26(2):91-108 [FREE Full text] [doi: 10.1111/j.1471-1842.2009.00848.x] [Medline: 19490148]
34. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
35. 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]
36. Lovari A, Bowen S. Social media in disaster communication: a case study of strategies, barriers, and ethical implications. J Public Aff 2019 Jun 13;20(1):e1967. [doi: 10.1002/pa.1967]
37. Digital tools and strategies in COVID-19 infodemic response: case studies and discussion. International Telecommunication Union. 2021. URL: https://www.itu.int/pub/D-STR-ICT_APP-2021-01 [accessed 2023-01-30]

38. Dupras C, Williams-Jones B. The expert and the lay public: reflections on influenza A (H1N1) and the risk society. Am J Public Health 2012 Apr;102(4):591-595. [doi: 10.2105/ajph.2011.300417]
39. Ćurković M, Košec A, Roje Bedeković M, Bedeković V. Epistemic responsibilities in the COVID-19 pandemic: is a digital infosphere a friend or a foe? J Biomed Inform 2021 Mar;115:103709 [FREE Full text] [doi: 10.1016/j.jbi.2021.103709] [Medline: 33571677]
40. Germani F, Biller-Andorno N. How to counter the anti-vaccine rhetoric: filling information voids and building resilience. Hum Vaccin Immunother 2022 Nov 30;18(6):2095825 [FREE Full text] [doi: 10.1080/21645515.2022.2095825] [Medline: 35802046]
41. Lor A, Thomas JC, Barrett DH, Ortmann LW, Herrera Guibert DJ. Key ethical issues discussed at CDC-sponsored international, regional meetings to explore cultural perspectives and contexts on pandemic influenza preparedness and response. Int J Health Policy Manag 2016 Nov 01;5(11):653-662 [FREE Full text] [doi: 10.15171/ijhpm.2016.55] [Medline: 27801360]
42. Ali S, Khalid A, Zahid E. Is COVID-19 immune to misinformation? A brief overview. Asian Bioeth Rev 2021 Jun 23;13(2):255-277 [FREE Full text] [doi: 10.1007/s41649-020-00155-x] [Medline: 33777228]
43. Knudsen J, Perlman-Gabel M, Uccelli IG, Jeavons J, Chokshi DA. Combating misinformation as a core function of public health. NEJM Catal Innov Care Deliv 2023 Jan 18;4(2). [doi: 10.1056/cat.22.0198]
44. Posada A, Iñigo RL, Sport J. Turning social listening data into action: barriers and recommendations observed through a COVID-19 rumor response, August 2022. Relief Web. 2022 Aug 31. URL: https://reliefweb.int/report/world/turning-social-listening-data-action-barriers-and-recommendations-observed-through-covid-19-rumor-response-august-2022 [accessed 2023-03-17]
45. Vanderslott S, Van Ryneveld M, Marchant M, Lees S, Nolna SK, Marsh V. How can community engagement in health research be strengthened for infectious disease outbreaks in Sub-Saharan Africa? A scoping review of the literature. BMC Public Health 2021 Apr 01;21(1):633 [FREE Full text] [doi: 10.1186/s12889-021-10348-0] [Medline: 33794820]
46. Charania NA, Tsuji LJ. A community-based participatory approach and engagement process creates culturally appropriate and community informed pandemic plans after the 2009 H1N1 influenza pandemic: remote and isolated First Nations communities of sub-arctic Ontario, Canada. BMC Public Health 2012 Apr 03;12:268 [FREE Full text] [doi: 10.1186/1471-2458-12-268] [Medline: 22472012]
47. Taylor J, Pagliari C. Mining social media data: how are research sponsors and researchers addressing the ethical challenges? Res Ethics 2017 Oct 26;14(2):1-39. [doi: 10.1177/1747016117738559]
48. Spitale G, Biller-Andorno N, Germani F. Concerns around opposition to the green pass in Italy: social listening analysis by using a mixed methods approach. J Med Internet Res 2022 Feb 16;24(2):e34385 [FREE Full text] [doi: 10.2196/34385] [Medline: 35156930]
49. Stevens G, O'Donnell VL, Williams L. Public domain or private data? Developing an ethical approach to social media research in an inter-disciplinary project. Educ Res Eval 2015 Apr 14;21(2):154-167. [doi: 10.1080/13803611.2015.1024010]
50. WHO guidelines on ethical issues in public health surveillance. World Health Organization. 2017 Jun 19. URL: https://www.who.int/publications/i/item/who-guidelines-on-ethical-issues-in-public-health-surveillance [accessed 2024-08-02]
51. Grande D, Mitra N, Marti XL, Merchant R, Asch D, Dolan A, et al. Consumer views on using digital data for COVID-19 control in the United States. JAMA Netw Open 2021 May 03;4(5):e2110918 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.10918] [Medline: 34009347]
52. Hendl T, Chung R, Wild V. Pandemic surveillance and racialized subpopulations: mitigating vulnerabilities in COVID-19 apps. J Bioeth Inq 2020 Dec;17(4):829-834 [FREE Full text] [doi: 10.1007/s11673-020-10034-7] [Medline: 32840858]
53. Abboah-Offei M, Gyasi Darkwa A, Ayim A, Ansah-Ofei AM, Dovlo D, Awoonor-Williams JK, et al. Adapting the Community-based Health Planning and Services (CHPS) to engage poor urban communities in Ghana: protocol for a participatory action research study. BMJ Open 2021 Jul 27;11(7):e049564 [FREE Full text] [doi: 10.1136/bmjopen-2021-049564] [Medline: 34315798]
54. Fahey RA, Hino A. COVID-19, digital privacy, and the social limits on data-focused public health responses. Int J Inf Manage 2020 Dec;55:102181 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102181] [Medline: 32836638]
55. Dash S, Parray AA, De Freitas L, Mithu MI, Rahman MM, Ramasamy A, et al. Combating the COVID-19 infodemic: a three-level approach for low and middle-income countries. BMJ Glob Health 2021 Jan 29;6(1):e004671 [FREE Full text] [doi: 10.1136/bmjgh-2020-004671] [Medline: 33514596]
56. Taylor L. The price of certainty: how the politics of pandemic data demand an ethics of care. Big Data Soc 2020 Jul 22;7(2). [doi: 10.1177/2053951720942539]
57. Cohen ER, Masum H, Berndtson K, Saunders V, Hadfield T, Panjwani D, et al. Public engagement on global health challenges. BMC Public Health 2008 May 20;8:168 [FREE Full text] [doi: 10.1186/1471-2458-8-168] [Medline: 18492256]
58. Green J, Petty J, Whiting L, Orr F, Smart L, Brown AM, et al. 'Blurred boundaries': when nurses and midwives give anti-vaccination advice on Facebook. Nurs Ethics 2022 May;29(3):552-568. [doi: 10.1177/09697330211041749] [Medline: 35142239]

59. Morley J, Cowls J, Taddeo M, Floridi L. Public health in the information age: recognizing the Infosphere as a social determinant of health. J Med Internet Res 2020 Aug 03;22(8):e19311 [FREE Full text] [doi: 10.2196/19311] [Medline: 32648850]
60. Spitale G, Merten S, Jafflin K, Schwind B, Kaiser-Grolimund A, Biller-Andorno N. A novel risk and crisis communication platform to bridge the gap between policy makers and the public in the context of the COVID-19 crisis (PubliCo): protocol for a mixed methods study. JMIR Res Protoc 2021 Nov 01;10(11):e33653 [FREE Full text] [doi: 10.2196/33653] [Medline: 34612823]
61. Oxman AD, Fretheim A, Lewin S, Flottorp S, Glenton C, Helleve A, et al. Health communication in and out of public health emergencies: to persuade or to inform? Health Res Policy Syst 2022 Mar 05;20(1):28 [FREE Full text] [doi: 10.1186/s12961-022-00828-z] [Medline: 35248064]
62. Lor P, Wiles B, Britz J. Re-thinking information ethics: truth, conspiracy theories, and librarians in the COVID-19 era. Libri 2021;71(1):1-14 [FREE Full text] [doi: 10.1515/libri-2020-0158]
63. Sambala EZ, Manderson L. Ethical problems in planning for and responses to pandemic influenza in Ghana and Malawi. Ethics Behav 2017 Feb 02;28(3):199-217. [doi: 10.1080/10508422.2016.1274993]
64. Marco-Franco JE, Pita-Barros P, Vivas-Orts D, González-de-Julián S, Vivas-Consuelo D. COVID-19, fake news, and vaccines: should regulation be implemented? Int J Environ Res Public Health 2021 Jan 16;18(2):744 [FREE Full text] [doi: 10.3390/ijerph18020744] [Medline: 33467179]
65. Finding the signal through the noise: a landscape review and framework to enhance the effective use of digital social listening for immunisation demand generation. Gavi, UNICEF, Vaccination Demand Hub and World Health Organization. 2021 Jun. URL: https://www.gavi.org/sites/default/files/2021-06/Finding-the-Signal-Through-the-Noise.pdf [accessed 2024-08-02]
66. Sniečkutė M, Gaižauskatė I. COVID-19 crisis: government's (dis)trust in the people and pitfalls of liberal democracies. Partecipazione e Conflitto 2021;14(1):152-175. [doi: 10.1285/i20356609v14i1p152]
67. Ummer O, Scott K, Mohan D, Chakraborty A, LeFevre AE. Connecting the dots: Kerala's use of digital technology during the COVID-19 response. BMJ Glob Health 2021 Jul 26;6(Suppl 5):e005355 [FREE Full text] [doi: 10.1136/bmjgh-2021-005355] [Medline: 34312152]
68. Stommel W, de Rijk L. Ethical approval: none sought. How discourse analysts report ethical issues around publicly available online data. Res Ethics 2021 Jan 19;17(3):275-297. [doi: 10.1177/1747016120988767]
69. Purnat TD, Ishizumi A, Yau B, White B, Bertrand-Ferrandis C, Briand S, et al. Delivering actionable infodemic insights and recommendations for the COVID-19 pandemic response. Eur J Public Health 2022 Oct;32(Supplement_3):ckac129.645 [FREE Full text] [doi: 10.1093/eurpub/ckac129.645]
70. Mills MC, Sivelä J. Should spreading anti-vaccine misinformation be criminalised? BMJ 2021 Feb 17;372:n272. [doi: 10.1136/bmj.n272] [Medline: 33597153]
71. Niemiec E. COVID-19 and misinformation: is censorship of social media a remedy to the spread of medical misinformation? EMBO Rep 2020 Nov 05;21(11):e51420 [FREE Full text] [doi: 10.15252/embr.202051420] [Medline: 33103289]
72. Shared UN considerations for online communications companies on the issues of countering disinformation and enhancing transparency. United Nations. 2021. URL: https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/general/UN_InteragencyDialogue2_v2.pdf [accessed 2024-08-02]
73. Cohen CB. The ethics of human reproductive cloning: when world views collide. Account Res 2004;11(3-4):183-199. [doi: 10.1080/08989620490891386] [Medline: 15812964]
74. Willingham DT. Ask the cognitive scientist: how can educators teach critical thinking? Education Healthcare Public Services. 2020. URL: https://www.aft.org/ae/fall2020/willingham#:~:text=Critical%20Thinking%20Can%20Be%20Taught,that%20are%20open%20to%20instruction [accessed 2023-11-20]
75. Spitale G, Germani F, Biller-Andorno N. Can AI disinform us better? OSF Home. 2022. URL: https://osf.io/9ntgf/ [accessed 2024-08-02]
76. Digital solutions to health risks raised by the COVID-19 infodemic: policy brief. World Health Organization. Regional Office for Europe. 2022. URL: https://iris.who.int/handle/10665/356315 [accessed 2023-01-30]
77. Monti C, Cinelli M, Valensise C, Quattrociocchi W, Starnini M. Online conspiracy communities are more resilient to deplatforming. PNAS Nexus 2023 Oct 31;2(10):pgad324 [FREE Full text] [doi: 10.1093/pnasnexus/pgad324] [Medline: 37920549]
78. Marco-Franco JE, Pita-Barros P, Vivas-Orts D, González-de-Julián S, Vivas-Consuelo D. COVID-19, Fake News, and Vaccines: Should Regulation Be Implemented? IJERPH 2021 Jan 16;18(2):744. [doi: 10.3390/ijerph18020744]
79. Spitale G, Biller-Andorno N. TopicTracker: a Python pipeline to search, download and explore PubMed entries. Zenodo. 2021 Feb 8. URL: https://zenodo.org/records/6373635 [accessed 2024-08-02]
80. Spitale G. Making sense in the flood. How to cope with the massive flow of digital information in medical ethics. Heliyon 2020 Jul;6(7):e04426 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e04426] [Medline: 32743090]
81. Roozenbeek J, Culloty E, Suiter J. Countering misinformation. Eur Psychol 2023 Jul;28(3):189-205. [doi: 10.1027/1016-9040/a000492]

82. Purnat T, Pundir P, Machiri S, Ishizumi A, Rajwar E, Murthy S, et al. Evidence and gap map of infodemic management interventions in emergencies: a case of COVID-19. Eur J Public Health 2023 Oct 24;33(Supplement_2):ckad160.1676. [doi: 10.1093/eurpub/ckad160.1676]
83. Spitale G, Germani F, Mandić-Rajčević S. Ethical considerations on social listening and infodemic management. OSF Home. 2023. URL: https://osf.io/28d73/ [accessed 2024-08-02]
84. Spitale G, Germani F. Records for systematic and scoping review on ethical considerations in infodemic management. Zotero. 2023. URL: https://www.zotero.org/groups/4931113/who_infodemics_ethics_review/library [accessed 2024-08-02]
85. RevMaster documentation. GitHub. URL: https://github.com/that-ugly-cat/revmaster [accessed 2023-11-02]
86. Sambala EZ, Kanyenda T, Iwu CJ, Iwu CD, Jaca A, Wiysonge CS. Pandemic influenza preparedness in the WHO African region: are we ready yet? BMC Infect Dis 2018 Nov 14;18(1):567 [FREE Full text] [doi: 10.1186/s12879-018-3466-1] [Medline: 30428846]
87. Gorman DR, Bielecki K, Larson HJ, Willocks LJ, Craig J, Pollock KG. Comparing vaccination hesitancy in Polish migrant parents who accept or refuse nasal flu vaccination for their children. Vaccine 2020 Mar 17;38(13):2795-2799. [doi: 10.1016/j.vaccine.2020.02.028] [Medline: 32089460]
88. Paakkari L, Okan O, Torppa M. "Missed" information: a moral failing that erodes efforts to tackle the COVID-19 pandemic. Int J Public Health 2021;66:1604667 [FREE Full text] [doi: 10.3389/ijph.2021.1604667] [Medline: 35095385]
89. Calabrese SK, Mayer KH, Marcus JL. Prioritising pleasure and correcting misinformation in the era of U=U. Lancet HIV 2021 Mar;8(3):e175-e180 [FREE Full text] [doi: 10.1016/S2352-3018(20)30341-6] [Medline: 33662266]
90. Hughes B, Miller-Idriss C, Piltch-Loeb R, Goldberg B, White K, Criezis M, et al. Development of a codebook of online anti-vaccination rhetoric to manage COVID-19 vaccine misinformation. Int J Environ Res Public Health 2021 Jul 15;18(14):7556 [FREE Full text] [doi: 10.3390/ijerph18147556] [Medline: 34300005]
91. Gupta L, Gasparyan AY, Misra DP, Agarwal V, Zimba O, Yessirkepov M. Information and misinformation on COVID-19: a cross-sectional survey study. J Korean Med Sci 2020 Jul 13;35(27):e256 [FREE Full text] [doi: 10.3346/jkms.2020.35.e256] [Medline: 32657090]
92. Antiochou K. Science communication: challenges and dilemmas in the age of COVID-19. Hist Philos Life Sci 2021 Jul 09;43(3):87 [FREE Full text] [doi: 10.1007/s40656-021-00444-0] [Medline: 34241709]
93. Ha BT, Ngoc Quang L, Quoc Thanh P, Duc DM, Mirzoev T, Bui TM. Community engagement in the prevention and control of COVID-19: insights from Vietnam. PLoS One 2021 Sep 8;16(9):e0254432 [FREE Full text] [doi: 10.1371/journal.pone.0254432] [Medline: 34495962]
94. Fafard P, Wilson LA, Cassola A, Hoffman SJ. Communication about COVID-19 from Canadian provincial chief medical officers of health: a qualitative study. CMAJ Open 2020 Sep 4;8(3):E560-E567 [FREE Full text] [doi: 10.9778/cmajo.20200110] [Medline: 32887695]
95. Armitage L, Lawson BK, Whelan ME, Newhouse N. Paying SPECIAL consideration to the digital sharing of information during the COVID-19 pandemic and beyond. BJGP Open 2020 Apr 28;4(2):bjgpopen20X101072. [doi: 10.3399/bjgpopen20x101072]
96. Boender S, Schneider P, Houareau C, Wehrli S, Purnat TD, Ishizumi A, et al. Establishing infodemic management in Germany: a framework for social listening and integrated analysis to report infodemic insights at the National Public Health Institute. JMIR Infodemiology 2023 Jun 01;3:e43646 [FREE Full text] [doi: 10.2196/43646] [Medline: 37261891]
97. Cheng IH, Lee ST. Strategic and ethical issues in antismoking message development: how US public health campaigners conceptualize efficacy and ethicality. Int J Health Promot Educ 2012 Sep;50(5):238-250. [doi: 10.1080/14635240.2012.723373]
98. Frank JW, Pagliari C, Geubbels E, Mtenga S. 'New forms of data for understanding low- and middle-income countries' health inequalities: the case of Tanzania. J Glob Health 2018;8(2):020302. [doi: 10.7189/jogh.08.020302]
99. Larson HJ. Blocking information on COVID-19 can fuel the spread of misinformation. Nature 2020 Apr;580(7803):306. [doi: 10.1038/d41586-020-00920-w] [Medline: 32231320]
100. Glanz JM, Kraus CR, Daley MF. Addressing parental vaccine concerns: engagement, balance, and timing. PLoS Biol 2015 Aug;13(8):e1002227 [FREE Full text] [doi: 10.1371/journal.pbio.1002227] [Medline: 26252770]
101. Chakraborti C. Pandemic management and developing world bioethics: bird flu in West Bengal. Dev World Bioeth 2009 Dec;9(3):161-166. [doi: 10.1111/j.1471-8847.2008.00240.x] [Medline: 18699841]
102. Jun J, Zain A, Chen Y, Kim S. Adverse mentions, negative sentiment, and emotions in COVID-19 vaccine tweets and their association with vaccination uptake: global comparison of 192 countries. Vaccines (Basel) 2022 May 08;10(5):735 [FREE Full text] [doi: 10.3390/vaccines10050735] [Medline: 35632491]
103. Carrera JS, Key K, Bailey S, Hamm JA, Cuthbertson CA, Lewis EY, et al. Community science as a pathway for resilience in response to a public health crisis in Flint, Michigan. Soc Sci 2019 Mar 13;8(3):94. [doi: 10.3390/socsci8030094]
104. Wilkinson TM. Counter-manipulation and health promotion. Public Health Ethics 2017;10(3):257-266.
105. Fleming N. Coronavirus misinformation, and how scientists can help to fight it. Nature 2020 Jul;583(7814):155-156. [doi: 10.1038/d41586-020-01834-3] [Medline: 32601491]
106. Verma N, Fleischmann KR, Zhou L, Xie B, Lee MK, Rich K, et al. Trust in COVID-19 public health information. J Assoc Inf Sci Technol 2022 Sep 20:10.1002/asi.24712 [FREE Full text] [doi: 10.1002/asi.24712] [Medline: 36246042]

107. Qureshi KA, Malick RA, Sabih M, Cherifi H. Complex network and source inspired COVID-19 fake news classification on Twitter. IEEE Access 2021;9:139636-139656. [doi: 10.1109/access.2021.3119404]
108. Unim B, Schutte N, Thissen M, Palmieri L. Innovative methods used in monitoring COVID-19 in Europe: a multinational study. Int J Environ Res Public Health 2022 Dec 29;20(1):564 [FREE Full text] [doi: 10.3390/ijerph20010564] [Medline: 36612884]
109. Farrell LC, Warin MJ, Moore VM, Street JM. Emotion in obesity discourse: understanding public attitudes towards regulations for obesity prevention. Sociol Health Illn 2016 May;38(4):543-558. [doi: 10.1111/1467-9566.12378] [Medline: 26564262]
110. Freiling I, Krause NM, Scheufele DA. Science and ethics of "curing" misinformation. AMA J Ethics 2023 Mar 01;25(3):E228-E237 [FREE Full text] [doi: 10.1001/amajethics.2023.228] [Medline: 36867171]
111. Islam MS, Kamal AH, Kabir A, Southern DL, Khan SH, Hasan SM, 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]
112. Petrini C, Ricciardi G. Ethical issues in public health surveillance: drawing inspiration from ethical frameworks. Ann Ist Super Sanita 2015;51(4):270-276 [FREE Full text] [doi: 10.4415/ANN_15_04_05] [Medline: 26783212]
113. Scott I. (Mis)communication? Social listening and the exclusion of marginalised voices. The Humanitarian Leader. 2022 Mar 14. URL: https://ojs.deakin.edu.au/index.php/thl/article/view/1558 [accessed 2024-08-02]
114. Sivaram S, Srikrishnan AK, Murgavel KG, Mayer KH, Anand S, Celentano DD, et al. An approach to addressing ethical issues in a community-based risk assessment for HIV: a case from Chennai, India. J Health Popul Nutr 2005 Jun;23(2):165-176. [Medline: 16117369]
115. Subbian V, Solomonides A, Clarkson M, Rahimzadeh VN, Petersen C, Schreiber R, et al. Ethics and informatics in the age of COVID-19: challenges and recommendations for public health organization and public policy. J Am Med Inform Assoc 2021 Jan 15;28(1):184-189 [FREE Full text] [doi: 10.1093/jamia/ocaa188] [Medline: 32722749]
116. Dar AA, Alam MZ, Ahmad A, Reegu FA, Rahin SA. Blockchain framework for secure COVID-19 pandemic data handling and protection. Comput Intell Neurosci 2022;2022:7025485 [FREE Full text] [doi: 10.1155/2022/7025485] [Medline: 36156957]
117. McGloin AF, Eslami S. Digital and social media opportunities for dietary behaviour change. Proc Nutr Soc 2015 May;74(2):139-148. [doi: 10.1017/S0029665114001505] [Medline: 25319345]
118. Vériter SL, Bjola C, Koops JA. Tackling COVID-19 disinformation: internal and external challenges for the European Union. Hague J Dipl 2020 Oct 21;15(4):569-582. [doi: 10.1163/1871191x-bja10046]
119. Valenti A, Mirabile M, Cannone E, Boccuni F, Dionisi P, Fortuna G, et al. The impact of COVID-19 pandemics on the development of health risk communication: challenges and opportunities. Int J Environ Res Public Health 2022 Dec 30;20(1):645 [FREE Full text] [doi: 10.3390/ijerph20010645] [Medline: 36612966]
120. Tadros E, Morgan AA, Durante KA. Criticism, compassion, and conspiracy theories: a thematic analysis of what Twitter users are saying about COVID-19 in correctional settings. Int J Offender Ther Comp Criminol 2024 Mar;68(4):370-388. [doi: 10.1177/0306624X221102847] [Medline: 35703315]
121. Hotez PJ. America's deadly flirtation with antiscience and the medical freedom movement. J Clin Invest 2021 Apr 01;131(7):e149072 [FREE Full text] [doi: 10.1172/JCI149072] [Medline: 33630759]
122. Hurford B, Rana A, Sachan RS. COMMENT: narrative-based misinformation in India about protection against Covid-19: not just another "moo-point". Indian J Med Ethics 2022;VII(1):1-10 [FREE Full text] [doi: 10.20529/IJME.2021.050] [Medline: 34730095]
123. Daud M, Abd Ghani Azmi IM. Digital disinformation and the need for internet co-regulation in Malaysia. Pertanika J Soc Sci Humanit 2021 May 17;29(S2). [doi: 10.47836/pjssh.29.s2.12]
124. Selection of data from online platforms that would enable better understanding of disinformation online and efforts to counter it. United Nations. URL: https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/general/UN_InteragencyDialogue1_v2_0.pdf [accessed 2024-08-02]
125. McNaughton D, Duong TT. Designing a community engagement framework for a new dengue control method: a case study from central Vietnam. PLoS Negl Trop Dis 2014 May 22;8(5):e2794 [FREE Full text] [doi: 10.1371/journal.pntd.0002794] [Medline: 24853391]
126. Florig HK. An analysis of public-interest group positions on radiation protection. Health Physics 2006 Nov;91(5):508-513. [doi: 10.1097/01.hp.0000232650.11844.74]
127. Simpson JK. Appeal to fear in health care: appropriate or inappropriate? Chiropr Man Therap 2017 Sep 20;25:27 [FREE Full text] [doi: 10.1186/s12998-017-0157-8] [Medline: 28932388]
128. Scott J. Managing the infodemic about COVID-19: strategies for clinicians and researchers. Acta Psychiatr Scand 2021 May;143(5):377-379 [FREE Full text] [doi: 10.1111/acps.13290] [Medline: 33861872]
129. Fontanin M. On fake news, gatekeepers and LIS professionals: the finger or the moon? Digit Libr Perspect 2021 Jan 26;37(2):168-178. [doi: 10.1108/dlp-09-2020-0097]

130. Jamison A, Broniatowski DA, Smith MC, Parikh KS, Malik A, Dredze M, et al. Adapting and extending a typology to identify vaccine misinformation on Twitter. Am J Public Health 2020 Oct;110(S3):S331-S339. [doi: 10.2105/AJPH.2020.305940] [Medline: 33001737]
131. Ben Messaoud M. Social media and the COVID-19 pandemic: the dilemma of fake news clutter vs. social responsibility. J Arab Muslim Media Res 2021 Apr;14(1):25-45. [doi: 10.1386/jammr_00023_1]
132. Vande Velde F, Hamed A, Lange JS, Sælid T, Bastien S. Assessing student perceptions of a Norwegian University's COVID-19 response strategy: a cross-sectional study. Front Public Health 2021 Aug 20;9:700542 [FREE Full text] [doi: 10.3389/fpubh.2021.700542] [Medline: 34490184]
133. Segev T, Harvey AP, Ajmani A, Johnson C, Longfellow W, Vandiver KM, et al. A case study in participatory science with mutual capacity building between university and tribal researchers to investigate drinking water quality in rural Maine. Environ Res 2021 Jan;192:110460 [FREE Full text] [doi: 10.1016/j.envres.2020.110460] [Medline: 33217437]
134. Modin PG, Hansson SO. Moral and instrumental norms in food risk communication. J Bus Ethics 2011 Jan 18;101:313-324. [doi: 10.1007/s10551-010-0724-6]
135. Slavik CE, Buttle C, Sturrock SL, Darlington JC, Yiannakoulias N. Examining tweet content and engagement of Canadian public health agencies and decision makers during COVID-19: mixed methods analysis. J Med Internet Res 2021 Mar 11;23(3):e24883 [FREE Full text] [doi: 10.2196/24883] [Medline: 33651705]
136. Okposen B. The evolution of Nigeria's COVID-19 response: how Nigeria have worked to promote vaccine demand, manage the infodemic and restore routine immunization. National Primary Health Care Development Agency. 2022. URL: https://demandhub.org/wp-content/uploads/2022/08/1.5-Nigeria-Update.pdf [accessed 2024-08-02]
137. Mahroum N, Watad A, Rosselli R, Brigo F, Chiesa V, Siri A, et al. An infodemiological investigation of the so-called "Fluad effect" during the 2014/2015 influenza vaccination campaign in Italy: ethical and historical implications. Hum Vaccin Immunother 2018 Mar 04;14(3):712-718 [FREE Full text] [doi: 10.1080/21645515.2017.1420448] [Medline: 29293392]
138. Marzo RR, Chen HW, Abid K, Chauhan S, Kaggwa MM, Essar MY, et al. Adapted digital health literacy and health information seeking behavior among lower income groups in Malaysia during the COVID-19 pandemic. Front Public Health 2022;10:998272 [FREE Full text] [doi: 10.3389/fpubh.2022.998272] [Medline: 36187682]
139. Stand on guard for thee: ethical considerations in preparedness planning for pandemic influenza. University of Toronto Joint Centre for Bioethics. 2005 Nov. URL: https://jcb.utoronto.ca/wp-content/uploads/2021/03/stand_on_guard.pdf [accessed 2024-08-02]
140. Oughton D, Forsberg EM, Bay I, Kaiser M, Howard B. An ethical dimension to sustainable restoration and long-term management of contaminated areas. J Environ Radioact 2004 Jan;74(1-3):171-183. [doi: 10.1016/j.jenvrad.2004.01.009] [Medline: 15063546]
141. Lindau ST, Makelarski JA, Chin MH, Desautels S, Johnson D, Johnson WEJ, et al. Building community-engaged health research and discovery infrastructure on the South Side of Chicago: science in service to community priorities. Prev Med 2011 Jan 12;52(3-4):200-207 [FREE Full text] [doi: 10.1016/j.ypmed.2011.01.001] [Medline: 21236295]
142. Piltch-Loeb R, Savoia E, Goldberg B, Hughes B, Verhey T, Kayyem J, et al. Examining the effect of information channel on COVID-19 vaccine acceptance. PLoS One 2021;16(5):e0251095 [FREE Full text] [doi: 10.1371/journal.pone.0251095] [Medline: 33979370]
143. Smith MJ, Upshur RE. Learning lessons from COVID-19 requires recognizing moral failures. J Bioeth Inq 2020 Dec;17(4):563-566 [FREE Full text] [doi: 10.1007/s11673-020-10019-6] [Medline: 32840847]
144. Purnat T. An ounce of prevention = a pound of cure: a case for investing in infodemic management. World Health Organization. URL: https://onedrive.live.com/view.aspx?resid=9F4BC3107673E8D3!47296&ithint=file%2cpptx&authkey=!AI1hUcN0Yi6sHwY [accessed 2023-03-17]
145. 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 01;24(2):e35552 [FREE Full text] [doi: 10.2196/35552] [Medline: 35007204]
146. Gisondi MA, Chambers D, La TM, Ryan A, Shankar A, Xue A, et al. A stanford conference on social media, ethics, and COVID-19 misinformation (INFODEMIC): qualitative thematic analysis. J Med Internet Res 2022 Feb 15;24(2):e35707 [FREE Full text] [doi: 10.2196/35707] [Medline: 35030089]
147. Keim ME, Noji E. Emergent use of social media: a new age of opportunity for disaster resilience. Am J Disaster Med 2011;6(1):47-54. [Medline: 21466029]
148. Adhikari B, Pell C, Cheah PY. Community engagement and ethical global health research. Glob Bioeth 2020 Dec 20;31(1):1-12 [FREE Full text] [doi: 10.1080/11287462.2019.1703504] [Medline: 32002019]
149. Milsom P, Cryan C, Sternthal S, Harmston C. P15-23. Community engagement strategies in the Canadian HIV vaccine initiative. Retrovirology 2009 Oct 22;6(S3):P224. [doi: 10.1186/1742-4690-6-s3-p224]
150. Pratt B, de Vries J. Community engagement in global health research that advances health equity. Bioethics 2018 Sep;32(7):454-463. [doi: 10.1111/bioe.12465] [Medline: 30035349]

151. McIntosh S, Pérez-Ramos JG, David T, Demment MM, Avendaño E, Ossip DJ, et al. A globally networked hybrid approach to public health capacity training for maternal health professionals in low and middle income countries. Glob Health Res Policy 2017 Mar 1;2(1):8 [FREE Full text] [doi: 10.1186/s41256-017-0027-x] [Medline: 29202076]
152. van der Linden S. We need a gold standard for randomised control trials studying misinformation and vaccine hesitancy on social media. BMJ 2023 May 05;381:1007. [doi: 10.1136/bmj.p1007] [Medline: 37146997]

Abbreviations

EG: expert group
OSF: Open Science Framework
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WHO: World Health Organization

Edited by A Mavragani; submitted 15.01.24; peer-reviewed by C Wardle, TS Boender, C Voegeli, SA Kumar; comments to author 03.05.24; revised version received 28.05.24; accepted 24.06.24; published 29.08.24.

Please cite as:
Germani F, Spitale G, Machiri SV, Ho CWL, Ballalai I, Biller-Andorno N, Reis AA
Ethical Considerations in Infodemic Management: Systematic Scoping Review
JMIR Infodemiology 2024;4:e56307
URL: https://infodemiology.jmir.org/2024/1/e56307
doi:10.2196/56307
PMID:

Review

The Use of Natural Language Processing Methods in Reddit to Investigate Opioid Use: Scoping Review

Alexandra Almeida[1,2,3*], PhD; Thomas Patton[3*], PhD; Mike Conway[4], PhD; Amarnath Gupta[5], PhD; Steffanie A Strathdee[3], PhD; Annick Bórquez[3], PhD

[1]Scientific Computing Program, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil
[2]San Diego State University, School of Social Work, San Diego, CA, United States
[3]Department of Medicine, University of California San Diego, San Diego, CA, United States
[4]School of Computing and Information Systems, The University of Melbourne, Melbourne, Australia
[5]San Diego Supercomputer Center, University of California San Diego, San Diego, CA, United States
[*]these authors contributed equally

Corresponding Author:
Alexandra Almeida, PhD
Scientific Computing Program
Oswaldo Cruz Foundation
Avenida Brasil, 4365, Antiga Residência Oficial Manguinhos
Rio de Janeiro, 21045900
Brazil
Phone: 55 213836 1212
Email: alexandra.almeida@fiocruz.br

Abstract

Background: The growing availability of big data spontaneously generated by social media platforms allows us to leverage natural language processing (NLP) methods as valuable tools to understand the opioid crisis.

Objective: We aimed to understand how NLP has been applied to Reddit (Reddit Inc) data to study opioid use.

Methods: We systematically searched for peer-reviewed studies and conference abstracts in PubMed, Scopus, PsycINFO, ACL Anthology, IEEE Xplore, and Association for Computing Machinery data repositories up to July 19, 2022. Inclusion criteria were studies investigating opioid use, using NLP techniques to analyze the textual corpora, and using Reddit as the social media data source. We were specifically interested in mapping studies' overarching goals and findings, methodologies and software used, and main limitations.

Results: In total, 30 studies were included, which were classified into 4 nonmutually exclusive *overarching goal* categories: methodological (n=6, 20% studies), infodemiology (n=22, 73% studies), infoveillance (n=7, 23% studies), and pharmacovigilance (n=3, 10% studies). NLP methods were used to identify content relevant to opioid use among vast quantities of textual data, to establish potential relationships between opioid use patterns or profiles and contextual factors or comorbidities, and to anticipate individuals' transitions between different opioid-related subreddits, likely revealing progression through opioid use stages. Most studies used an embedding technique (12/30, 40%), prediction or classification approach (12/30, 40%), topic modeling (9/30, 30%), and sentiment analysis (6/30, 20%). The most frequently used programming languages were Python (20/30, 67%) and R (2/30, 7%). Among the studies that reported limitations (20/30, 67%), the most cited was the uncertainty regarding whether redditors participating in these forums were representative of people who use opioids (8/20, 40%). The papers were very recent (28/30, 93%), from 2019 to 2022, with authors from a range of disciplines.

Conclusions: This scoping review identified a wide variety of NLP techniques and applications used to support surveillance and social media interventions addressing the opioid crisis. Despite the clear potential of these methods to enable the identification of opioid-relevant content in Reddit and its analysis, there are limits to the degree of interpretive meaning that they can provide. Moreover, we identified the need for standardized ethical guidelines to govern the use of Reddit data to safeguard the anonymity and privacy of people using these forums.

(JMIR Infodemiology 2024;4:e51156) doi:10.2196/51156

KEYWORDS

opioid; Reddit; natural language processing; NLP; machine learning

Introduction

Background

Opioid use disorder (OUD) is a chronic condition that affects more than 40 million people worldwide [1]. In 2019, the Global Burden of Disease study estimated that 128,000 deaths were attributed to drug use disorders [2], and the United States accounted for a large proportion of these deaths, with >70,000 deaths directly attributable to overdose [3]. US overdose deaths reached a record high in 2021, with 107,000 deaths, of which 75% were opioid related [4]. The magnitude of these numbers highlights the need for a strong surveillance infrastructure, including systems that support tracking trends in drug use and emerging patterns in the drug supply chain, as well as a good understanding of the experiences of, challenges of, and sources of support for people who use opioids, to inform a timely and effective response to this ongoing epidemic.

Social media platforms represent an important and accessible source of community support for people who use opioids, given they facilitate getting technical know-how and peer support. Nonetheless, analyzing such large textual data sets is challenging. Beyond the traditional qualitative approach to analyzing textual data, natural language processing (NLP) allows the analysis of textual data using computational methods and artificial intelligence. NLP lexicon-based methods and both supervised and unsupervised machine learning tools have been used to explore substance-related research questions on platforms such as Twitter (subsequently rebranded X; X Corp) [5-10], Reddit (Reddit Inc) [11-13], and web-based health communities [11,13].

In contrast with Reddit, which allows for long-form narratives, Twitter provides very condensed information for each data point since, historically, posts have been limited to 280 characters. An additional characteristic of tweets as a data type is that (at least in some cases) tweets contain metadata geolocational information, allowing for the ecological study of associations between the volume of mentions of a particular topic and a health outcome of interest in a given setting [14,15]. Facebook (Meta Platforms Inc) offers the advantage of providing information on social networks and the dissemination of information within these networks [16]. However, Facebook data are not necessarily public in the same way that Twitter and Reddit data are public (ie, on the open web). Facebook users typically form closed and semiprivate communities in which there is an expectation of data privacy, rendering the data unsuitable (both practically and ethically) for research purposes, as a specific consent model and process is absent. Reddit is a publicly accessible social media platform, with forums created and moderated to discuss specific themes. Participants (ie, redditors) use the platform anonymity and "throwaway" user accounts to share news, content, and thoughts through posting. Redditors' anonymity is key to authentic accounts of both positive and negative experiences with drugs (including in the context of treatment) and daily life situations that may impact the physical and mental health of people who use opioids, without fearing stigmatization. This platform is one of the most popular web-based social media platforms and has provided a space for exchange and discussion since 2005. It is mainly used by English speakers [17].

Despite Twitter being the platform of choice for most studies [18], Reddit thematic forums have shown to be an advantage for research on substance use, skipping the filtering stage of posts related to the research topic. NLP techniques enable the use of massive text corpora and the exploration of a multitude of research questions related to opioid use based on the perspectives of people sharing their experiences on social media platforms, such as Reddit.

Reddit ranked as the third most visited website in the United States [19], and in October 2020, it achieved >52 million daily active users [20]. Reddit has been used as a source of information to understand drug-related epidemics mainly due to the lower risk of social desirability bias, the real-time aspect of the data, and lower noise due to its thematic forums structure [21-23]. Given that the majority of its users are from the United States (47.82%), United Kingdom (7.6%), Canada (7.45%), and Australia (3.89%) and that these countries are all undergoing severe opioid-related epidemics, Reddit represents a valuable source of data to better understand the daily experiences of people who use opioids in these settings [24].

Objective

This scoping review aimed to understand how Reddit data have been leveraged to study the opioid epidemic using NLP to provide an overview of how these novel data sources and tools can serve the field of substance use research. More specifically, we were interested in mapping the studies' (1) overarching goals and findings, (2) methodologies and software used, and (3) main limitations. In addition, to help future research in this area to deal with abbreviations, slang, and common misspellings, we systematically collected the papers' list of synonyms and built a comprehensive synset with semantically equivalent terms within the opioid use field. This list could help the broader research community to rigorously investigate opioid use in Reddit using text processing algorithms.

Methods

Overview

This scoping review followed the Joanna Briggs Institute (JBI) methodology for scoping reviews based on the studies by Arksey and O'Malley [25] and Levac et al [26]. We used the JBI population (or participants)/concept/context framework to guide our research questions and ultimately inform our search strategy to understand how NLP methods have been applied to Reddit data to inform research on opioid use. Our study protocol was registered on the Open Science Framework [27] on June 29, 2022. Multimedia Appendix 1 details the JBI population (or participants)/concept/context framework rationale behind each research question of this scoping review.

Search Strategy

As the research topic encompasses studies targeting opioid use, we selected the following search engines: (1) PubMed to cover the biomedical literature; (2) Scopus to cover life, social, physical, and health sciences; (3) PsycINFO to cover the psychology field; and (4) ACL Anthology to cover NLP and computational linguistics. As the research team was aware of recent studies using NLP in the substance use field led mainly by authors from the engineering field, we also included IEEE Xplore and Association for Computing Machinery as data repositories for conference abstracts.

Our search was built based on team members' expertise in the topic and aimed to identify studies jointly covering three essential aspects: (1) opioid use, as the subject under analysis; (2) NLP techniques, as the methods used to analyze the textual corpora; and (3) Reddit, as the data source used in the study's analyses. The terms used to capture opioid-related studies were based on standard terms describing opioids, including "heroin" and "fentanyl" as well as the National Institute on Drug Abuse's list of most commonly prescribed opioids [28] and medications for OUD (MOUD) [29]. To ensure our search would only select studies that used a quantitative methodology to analyze the textual data (ie, NLP), we leveraged the "artificial intelligence" methods included in the Medical Subject Headings vocabulary thesaurus [30], which we supplemented with members' expertise in the topic to broaden our scope. As many essential papers in this field do not state "Reddit" as the social media environment under analysis in the abstract, and many search engines only consider the abstract (and not the full paper), we included the broad term "social media" in the search. Multimedia Appendix 2 details the search strategy, presenting the Medical Subject Headings hierarchy and the final search query used to obtain relevant studies.

We included papers that appeared in the search until July 19, 2022, with no cutoff for earlier studies. We selected original research papers published in English that stated that they collected textual data on opioid use from Reddit forums (even if other social media were also used) and specified they used NLP techniques to analyze these data.

Our focus was on studies that applied quantitative techniques to textual data. NLP can use a variety of different methods, including rule-based methods (eg, linguistic rules, deployment of lexicons, and development of bespoke regular expressions) and machine learning–based methods (eg, supervised classification and sequence labeling–based methods and unsupervised methods). Therefore, studies that used manual annotation were only included if they also applied quantitative methods to analyze the textual data.

Two researchers screened papers considering the title and abstract. If opioid use was not stated but the authors investigated substance use (eg, "...we evaluated attitudes of people who use substances..."), we included the papers for the second round of screening. If the methodology indicated any attempt to use NLP in any step of the study or if the authors were not specific about using only qualitative techniques (eg, "...we manually classified a set of posts..."), we included the paper. Finally, if the paper did not mention Reddit but was not specific about using another social media platform, we included it (eg, papers saying "...we evaluated Twitter data..." were excluded, whereas papers stating "...we used social media discussions to understand..." were included). The full-text review excluded the following: (1) review papers (to avoid double-counting); (2) papers solely using qualitative techniques or approaching substance use in general without any particular (or stratified) analysis for opioid-related information; and (3) papers not using Reddit as the data source. AA and TP independently screened for the title, abstract, and full paper. In both stages, when AA and TP diverged in their opinions about the papers' eligibility, AB (a third reviewer blinded to their classification) decided on eligibility as a tiebreaker. Each reviewer was responsible for the data extraction of critical outcomes of interest.

Data Extraction

We systematically extracted the following information: study's year of publication; first author's country of affiliation; authors' areas of expertise; social media platforms investigated (in addition to Reddit); subreddit investigated; period under analysis; types of posts analyzed; number of posts analyzed; number of Reddit users represented; whether the analysis was limited to opioids and whether it distinguished between different types of opioids to infer the scope of the population investigated; whether the study attempted to have a geographic focus; whether the study described their semantic analysis to identify opioid-related terms and whether they provided a synset; study's overarching goal, objectives, main findings, limitations, NLP methods and software used; and ethical approval information.

We followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) [31] guidelines to ensure methodological quality and clarity of the findings. The detailed PRISMA-ScR checklist describing the 20 essential and 2 optional reporting items is presented in Multimedia Appendix 3.

Semantic Analysis and "Opioid Synset" Development

To help further studies more efficiently address opioid-related research questions using social media data, we identified the different strategies used to carry out semantic analyses of opioid-related content, compiled their lists of opioid-related terms (produced either by automatic variant generator tools or by word embedding) or similar words, and curated their content, creating a refined version where the terms are aggregated by abbreviation, misspelling, variants and their misspellings, brand names and their misspellings, slang, analogues, and mixed substances.

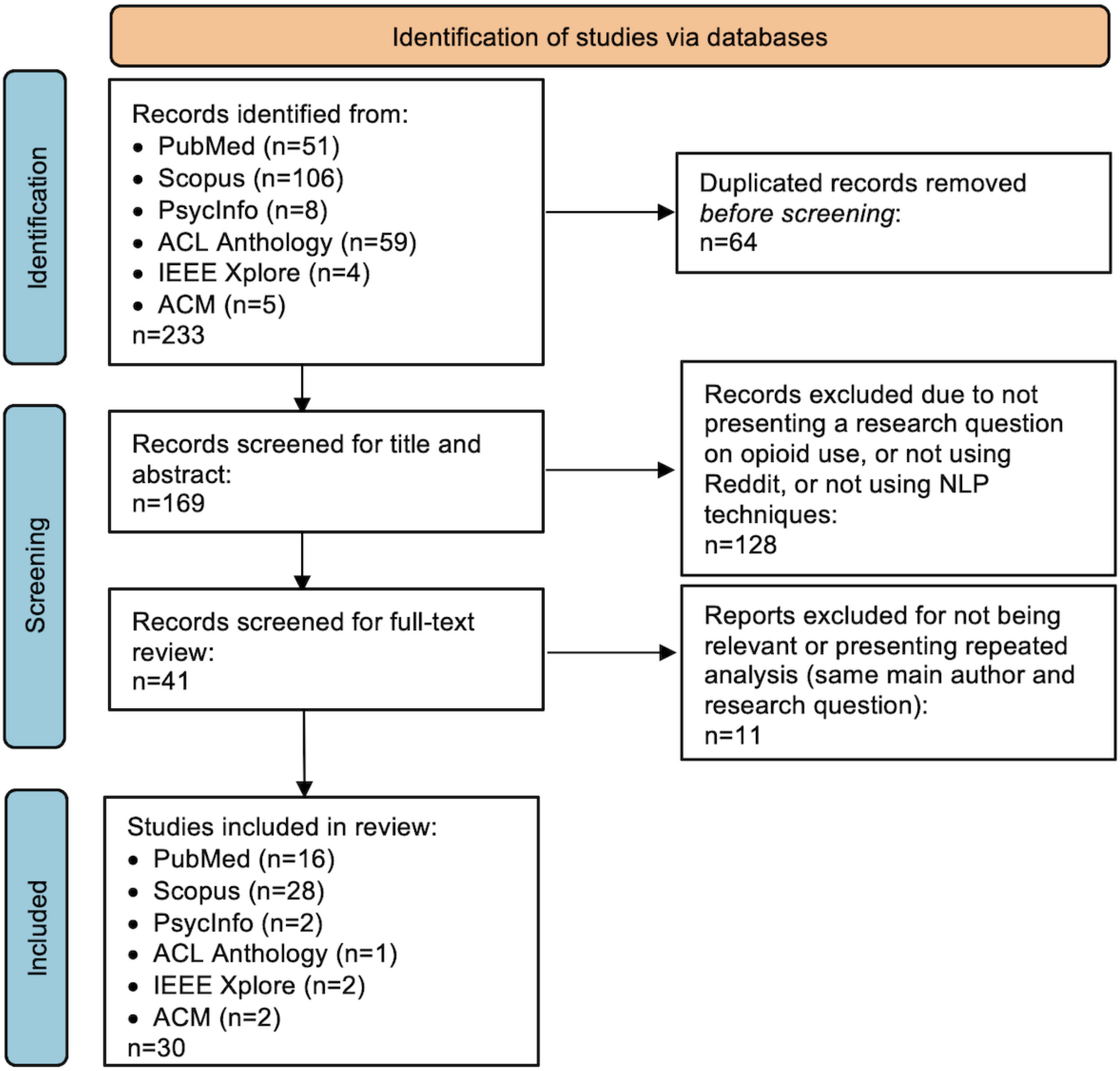
Results

Selection of Sources of Evidence

The literature search returned 233 papers, of which 64 (27.4%) were duplicates, 128 (54.9%) were considered ineligible and were thus removed after the title and abstract screening, and 11 (4.7%) were similarly removed after the full-text screening, resulting in a final list of 30 papers. Figure 1 displays the search process, following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [32]. A table with a summary of all study findings is provided in

Multimedia Appendix 4. Graphs illustrate the main results to give an overview of the literature studied.

Figure 1. Diagram of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)–oriented review process. ACM: Association for Computing Machinery; NLP: natural language processing. Note: Under "Studies included in Review," some of the papers were available on more than one database.



Studies' Characteristics

As detailed in Table 1, the number of studies applying NLP methods to data from Reddit forums to study opioid use has increased over time. Over half of the studies (19/30, 64%) were published in peer-reviewed journals, over a third (10/30, 33%) were published in conference proceedings, and 1 (3%) was published as a chapter in a book. Most studies (28/30, 93%) were disseminated in platforms with subject areas in computer science and health. Half of the studies (16/30, 53%) were exclusively concerned with opioids and OUD, while the remaining studies (14/30, 47%) were broader in scope (eg, substance use). However, content relating to other substances was sometimes referenced in the studies focused on opioids. For instance, the study by Preiss et al [33] identified a broader range of substances used for self-medication among people experiencing opioid withdrawal symptoms.

Table 1. Summary of characteristics of studies using natural language processing methods to investigate opioid use (N=30).

Study characteristics	Studies, n (%)
Year of publication	
2017	1 (3)
2018	1 (3)
2019	5 (17)
2020	6 (20)
2021	5 (17)
2022	12 (40)
Dissemination platform	
Peer-reviewed journal	19 (64)
Conference proceedings	10 (33)
Book chapter	1 (3)
Subject area of the dissemination platform	
Computer science	10 (33)
Health and medicine	8 (27)
Both computer science and health and medicine	10 (33)
Other	2 (7)
Range of substances considered	
Limited to opioids	16 (53)
Distinction between different opioid types	11 (37)
Software used for analysis[a]	
Not reported	6 (20)
Python	20 (66)
The Python Reddit API Wrapper	11 (37)
Gensim	4 (13)
spaCy	3 (10)
Linguistic Inquiry and Word Count	2 (7)
R	2 (7)
IBM Watson Natural Language Understanding	1 (3)
Reported sources of research funding[b]	
Financial support not reported	8 (27)
National Science Foundation	7 (23)
National Institute on Drug Abuse	5 (17)
Other (eg, institute-specific funding)	4 (13)
Other institutes within the National Institutes of Health	4 (13)
Centers for Disease Control and Prevention	3 (10)
No funding associated with the study	2 (7)
Natural Sciences and Engineering Research Council of Canada Discovery Grant	1 (3)

[a]Some studies used multiple software packages.

[b]Some studies had multiple research grants.

Over a third (11/30, 37%) of the reviewed studies explicitly referred to specific opioid types in the research methodology. The content under analysis varied as some studies focused on specific elements of posts, such as the title, the first submission, or the comments, with most of the papers that explained the unit of analysis using a combination of the first submission and comments (Multimedia Appendix 4). Python was the most common programming language used to analyze posting content, although alternative analysis methods and languages were also reported, including R, IBM Watson's Natural Language Understanding application, and Linguistic Inquiry and Word Count software.

In total, 8 (N=30, 27%) studies did not report any sources of funding and another 2 (7%) studies reported that there was no funding associated with the research. Among the remaining studies, the most common source of funding was the National Science Foundation (7/30, 23% of studies) followed by the National Institute on Drug Abuse (5/30, 17% of studies). Furthermore, 5 sponsored grants were found to have supported multiple studies in the review. Table 2 describes the authors from studies identified in the review, almost all (99/104, 95%) of whom were based in North America. The most common field of education among the authors, based on their terminal degree, was computer science (43/104, 41%). The studies analyzed more than 20 subreddits, most of which covered specific drugs and treatments, such as r/methadone and r/suboxone. However, the rooms with a broader scope and higher number of redditors, such as r/opiates and r/opiatesrecovery, were more frequently used as data sources, in 19 (63%) and 18 (60%) studies, respectively (Multimedia Appendix 4).

Table 2. Summary of researcher characteristics authoring studies using natural language processing methods to investigate opioid use (N=104).

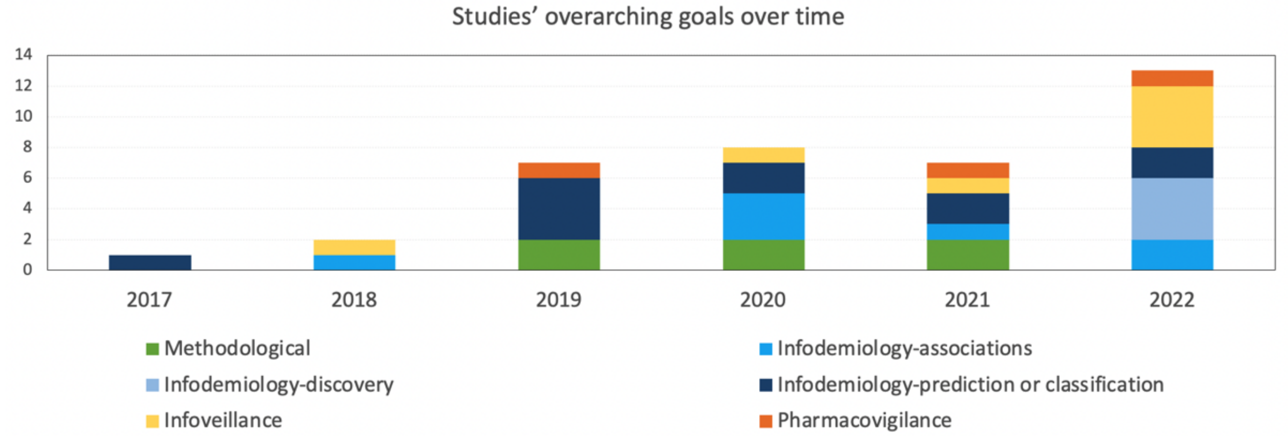
Researcher characteristics	Researchers, n (%)
Number of studies per researcher	
1	90 (86.5)
2	9 (8.7)
4	5 (4.8)
Location	
United States	92 (88.5)
Canada	7 (6.7)
Italy	4 (3.8)
Multiple locations (Australia and the United States)	1 (0.9)
Area of expertise (based on terminal degree)	
Computer science	43 (41.3)
Medicine	12 (11.5)
Health-related (eg, public health)	11 (10.6)
Unclear	8 (7.7)
Data science or analytics or machine learning	5 (4.8)
Multidisciplinary	5 (4.8)
Psychology	5 (4.8)
Engineering	5 (4.8)
Mathematics and statistics	3 (2.9)
Physics	2 (1.9)
Biology	1 (0.9)
Chemistry	1 (0.9)
Linguistics	1 (0.9)
Sociology	1 (0.9)
Social work	1 (0.9)

Overarching Goal, Methods Used, and Main Findings

To broadly characterize opioid-related research using Reddit data, we classified the studies according to their overarching goal. We divided them into four groups based on their contribution to the field: (1) methodological, (2) infodemiology, (3) infoveillance, and (4) pharmacovigilance. Methodological papers proposed novel NLP methods to analyze Reddit data or provide scientific tools. They apply these methods or tools to opioid use as an illustrative example but are not focused on this public health problem. Infodemiology, as defined by Eysenbach [34], is a science analogous to traditional epidemiology but with information coming from electronic media and with the ultimate goal of informing public health actions. Infoveillance papers contemplate the "longitudinal tracking of infodemiology metrics for surveillance and trend analysis" [35]. Pharmacovigilance covers the detection of new drugs and the surveillance of drugs and their adverse effects [36]. The papers' classification followed the procedures adopted in the study's screening process: AA and TP independently classified the papers according to their overarching goal, and any divergence was discussed and resolved by AB as a tiebreaker.

The number of studies published each year by overarching goals is presented in Figure 2, where we can see the increased volume and diversity of overarching goals over time. Studies could address more than one overarching goal; therefore, the cumulative number of studies represented is higher than the total number of studies included in the review.

Figure 2. Number of studies investigating opioid use per year using Reddit data and studies' overarching goals (each study could address more than 1 overarching goal).



Methodological Studies

In total, 6 (20%) studies were classified as methodological studies. Adams et al [37] compared Reddit and Twitter as sources of social media information relevant to opioid use and to improve keyword synonym lists for drug-term exploration using embedding techniques. They found that Reddit is better than Twitter as a source for discovering synonyms to be included in a keyword filter list for opioids. Akioyamen et al [38] provided ways for users to navigate Reddit posts to find the content of their interest using topic modeling and showed that text mining could play a fundamental role in supporting the identification of similar documents. Davis et al [39] proposed an archetype-based modeling and search method to identify messages from people who use opioids through learning from their usual vocabulary, which solicits documents from a user and finds similar documents posted by new authors based on the vocabulary. Jha and Singh [40] used a set of NLP techniques, including topic modeling, sentiment analysis, and propensity score matching to extract, treat, and analyze social media textual data on legal and illegal drugs to make it available to the academic community through a web application called SMARTS. Wright et al [19] developed new measures to detect changes in words' meaning and their association with overdose by applying word embedding and neural network techniques and found that the semantic proximity of fentanyl moved more closely to overdose and overdose-related terms over time. Zhu and Bhat [41] proposed a novel method to identify euphemisms or nicknames for drugs using single and multiwords embedding techniques. The new approach, known as euphemistic phrase detection, was reported to have higher detection accuracies compared to alternative methods.

Infodemiology Studies

Of the 22 studies classified as infodemiology, 7 (31%) explored associations: Andy [42] investigated correlations between different types of self-disclosures (positive and negative) and social support seeking (emotional and informational). The results showed that correlations between positive self-disclosure and emotional support seeking were moderate to strong, while those between negative self-disclosure and informational support seeking were low to moderate. Balsamo et al [43] estimated the odds ratio between opioid use and (1) routes of administration and (2) drug-tampering methods and routes of administration and drug-tampering procedures. Their findings suggest that substances were consumed in multiple, nonexclusive ways. Jha and Singh [44] explored associations between various latent constructs, addiction recovery, and relapse (defined as posted by users in drug addiction recovery [DAR] forums). Latent constructs capturing emotional distress, physical pain, self-development activities, and social relationships were all significantly associated with addiction recovery (ie, if a user posted in a DAR subreddit). Latent constructs capturing social activities and physical exercise were significantly associated with addiction relapse (ie, if a user posted in a recreational drug use subreddit after posting in a DAR subreddit).

Spadaro et al [45] explored associations between buprenorphine induction, fentanyl use, and precipitated opioid withdrawal. The paper found a relationship between increased mentions of fentanyl and its analogues with (1) increased mentions of precipitated opioid withdrawal and (2) increased mentions of the Bernese method (a microdosing strategy for buprenorphine induction). Pandrekar et al [46] evaluated the association between posts of social support (measured by the number of comments and the difference between upvotes and downvotes received) and the attributes of those posts, including the use of specific terms, the semantic categories observed, and posts' length. These analyses suggested that posts on personal issues like family, death, and home are positively associated with increased social support. Pandrekar et al [46] also explored the differences between posts from anonymous and nonanonymous users, with the former being more likely to discuss negative emotions like anxiety and sadness and containing words related to health and risk. Andy and Guntuku [47] explored the relationship between social support and the number of comments received using a mix of topic modeling and sentiment and correlation analysis. Their results showed that the average number of comments received was (1) positively associated with emotional support seeking and (2) negatively associated with information support seeking. Finally, Alambo et al [48] explored topical correlations between postings on substance use disorders and COVID-19 between May 2020 and Sep 2020. The results showed that these correlations fluctuated over time.

In total, 4 (18%) studies used discovery techniques to unveil patterns in the data. Ramachandran et al [49] used sentiment analysis and ANOVA to explore the public perceptions of the opioid epidemic. The results showed an overall negative sentiment with slight variation in emotional tones across different subreddits. Gauthier et al [50] tried to understand how Reddit communities support recovery using a topic-guided thematic analysis. The findings showed that community members often raised concerns about sensitive issues, such as withdrawal symptoms, body weight, legal troubles, and personal finances, and they encouraged others to seek help and navigate 12-step programs. Chen et al [51] used topic modeling to explore stigma-related experiences associated with substance use. This study found topics related to negative feelings, the challenges of coping with withdrawal symptoms, dealing with others during the recovery process, and dealing with chronic pain. Graves et al [52] used manual annotation and lexical similarity filter to yield information about firsthand experiences with buprenorphine-naloxone. They showed that the most frequently discussed topics included advice on the use of buprenorphine-naloxone as pharmacological treatment for OUD, information and guidance on dosage, information about tapering of dosage, its side effects and withdrawal information, and specific questions about use.

In total, 11 (50%) infodemiology studies targeted the prediction or classification of people along the OUD continuum. A common theme in these studies was the prediction of redditors' transitions between drug use, help seeking, and recovery stages. Furthermore, 2 (18%) studies developed models to predict whether redditors would post in a recovery subreddit based on their nonrecovery subreddit activities [53,54]. Another 2 (18%) studies developed models to predict relapse, albeit in different ways. Jha et al [55] predicted the likelihood of changes in the content of users' posts indicative of relapse. Yang et al [56] predicted substance use relapse in the following week using manually labeled data to identify instances where redditors self-reported experiencing a relapse. Relapse prediction was based on emotions detected in redditors' previous posting activities.

Among the remaining prediction or classification studies, 4 (36%) assessed the predictive performance of classification models for sorting Reddit content into groups. Davis et al [39] compared various classification models to predict whether redditors' posts indicated opioid use and found that the linear support vector machine model yielded the best performance. Yao et al [57] developed 2 sets of classifiers aiming to extract suicide risk posts from an opioid use context and extract opioid addiction posts from a suicidal ideation context. The former analysis yielded results showing (1) prediction accuracy and optimal model specification of classifiers varying with the percentage of suicide risk labels and (2) classifiers with better accuracy compared to those extracting opioid addiction posts from a suicidal ideation context. Chancellor et al [58] developed a classification model to determine whether a post was about OUD recovery. Another study [59] presented two classifiers that (1) sorted redditors into OUD and non-OUD groups according to their posting content and (2) determined whether those in the OUD group exhibited evidence of being in a positive recovery process. ElSherief et al [60] developed a classifier to identify posts related to a common myth that using MOUD is simply replacing one drug with another, which may discourage people from seeking this evidence-based treatment.

A study by Andy [42] assessed the predictive performance of models developed to measure the degree of positive and negative self-disclosure of post titles. The model outputs were compared to annotator ratings. Another study by Andy and Guntuku [47] compared the predictive performance of random forest models developed to measure the degree of social and informational support seeking expressed in post titles. They showed that the type of social support in post titles varies according to the substance use recovery forum.

Infoveillance Studies

In total, 7 (23%) papers were classified as an infoveillance study. Alambo et al [48] estimated the longitudinal topical correlation between substance use or mental health and COVID-19–related posts through the COVID-19 pandemic, finding that this correlation peaked after August 2020. Balsamo et al [43] traced the temporal evolution of posts from 2014 to 2018 relating to nonmedical consumption patterns, administration routes, and drug tampering. Their findings showed that mentions of heroin and hydrocodone decreased over time, mentions of buprenorphine and oxycodone remained relatively static, and mentions of fentanyl and codeine increased. El-Bassel et al [22] explored the COVID-19 period, from March to May 2020, and revealed the topics discussed in opioid-related subreddits and how they changed over time. There was a decrease in conversation on the topic referred to as "supply shutdown," gradual reductions in discussions on MOUD experiences and access issues, and increases in conversations on the negative consequences of OUD. Sarker et al [61] used the lexical variant generation model to examine stimulant co-mention trends among people who use opioids and people using MOUD and found that the number and proportion of redditors mentioning both increased steadily over time.

Pandrekar et al [46] analyzed the main psychological categories (ie, relativity, cognitive processes, social words, drives, affect words, biological processes, percept, and informal speech) of the posts from the r/opiates subreddit between 2014 and 2017 and showed that posts on personal issues tend to receive more social support. Sarker et al [23] compared pre– and post–COVID-19 monthly discussions on drug use, treatment access, care, and withdrawal. The results showed an increase in posts discussing withdrawal, treatment, and access to care from pre– to post–COVID-19 periods. Finally, a study by Sumner et al [62] used a time series regression with the least absolute shrinkage and selection operator approach to predict opioid overdose deaths (from CDC WONDER), using Twitter and Reddit post volume as predictor variables.

Pharmacovigilance Studies

In total, 3 (10%) studies were classified as a pharmacovigilance study. Chancellor et al [58] identified alternative drugs for treating OUD using word embeddings. The top 5 most frequently observed potential treatments were kratom, loperamide, Xanax, Valium, and Klonopin. Preiss et al [33] aimed to find medications to cope with opioid withdrawal symptoms based on observed correlations between text entities identified within posts relating to withdrawal symptoms and substances. Besides the ones already approved by the Food and Drug Administration or commonly used to treat symptoms, they found additional medicines considered potentially helpful (eg, gabapentin for body aches) and natural or home remedies (eg, ginger for nausea). Wright et al [19] evaluated the semantic proximity of individual substances to "overdose" through the years, with fentanyl being found to experience the most significant changes as it moved more closely to overdose and overdose-related terms.

Table 3 presents the selected studies, their overarching goal, the research question they address, and the NLP methods used. This can serve as a guide to help researchers studying substance use in formulating research questions that could be answered by applying NLP methods to Reddit data and to identify which NLP methods (and miscellaneous methods) to use in each case. To facilitate this process among researchers studying substance use with no expertise in NLP methods, we provide a brief description of these methods alongside examples of research questions they can answer in Multimedia Appendix 5. Full details on the data extracted for each study during the review process are provided in Multimedia Appendix 4.

Table 3. Overarching goals, research questions, and methods for studies applying natural language processing methods to Reddit data to investigate opioid use.

Study	Overarching goal	Research question	Methods
Adams et al [37]	Methodological	Drug term discovery	Word embedding
Akioyamen et al [38]	Methodological	Combine methods to allow users to actively navigate through topics or posts of interest	Topic modeling (LDA[a]) and sentiment analysis
Jha and Singh [40]	Methodological	Develop a tool to analyze data from Reddit and Twitter and make it available to the academic community for use	Topic modeling (LDA), sentiment analysis, and propensity score matching
Zhu and Bhat [41]	Methodological	Produce a list of euphemistic phrase candidates that are used as substitutes for target keywords corresponding to drug names	Phrase mining on a raw text corpus
Davis et al [39]	Methodological and infodemiology—prediction or classification	Predict archetypes	Archetype-based modeling and search
Wright et al [19]	Methodological and pharmacovigilance	Measure movements over time of the semantic proximity of substance-related terms to "overdose"	Diachronic word embedding
Jha and Singh [44]	Infodemiology—associations	Identify and quantify the relationship between emotional distress, physical pain, self-development, relationships, and geographic disparities versus drug addiction recovery and relapse	Semantic-based analysis and structural equation modeling
Spadaro et al [45]	Infodemiology—associations	Study potential associations between fentanyl, buprenorphine induction, and precipitated opioid withdrawal	Annotation and term-frequency matrix
Andy [42]	Infodemiology-associations and infodemiology—prediction or classification	Measure types of self-disclosures (ie, positive and negative) and social supports sought (ie, emotional and informational)	Random forest and correlation analysis
Andy and Guntuku [47]	Infodemiology-association and infodemiology—prediction or classification	Determine the relationship between the social supports expressed in the titles of posts and the number of comments they receive	Topic modeling (LDA) and sentiment analysis
Alambo et al [48]	Infodemiology—associations and infoveillance	Monitor trends of correlation between depression or substance use disorder and coronavirus posts	Word embedding, topic modeling (LDA), and correlation analysis
Balsamo et al [43]	Infodemiology—associations and infoveillance	Characterize patterns and estimate correlations between routes of administration and drug tampering	Word embedding and correlation analysis
Pandrekar et al [46]	Infodemiology—associations and infoveillance	Understand posts' psychological categories and examine the association between posts' attributes and the social support received	Topic modeling (LDA), semantic-based analysis, negative binomial, and Mann-Whitney U tests
Ramachandran et al [49]	Infodemiology—discovery	Unveil public opinion on the opioid epidemic	Sentiment analysis and ANOVA
Gauthier et al [50]	Infodemiology—discovery	Understand how web-based communities support recovery	Topic modeling (LDA) and thematic analysis
Chen et al [51]	Infodemiology—discovery	Examine the nature of stigma-related experience related to substance use and the salient affective and temporal factors in the use of 3 substances (including opioids)	Manual annotation and topic modeling (NMF[b])
Graves et al [52]	Infodemiology—discovery	Identify topics discussing firsthand experiences with buprenorphine-naloxone	Manual annotation and lexical similarity filter
ElSherief et al [60]	Infodemiology—prediction or classification	Identify misinformation related to medications for opioid use disorder	Bidirectional long short-term memory
Eshleman et al [53]	Infodemiology—prediction or classification	Predict users' propensity for seeking drug recovery interventions	Word embedding and prediction or classification: K-NN[c], K-NN, random forests, logistic regression, and naive Bayes.

Study	Overarching goal	Research question	Methods
Jha et al [55]	Infodemiology—prediction or classification	Characterize addiction stages of opioid use from users' social media posts	Word embedding and a combination of bidirectional long short-term-memory networks and conditional random fields
Lu et al [54]	Infodemiology—prediction or classification	Predict users' transitions from casual drug discussion forums to drug recovery forums	Word embedding, binary classifier, and Cox regression
Yang et al [59]	Infodemiology—prediction or classification	Predict relapse among people who use opioids	Topic modeling (LDA), correlation analysis, sentiment analysis, and support vector machine
Yang et al [56]	Infodemiology—prediction or classification	Predict opioid use disorder and recovery among people who use opioids	Sentiment analysis, generative adversarial networks, and correlation analysis
Yao et al [57]	Infodemiology—prediction or classification	Identify posts of suicidality among people who use opioids	Word embedding, convolutional neural network (also tested logistic regression, random forest, support vector machines, FastText, recurrent neural network, and attention-based bidirectional recurrent neural network)
Chancellor et al [58]	Infodemiology—prediction or classification and pharmacovigilance	Identify messages related to opioid use recovery and alternative treatments	Binary transfer learning classifier and word embeddings
El-Bassel et al [22]	Infoveillance	Identify challenges faced by people who use opioids and how these change over time	Word embedding and topic modeling (LDA)
Sarker et al [23]	Infoveillance	Identify prescription or illegal opioid use, describe opioid treatment access and care, and withdrawal	Annotation and term-frequency matrix
Sarker et al [61]	Infoveillance	Examine stimulant comention trends among people who use opioids or receive medications for opioid use disorder.	Generate lexical variants and negatives (LexExp[d] and NegEx[e])
Sumner et al [62]	Infoveillance	Build a statistical model for estimating national opioid overdose deaths using multiple predictors, including data on the volume of Reddit posts mentioning heroin and synthetic opioids.	Time-series analysis (LASSO[f])
Preiss et al [33]	Pharmacovigilance	Identify symptoms and remedies for opioid withdrawal	Word embedding and named entity recognition

[a]LDA: latent Dirichlet allocation.

[b]NMF: nonnegative matrix factorization.

[c]K-NN: K-nearest neighbors.

[c]LexExp: unsupervised lexicon expansion system.

[d]NegEx: negation detection algorithm.

[e]LASSO: least absolute shrinkage and selection operator.

Scope of Substance Use Populations Considered and Distinction of Opioid Types

The anonymity characteristic of Reddit data prevents any definitive description of the study participants, composed of individuals contributing and responding to the posts selected for analysis. Nonetheless, participation in subreddits implies an active personal engagement with the topic, and therefore most studies in this review presuppose that the data from posts represent the language, views, and experiences of people who use opioids. For instance, the study by Andy and Guntuku [47] et al analyzed data to investigate the types of support sought by people recovering from OUD, which was assumed based on their activity in the r/OpiatesRecovery and r/suboxone subreddits. Furthermore, 4 (13%) other studies also made inferences about the drug use and recovery status of individual redditors based on the subreddits they were found to post in [53-55]. In total, 5 (17%) studies involved the manual inspection and labeling of a sample of posts to verify whether they were indicative of ongoing opioid use, recovery efforts, or a recent relapse [55-59]. Similarly, a study by Graves et al [52] involved a manual thematic categorization of 200 posts by 3 separate researchers to determine whether the posts' contents were indicative of personal experiences of buprenorphine-naloxone use. Another study by El-Bassel et al [22] was more cautious in acknowledging that their data might not reflect the firsthand experiences of people using drugs, despite this being the purpose of their study. In all 5 studies, the posts were labeled by health or substance use researchers except for 1 study, relying on nonspecialist workers from Amazon Mechanical Turk [57].

Limitations Mentioned by the Authors

In total, 20 (67%) studies reported limitations associated with the methods used or the results produced. The most commonly cited limitation was the uncertainty regarding the representativeness of the community of people discussing opioid use on Reddit to the broader population of people who use opioids [23,43,45,52,54,58,61,62]. A related set of limitations was the lack of information on the demographics and location of the study participants [22,46,50] and the inability to ascertain whether study participants had a clinical diagnosis of OUD [22,43,52,54,61]. Another limitation was the presumption that subreddits dedicated to opioid use and recovery would consistently provide relevant content. A manual review undertaken in the study by Yao et al [57] showed that this was not true in many instances. Another study acknowledged that relying on thematic subreddits addressing substance use disorders disregards relevant content in nondrug use subreddits [54].

Semantic Analysis and "Opioid Synset" Development

As Reddit forums are internet-based spaces used by people looking for information, advice, or simply to share experiences, the textual variation by colloquial terms (slang), misspellings, and abbreviations is frequent. This heterogeneity imposes a challenge for NLP methods.

More recent studies have used NLP tools for automatic variant generation [23,45,61]. Some studies have used existing lists of slang terms derived from the Drug Enforcement Administration Drug Slang Code Words [63], the Drug Abuse Ontology [64], and the Urban Dictionary [37,43,48,65]. However, many of the papers have explored the potentialities of embedding words to deal with semantic variation. While some studies incorporated word embedding methods into the analysis [22,48,53-55,57], others used this methodology with the ultimate (or intermediate) goal of expanding opioid-related vocabulary [19,33,37,43,58]. In this spirit, Zhu and Bhat [41] proposed a multiword method to identify nonsingle word euphemisms for the drug use field.

Despite the accessibility of word embeddings for document identification, it is important to emphasize that the outputs require expert curation to ensure precision in the meaning of the words used. This is because word embedding outcomes do not provide synonyms but words with similar representations. For example, in the study of Chancellor et al [58], word embedding outcomes for fentanyl included morphine, heroin, or even ketamine. We systematically categorized opioid-related terms used and generated across selected studies to produce a comprehensive "opioid synset" and support research in this area (Multimedia Appendix 6).

Ethical Considerations

Ethics approval was reported for 2 (7%) of the studies [50,52], while a further 4 (13%) studies were granted exemptions from ethics reviews [23,47,61,62]. The studies' dissemination vehicle did not seem to determine whether they included ethical considerations. About 45% (5/11) of the conference proceedings and books and 63% (12/19) of the peer-reviewed papers did not report obtaining ethical clearance or any type of action to protect participants. From the remaining 55% (6/11) of the conference proceedings and books, 2 (33%) submitted their projects to the Institutional review board (IRB) and had them approved or exempted, 2 (33%) used strategies to protect participants, such as paraphrasing, and 2 (33%) argued the studies did not qualify for the ethics board review. Of the 37% (7/19) of the peer-reviewed papers that reported some ethical considerations to protect participants' privacy, 4 (57%) submitted their projects to the IRB and had them approved or exempted, and 3 (43%) argued the studies did not qualify for ethics board review. The papers' ethical considerations are available in Multimedia Appendix 4.

Discussion

Contributions and Potential of This Research

This paper describes the small but growing body of research applying NLP methods to analyze content from Reddit relating to opioid use. We show that the existing literature is diverse in scope, as indicated by the combination of infodemiology, infoveillance, pharmacovigilance, and methodological studies' overarching goals. The literature is also varied in terms of NLP methods used, which include word expansion techniques, topic models, and prediction or classification methods. We found that most studies have been led by researchers with training in computer science, which was also the leading subject area associated with the platforms used for their dissemination or publication. This trend might explain the high degree of emphasis placed on the conduct of methods-driven research in the literature that demonstrates the application of NLP methods in the context of OUD, as opposed to research being primarily motivated by a theory or hypotheses specific to OUD. Indeed, while some studies showcase NLP applications to Reddit data that could directly be used toward routine public health surveillance or interventions in the context of OUD, many studies presented NLP methodological development and used OUD as an opportune case study without considering their practical use or operationalization.

Given our focus on leveraging big data and NLP methods to improve public health responses to the opioid crisis, we provide an overview of the key contributions that have been and could be achieved through further engagement with this research. Reddit has been increasingly used as a source of information to understand opioid use behaviors and their connections with other conditions. Significant problems inherent to the substance use field, including the identification and classification of participants' membership along the OUD continuum and the longitudinal tracking of OUD-related metrics, have been robustly explored in studies, such as Jha et al [55] and Alambo et al [48].

Compared to traditional qualitative studies that rely on recruiting and interviewing participants, studies using Reddit data provide temporal flexibility, enabling the retrospective investigation of particular periods of interest, bypassing recall bias, as well as longitudinal perspectives (which are rare in the context of qualitative studies actively engaging participants). The latter can include following redditor cohorts (ie, following the same individuals over time) [53] or specific topics (including a mix of old and new redditors participating in that specific subreddit at each time point) [23]. Indeed, the studies using Reddit data use Reddit data opportunistically, while traditional qualitative studies that actively engage participants are designed to answer specific research questions. Compared to studies using traditional qualitative methods to analyze Reddit data, we found that those studies using NLP methods identified similar topics. We identified 3 studies using qualitative thematic analysis on posts of the subreddits r/Opiates and r/OpiatesRecovery, during the peak months of the COVID-19 pandemic, March to May 2020 [66-68], which could be compared with the NLP-based study included in this scoping review using data from March to May 2020, from the subreddits r/opiates, r/OpiatesRecovery, r/suboxone, and r/Methadone [22].

The qualitative study conducted by Krawczyk et al [68], which explored the impact of changes imposed by the pandemic on OUD treatment access, found that people who use opioids were concerned about OUD treatment facilities' closure, the transition to telehealth, inconsistency between methadone daily clinic requirements and exposure to COVID-19 risks, impressions on regulation changes for MOUD, and how the pandemic was impacting treatment motivation and progress. Bunting et al [67] also qualitatively explored COVID-19 restrictions' consequences on the daily lives of people who use opioids, such as social isolation and the consequent change in the social network, and how Reddit was used to ask for or offer advice. Finally, Arshonsky et al [66] described psychological and behavioral coping strategies for withdrawal symptoms and cessation or reduction of opioid use.

Comparatively, the paper by El-Bassel et al [22] on NLP provided an overview of the main topics representing different concerns around OUD in the context of COVID-19 and its trend and changes as the pandemic evolved. It found similar access issues during the COVID-19 months, such as closed and overcrowded treatment centers, limited take-home doses, switching from methadone to buprenorphine treatment to avoid daily clinic visits, or switching from oral to injectable buprenorphine to avoid access barriers, financial barriers to access MOUD, helpfulness of telehealth, withdrawal issues originated by diminished drug supply and MOUD facilities' closures, confusion if their symptoms were because of withdrawal or COVID-19 infection, stigmatization by health care providers, self-medication to keep in the recovery track, and home detox. Sarker et al [23] also found similar topics during the COVID-19 pandemic to those reported by El-Bassel et al [22] and by the 3 qualitative studies mentioned [66-68], with the bonus of following the relative frequency of themes associated with drug use, treatment and access, and withdrawal over time (ie, infoveillance). This suggests NLP methods can effectively identify key themes while saving time and covering larger textual corpora. A study purposefully designed to compare NLP versus traditional qualitative methods is needed to rigorously investigate the advantages and drawbacks of NLP methods.

Given the complexity of the research questions, in general, researchers used a mix of NLP methods to address opioid-related problems. Traditional NLP methods, such as topic modeling, were used more frequently to discover aspects of opioid use behavior and follow them over time (infodemiology and infoveillance papers), whereas word embedding was often used to discover new drugs (pharmacovigilance) and in tandem with classification and prediction methods to detect people's stage in the OUD continuum (infodemiology). In contrast, sentiment analysis was mostly used to infer redditor's opinions and positioning about opioid use. Another important aspect of the studies analyzed in this scoping review is its implications for public health responses. While none of the studies described the implementation of Reddit-based interventions, several studies have the potential to be used toward such purposes. For example, the study by ElSherief et al [60] aimed to identify misinformation related to MOUD. It is natural to foresee how such a tool could be used on a routine basis to detect existing and new myths, and both respond to those posts providing evidence-based information as well as to compile these myths to guide public health messaging by providers and relevant agencies. Studies presenting methods to identify users at high risk of relapse [55] or suicide [57] could also be used to reach out to these individuals with helpful resources, including crisis lines and other types of support, as has been done in other studies [69-71]. Similarly, those identified as having a high probability of transitioning to an opioid recovery subreddit could benefit from information and resources to engage in treatment, counseling, support groups, or activities to facilitate this process [53].

Methodological Remarks and Recommendations

The NLP methods used by the selected studies were diverse. We provided a map to identify appropriate methods depending on the research problem addressed. In addition, we have endeavored to bring clarity regarding the different types of data used in the studies. The commonly used term *post*, used across many studies is not specific enough as it does not fully characterize the data considered (first submission, comment, or whether it includes the title). Similarly, it is not always clear how many posts were included in the analyses (eg, all those available over a period versus a sample of the posts submitted) and how many different users these posts represented. Importantly, details about the subreddits considered and the period covered are not systematically specified. Best practices for Reddit or other social media platforms' data description should include these key characteristics to improve transparency about the representativity of the data and to enable its reproducibility.

We also found important variation in terms of the use of "dictionaries" to enable the identification of opioid use–related content, which can have an impact on the results through modifying the body of data examined. We, therefore, offer an important resource to increase the quality of new research on opioid use undertaken in this space through the synset we compiled and shared in Multimedia Appendix 6.

Regarding the use of NLP on Reddit data, some considerations should be highlighted. Approaching Reddit data from an exploratory perspective, at least initially, is useful to ascertain the extent to which Reddit data are suitable for your particular research question. Careful selection of appropriate subreddits is important given that each subreddit has a distinct culture and associated norms. For example, the r/trees subreddit is devoted to the discussion of cannabis use, whereas the r/marijuanaenthusiasts subreddit is devoted to the discussion of arboreal matters. Reddit-based research—particularly in health-related areas—should be assessed by research ethics committees or IRBs and follow research ethics protocols related to privacy protection. For example, Benton et al [72] recommend adopting policies related to privacy protection (eg, refraining from using direct quotations in presentations and publications) and—where appropriate—deidentifying data (eg, using anonymous numerical codes in the place of social media usernames).

Recent changes in Reddit's application programming interface access have the potential to affect future research that relies on real-time, high-throughput data access. However, historical data (before April 2023) remains publicly available via PushShift and the free Reddit application programming interface (accessed via PRAW in Python) can be used to collect recent posts containing specific keywords. Furthermore, it is possible to purchase data from Reddit, with the study by Poudel and Weninger [73] estimating a potential cost of US $240 per 1 million posts. Finally, as of late May 2024, Reddit is planning on implementing a new service aimed at providing academic researchers with affordable data [74].

Given the rapid changes in NLP technology, it is difficult to predict the future directions of application of NLP methods to address substance use disorders–related research questions, particularly concerning the future performance of large language models such as Gemini and ChatGPT (OpenAI LP). However, 2 broad themes are emerging. First, a growing democratization of NLP methods, given that generative NLP methods allow for the creation of baseline NLP tools by public health experts with relatively little computer science expertise (for example, by using prompt engineering approaches). Second, potentially increasing costs associated with accessing data for research given the industry-wide drive to monetize social media data.

Limitations

Our findings highlight 2 fundamental paradoxes associated with NLP-based research on opioids using content from Reddit. The first paradox relates to the anonymity afforded to Reddit users to disclose their own experiences and viewpoints related to opioid use. While this anonymity is advantageous as far as it supports the availability of evidence offering unique insights compared to other sources (eg, routine household surveys), it also represents a limitation because it implies a lack of demographic and clinical information on study participants. This may cast doubts upon the authenticity of the findings as there is no way of verifying whether the content reflects firsthand experiences or whether study participants meet the clinical criteria for a diagnosis of OUD. The second paradox relates to the methodological capabilities of NLP techniques to draw inferences from user-generated textual data. These computational methods have facilitated the large-scale computation of specific tasks, such as content identification and detection of textual patterns involving huge quantities of data. Despite these advances, there appear to be limits on the types of inferences that NLP methods can make, particularly when it comes to analyses that go beyond content identification (eg, posts related to opioid use) and attempt to interpret meaning. This observation corresponds with findings elsewhere, showing that NLP methods are valuable adjuncts to traditional qualitative research methods but do not yet represent an adequate alternative to human analyses [75].

Despite the extensive use of bespoke lexica to perform analysis of social media text, these lexicon-based approaches remain error-prone due to ambiguities characteristic of natural language. For example, different drugs may be referred to by the same slang term in different contexts (eg, "junk" can be used to refer to either heroin or cocaine).

A methodological limitation of this scoping review is the inclusion of papers exclusively written in English. Although more than half of Reddit users are from countries with the ongoing opioid epidemic and that speak predominantly English, such as the United States, United Kingdom, Canada, and Australia [24], this imposes a potential constrain on the understanding of how researchers are using NLP methods in Reddit to investigate opioid use.

Conclusions

This comprehensive review explores the expanding application of NLP methodologies to analyze text data sourced from Reddit, with a particular focus on opioid-related subreddits. A wide variety of NLP techniques and applications can be observed in the literature that demonstrate the potential for NLP use to support surveillance and social media interventions addressing the opioid crisis. Although we found that these methods offer useful insights into the behavior and attitudes of people who use opioids, there are limits to the utility of these automated approaches, with current methods best thought of as supplementary to current, established epidemiological methods.

Acknowledgments

This study was funded by the National Institute on Drug Abuse DP2 DA049295 grant (principal investigator AB). TP acknowledges funding from the T32 DA023356 grant, and MC acknowledges funding from the National Institute on Drug Abuse R21DA05668 grant.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Joanna Briggs Institute population (or participants)/concept/context (PCC) framework rationale behind each research question.
[DOCX File , 14 KB - infodemiology_v4i1e51156_app1.docx]

Multimedia Appendix 2
Search terms strategy.
[DOCX File , 200 KB - infodemiology_v4i1e51156_app2.docx]

Multimedia Appendix 3
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.
[DOCX File , 18 KB - infodemiology_v4i1e51156_app3.docx]

Multimedia Appendix 4
Data extraction table.
[XLS File (Microsoft Excel File), 144 KB - infodemiology_v4i1e51156_app4.xls]

Multimedia Appendix 5
Brief definitions of the natural language processing methods most used by the scoping review papers, according to Wikipedia (accessed on January 20, 2023).
[DOCX File , 19 KB - infodemiology_v4i1e51156_app5.docx]

Multimedia Appendix 6
Synset.
[XLS File (Microsoft Excel File), 42 KB - infodemiology_v4i1e51156_app6.xls]

References

1. Degenhardt L, Grebely J, Stone J, Hickman M, Vickerman P, Marshall BD, et al. Global patterns of opioid use and dependence: harms to populations, interventions, and future action. Lancet 2019 Oct 26;394(10208):1560-1579 [FREE Full text] [doi: 10.1016/S0140-6736(19)32229-9] [Medline: 31657732]
2. UNODC world drug report 2019. UN Office on Drugs and Crime. URL: https://tinyurl.com/zbxdwx7k [accessed 2024-04-29]
3. Drug overdose death rates. National Institute on Drug Abuse. 2023. URL: https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates [accessed 2024-04-29]
4. U.S. overdose deaths in 2021 increased half as much as in 2020 – but are still up 15%. National Center for Health Statistics. 2022. URL: https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/202205.htm#:~:text=Provisional%20data%20from%20CDC's%20National,93%2C655%20deaths%20estimated%20in%202020 [accessed 2024-04-18]
5. Allem JP, Dharmapuri L, Unger JB, Cruz TB. Characterizing JUUL-related posts on Twitter. Drug Alcohol Depend 2018 Sep 01;190:1-5 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.05.018] [Medline: 29958115]
6. Curtis B, Giorgi S, Buffone AE, Ungar LH, Ashford RD, Hemmons J, et al. Can Twitter be used to predict county excessive alcohol consumption rates? PLoS One 2018;13(4):e0194290 [FREE Full text] [doi: 10.1371/journal.pone.0194290] [Medline: 29617408]
7. Lazard AJ, Saffer AJ, Wilcox GB, Chung AD, Mackert MS, Bernhardt JM. E-cigarette social media messages: a text mining analysis of marketing and consumer conversations on Twitter. JMIR Public Health Surveill 2016 Dec 12;2(2):e171 [FREE Full text] [doi: 10.2196/publichealth.6551] [Medline: 27956376]
8. Mackey T, Kalyanam J, Klugman J, Kuzmenko E, Gupta R. Solution to detect, classify, and report illicit online marketing and sales of controlled substances via Twitter: using machine learning and web forensics to combat digital opioid access. J Med Internet Res 2018 Apr 27;20(4):e10029 [FREE Full text] [doi: 10.2196/10029] [Medline: 29613851]
9. Mackey TK, Kalyanam J, Katsuki T, Lanckriet G. Twitter-based detection of illegal online sale of prescription opioid. Am J Public Health 2017 Dec;107(12):1910-1915. [doi: 10.2105/AJPH.2017.303994] [Medline: 29048960]
10. Simpson SS, Adams N, Brugman CM, Conners TJ. Detecting novel and emerging drug terms using natural language processing: a social media corpus study. JMIR Public Health Surveill 2018 Jan 08;4(1):e2 [FREE Full text] [doi: 10.2196/publichealth.7726] [Medline: 29311050]
11. Chen AT, Zhu SH, Conway M. What online communities can tell us about electronic cigarettes and hookah use: a study using text mining and visualization techniques. J Med Internet Res 2015 Sep 29;17(9):e220 [FREE Full text] [doi: 10.2196/jmir.4517] [Medline: 26420469]
12. Meacham MC, Paul MJ, Ramo DE. Understanding emerging forms of cannabis use through an online cannabis community: an analysis of relative post volume and subjective highness ratings. Drug Alcohol Depend 2018 Jul 01;188:364-369 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.03.041] [Medline: 29883950]

13. Zhan Y, Liu R, Li Q, Leischow SJ, Zeng DD. Identifying topics for e-cigarette user-generated contents: a case study from multiple social media platforms. J Med Internet Res 2017 Jan 20;19(1):e24 [FREE Full text] [doi: 10.2196/jmir.5780] [Medline: 28108428]
14. Cuomo R, Purushothaman V, Calac AJ, McMann T, Li Z, Mackey T. Estimating county-level overdose rates using opioid-related twitter data: interdisciplinary infodemiology study. JMIR Form Res 2023 Jan 25;7:e42162 [FREE Full text] [doi: 10.2196/42162] [Medline: 36548118]
15. 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]
16. Kilgo DK, Midberry J. Social media news production, emotional Facebook reactions, and the politicization of drug addiction. Health Commun 2022 Mar;37(3):375-383 [FREE Full text] [doi: 10.1080/10410236.2020.1846265] [Medline: 33213217]
17. Potter M. Bad actors never sleep: content manipulation on Reddit. Continuum 2021 Sep 28;35(5):706-718 [FREE Full text] [doi: 10.1080/10304312.2021.1983254]
18. Conway M, Hu M, Chapman WW. Recent advances in using natural language processing to address public health research questions using social media and consumergenerated data. Yearb Med Inform 2019 Aug;28(1):208-217 [FREE Full text] [doi: 10.1055/s-0039-1677918] [Medline: 31419834]
19. Wright AP, Jones CM, Chau DH, Matthew Gladden R, Sumner SA. Detection of emerging drugs involved in overdose via diachronic word embeddings of substances discussed on social media. J Biomed Inform 2021 Jul;119:103824 [FREE Full text] [doi: 10.1016/j.jbi.2021.103824] [Medline: 34048933]
20. Patel S. Reddit claims 52 million daily users, revealing a key figure for social-media platforms. The Wall Street journal. 2020. URL: https://www.wsj.com/articles/reddit-claims-52-million-daily-users-revealing-a-key-figure-for-social-media-platforms-11606822200 [accessed 2024-04-29]
21. Tang LU, Bie B, Park SE, Zhi D. Social media and outbreaks of emerging infectious diseases: a systematic review of literature. Am J Infect Control 2018 Sep;46(9):962-972 [FREE Full text] [doi: 10.1016/j.ajic.2018.02.010] [Medline: 29628293]
22. El-Bassel N, Hochstatter KR, Slavin MN, Yang C, Zhang Y, Muresan S. Harnessing the power of social media to understand the impact of COVID-19 on people who use drugs during lockdown and social distancing. J Addict Med 2022;16(2):e123-e132 [FREE Full text] [doi: 10.1097/ADM.0000000000000883] [Medline: 34145186]
23. 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]
24. Dean B. Reddit user and growth stats. Backlinko. 2021. URL: https://backlinko.com/reddit-users [accessed 2024-04-29]
25. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
26. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
27. The use of artificial intelligence methods in Reddit to investigate opioid use: a scoping review. Open Science Framework. URL: https://osf.io/ftqj3 [accessed 2024-04-29]
28. Prescription opioids DrugFacts. National Institute on Drug Abuse. 2021. URL: https://nida.nih.gov/publications/drugfacts/prescription-opioids [accessed 2024-04-29]
29. Effective treatments for opioid addiction. National Institute on Drug Abuse. 2016. URL: https://tinyurl.com/mry9kknr [accessed 2024-07-20]
30. MeSH: artificial intelligence. National Library of Medicine. 2022. URL: https://www.ncbi.nlm.nih.gov/mesh/68001185 [accessed 2024-04-29]
31. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
32. 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]
33. Preiss A, Baumgartner P, Edlund MJ, Bobashev GV. Using named entity recognition to identify substances used in the self-medication of opioid withdrawal: natural language processing study of reddit data. JMIR Form Res 2022 Mar 30;6(3):e33919 [FREE Full text] [doi: 10.2196/33919] [Medline: 35353047]
34. 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 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
35. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158 [FREE Full text] [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]

36. Regulation and prequalification - what is pharmacovigilance? World Health Organization. URL: https://www.who.int/teams/regulation-prequalification/regulation-and-safety/pharmacovigilance [accessed 2024-04-29]
37. Adams N, Artigiani EE, Wish ED. Choosing your platform for social media drug research and improving your keyword filter list. Journal of Drug Issues 2019 Mar 13;49(3):477-492 [FREE Full text] [doi: 10.1177/0022042619833911]
38. Akioyamen P, Nicklas LC, Sanchez-Arias R. A framework for intelligent navigation using latent Dirichlet allocation on reddit posts about opiates. In: Proceedings of the 2020 4th International Conference on Compute and Data Analysis. 2020 Presented at: ICCDA '20; March 9-12, 2020; Silicon Valley, CA p. 190-196 URL: https://doi.org/10.1145/3388142.3388156 [doi: 10.1145/3388142.3388156]
39. Davis BD, Sedig K, Lizotte DJ. Archetype-based modeling and search of social media. Big Data Cogn Comput 2019 Jul 24;3(3):44 [FREE Full text] [doi: 10.3390/bdcc3030044]
40. Jha D, Singh R. SMARTS: the social media-based addiction recovery and intervention targeting server. Bioinformatics 2019 Oct 24:1970 [FREE Full text] [doi: 10.1093/bioinformatics/btz800] [Medline: 31647520]
41. Zhu W, Bhat S. Euphemistic phrase detection by masked language model. In: Proceedings of the 2021 Association for Computational Linguistics. 2021 Presented at: EMNLP '21; November 7-11, 2021; Punta Cana, Dominican Republic p. 163-168 URL: https://aclanthology.org/2021.findings-emnlp.16.pdf [doi: 10.18653/v1/2021.findings-emnlp.16]
42. Andy A. Self-disclosure in opioid use recovery forums. In: Shaban-Nejad A, Michalowski M, Bianco S, editors. AI for Disease Surveillance and Pandemic Intelligence: Intelligent Disease Detection in Action. Cham, Switzerland: Springer; 2022:65-74.
43. 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]
44. Jha D, Singh R. Analysis of associations between emotions and activities of drug users and their addiction recovery tendencies from social media posts using structural equation modeling. BMC Bioinformatics 2020 Dec 30;21(Suppl 18):554 [FREE Full text] [doi: 10.1186/s12859-020-03893-9] [Medline: 33375934]
45. Spadaro A, Sarker A, Hogg-Bremer W, Love JS, O'Donnell N, Nelson LS, et al. Reddit discussions about buprenorphine associated precipitated withdrawal in the era of fentanyl. Clin Toxicol (Phila) 2022 Jun;60(6):694-701 [FREE Full text] [doi: 10.1080/15563650.2022.2032730] [Medline: 35119337]
46. Pandrekar S, Chen X, Gopalkrishna G, Srivastava A, Saltz M, Saltz J, et al. Social media based analysis of opioid epidemic using Reddit. AMIA Annu Symp Proc 2018;2018:867-876 [FREE Full text] [Medline: 30815129]
47. Andy A, Guntuku SC. Does social support (expressed in post titles) elicit comments in online substance use recovery forums? In: Proceedings of the 4th Workshop on Natural Language Processing and Computational Social Science. 2020 Presented at: NLPCSS '20; November 20, 2020; Virtual Event p. 35-40 URL: https://doi.org/10.18653/v1/2020.nlpcss-1.4 [doi: 10.18653/v1/2020.nlpcss-1.4]
48. Alambo A, Padhee S, Banerjee T, Thirunarayan K. COVID-19 and mental health/substance use disorders on reddit: a longitudinal study. In: Proceedings of the 2021 Conference on Pattern Recognition, ICPR International Workshops and Challenges. 2021 Presented at: ICPR '21; January 10-15, 2021; Virtual Event p. 20-27 URL: https://doi.org/10.1007/978-3-030-68790-8_2 [doi: 10.1007/978-3-030-68790-8_2]
49. Ramachandran S, Brown L, Ring D. Tones and themes in Reddits posts discussing the opioid epidemic. J Addict Dis 2022;40(4):552-558 [FREE Full text] [doi: 10.1080/10550887.2022.2049170] [Medline: 35274598]
50. Gauthier RP, Costello MJ, Wallace JR. “I will not drink with you today”: a topic-guided thematic analysis of addiction recovery on Reddit. 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/10.1145/3491102.3502076 [doi: 10.1145/3491102.3502076]
51. Chen AT, Johnny S, Conway M. Examining stigma relating to substance use and contextual factors in social media discussions. Drug Alcohol Depend Rep 2022 Jun;3:100061 [FREE Full text] [doi: 10.1016/j.dadr.2022.100061] [Medline: 36845987]
52. Graves RL, Perrone J, Al-Garadi MA, Yang YC, Love J, O'Connor K, et al. Thematic analysis of Reddit content about buprenorphine-naloxone using manual annotation and natural language processing techniques. J Addict Med 2022;16(4):454-460 [FREE Full text] [doi: 10.1097/ADM.0000000000000940] [Medline: 34864788]
53. Eshleman R, Jha D, Singh R. Identifying individuals amenable to drug recovery interventions through computational analysis of addiction content in social media. In: Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine. 2017 Presented at: BIBM '17; November 13-6, 2017; Kansas City, MO p. 849-854 URL: https://doi.org/10.1109/BIBM.2017.8217766 [doi: 10.1109/bibm.2017.8217766]
54. 51 LJ, Sridhar S, Pandey R, Hasan MA, Mohler G. Investigate transitions into drug addiction through text mining of reddit data. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mini. 2019 Presented at: KDD '19; August 4-8,2019; Anchorage, AK p. 2367-2375 URL: https://dl.acm.org/doi/10.1145/3292500.3330737 [doi: 10.1145/3292500.3330737]
55. Jha D, La Marca SR, Singh R. Identifying and characterizing opioid addiction states using social media posts. In: Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine. 2021 Presented at: BIBM '21; December

9-12, 2021; Houston, TX p. 913-918 URL: https://doi.org/10.1109/BIBM52615.2021.9669628 [doi: 10.1109/bibm52615.2021.9669628]
56. Yang Z, Nguyen LH, Jin F. Opioid relapse prediction with GAN. In: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2019 Presented at: ASONAM '19; August 27-30, 2019; Vancouver, BC p. 560-567 URL: https://dl.acm.org/doi/10.1145/3341161.3342951 [doi: 10.1145/3341161.3342951]
57. Yao H, Rashidian S, Dong X, Duanmu H, Rosenthal RN, Wang F. Detection of suicidality among opioid users on Reddit: machine learning-based approach. J Med Internet Res 2020 Nov 27;22(11):e15293 [FREE Full text] [doi: 10.2196/15293] [Medline: 33245287]
58. Chancellor S, Nitzburg G, Hu A, Zampieri F, Choudhury MD. 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 p. 1-15 URL: https://dl.acm.org/doi/10.1145/3290605.3300354 [doi: 10.1145/3290605.3300354]
59. Yang Z, Bradshaw S, Hewett R, Jin F. Discovering opioid use patterns from social media for relapse prevention. Computer 2022 Feb;55(2):23-33 [FREE Full text] [doi: 10.1109/mc.2021.3095826]
60. ElSherief M, Sumner SA, Jones CM, Law RK, Kacha-Ochana A, Shieber L, et al. Characterizing and identifying the prevalence of web-based misinformation relating to medication for opioid use disorder: machine learning approach. J Med Internet Res 2021 Dec 22;23(12):e30753 [FREE Full text] [doi: 10.2196/30753] [Medline: 34941555]
61. Sarker A, Al-Garadi MA, Ge Y, Nataraj N, Jones CM, Sumner SA. Signals of increasing co-use of stimulants and opioids from online drug forum data. Harm Reduct J 2022 May 25;19(1):51 [FREE Full text] [doi: 10.1186/s12954-022-00628-2] [Medline: 35614501]
62. Sumner SA, Bowen D, Holland K, Zwald M, Vivolo-Kantor A, Guy G, et al. Estimating weekly national opioid overdose deaths in near real time using multiple proxy data sources. JAMA Netw Open 2022 Jul 01;5(7):e2223033 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.23033] [Medline: 35862045]
63. Drug slang code words. Drug Enforcement Administration. 2017. URL: https://ndews.umd.edu/sites/ndews.umd.edu/files/dea-drug-slang-code-words-may2017.pdf [accessed 2024-04-29]
64. Lokala U, Lamy F, Daniulaityte R, Gaur M, Gyrard A, Thirunarayan K, et al. Drug abuse ontology to harness web-based data for substance use epidemiology research: ontology development study. JMIR Public Health Surveill 2022 Dec 23;8(12):e24938 [FREE Full text] [doi: 10.2196/24938] [Medline: 36563032]
65. Home page. The Urban Dictionary. URL: https://www.urbandictionary.com/define.php?term=a%7Cnd [accessed 2024-04-29]
66. Arshonsky J, Krawczyk N, Bunting A, Frank D, Friedman SR, Bragg MA. Informal coping strategies among people who use opioids during COVID-19: thematic analysis of reddit forums. JMIR Form Res 2022 Mar 03;6(3):e32871 [FREE Full text] [doi: 10.2196/32871] [Medline: 35084345]
67. Bunting AM, Frank D, Arshonsky J, Bragg MA, Friedman SR, Krawczyk N. Socially-supportive norms and mutual aid of people who use opioids: an analysis of Reddit during the initial COVID-19 pandemic. Drug Alcohol Depend 2021 May 01;222:108672 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108672] [Medline: 33757708]
68. Krawczyk N, Bunting AM, Frank D, Arshonsky J, Gu Y, Friedman SR, et al. "How will I get my next week's script?" Reactions of Reddit opioid forum users to changes in treatment access in the early months of the coronavirus pandemic. Int J Drug Policy 2021 Jun;92:103140 [FREE Full text] [doi: 10.1016/j.drugpo.2021.103140] [Medline: 33558165]
69. Keasar V, Sznitman S, Baumel A. Suicide prevention outreach on social media delivered by trained volunteers. Crisis 2023 May;44(3):247-254 [FREE Full text] [doi: 10.1027/0227-5910/a000864] [Medline: 35656647]
70. Liu X, Liu X, Sun J, Yu N, Sun B, Li Q, et al. Proactive suicide prevention online (PSPO): machine identification and crisis management for Chinese social media users with suicidal thoughts and behaviors. J Med Internet Res 2019 May 08;21(5):e11705 [FREE Full text] [doi: 10.2196/11705] [Medline: 31344675]
71. Notredame CE, Grandgenèvre P, Pauwels N, Morgiève M, Wathelet M, Vaiva G, et al. Leveraging the web and social media to promote access to care among suicidal individuals. Front Psychol 2018;9:1338 [FREE Full text] [doi: 10.3389/fpsyg.2018.01338] [Medline: 30154742]
72. 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 URL: https://doi.org/10.18653/v1/W17-1612 [doi: 10.18653/v1/w17-1612]
73. Poudel A, Weninger T. Navigating the post-API dilemma search engine results pages present a biased view of social media data. arXiv Preprint posted online March 11, 2024 [FREE Full text] [doi: 10.1145/3589334.3645503]
74. r/research. Reddit. 2024. URL: https://www.reddit.com/r/research/ [accessed 2024-04-29]
75. Leeson W, Resnick A, Alexander D, Rovers J. Natural language processing (NLP) in qualitative public health research: a proof of concept study. Int J Qual Methods 2019 Nov 13;18:160940691988702 [FREE Full text] [doi: 10.1177/1609406919887021]

Abbreviations

DAR: drug addiction recovery

IRB: institutional review board
JBI: Joanna Briggs Institute
MOUD: medication for opioid use disorder
NLP: natural language processing
OUD: opioid use disorder
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by T Mackey; submitted 23.07.23; peer-reviewed by M Haupt, R Chandrasekaran, T Nguyen; comments to author 20.03.24; revised version received 01.06.24; accepted 18.06.24; published 13.09.24.

Please cite as:
Almeida A, Patton T, Conway M, Gupta A, Strathdee SA, Bórquez A
The Use of Natural Language Processing Methods in Reddit to Investigate Opioid Use: Scoping Review
JMIR Infodemiology 2024;4:e51156
URL: https://infodemiology.jmir.org/2024/1/e51156
doi:10.2196/51156
PMID:

Original Paper

Dynamic Associations Between Centers for Disease Control and Prevention Social Media Contents and Epidemic Measures During COVID-19: Infoveillance Study

Shuhua Yin[1], PhD; Shi Chen[1], PhD; Yaorong Ge[1], PhD

University of North Carolina at Charlotte, Charlotte, NC, United States

Corresponding Author:
Shi Chen, PhD
University of North Carolina at Charlotte
9201 University City Blvd
Charlotte, NC, 28223
United States
Phone: 1 8148800738
Email: schen56@charlotte.edu

Abstract

Background: Health agencies have been widely adopting social media to disseminate important information, educate the public on emerging health issues, and understand public opinions. The Centers for Disease Control and Prevention (CDC) widely used social media platforms during the COVID-19 pandemic to communicate with the public and mitigate the disease in the United States. It is crucial to understand the relationships between the CDC's social media communications and the actual epidemic metrics to improve public health agencies' communication strategies during health emergencies.

Objective: This study aimed to identify key topics in tweets posted by the CDC during the pandemic, investigate the temporal dynamics between these key topics and the actual COVID-19 epidemic measures, and make recommendations for the CDC's digital health communication strategies for future health emergencies.

Methods: Two types of data were collected: (1) a total of 17,524 COVID-19–related English tweets posted by the CDC between December 7, 2019, and January 15, 2022, and (2) COVID-19 epidemic measures in the United States from the public GitHub repository of Johns Hopkins University from January 2020 to July 2022. Latent Dirichlet allocation topic modeling was applied to identify key topics from all COVID-19–related tweets posted by the CDC, and the final topics were determined by domain experts. Various multivariate time series analysis techniques were applied between each of the identified key topics and actual COVID-19 epidemic measures to quantify the dynamic associations between these 2 types of time series data.

Results: Four major topics from the CDC's COVID-19 tweets were identified: (1) information on the prevention of health outcomes of COVID-19; (2) pediatric intervention and family safety; (3) updates of the epidemic situation of COVID-19; and (4) research and community engagement to curb COVID-19. Multivariate analyses showed that there were significant variabilities of progression between the CDC's topics and the actual COVID-19 epidemic measures. Some CDC topics showed substantial associations with the COVID-19 measures over different time spans throughout the pandemic, expressing similar temporal dynamics between these 2 types of time series data.

Conclusions: Our study is the first to comprehensively investigate the dynamic associations between topics discussed by the CDC on Twitter and the COVID-19 epidemic measures in the United States. We identified 4 major topic themes via topic modeling and explored how each of these topics was associated with each major epidemic measure by performing various multivariate time series analyses. We recommend that it is critical for public health agencies, such as the CDC, to update and disseminate timely and accurate information to the public and align major topics with key epidemic measures over time. We suggest that social media can help public health agencies to inform the public on health emergencies and to mitigate them effectively.

(JMIR Infodemiology 2024;4:e49756) doi:10.2196/49756

KEYWORDS

infoveillance; social media; COVID-19; US Centers for Disease Control and Prevention; CDC; topic modeling; multivariate time series analysis

Introduction

The COVID-19 pandemic caused more than 760 million cases and 6.8 million deaths globally as of April 2023 [1]. Therefore, it is crucial for public health agencies, such as the US Centers for Disease Control and Prevention (CDC), to quickly and effectively disseminate up-to-date and reliable health information to the public to curb the pandemic. Over the past years, social media has been widely used by various public health agencies to make announcements, disseminate information, and deliver guidelines of effective interventions to the public. The CDC is among the early adopters of social media to engage with the public, increase health literacy in the society, and promote healthy behaviors [2]. Moreover, the CDC's social media team has developed the Health Communicator's Social Media Toolkit to efficiently use social media platforms; map health strategies; listen to health concerns from the public; and deliver evidence-based, credible, and timely health communications in multiple formats such as texts, images, and videos. The CDC's digital health communication efforts have been especially established on various social media platforms such as Twitter, Facebook, and Instagram.

Building successful interactions with the public relies on people understanding the content and raising awareness of it. The CDC has been heavily engaging in social media presence [3]. For example, during the COVID-19 pandemic since 2019, it has been responsive and proactive on Twitter to continuously tweet about reliable health-related messages and quickly diffuse public engagement by responding to user comments, retweeting credible sources, and monitoring online conversations in real time. Hence, it is meaningful to recognize the COVID-19 pandemic information disseminated by the CDC on social media, characterize various contents and topics, and evaluate posting patterns with regard to the actual epidemic dynamics. Monitoring the content, topics, and trends will help identify current issues or interests and the levels of interventions. It is critical to evaluate the associations between various COVID-19 content topics tweeted by the CDC and the actual COVID-19 epidemic measures (eg, cases, deaths, testing, and vaccination records). Knowing the underlying associations between the CDC's digital health communication contents on social media and the actual COVID-19 epidemics will help in understanding and evaluating the CDC's tweeting patterns with changes in the epidemic, and will further help in recommending more effective social media communication strategies for public health agencies accordingly.

Infodemiology and infoveillance studies tackle health challenges, generate insights, and predict patterns and trends of diseases using previously neglected online data. Infodemiology, which is the conjunction of "information" and "epidemiology," defined by Gunther Eysenbach, is the field of distribution and determinants of information of a population through the internet or other electronic media [4]. Infoveillance takes surveillance as the primary aim and generates automated analysis from massive online data. It employs innovative computational approaches to mine and analyze unstructured online text information, such as analyzing patterns and trends, predicting potential outbreaks, and addressing current issues of public health. Unlike traditional epidemiological surveillance systems, which include cohort studies, disease registries, population surveys, and health care records, infoveillance studies discover a wide range of health topics, monitor health issues including outbreaks and pandemics, and forecast epidemiological trends in real time. A large amount of anonymous online data can be obtained in a more timely manner with these approaches than with traditional surveillance systems, and this will help researchers and public health agencies to prepare for and tackle public health emergencies and issues more efficiently and effectively.

Social media platforms have been having impacts on the community education of COVID-19 and delivering various health information about the disease. Many studies have also incorporated the concept of infoveillance by analyzing unstructured textual data obtained from social media. Liu et al [5] collected and analyzed media reports and news articles on COVID-19 to derive topics and useful information. They aimed to investigate the relationship between media reports and the COVID-19 outbreak, and the patterns of health communication on the coronavirus through mass media to the general audience. They obtained media reports and articles related to the pandemic and studied prevalent topics. There had been prevalent public discussions of attitudes and perspectives on mask-wearing on social media. Therefore, it is important for public health agencies to disseminate the supporting evidence and benefits of masking to mitigate the spread of COVID-19. Al-Ramahi et al [6] studied the topics associated with the public discourse against wearing masks in the United States on Twitter. They identified and categorized different topics in their models. These studies all applied infoveillance to investigate the potential impacts of diseases, health behaviors, or interventions on target populations, communities, and the society. However, mass media and social media are also prone to the spreading of misinformation and conspiracy theories, especially from unreliable sources [7]. Hence, the sources of information obtained from social media are crucial as misinformation could potentially create bias, mislead public perceptions, and provoke negative emotions. Official accounts of public health agencies are usually sources of unbiased and reliable health information. Although there have been several studies that collectively explored the topics discussed by the general public on social media during the pandemic, no investigations have been performed so far to identify various topics from health agencies, such as the CDC, during a large health emergency.

Furthermore, information discrepancies and delays could occur between topics posted by health agencies and real-time epidemic trends. Such discrepancies could cause confusion among the public on interventions for health emergencies. Therefore, quantifying their associations is important to reduce knowledge gaps. Chen et al [8] studied correlations between the Zika epidemic in 2016 and the CDC's responses on Twitter. They quantified the association between the 2 types of data through multivariate time series analyses and information theory measurements. The study discovered the CDC's varying degrees of efforts in disseminating health-related information to the public during different phases of the Zika pandemic in 2016. However, no study so far has investigated such dynamic associations, more specifically, the CDC's COVID-19 content topic tweeting patterns and the actual COVID-19 epidemic metrics.

While still being investigated, it is imperative to understand the dynamic associations between various content topics on social media and actual epidemic outcome metrics, which will guide health agencies to identify driving factors between the 2 and help in disseminating helpful knowledge to the public accordingly. In this study, we aimed to discover the underlying COVID-related topics posted by the CDC during different phases of the COVID-19 pandemic. We also aimed to further quantify and evaluate the dynamic associations between content topics of the pandemic and multiple COVID-19 epidemic metrics. The findings of this study will significantly increase our knowledge about the efficiency of the CDC's health communications during the pandemic and help make further recommendations for the CDC's social media communication strategies with the public in the future.

Methods

Data Acquisition and Preprocessing

Using the Twitter academic API (application programming interface) and search query (see search query in Multimedia Appendix 1), we retrieved a total of 17,524 English tweets posted by 7 official CDC-affiliated Twitter accounts up to January 15, 2022 (for the detailed acquisition process for CDC tweets, see Multimedia Appendix 1). We also acquired the COVID-19 epidemic metric data in the United States from the Johns Hopkins University – Center for Systems Science and Engineering (CSSE) public GitHub repository [9-11]. Four sets of important COVID-19 time series data were retrieved, including daily cumulative confirmed cases, deaths, testing, and vaccination. The data were all at the US national level. The 4 sets of original COVID-19 time series data consisted of dates and their cumulative targeted measurements. The case series set included the daily cumulative number of confirmed COVID-19 reported cases, and it had 751 records, ranging from January 22, 2020, to February 10, 2022. The death series set reported the daily cumulative number of confirmed COVID-19 death cases, and it had 908 records, ranging from January 22, 2020, to July 17, 2022. The testing data set reported the daily cumulative number of completed polymerase chain reaction (PCR) tests or other approved nucleic acid amplification tests, and it had 760 records, ranging from January 13, 2020, to February 10, 2022. The vaccination data set included the daily cumulative number of people who received a complete primary series of vaccine doses from the CDC Vaccine Tracker, and it had 428 records, ranging from December 10, 2020, to February 10, 2022.

For consistency in subsequent analyses, all CDC tweet time series and US COVID-19 variable time series were standardized to the same time span in this study, ranging from the start date of reported case data (January 22, 2020) to the end date of CDC tweet collection (January 15, 2022), with a total of 725 records for each data type. Since vaccination data were not available until late 2020, missing values were filled with zeros. In summary, we had 4 time series from 4 different COVID-19 US epidemic metrics and another time series of number of tweets from all 7 CDC-associated Twitter accounts.

Natural Language Processing

In order to identify major topics in the CDC's COVID-19 tweets, we performed various natural language processing (NLP) steps. NLP, especially topic modeling, provides granular characterization of textual inputs such as the CDC's COVID-19 communications.

Regular expressions were first applied to process tweet texts by removing @mentions, hashtags, special characters, emails, punctuations, URLs, and hyperlinks. Tokenization was performed to break down sentences into individual tokens, which can be individual words or punctuations. For example, the sentence "As COVID19 continues to spread, we must remain vigilant" becomes tokens of "*As*," "*COVID19*," "*continues*," "*to*," "*spread*," "*,*" "*we*," "*must*," "*remain*," and "*vigilant*" after tokenization. Next, lemmatization, a structural transformation where each word or token is turned to its base or dictionary form of the morphological information, was performed. For example, for words "studies" and "studying," the base form, or lemma, was the same "study." In addition to stop word removal via the Python NLTK library, we created our own list of stop words and removed them from the texts (see the stop words list in Multimedia Appendix 1). With help from domain experts, we excluded stop words that did not contribute to topic mapping.

N-grams, phrases with *n* words, were developed with a threshold value of 1 to form phrases from tweets. Phrase-level *n*-grams were applied here because phrases offer more semantic information than individual words [12]. A higher threshold value resulted in fewer phrases to be formed. The texts were mapped into a dictionary of word representations, which was a list of unique words, and it was then used to create bag-of-words presentations of the texts. A term frequency-inverse document frequency (TF-IDF) model was implemented to evaluate the importance and relevancy of the words to a document. It was calculated by multiplying term frequency, which is the relative frequency of a word within a document, with inverse document frequency, which measures how common or rare a word is across a corpus. A higher TF-IDF value indicates that the word is more relevant to the document it is in [13,14]. Words that were missing and lower than the threshold value of 0.005 from the TF-IDF model were excluded. Table 1 shows the process of data collection and preprocessing, and Table 2 shows the steps of subsequent NLP and statistical analyses.

Table 1. Data collection and preprocessing.

Variable	Data collection	Data preprocessing
CDC[a] tweets	• Twitter API[b] using a search query • 17,524 English tweets by January 15, 2022	• Remove @mentions, hashtags, special characters, emails, punctuations, URLs, and hyperlinks • Tokenization: break down sentences into individual tokens • Lemmatization: each word or token is turned to its base or dictionary form • Remove a list of stop words created by research experts • N-grams: form phrases from the tweets • Modify the date range: January 22, 2020 (the start date of reported case data) to January 15, 2022 (the end date of CDC tweets)
COVID-19 epidemic metrics	• Public GitHub repository of the CSSE[c] at Johns Hopkins University • Confirmed case count: 751 records; January 22, 2020, to February 10, 2022 • Death count: 908 records; January 22, 2020, to July 17, 2022 • Completed COVID-19 tests: 760 records; January 13, 2020, to February 10, 2022 • Complete vaccination: 428 records; December 10, 2020, to February 10, 2022	• Standardize metric time series to be the same as that of CDC tweets • Fill missing values in the vaccination data with zeros • 725 records for each of the 4 metric series • Turn cumulative records to daily records

[a]CDC: Centers for Disease Control and Prevention.

[b]API: application programming interface.

[c]CSSE: Center for Systems Science and Engineering.

Table 2. Subsequent analyses.

Variable	Topic modeling	Data analysis
CDC[a] tweets and COVID-19 metrics	• Construct an LDA[b] topic model using CDC tweets assigning 4 topics • Extract generated topics with their top 10 unique associated keywords • Produce interactive visualizations using pyLDAvis	• Domain experts examine topic keywords with randomly sampled tweets in iteration • Domain experts determine the theme of each topic • Perform multivariate time series analyses between each topic time series and each COVID-19 metric time series: 1. Visualization 2. Cross-correlation function (CCF) 3. Mutual information (MI) 4. Autoregressive integrated moving average with external variable (ARIMAX) model

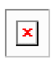
[a]CDC: Centers for Disease Control and Prevention.

[b]LDA: latent Dirichlet allocation.

Topic Modeling With Latent Dirichlet Allocation

To identify more specific topics from all the COVID-19 tweets posted by the CDC, we performed topic modeling via latent Dirichlet allocation (LDA). LDA automatically generates nonoverlapping clusters of words (ie, clusters of words based on their distributions in their corresponding topics) that represent different topics based on probabilistic distributions across the whole corpus (ie, all CDC tweets in this study). LDA was developed to find latent, hidden topics from a collection of unstructured documents or a corpus with text data. Topic models are probabilistic models that perform at 3 levels of documents: a word, a document, and a corpus consisting of multiple documents. The details of LDA and topic models are provided in Multimedia Appendix 1. We investigated and compared across 3 to 8 potential topics and determined the optimal number of topics based on both topic model evaluation and domain expert interpretations of the identified topic clusters.

Model perplexity and topic coherence scores were calculated as performance metrics of LDA. Perplexity is a decreasing "held-out log-likelihood" function that assesses LDA performance using a set of training documents. The trained LDA model is then used to test documents (held-out set). The perplexity of a probability model *q* on how well it predicts a set of samples x_1, x_2, ..., x_N drawn from an unknown probability distribution *p*, is defined as follows [15]:



An ideal *q* should have high probabilities $q(x_i)$ for the new data. Perplexity decreases as the likelihood of the words in new data increases. Therefore, lower perplexity indicates better predictability of an LDA model.

Topic coherence assesses the quality of the topics, which is measured as the understandability and semantic similarities between high scoring words (ie, the words that have a high probability of occurring within a particular topic) in topics generated by LDA [16]. We used the UMass coherence score [17], which accounts for the order of a word appearing among the top words in a topic. It is defined as follows [18]:

where N is the number of top words of a topic of a sliding window, $P(w_i)$ is the probability of the ith word w appearing in the sliding window that moves over a corpus to form documents, and $P(w_i, w_j)$ is the probability of words w_i and w_j appearing together in the sliding window. According to the study from UMass, coherence decreases initially and becomes stationary as the number of topics increases [16].

Representations of all topics were presented in word-probability pairs for the most relevant words grouped by the topics. Interactive visualizations were produced using the pyLDAvis package in Python 3.7 to examine the topics generated by LDA and their respective associated keywords. A data frame of all dominant key topics was created. The original unprocessed full texts of the CDC tweets, IDs, and posting dates were combined into a data frame along with their corresponding key topic number labels and topic keywords. In addition, the daily percentage of each topic from LDA was calculated for further time series analysis. For instance, vaccine/vaccination is an identified key topic, so the percentage of vaccine-related CDC tweets on each day was calculated for the entire study period to construct the vaccine/vaccination-specific topic time series. Since LDA is technically an unsupervised clustering method, after the topics or clusters of word distributions from the CDC's tweets were generated using LDA, domain experts were involved to further label and interpret the content of the topics using domain knowledge. We randomly generated 20 sample tweets from each topic using Python for domain experts to examine, analyze, and determine the themes of the topics. For each topic, LDA provided a list of the top keywords associated with that topic, and we selected the top 10 keywords. We examined these keywords and referred to the 20 sample tweets, and then derived a theme or context that encompasses these keywords and the original tweets through further discussions, which was important for understanding the context in which these words were used. The final agreement on the interpretation of LDA-generated topics was reached after multiple iterations and discussions of the above process.

Multivariate Time Series Analyses Between Identified CDC Tweet Topics and COVID-19 Epidemic Metrics

Data Preparation

Key topic time series data were derived from the previous NLP and LDA processes. We constructed a multivariate data frame with posting dates and number of tweets for each key topic at a daily resolution. Since LDA identified 4 key topics, a total of 4 CDC key topic time series were developed. There were also 4 US COVID-19 epidemic metric time series: daily cumulative reported cases, cumulative confirmed deaths, cumulative number of completed PCR tests or other approved nucleic acid amplification tests, and cumulative number of people who received a complete primary series of vaccines. These 4 sets of COVID-19 epidemic metric time series were then converted to daily measures via first order differencing. Multivariate time series analyses were implemented to investigate the associations between time series of key CDC tweet topics and US COVID-19 epidemic metrics.

Visualizations

Both types of time series, CDC key topics and COVID-19 metrics, were visually inspected in the same plot on double y-axes, with the left y-axis displaying the daily COVID-19 metric and right y-axis displaying the daily CDC tweet topic count. In addition, each plot was further divided based on COVID-19 phases with different dominant variants: the original, Alpha, Delta, and Omicron variants, with their corresponding starting dates: March 11, 2020; December 29, 2020; June 15, 2021; and November 30, 2021, respectively. This helps further observe and identify dynamic changes of time series and their associations during different phases of the pandemic.

Cross-Correlation Function

Between 2 time series (also known as signals x and y), the cross-correlation function (CCF) [19] quantifies their levels of similarities (ie, how similar the 2 series are at different times), their associations (ie, how values in one series can provide information about the other series), and when they occur [20]. The CCF takes the sum of the product for each of the x and y data points at time lag l, defined as follows [19]:

where N is the number of observations in each time series, and x_i and y_i are the observations at the ith time step in each of the time series. The CCF ranges from −1 to 1, and a larger absolute value of the CCF is related to a greater association shared by the 2 time series at a given time lag l [21]. In this study, each of the 4 CDC tweet topic time series was compared with each of the 4 COVID-19 epidemic metric time series to calculate their respective CCFs. All CCF values were calculated with a maximum lag of 30 days, as we assumed that the real-world epidemic could not influence online discussions for more than a month and vice versa.

Mutual Information

Mutual information (MI) was calculated by computing the entropy of the empirical probability distribution to further quantify the association between each of the 4 key CDC tweet topics and each of the 4 US COVID-19 epidemic metrics. MI measures the amount of mutual dependence or average dependency between 2 random variables X and Y. It is defined as follows [22]:

where x_i and y_i are the ith elements of the variables X and Y, respectively. When applied to time series data, X and Y are 2 individual time series and x_i and y_i are their respective observations at the *i*th time step. Note that MI is a single value instead of a function over lag *l* as in the CCF. A larger MI value indicates a higher shared mutual dependency between the 2 time series.

Autoregressive Integrated Moving Average With External Variable

Neither the CCF nor MI differentiate dependent and independent variables, that is, the formula was symmetric with regard to *X* and *Y* variables. We further evaluated whether the CDC tweeting topics were influenced by real-world COVID-19 epidemic outcomes. An autoregressive integrated moving average with external variable (ARIMAX) model was constructed to fit each of the 4 CDC topics with each of the 4 COVID-19 epidemic metrics during the entire study period. A univariate autoregressive integrated moving average (ARIMA) model fits and forecasts time series data with the integration of an autoregressive (AR) component and a moving average (MA) component with their respective orders/lags (see Multimedia Appendix 1 for detailed information about the AR model). The ARIMA model consists of both AR(*p*) and MA(*q*) as well as an order *d* differencing term, resulting in the following ARIMA (*p*, *d*, *q*) model [23, 24]:

or in backward shift operator form:

See Multimedia Appendix 1 for details on the parameters.

The ARIMAX model further extends ARIMA to the multivariate time series by incorporating at least one exogenous independent variable x_t. ARIMAX (*p*, *d*, *q*) is specified as follows [25]:

or in backward shift operator form [26]:

where contributes to the exogeneous independent variable that could potentially influence the dependent variable y_t.

In this study, ARIMAX was developed to evaluate how real-world epidemic metrics, modeled as exogeneous variables, impact CDC tweet topic dynamics as dependent variables. Each of the 4 CDC tweet topics was modeled as a dependent variable (y_t) and each of the 4 COVID-19 epidemic measures was an independent exogeneous variable (x_t). The optimal ARIMA and ARIMAX model parameter set (*p*, *d*, *q*) was determined by the *R* ARIMA model package.

In addition to reporting the values of the ARIMAX model parameter set (*p*, *d*, *q*), difference in Akaike information criterion (dAIC), root mean square error (RMSE), and mean absolute error (MAE) were also computed to compare different ARIMAX performances. The optimal model was the one with the lowest AIC score. dAIC was computed in between 2 models (see Multimedia Appendix 1 for detailed information on AIC). We had an ARIMA model of a single topic time series and an ARIMAX model of that topic time series with an exogeneous variable. Negative dAIC values indicated that the ARIMAX model showed improvement in model performance over the ARIMA counterpart that did not include an exogenous variable.

The commonly used RMSE and MAE were adopted as performance metrics. They are defined as follows [27]:

where *n* is the number of data points in a sample y (y_i, where *i*=1, 2, …, *n*). RMSE and MAE are Euclidean distance and Manhattan distance in high-dimensional space, respectively.

Results

Topic Modeling and Content Results

A total of 17,524 English tweets posted by the CDC were retrieved and analyzed. Four key topics were generated via LDA based on evaluation metrics including perplexity and coherence score. These topics were then examined and categorized to themes by domain experts (Textbox 1 with example tweets with their respective topics). The themes of the topics and their top 10 unique associated keywords are presented in Table 3.

Topics were plotted as circles and displayed on the left panel; the most relevant terms or associated keywords with their corresponding topics were displayed in frequency bars on the right panel, which showed each term's frequency from each topic across the corpus (ie, all CDC COVID-19 tweets sampled) [28] (see Multimedia Appendix 1 for more detailed information about visualizations in the pyLDAvis package). The size of the circle indicated the prevalence of that topic in the corpus. Visualizations for all topics, displayed in circles on the left panel, and their top 15 corresponding relevant terms or associated keywords, displayed in frequency bars on the right panel, are provided in Figures S1-S5 in Multimedia Appendix 1.

Based on the LDA visualization results, these 4 identified key topics had the largest distances and minimal dimensional overlap in the reduced 2D plane. From a public health perspective, the CDC's online health communication of COVID-19, the largest health emergency in the 21st century, has been relatively cohesive and comprehensive. Therefore, the 4 key topics identified via LDA were not completely mutually exclusive. In addition, the 4-topic model had the balance of separation of topics from a computational perspective and clear interpretability from a health perspective. Increasing the number of topics yields a substantial amount of topic overlap, which was also challenging to provide explicit and clear interpretations. Figure 1 illustrates an example of topic 4 [29,30]. A list of associated terms of topic 4 and the overall frequency of the terms in the corpus have been displayed in the right panel. The 5 key terms from topic 4 based on overall frequency across all tweets were "booster," "school," "increase," "parent," and "country."

Textbox 1. Example tweets from each topic theme.

Topic 1: General vaccination information and education, especially preventing adverse health outcomes of COVID-19

- "Even as the world's attention is focused on #COVID19, this week we are taking time to highlight how #VaccinesWork and to thank the heroes who help develop and deliver lifesaving vaccines. #WorldImmunizationWeek message"
- "CDC's #COVID19 Vaccine Webinar Series is a great place to start learning about a variety of topics around COVID-19 vaccination."
- "The #DeltaVariant of the virus that causes #COVID19 is more than two times as contagious as the original strain. Wear a mask indoors in public, even if vaccinated and in an area of substantial or high transmission. Get vaccinated as soon as you can."

Topic 2: Pediatric intervention, pediatric vaccination information, family safety, and school and community protection

- "Make #handwashing a family activity! Explain to children that handwashing can keep them healthy. Be a good role model—if you wash your hands often, your children are more likely to do the same. #COVID19"
- "Parents: During #COVID19, well-child visits are especially important for children under 2. Schedule your child's routine visit, so the healthcare provider can check your child's development & provide recommended vaccines."
- "It is critically important for our public health to open schools this fall. CDC resources will help parents, teachers and administrators make practical, safety-focused decisions as this school year begins."

Topic 3: Updates on COVID-19 testing, case, and death data, and relevant information of the disease

- "CDC tracks 12 different forecasting models of possible #COVID19 deaths in the US. As of May 11, all forecast an increase in deaths in the coming weeks and a cumulative total exceeding 100,000 by June 1. See national & state forecasts."
- "The latest CDC #COVIDView report shows that the percentage of #COVID19-associated deaths has been on the rise in the United States since October and has now surpassed the highest percentage seen during summer."
- "#COVID19 cases are going up dramatically. This increase is not due to more testing. As the number of cases rise, so does the percentage of tests coming back positive, which shows that COVID-19 is spreading."

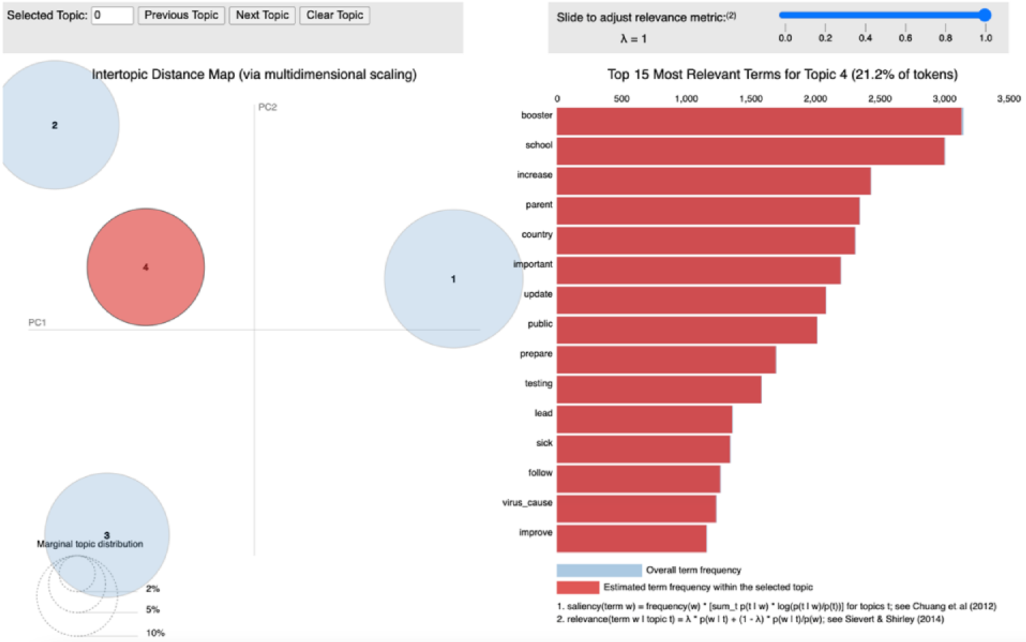
Topic 4: Research, study, health care, and community engagement to curb COVID-19

- "Our Nation's medical community has been vigilant and their help in identifying confirmed cases of #COVID19 in the United States to date has been critical to containing the spread of this virus."
- "In a new report using data from Colombia, scientists found that pregnant women with symptomatic #COVID19 were at higher risk of hospitalization & death than nonpregnant women with symptomatic COVID-19. HCPs can inform pregnant women about how to stay safe."
- "A new study finds masking and fewer encounters or less time close to persons with #COVID19 can limit the spread in university settings. #MaskUp when inside indoor public places regardless of vaccination status."

Table 3. Identified key topics of Centers for Disease Control and Prevention tweets with unique focal keywords.

Key topics	Top 10 unique keywords
1. General vaccination information and education, especially preventing adverse health outcomes of COVID-19 (including cases, severe conditions/hospitalization, and death)	learn, time, safe, fully vaccinate, prevent, child age, old, share, flu, month
2. Pediatric intervention, pediatric vaccination information, family safety, and school and community protection	work, school, datum, test, infection, family, free, home, public, check
3. Updates on COVID-19 testing, case, and death data, and relevant information of the disease	patient, update, booster, cause, recommend, increase, day, program, important, read
4. Research, study, health care, and community engagement to curb COVID-19	vaccination, vaccinate, child, protect, protection, report, visit, risk, community, travel

Figure 1. Interactive mapping of topic 4 generated by latent Dirichlet allocation.



Multivariate Time Series Analysis Results

CCF Results

The time series of CDC tweet topics and COVID-19 metrics were plotted to visually examine patterns and potential associations. A total of 16 time series plots (4 topics × 4 COVID-19 epidemic metrics) were generated (Figures S14-S29 in Multimedia Appendix 1). CCFs were computed to quantify the dynamic association between each CDC key topic series and each of the 4 COVID-19 epidemic metrics. Quantitative results have been presented (Tables S3-S6 in Multimedia Appendix 1). Visualizations (Figures S30-S44 in Multimedia Appendix 1) illustrated CCFs between both types of time series. CCF values and plots showed that the CDC's key COVID-19 tweet topic series was not substantially correlated with the confirmed COVID-19 case count series. As an example, there were no specific patterns between topic 2 and daily confirmed COVID-19 cases (Figure 2A).

COVID-19 confirmed cases and the death time series had very similar dynamic patterns in the United States across the time span (Figure 2B). Consequently, they also showed similar CCFs with the CDC key topic series (Figure S45 in Multimedia Appendix 1). COVID-19 deaths had no substantial correlations with any of the 4 CDC key topics (Figures S18-S21 in Multimedia Appendix 1) based on CCFs. There were no substantial correlations between any of the 4 key topics and the COVID-19 testing series as well as the fully vaccinated rate series. Examples showed the CCFs between those and topic 2 (Figures 3 and 4). These results indicated a potential discrepancy between the CDC's health communication focus and the actual COVID-19 epidemic dynamics in the United States during the pandemic.

Figure 2. Time series of topic 2 against 2 COVID-19 metrics: (A) case counts, (B) death counts. CDC: Centers for Disease Control and Prevention; US: United States.

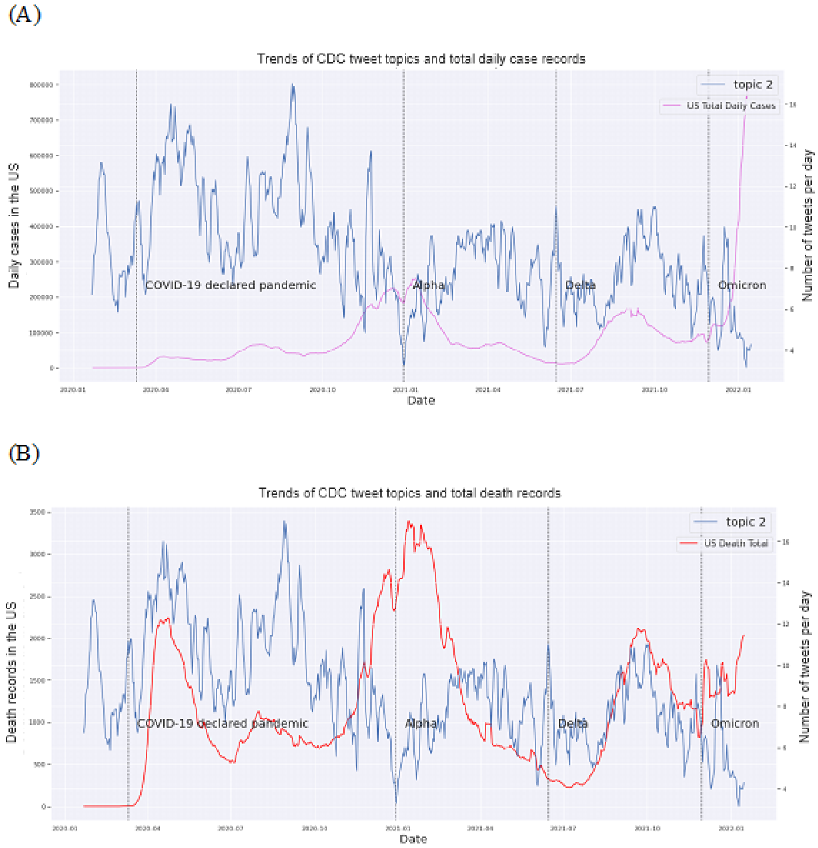


Figure 3. Cross-correlation function (CCF) between the completed COVID-19 test series and topic 2 tweets. (A) Trends of CDC tweet topics and number of completed tests; (B) CCF between COVID-19 confirmed cases and topic 2 tweets. CDC: Centers for Disease Control and Prevention.

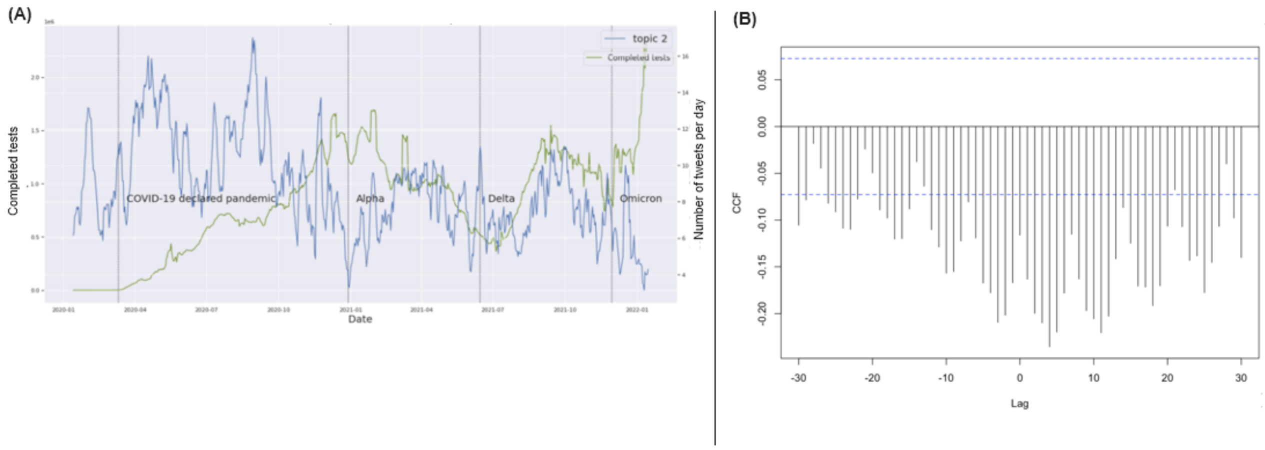
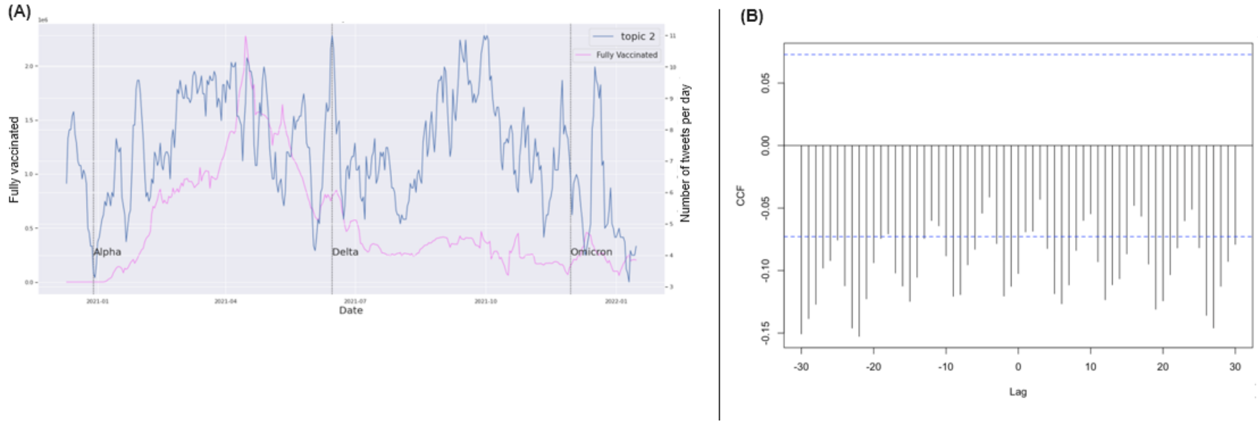


Figure 4. Cross-correlation function (CCF) between the completed COVID-19 vaccination series and topic 2 tweets. (A) Trends of CDC tweet topics and vaccination records; (B) CCF between records of fully vaccinated people and topic 2 tweets. CDC: Centers for Disease Control and Prevention.



MI Results

MI values between each CDC tweet topic and each COVID-19 metric were calculated, and they are shown in Table 4. Confirmed case counts and topic 4 (research, health care, and community engagement to restrain COVID-19) had the highest MI value (3.21), indicating that there was a strong dependency in COVID-19 cases and topic 4. On the other hand, the vaccination rate and topic 3 had the lowest MI value (0.56), indicating almost independence between the 2 series. Among all 4 key topics, topic 4 showed the highest MI values (3.21, 3.02, 3.21, and 1.65) with the 4 COVID-19 metrics. Topic 2 (pediatric intervention, family safety, and school and community protection) had consistently lower MI values with the COVID-19 metric than topic 4. The MI of topic 1 (information on COVID-19 vaccination and education on preventing its adverse health outcomes) and topic 3 (updates on COVID-19 testing, case, and death metrics, and relevant information of the disease) showed similar values with all 4 COVID-19 metrics, although the MI values of topic 1 were slightly higher. Vaccination and educational information on the adverse health outcomes of COVID-19 appeared to not be substantially correlated with COVID-19 epidemic metrics, including testing, cases, and deaths. We speculated that the CDC considered both vaccination and preventing adverse health outcomes of COVID-19 critical to public health and disseminated these topics regardless of the current COVID-19 situation at the time of posting.

In addition, MI values between all pairs of CDC topics were calculated (Table S7 in Multimedia Appendix 1). The resulting MI values, ranked from the largest to smallest, were for topics 2 and 4, topics 3 and 4, topics 1 and 2, topics 2 and 3, topics 1 and 4, and topics 1 and 3. Based on the CDC's COVID-19 tweeting patterns, pediatric intervention and family and community safety were strongly associated with health care research studies and public engagement to curb the spread of COVID-19.

Table 4. Mutual information values between Centers for Disease Control and Prevention key topics and COVID-19 metrics in the United States.

COVID-19 daily measurements in the United States	Topic 1[a]	Topic 2[b]	Topic 3[c]	Topic 4[d]
Confirmed case counts	1.25	2.93	1.18	3.21
Death counts	1.12	2.74	1.06	3.02
Completed COVID-19 test counts	1.24	2.91	1.18	3.21
Fully vaccinated counts	0.60	1.49	0.56	1.65

[a]Topic 1: General vaccination information and education, especially preventing adverse health outcomes of COVID-19.

[b]Topic 2: Pediatric intervention, pediatric vaccination information, family safety, and school and community protection.

[c]Topic 3: Updates on COVID-19 testing, case, and death data, and relevant information of the disease.

[d]Topic 4: Research, study, health care, and community engagement to curb COVID-19.

ARIMAX Results

ARIMAX performance measures, including values of ARIMAX parameters (*p*, *d*, *q*), dAIC, RMSE, and MAE, are reported in Table 5. As an external input, the vaccination rate time series significantly improved the performances of the original ARIMA models for all CDC key topics (dAIC = −108.15, −69.79, −90.54, and −91.53 for topics 1 to 4, respectively). This was the largest increase in model performance across all topics with the exogeneous variable in the ARIMAX model. The COVID-19 case series improved the ARIMA model performance for CDC topics 1 and 3 (dAIC = −104.76 and −1.53 for topics 1 and 3, respectively). Including the death or testing series did not result in substantial improvements to the ARIMA model performance for all CDC key topics.

ARIMAX models with lower RMSE and MAE values indicated higher accuracy of the time series models (Table 5). Overall, ARIMAX models for topics 1 and 3 with all COVID-19 metrics delivered the smallest RMSE values (lowest [1.10] for topic 3 with death counts and highest [1.21] for topic 1 with full vaccination records), while those of topic 4 delivered the largest

RMSE values (lowest [6.25] with death counts and highest [6.93] with full vaccination records). Similarly, MAE values were the lowest for ARIMAX models for topics 1 and 3 with the epidemic metrics (lowest [0.82] for topic 3 with death counts and highest [0.91] for topic 1 with full vaccination records), and they were the largest for topic 4 with the epidemic metrics (lowest [4.97] with death counts and highest [5.56] with full vaccination records). These ARIMAX performance results showed significant variabilities between the 2 types of time series (CDC key tweet topics and actual COVID-19 metrics in the United States).

We performed an exhaustive search to identify the optimal ARIMAX parameters (*p*, *d*, *q*). For example, topic 1 with death counts and completed testing records had the same parameter set (*p*, *d*, *q*=2, 1, 3), indicating that the optimal ARIMAX model between these time series needed first-order differencing (*d*=1) to achieve stationarity and minimal AIC values, its AR time lag was 2 (*p*=2), and its MA time lag was 3 (*q*=3). The topic 1 series with case counts and complete vaccination had the same parameter values (*p*, *d*, *q*=5, 1, 0), indicating that the model was simply an AR model (*q*=0 with no MA terms) with a time lag of 5 (*p*=5) after first-order differencing (*d*=1). The complete ARIMAX parameters are shown in Table 5. All ARIMAX models needed first-order differencing (*d*=1) to be stationary and to minimize AIC values.

Table 5. Autoregressive integrated moving average with external variable performance measures of each Centers for Disease Control and Prevention topic and COVID-19 epidemic metric pair.

COVID-19 epidemic measures and ARIMAX[a] metrics	Topic 1[b]	Topic 2[c]	Topic 3[d]	Topic 4[e]
Case counts				
ARIMAX par[f]	(5, 1, 0)	(4, 1, 1)	(2, 1, 1)	(3, 1, 2)
dAIC[g]	−104.76[h] (2240.19, 2344.95)[i]	0.45 (4304.09, 4303.64)	−1.53[h] (2227.59, 2229.12)	11.97 (4785.89, 4773.92)
RMSE[j]	1.21	4.66	1.12	6.45
MAE[k]	0.90	3.66	0.86	5.10
Death counts				
ARIMAX par	(2, 1, 3)	(4, 1, 1)	(2, 1, 1)	(3, 1, 2)
dAIC	6.72 (2240.19, 2233.47)	36.60 (4304.09, 4267.49)	20.43 (2227.59, 2207.16)	60.14 (4785.89, 4725.75)
RMSE	1.12	4.56	1.10	6.25
MAE	0.84	3.57	0.82	4.97
Testing				
ARIMAX par	(2, 1, 3)	(4, 1, 1)	(0, 1, 2)	(3, 1, 2)
dAIC	0.13 (2240.19, 2240.06)	19.56 (4304.09, 4284.53)	1.83 (2227.59, 2225.76)	36.97 (4785.89, 4748.92)
RMSE	1.13	4.60	1.11	6.34
MAE	0.84	3.61	0.85	4.99
Vaccination				
ARIMAX par	(5, 1, 0)	(5, 1, 0)	(5, 1, 0)	(5, 1, 0)
dAIC	−108.15[h] (2240.19, 2348.34)	−69.79[h] (4304.09, 4373.88)	−90.54[h] (2227.59, 2318.13)	−91.53[h] (4785.89, 4877.42)
RMSE	1.21	4.90	1.18	6.93
MAE	0.91	3.81	0.89	5.56

[a]ARIMAX: autoregressive integrated moving average with external variable.

[b]Topic 1: General vaccination information and education, especially preventing adverse health outcomes of COVID-19.

[c]Topic 2: Pediatric intervention, pediatric vaccination information, family safety, and school and community protection.

[d]Topic 3: Updates on COVID-19 testing, case, and death data, and relevant information of the disease.

[e]Topic 4: Research, study, health care, and community engagement to curb COVID-19.

[f]ARIMAX parameters (p, d, q).

[g]dAIC: delta Akaike information criterion (AIC) or difference in AIC.

[h]Negative dAIC: indicates improvement of performance in the ARIMAX model compared with its autoregressive integrated moving average (ARIMA) model.

[i]AIC values of ARIMA and its corresponding ARIMAX models.

[j]RMSE: root mean square error.

[k]MAE: mean absolute error.

Discussion

Principal Findings

In this study, we systematically investigated and comprehensively identified the CDC's key topics, COVID-19 epidemic metrics, and dynamic associations between the 2 types of data series based on 17,524 COVID-related English tweets from the CDC since January 2022. The LDA topic model was built to characterize and identify the dynamic shifts of topics in the CDC's COVID-19 communication over a period of more than 2 years. For the first time, we were able to identify the following 4 key topics: (1) general vaccination information and education; (2) pediatric intervention that also involved family and school safety; (3) updates on the COVID-19 epidemic situation, such as numbers of cases, deaths, etc; and (4) research studies that were able to curb the pandemic.

Our study took a unique approach of infoveillance by identifying potential associations between COVID-19 epidemic outcome metrics in the United States and the CDC's key topic dynamics during different stages of the pandemic. This innovative framework significantly expanded the original infoveillance approach that generally relied on the number of posts (ie, posting dynamics) without further extracting more detailed and meaningful content topics and sentiments from the textual data. Our study was able to further provide practical and useful health communication strategies for public health agencies to effectively communicate timely and accurate information to the public. It is important to investigate the dynamic associations between the CDC's tweets on COVID-19 and the progression of the pandemic for several reasons:

1. Understanding their relationship can reveal how public health messaging impacts public perception and engagement at different stages of a major health emergency. A strong association between the CDC's tweets and epidemic measures indicates that public health messaging is effective. Weak associations might indicate that messaging from the CDC to the public over time is not effective; however, it will lead us to further explore the influential factors and provide health communication strategies for public health agencies.
2. It can also show if the CDC's messaging on Twitter is proactive or reactive to the actual epidemic, informing strategies for future public health communication.
3. It helps public health agencies better allocate resources. For example, if tweets related to educating the public on monitoring COVID-19 symptoms and updating certain metrics lead to an increase in the number of people trying to get COVID tests, then resources could be directed toward opening testing centers and sending free test kits to homes.

Our study is the first of its kind to comprehensively evaluate the impact of online public health communication, especially on Twitter, which is one of the major social media platforms, during different phases of a large health emergency. We studied the overall daily volume of COVID-19–related tweets posted by the CDC over time as a baseline (Figure 5), and the volume of tweets was higher in the early phase of the pandemic, indicating a strong effort at the CDC to disseminate important information to the public. We did not observe visually clear patterns of an association with COVID-19 epidemic measures. We further applied novel NLP to significantly reduce the gap of previous studies that overlooked the dynamic association between detailed topics discussed by public health agencies on social media and real-world epidemic metrics.

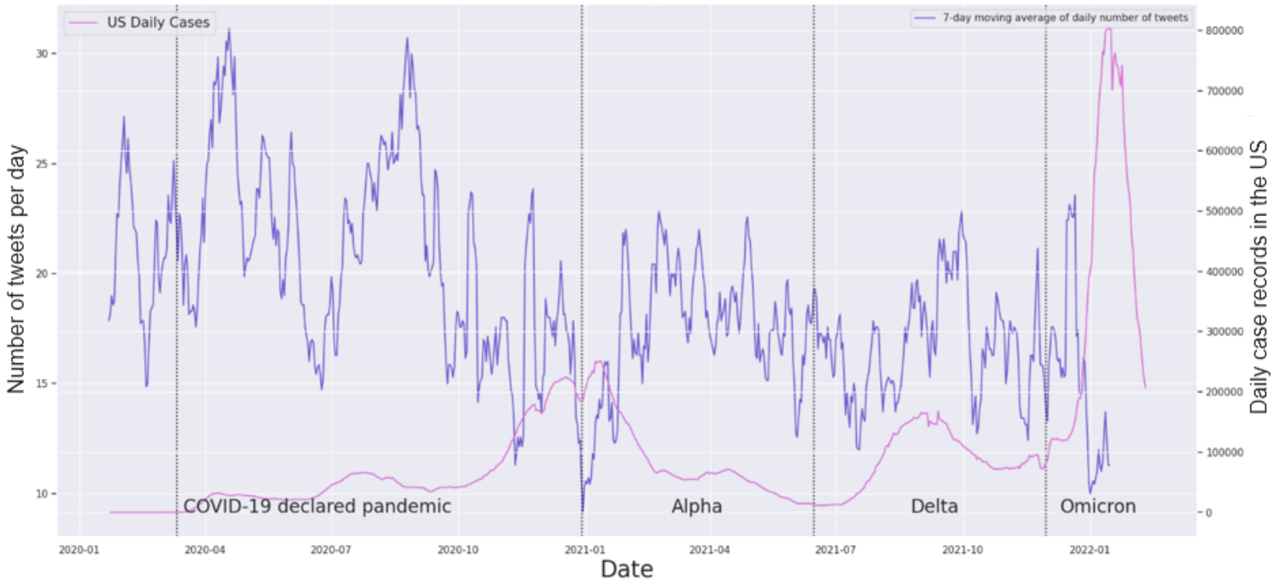
We then examined the dynamic associations between the 4 identified key topics and 4 COVID-19 epidemic outcome metrics. Among the 4 major topics, topic 1, which covered information on vaccination and adverse health outcomes of COVID-19, had substantially strong associations with death counts and testing records during the Alpha phase (December 29, 2020, to June 14, 2021). We found that during this phase, when the overall vaccination-related CDC tweets were decreasing, the daily vaccination rate (number of people who received a complete primary series of the COVID-19 vaccine based on the CDC Vaccine Tracker) was increasing, which aligned with the CDC's effort in emphasizing the importance of vaccination to the public on social media. When discussions from the CDC about vaccination were increasing after the Alpha phase, the vaccination rate started to decrease. The reasons could be but are not limited to the following:

1. Ineffective messaging from the CDC on social media to the public during later stages of the pandemic.
2. Lack of engagement from the public, since not everyone follows or engages with official accounts and might miss or overlook them amidst other content.
3. Fatigue from information overload where frequent data updates on social media platforms can lead to desensitization, making it less likely for users to pay attention over time and act on the information.
4. Temporal delays create time lag, which can impact the associations between the topics and the real epidemic measures.
5. Political factors such as antivaccination groups.

Therefore, with all possible influential factors, the CDC could not fully impact the public's responses and actions on getting vaccinated even though they had been making efforts on sharing educational information about vaccination. This finding showed that the CDC had been making efforts to emphasize the importance of vaccination on Twitter, but the public response was weak. Thus, it is important to further study the influential factors for the CDC's social media strategies. Topic 3, which provided updates on 3 of the COVID-19 measures (testing, cases, and deaths) and their relevant information, aligned better with the case series during the Delta phase (June 15, 2021, to November 29, 2021). It also matched with the death series during the declared pandemic phase (original variant: March 11, 2020, to December 28, 2020) and Delta phase, classified by the World Health Organization on May 11, 2021. Furthermore, even though topic 3 did not demonstrate a visible association with the testing series, timely communication from the CDC was actually strongly associated with the testing time series over the entire study period based on the multivariate time series analysis.

According to these key findings, we suggest that aligning the content topics of health communication from public health agencies with the temporal dynamics of COVID-19 or other emerging public health emergencies (eg, major epidemic outcome metrics) can help provide more timely and relevant information to the public. Therefore, we recommend that the CDC and other public health agencies monitor the epidemic outcome metrics in real time. Health agencies can then post timely updates about the emergency, most recent findings, and interventions on social media according to the dynamic changes of these outcome metrics. Public health agencies can regain trust from the public by not only helping the public better understand the complex dynamics of the health emergency, but also informing the public with evidence-based guidance and recommendations more effectively.

Figure 5. Time series of the daily number of Centers for Disease Control and Prevention (CDC) tweets and COVID-19 case counts. US: United States.



Limitations and Future Work

There are several limitations in this infoveillance study that could be improved in future work. First, while we focused on probabilistic-based LDA for topic modeling, there are other alternative NLP approaches such as deep learning–based bidirectional encoder representations from transformers (BERT). Hence, we will explore BERT and other state-of-the-art NLP techniques for content topic modeling and sentiment analysis in the future. Second, given the complexity of this study, we will incorporate subthemes to further help contextualize the clusters in future work. Third, the CDC does not have the sole power of controlling people's responses and actions over time (eg, getting tested and receiving full vaccine doses), even with consistent effort on Twitter to educate the public and mitigate the pandemic. There are other factors that could affect the associations between the CDC's messages and the COVID-19 measures:

1. Time lags: What is posted might not reflect real-time situations, which can impact the association strength between the posted measures and real-world metrics; thus, we suggest aligning the content topics of health communication with up-to-date epidemic outcome metrics.
2. Discrepancies in posting methods: The CDC simplifies the data in their posts to make the information more comprehensible for the audience, which might not align with the detailed epidemic metrics posted from other sources with different interpretations of the reported metrics.
3. Variability in the data source: The data open to the public might come from sources and reporting standards that are different from the CDC's protocol, which could weaken potential associations.
4. Audience: As a government health agency, the CDC prioritizes certain data for social media to cater to the public for relevancy. For example, posting daily epidemic measures could lead to strong associations with COVID-19 metrics, but an association does not mean causality, and we assume that the CDC does not generate their tweets with the intention to improve associations of any kind and their priority is to present a variety of reliable information to the public.
5. Fatigue from information overload: Frequent data updates on social media can lead to desensitization, making it less likely for users to pay attention and react to the information over time, for example, tweeting about daily epidemic measures decreases the public's attention over time.
6. Political and societal factors, for example, antivaccination groups and conspiracy theories about the pandemic.

In addition, it is important for us to continue to examine the validity of the underlying assumption that the CDC's health communication makes an impact during a pandemic. In this infodemiology study, we focused on the national effects of these tweets. Future studies should further examine geospatial factors and other confounding factors to help understand whether and how much the CDC's tweets impact pandemic outcomes.

Lastly, public engagement (ie, retweets, likes, replies, etc) of the CDC's health communication is an important indicator of the effectiveness of online health communication efforts. There have been studies that analyzed public sentiments and attitudes [31-34] toward various health-related topics. However, very few studies have investigated the associations of public sentiment shifts along disease-related metrics. In addition, public sentiments and attitudes are heavily influenced by health agencies' messages and should not be misled by misinformation. Public opinions also influence health practices and interventions, which have a significant impact on the actual epidemic outcomes (eg, case, death, vaccination, etc). Thus, it is important to further investigate the underlying association between public health communication topics and actual epidemic measures. The insights can help public health agencies develop better social media strategies to address public concerns at different stages of the emergency. Therefore, we suggest that examining the dynamics and patterns of public responses to health agencies' original communications can provide valuable insights on public perceptions and attitudes around various issues during the pandemic, such as pharmaceutical interventions (eg, vaccination) and nonpharmaceutical interventions. Detailed content analysis can be applied to explicitly identify public concerns in response to the CDC's health communications. In addition, sentiment analysis can be applied to extract public sentiments (ie, positive, neutral, or negative) toward the CDC's health communications, and further help identify public attitudes and reactions to various content topics that the CDC has communicated. Public attitudes will ultimately determine individual health behavior and decision-making, such as vaccination acceptance and compliance with nonpharmaceutical interventions, which in turn drive the overall epidemic dynamics. Therefore, it is critical to investigate real-time public engagement, such as retweeting or replying on social media, toward public health agencies' communications to better inform health agencies about prioritizing their communications and addressing public concerns about specific content topics.

Conclusions

This study investigated the dynamic associations between the CDC's detailed COVID-19 communication topics on Twitter and epidemic metrics in the United States for almost 2 years during the pandemic. Using LDA topic modeling, we were the first to comprehensively identify and explore various COVID-related topics tweeted by the federal public health agency during the pandemic. We also collected daily COVID-19 epidemic metrics (confirmed case counts, death counts, completed tests records, and fully vaccinated records) and performed various multivariate time series analyses to unravel the temporal patterns and associations with the CDC's COVID-19 communication patterns (ie, investigated the dynamic associations between the time series of each topic generated by the LDA model and the time series of each epidemic metric). The results suggested that some topics were strongly associated with certain COVID-19 epidemic metrics, indicating that advanced social media analytics (eg, NLP) could be a valuable tool for effective infoveillance. Based on our findings, we recommend that the CDC, along with other public health agencies, could further optimize their health communications on social media platforms by posting contents and topics that align with the temporal dynamics of key epidemic metrics. While the CDC had been making efforts to share information on social media platforms to educate the public throughout the pandemic, the public responses to these messages were relatively weak. It is important to further explore the potential factors that played a role in the effectiveness of the CDC's social media performance in future studies. As such, we suggest increasing online health communication on health practices and interventions during high-level epidemic periods with large numbers of cases and deaths. Our findings also highlighted the importance of health communication on social media platforms to better respond to and tackle future health emergencies and issues.

Acknowledgments

We thank Naomi Nikita Thammadi, former graduate student of the University of North Carolina at Charlotte, who helped with data collection through the Twitter application programming interface and initial data preprocessing. This project was partially supported by the Models of Infectious Disease Agent Study (MIDAS) Network grant MIDASUP-05 to SC.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary information.
[PDF File (Adobe PDF File), 3131 KB - infodemiology_v4i1e49756_app1.pdf]

References

1. WHO Coronavirus (COVID-19) Dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2023-11-11]
2. The Health Communicator's Social Media Toolkit. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthcommunication/toolstemplates/socialmediatoolkit_bm.pdf [accessed 2023-11-11]
3. Yun GW, Morin D, Park S, Joa CY, Labbe B, Lim J, et al. Social media and flu: Media Twitter accounts as agenda setters. Int J Med Inform 2016 Jul;91:67-73. [doi: 10.1016/j.ijmedinf.2016.04.009] [Medline: 27185510]
4. 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 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
5. Liu Q, Zheng Z, Zheng J, Chen Q, Liu G, Chen S, et al. Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China: Digital Topic Modeling Approach. J Med Internet Res 2020 Apr 28;22(4):e19118 [FREE Full text] [doi: 10.2196/19118] [Medline: 32302966]
6. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public Discourse Against Masks in the COVID-19 Era: Infodemiology Study of Twitter Data. JMIR Public Health Surveill 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: 10.2196/26780] [Medline: 33720841]

7. 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]
8. Chen S, Xu Q, Buchenberger J, Bagavathi A, Fair G, Shaikh S, et al. Dynamics of Health Agency Response and Public Engagement in Public Health Emergency: A Case Study of CDC Tweeting Patterns During the 2016 Zika Epidemic. JMIR Public Health Surveill 2018 Nov 22;4(4):e10827 [FREE Full text] [doi: 10.2196/10827] [Medline: 30467106]
9. COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. GitHub. URL: https://github.com/CSSEGISandData/COVID-19/blob/master/README.md [accessed 2021-12-31]
10. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases 2020 May;20(5):533-534. [doi: 10.1016/s1473-3099(20)30120-1]
11. Dong E, Ratcliff J, Goyea TD, Katz A, Lau R, Ng TK, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: data collection process, challenges faced, and lessons learned. The Lancet Infectious Diseases 2022 Dec;22(12):e370-e376. [doi: 10.1016/s1473-3099(22)00434-0]
12. A Note on Topical N-grams. University of Massachusetts Amherst. URL: https://people.cs.umass.edu/~mccallum/papers/tng-tr05.pdf [accessed 2023-11-11]
13. Rajaraman A, Ullman JD. Data Mining. In: Mining of Massive Datasets. Cambridge, United Kingdom: Cambridge University Press; 2011.
14. Bafna P, Pramod D, Vaidya A. Document clustering: TF-IDF approach. 2016 Presented at: International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT); March 03-05, 2016; Chennai, India. [doi: 10.1109/ICEEOT.2016.7754750]
15. Perplexity. Wikipedia. URL: https://en.wikipedia.org/wiki/Perplexity [accessed 2023-11-11]
16. Stevens K, Kegelmeyer P, Andrzejewski D, Buttler D. Exploring topic coherence over many models and many topics. In: EMNLP-CoNLL '12: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. 2012 Presented at: 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning; July 12-14, 2012; Jeju Island, Korea. [doi: 10.5555/2390948.2391052]
17. Topic coherence pipeline. Gensim: topic modeling for humans. Radim Řehůřek. URL: https://radimrehurek.com/gensim/models/coherencemodel.html#gensim.models.coherencemodel.CoherenceModel [accessed 2021-12-31]
18. Tijare P, Jhansi Rani P. Exploring popular topic models. J. Phys.: Conf. Ser 2020 Dec 01;1706(1):012171. [doi: 10.1088/1742-6596/1706/1/012171]
19. Derrick TR, Thomas JM. Time-Series Analysis: The cross-correlation function. In: Stergiou N, editor. Innovative Analyses of Human Movement. Champaign, Illinois: Human Kinetics Publishers; 2004:189-205.
20. statsmodels.tsa.stattools.adfuller. statsmodels. URL: https://www.statsmodels.org/dev/generated/statsmodels.tsa.stattools.adfuller.html [accessed 2021-12-31]
21. Safarnejad L, Xu Q, Ge Y, Bagavathi A, Krishnan S, Chen S. Identifying Influential Factors in the Discussion Dynamics of Emerging Health Issues on Social Media: Computational Study. JMIR Public Health Surveill 2020 Jul 28;6(3):e17175 [FREE Full text] [doi: 10.2196/17175] [Medline: 32348275]
22. Time Series Measures. GitHub. URL: https://elife-asu.github.io/PyInform/timeseries.html [accessed 2021-12-31]
23. Non-seasonal ARIMA models. OTexts. URL: https://otexts.com/fpp2/non-seasonal-arima.html [accessed 2023-04-09]
24. ARIMA modelling in R. OTexts. URL: https://otexts.com/fpp2/arima-r.html [accessed 2023-04-11]
25. Rabbi F, Tareq SU, Islam MM, Chowdhury MA, Kashem MA. A Multivariate Time Series Approach for Forecasting of Electricity Demand in Bangladesh Using ARIMAX Model. 2020 Presented at: 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI); December 19-20, 2020; Dhaka, Bangladesh. [doi: 10.1109/STI50764.2020.9350326]
26. Yang M, Xie J, Mao P, Wang C, Ye Z. Application of the ARIMAX Model on Forecasting Freeway Traffic Flow. 2017 Presented at: 17th COTA International Conference of Transportation Professionals; July 7-9, 2017; Shanghai, China. [doi: 10.1061/9780784480915.061]
27. Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci. Model Dev 2022 Jul 19;15(14):5481-5487. [doi: 10.5194/gmd-15-5481-2022]
28. TMO Guide (3) – pyLDAvis. WE1S project. URL: https://we1s.ucsb.edu/research/we1s-tools-and-software/topic-model-observatory/tmo-guide/tmo-guide-pyldavis/ [accessed 2023-04-09]
29. Chuang J, Manning CD, Heer J. Termite: Visualization Techniques for Assessing Textual Topic Models. 2012 Presented at: International Working Conference on Advanced Visual Interfaces; May 21-25, 2012; Capri Island, Italy.
30. Sievert C, Shirley K. LDAvis: A method for visualizing and interpreting topics. 2014 Presented at: Workshop on Interactive Language Learning, Visualization, and Interfaces; June 2014; Baltimore, MA, USA. [doi: 10.3115/v1/W14-3110]
31. Mohamed Ridhwan K, Hargreaves CA. Leveraging Twitter data to understand public sentiment for the COVID‐19 outbreak in Singapore. International Journal of Information Management Data Insights 2021 Nov;1(2):100021 [FREE Full text] [doi: 10.1016/j.jjimei.2021.100021]

32. 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]
33. 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]
34. Shen L, Yao R, Zhang W, Evans R, Cao G, Zhang Z. Emotional Attitudes of Chinese Citizens on Social Distancing During the COVID-19 Outbreak: Analysis of Social Media Data. JMIR Med Inform 2021 Mar 16;9(3):e27079 [FREE Full text] [doi: 10.2196/27079] [Medline: 33724200]

Abbreviations

AIC: Akaike information criterion
AR: autoregressive
ARIMA: autoregressive integrated moving average
ARIMAX: autoregressive integrated moving average with external variable
BERT: bidirectional encoder representations from transformers
CCF: cross-correlation function
CDC: Centers for Disease Control and Prevention
dAIC: difference in Akaike information criterion
LDA: latent Dirichlet allocation
MA: moving average
MAE: mean absolute error
MI: mutual information
NLP: natural language processing
PCR: polymerase chain reaction
RMSE: root mean square error
TF-IDF: term frequency-inverse document frequency

Edited by T Mackey; submitted 09.06.23; peer-reviewed by W Zhang, R Sun, P Dobbs; comments to author 23.08.23; revised version received 02.10.23; accepted 14.10.23; published 23.01.24.

Please cite as:
Yin S, Chen S, Ge Y
Dynamic Associations Between Centers for Disease Control and Prevention Social Media Contents and Epidemic Measures During COVID-19: Infoveillance Study
JMIR Infodemiology 2024;4:e49756
URL: https://infodemiology.jmir.org/2024/1/e49756
doi:10.2196/49756
PMID:38261367

Original Paper

Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users

Ryuichiro Ueda[1], MHA; Feng Han[2], MHA; Hongjian Zhang[1], PhD; Tomohiro Aoki[1], MHA; Katsuhiko Ogasawara[1], MBA, PhD

[1]Faculty of Health Sciences, Hokkaido University, Sapporo, Japan
[2]Graduate School of Medicine, Hokkaido University, Sapporo, Japan

Corresponding Author:
Katsuhiko Ogasawara, MBA, PhD
Faculty of Health Sciences
Hokkaido University
N12-W5, Kita-ku
Sapporo, 060-0812
Japan
Phone: 81 11 706 3409
Email: oga@hs.hokudai.ac.jp

Related Article:

This is a corrected version. See correction statement: https://infodemiology.jmir.org/2024/1/e57880

Abstract

Background: The COVID-19 pandemic prompted global behavioral restrictions, impacting public mental health. Sentiment analysis, a tool for assessing individual and public emotions from text data, gained importance amid the pandemic. This study focuses on Japan's early public health interventions during COVID-19, utilizing sentiment analysis in infodemiology to gauge public sentiment on social media regarding these interventions.

Objective: This study aims to investigate shifts in public emotions and sentiments before and after the first state of emergency was declared in Japan. By analyzing both user-generated tweets and retweets, we aim to discern patterns in emotional responses during this critical period.

Methods: We conducted a day-by-day analysis of Twitter (now known as X) data using 4,894,009 tweets containing the keywords "corona," "COVID-19," and "new pneumonia" from March 23 to April 21, 2020, approximately 2 weeks before and after the first declaration of a state of emergency in Japan. We also processed tweet data into vectors for each word, employing the Fuzzy-C-Means (FCM) method, a type of cluster analysis, for the words in the sentiment dictionary. We set up 7 sentiment clusters (negative: anger, sadness, surprise, disgust; neutral: anxiety; positive: trust and joy) and conducted sentiment analysis of the tweet groups and retweet groups.

Results: The analysis revealed a mix of positive and negative sentiments, with "joy" significantly increasing in the retweet group after the state of emergency declaration. Negative emotions, such as "worry" and "disgust," were prevalent in both tweet and retweet groups. Furthermore, the retweet group had a tendency to share more negative content compared to the tweet group.

Conclusions: This study conducted sentiment analysis of Japanese tweets and retweets to explore public sentiments during the early stages of COVID-19 in Japan, spanning 2 weeks before and after the first state of emergency declaration. The analysis revealed a mix of positive (joy) and negative (anxiety, disgust) emotions. Notably, joy increased in the retweet group after the emergency declaration, but this group also tended to share more negative content than the tweet group. This study suggests that the state of emergency heightened positive sentiments due to expectations for infection prevention measures, yet negative information also gained traction. The findings propose the potential for further exploration through network analysis.

(JMIR Infodemiology 2024;4:e37881) doi:10.2196/37881

KEYWORDS

COVID-19; sentiment analysis; Twitter; infodemiology; NLP; Natural Language Processing

Introduction

Background

The COVID-19 outbreak that occurred in December 2019 in Wuhan City, Hubei Province, China, spread rapidly in other countries after January 2020. Lockdowns were implemented primarily in Europe after March 2020 as infection prevention measures. The use of lockdowns as a quarantine measure varied from country to country; however, in the United States, the United Kingdom, France, and other countries, strict measures to regulate behavior were implemented, such as curfews and total school closures, with penalties imposed for violations.

COVID-19 spread rapidly in Japan after the first infection was confirmed on January 16, 2020, with incidents such as the mass infection on the Diamond Princess cruise ship in early February [1]. On April 7, the Japanese government declared a state of emergency in 7 prefectures—Tokyo, Kanagawa, Saitama, Chiba, Osaka, Hyogo, and Fukuoka—owing to the rapid spread of the infection by mass infection in medical facilities and elsewhere [2]. Although the restrictions imposed by the emergency declaration (eg, requests to remain inside and limitations on large-scale events) were less enforceable than those imposed by the lockdown, they did result in a significant decrease in travel rates throughout Japan. However, previous studies have shown that such strong behavioral restrictions may have a negative psychological impact on the public [3]. The emergency declaration was extended to all prefectures, and the restrictions imposed by the emergency declaration were subsequently lifted on May 25. Table 1 summarizes the major developments in the early stages of the COVID-19 outbreak in Japan in chronological order.

Table 1. Japan's response to the initial spread of COVID-19.

Date	Events	References
2020/1/16	The first case of COVID-19 infection is confirmed in Kanagawa Prefecture, Japan.	[4]
2020/2/4	COVID-19 infection is confirmed in passengers on the Diamond Princess, a large cruise ship, returning to Hong Kong.	[1]
2020/2/27	The Japanese government requests the temporary closure of all elementary schools, junior high schools, and high schools in Japan from March 2 to spring break.	[5]
2020/3/10	The Japanese government declares the new coronavirus infection a historical emergency.	[6]
2020/3/13	The prime minister can now declare a "state of emergency."	[7]
2020/3/26	The prime minister also orders the establishment of a government task force based on the act on special measures.	[8]
2020/4/7	The Japanese government declares a state of emergency. Seven prefectures (Tokyo, Kanagawa, Saitama, Chiba, Osaka, Hyogo, and Fukuoka), including the Tokyo metropolitan area, are designated as target areas.	[9]
2020/4/16	An emergency declaration is extended to cover all prefectures until May 6.	[10]
2020/5/4	A decision is made to extend the period of the state of emergency until May 31.	[9]
2020/5/14	The Japanese government decides to lift the state of emergency for 39 prefectures, excluding 8 prefectures on special alert (Tokyo, Kanagawa, Saitama, Chiba, Hokkaido, Kyoto, Osaka, and Hyogo).	[9]
2020/5/21	The Japanese government decides to the lift state of emergency for Kyoto, Osaka, and Hyogo.	[9]
2020/5/25	The Japanese government decides to lift the state of emergency for all prefectures.	[9]

Prior Work in Infodemiology

Following the spread of COVID-19, social networking services (SNSs) were used to transmit information about the virus, accelerating activity in the field of infodemiology, which utilizes this data. Infodemiology is a relatively new research field that combines health informatics and public health with data analysis. It is a scientific discipline that studies the distribution of information and its determinants in information media, particularly the internet, to provide reliable information on public health [11]. Infodemiology became widely known after the World Health Organization (WHO) used the term at the first WHO Infodemiology Conference in response to the spread of COVID-19 and stated the need to promote research activities in this field worldwide [12]. In a previous study, Su et al [13] used sentimental analysis of text information from SNS data to reflect public concerns and psychological changes in individuals, providing information to promote public health. In particular, a sentiment analysis of the Italian region of Lombardy, where the lockdown was enforced, indicated that the number of SNS users with feelings of "anxiety" decreased after the lockdown. In addition, Heras-Pedrosa et al [14] observed through sentiment analysis that "anxiety" and "anger" toward government policies were the top feelings in Spain in the early stages of the infection. Furthermore, in Japan, Niu et al [15] conducted a sentiment analysis from SNS text data on the reasons for the delay in COVID-19 vaccine uptake compared to other countries, suggesting that concerns about side effects may have outweighed the fear of infection in the initial vaccination process. Thus, social media–based analysis reflects the psychological changes in individuals and enables the provision of real-time information to the government enacting public health policies and infection prevention measures.

SNS Usage in Japan

The importance of social media has been increasing in Japan as well, with social media being utilized in public health countermeasures against recent pandemics. The usage rate of SNSs in Japan is still on the rise, with the Ministry of Internal Affairs and Communications' 2020 Survey on Communications

Usage Trends [16] showing that the percentage of people using SNSs was 73.8%, an increase of 4.8% from the previous year. It also points out that the growth is particularly large in the age groups comprising people 19 years and below and 60 years and above, indicating that the usage rate of SNSs by age group is increasing for all generations. In terms of the purpose of use, the second-highest percentage of respondents chose “to search for information I want to know,” followed by “to communicate with acquaintances,” suggesting that social media is used by all generations in Japan as an important means of obtaining information. However, while the research field of infodemiology is being actively promoted, there are limited reports on infodemiology in Japan, even though social media is used by a wide range of generations.

Study Purpose

In this study, we investigated psychological changes in individuals after the initial spread of COVID-19 in Japan and public sentiment changes following state-of-emergency declarations by conducting sentiment analysis using SNS data in infodemiology.

Methods

Research Data

We extracted geocoded Twitter data using “Nazuki no Oto,” a service provided by NTT Data Corporation [17]. The target period was from midnight on March 23, 2020, 2 weeks before the first declaration of a state of emergency in Japan, to April 21, 2020. We selected tweets containing the keywords “コロナ(corona),” “COVID-19,” and “新型肺炎 (new pneumonia)” by random sampling of 4,997,353 tweets. In addition, the data used in this study include retweets, a function that allows users to repost other users’ or their own tweets. Duplicate tweets were removed from the Twitter data extracted for this study, and only unique Twitter data were used.

Data Preprocessing

Before conducting the sentiment analysis on the extracted Twitter data, we preprocessed the data. For preprocessing, we deleted Twitter data that contained symbols that could not be analyzed by morphological analysis, hashtags (eg, #COVID-19), and URLs only. Consequently, a total of 4,965,100 tweets were used as the target data for sentiment analysis.

Morphological Analysis

In contrast to structured and quantitative data, which can be easily analyzed by a computer, qualitative text data, which are often used in sentiment analysis, require processing to extract the data objectively. Therefore, unstructured data are analyzed to convert them from qualitative to quantitative data. However, thus far, analyzing qualitative data in Japanese has been considered a difficult task. One reason for this is that Japanese grammar is more complex than English and other languages [18]. However, with the recent development of natural language processing, it is possible to separate sentences naturally and convert them into quantitative data on a practical level by preparing Japanese dictionary functions for Japanese text data. Morphological analysis determines the smallest grammatically meaningful unit that constitutes a sentence by demarcating the boundaries of words and phrases in the text data. Following decomposition, the part of speech and the type of conjugation are determined by referring to a registered dictionary. In this study, we used a morphological analyzer, MeCab (version 0.996; Kyoto University).

The International Phonetic Alphabet (IPA) dictionary, integrated within the Japanese morphological analysis system Chasen, is widely used for performing morphological analysis in MeCab [19]. However, conventional IPA dictionaries are limited in their ability to support conventional Japanese words and phrases and do not support neologisms and phrases unique to Japanese. To solve this problem, a new system dictionary called mecab-ipadic-NEologd was introduced [20]. This dictionary is updated every Monday and Sunday and can be automatically updated and registered from websites, such as news sites and social media. Therefore, the dictionary can handle text data on the web where unique expressions and new words are frequently used. In this study, we registered mecab-ipadic-NEologd and performed morphological analysis on text data from the SNS Twitter because many unique expressions and new words are used there.

Japanese Sentiment Dictionary

We utilized the Japanese Linguistic Inquiry and Word Count (JIWC) dictionary (Nara Institute of Science and Technology) for the sentiment analysis, employing cloud sourcing to access the latest corpus. This Japanese emotional dictionary was used for determining emotions in sentiment analysis, encompassing 7 categories: “anger,” “concern,” “disgust,” “sadness,” “surprise,” “trust,” and “joy” [21]. Examples of words in the Japanese emotion expression dictionary are shown in Table 2. Among the emotions, “trust” and “joy” were selected as positive emotions, and “anger,” “anxiety,” “disgust,” and “sadness” were selected as negative emotions based on previous studies [22].

Table 2. Examples of words included in the JIWC[a] dictionary.

Sentiment	Examples of words
Anger	怒った (angry), 怒り(rage), 悪い(bad), 嫌がらせ(harassment), イライラ(irritation), うるさい (noisy), ゴミ (garbage), 暴言(rant), 煽り (aggravation), 理不尽な (unreasonable), 騒音 (noise), 迷惑 (annoyance), 被害 (damage), 虐待(abuse), 裏切り (betrayal)
Anxiety	不安(anxious), 不安だ (worrying), 不安な(anxiety), 病 (illness), 症状(symptom), このまま (at this rate), この先 (from now on), 考える (thinking)
Disgust	嫌いな (dislike), 嫌がらせ (harass), 嫌な (disgust), うるさい (loud), テロ (terror), 犯罪 (crime), 犯人 (criminal), ひどい (terrible), 悪 (evil), 悪かった (bad), 批判 (criticize), 無い (no), 無し (none), 無視 (ignore), 嘘 (lie), 汚い (dirty)
Sad	悲しい (sad), 悲観 (pessimistic), 悲愁 (melancholy), 哀感 (sorrowful), 哀傷 (piteous), 泣き (weeping), 泣き叫ぶ (wailing), 嘆き (lamenting), 涙 (tears), 涙声 (tearful), 追悼 (mourning), 痛嘆 (painful), センチメンタル (sentimental)
Surprise	いきなり (suddenly), サプライズ (surprise), びっくり (surprised), 偶然 (accidentally), 知った (learned), 知って (knew), 解散 (dissolved), 詐欺 (fraud), 発見 (discovered)
Trust	仲間 (companion), 任せ (entrust), 依頼 (request), 信用 (trust), 頼り (rely), 頼んで (ask), 助け (help), 守って (protect), 親友 (friend), 親身に (friendly), 関係 (relationship), サポート (support), フォロー (follow)
Joy	遊び (play), 遊んで (playing), 楽しい (fun), 出かけた (went out), おいしい (delicious), 食事 (meal), できた (could), 会って (meet), 会話 (conversation), 笑い (laugh), 笑顔 (smile), 好きな (like)

[a]JIWC: Japanese Linguistic Inquiry and Word Count.

Data Clustering

The sentiment analysis conducted in this study involved determining emotions in Twitter data by comparing the words in the text with those found in the JIWC dictionary. However, since the words after the morphological analysis were unstructured data, it was not possible to perform numerical calculations to assess their similarity to the words in the dictionary. To address this issue, we used Word2Vec processing to vectorize the text data for both Twitter data and the Japanese emotional dictionary.

Word2Vec is a model proposed by Mikolov et al [23,24] that represents word meanings using low-dimensional vectors, enabling semantic calculations in natural language processing. When vectorizing a large amount of text data, as in this study, individually vectorizing each word can result in an enormous number of dimensions, making it impractical in terms of computation time. Therefore, Word2Vec enables the vectorization of large text data through an inference-based approach using neural networks. Inference-based methods involve making predictions about what goes into a word when given its context (the surrounding words in a sentence). For example, when given the sentence “You ??? goodbye, and I say hello,” we can easily infer that the missing word is “say.” In this case, the context for “???” consists of 2 words: “you” and “goodbye.” The challenge is to infer what fits into that word based on the surrounding context, and thus learn word occurrence patterns. This approach is based on the distributional hypothesis, which suggests that word meanings are formed by the context of the surrounding words rather than inherent in the words themselves. Word2Vec includes 2 models, namely, the continuous bag-of-words (CBOW) model and the skip-gram model, to solve this inference issue. Generally, the skip-gram model is considered to have higher model accuracy after training, but it incurs higher computational costs since it needs to calculate losses for each context. This study’s text data comprises millions of individual pieces, and due to the added morphological analysis, a higher number of words per sentence was anticipated. Therefore, we anticipated that the computational cost for predictions would become immense. As a result, we employed the CBOW model for word embedding processing. After the data collected from Twitter and the terms registered from each Japanese sentiment dictionary were vectorized, Fuzzy-C-Means (FCM) was used to cluster each of the 7 sentiments.

The FCM method is a nonhierarchical soft clustering technique based on fuzzy logic theory. Fuzzy logic theory, originating from the concept of fuzzy sets proposed by LA Zadeh in 1965, provides a framework for quantitatively handling uncertainty and ambiguity in human subjective thinking and decision-making. FCM is a soft clustering method that applies fuzzy logic theory to cluster data [25]. In traditional hard clustering, data are assigned to clusters by being represented as either belonging (1) or not belonging (0) to a specific cluster. In contrast, because FCM is a soft clustering method, it allows data to partially belong to multiple clusters, such as 0.8 belonging to one cluster and 0.2 belonging to another. FCM clustering is carried out using the following algorithm. The membership values, representing the degree to which data points belong to different clusters, are considered:

In this case, the following conditions are satisfied:

The matrix *U*, denoted as [*u_it*], represents an n×c matrix with the membership value *u_it* as an element. Meanwhile, the matrix *V*, represented as [*v_t*], is an n×c matrix with cluster center *v_t* as an element.

Bezdek proposed the following formula for the FCM model that minimizes the objective function by the weighted sum of the Euclidean squared distances between each data and the center of each cluster under the condition of (1) [26]:

Here, *m* is a fuzzy coefficient parameter (*m*>1) that adjusts the strength of ambiguity. When *m*=1, the FCM model corresponds to the hard clustering k-means model. In this case, the objective function *J(U, V)* is linearized with respect to *u_it*, eliminating soft clustering. FCM clustering is carried out through the following steps. First, given a data set {x_1, , x_n}, we determine the number of clusters *t* (2 ≤ t ≤ c) and the parameter m ∈ (1, ∞). Next, we initialize the membership values *u_it* with U^(0) = {u_it^0} randomly. We provide a sufficiently small positive number ε to determine the termination of the loop. Second, we use the current membership values u_it to calculate the cluster centers v_t^p using the following formula:

Third, we update the membership values from *u_it^p* to *u_it^*(*p*+1) using the following formula:

Finally, if ‖ *u_it^*(*p*+1) - *u_it^p* ‖ < ε holds for all *i* and *t*, we terminate the loop. Otherwise, we increment *p* by 1 and return to the second step. Once the loop terminates, we obtain the center points for each cluster and the membership values for each sample data, completing the clustering process. In this study, FCM was used on text data to reduce the number of words included in an emotion dictionary and construct the emotion dictionary, allowing for more accurate sentiment analysis of the text data due to the influence of a single word on multiple emotions. Both tweets and retweets of Twitter data were used, and quoted retweets, which are retweets of others' posts with additional comments, were also included.

After vectorization using Word2Vec and clustering using the FCM method, the distance between the vector coordinates of each tweet and the center-of-gravity vector of each written sentiment was calculated. Next, the value with the shortest vector distance was determined as the sentiment of that tweet. The entire sentiment analysis in this study was performed using the Python programming language (version 3.9.4). A path diagram of the overall sentiment analysis is shown in Figure 1, and a summary diagram of the sentiment determination method is shown in Figure 2.

Figure 1. Sentiment analysis flowchart.

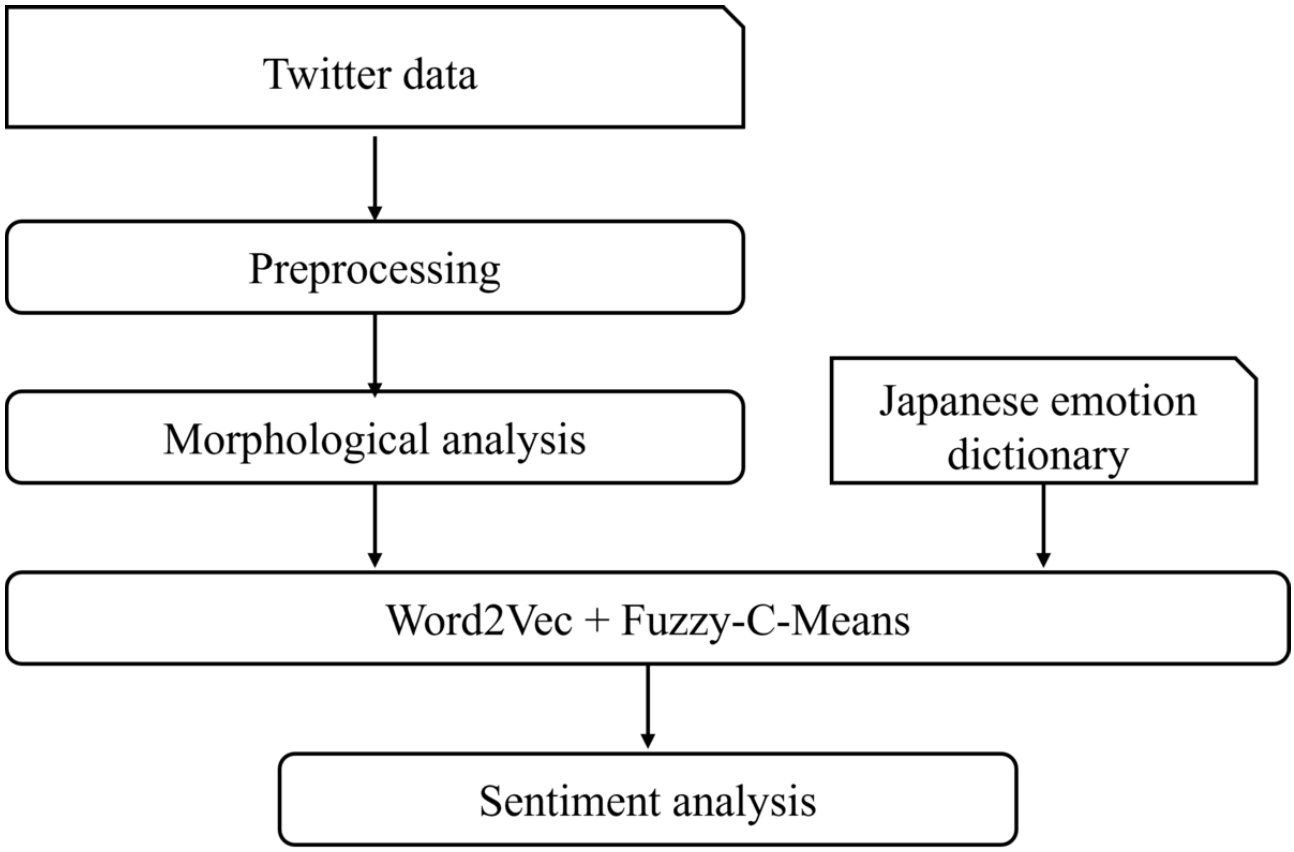
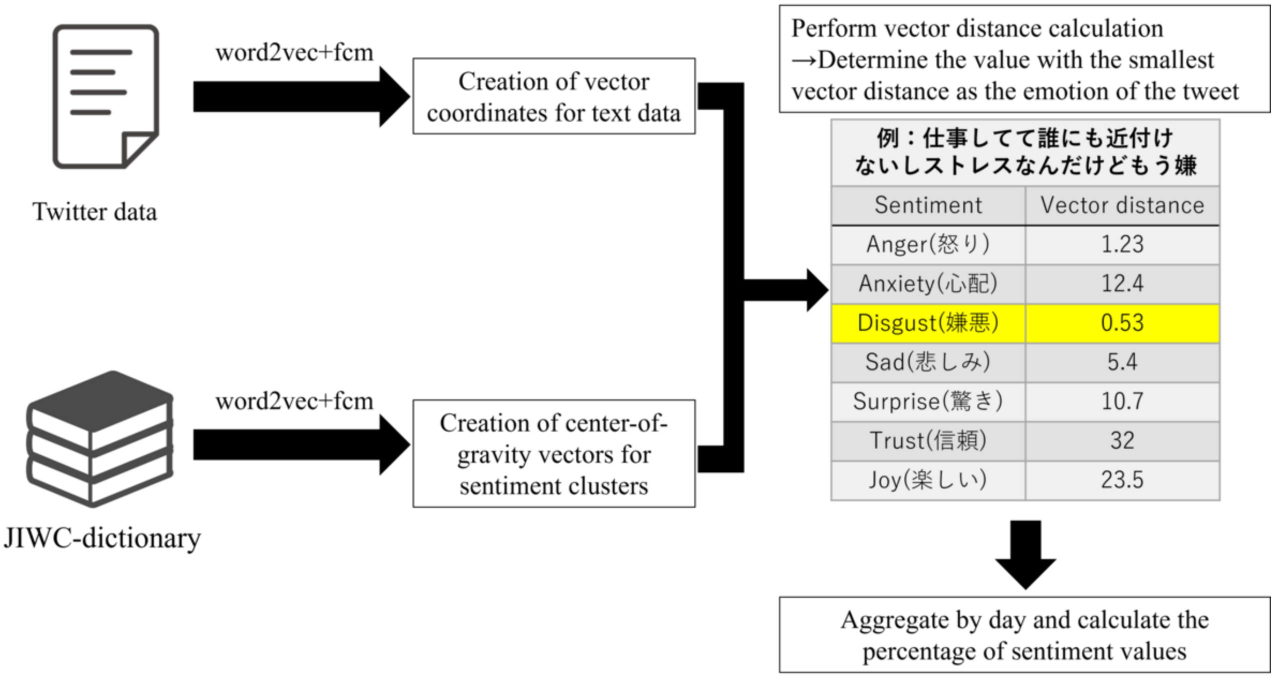


Figure 2. Diagram of the sentiment determination method. FCM: Fuzzy-C-Means; JIWC: Japanese Linguistic Inquiry and Word Count.



Examining Sentiment Changes Before and After the State of Emergency Declaration

The Twitter data were categorized into 2 groups: the tweet group and the retweet group. The study period was divided into “before the declaration of a state of emergency,” which ranged from midnight on March 23, 2020, until PM 11:59:59 on April 6, 2020, and “after the declaration of a state of emergency,” which ranged from midnight on April 7, 2020, to PM 11:59:59 until April 21, 2020. We calculated the proportion of emotions before and after the declaration of a state of emergency in both the tweet and retweet groups. The sentiment analysis results were validated using 2 methods. The first method involved comparing emotions using a between-group comparison of 7 emotions over approximately 2 weeks before and after the declaration of a state of emergency. This comparison was based on daily average values for each emotion. The second method involved dividing the data into two groups: (1) the tweet group, consisting of posts made by the users themselves, and (2) the retweet group, consisting of posts shared for the purpose of dissemination. Sentiment analysis results were aggregated daily, classifying the data as either positive (“trust” and “joy”) or negative (“anger,” “concern,” “disgust,” and “sadness”) and then comparing the tweet and retweet groups. Both methods conducted a median difference examination using the Mann-Whitney U test, with statistical significance set at $P<.05$, utilizing the statistical software JMP (version 16.0; SAS).

Ethical Considerations

This study was conducted while adhering to strict ethical considerations and did not require ethics approval. To avoid identification of personal information, the Twitter data used were limited to the type of post (tweet or retweet), text, and the date and time of the post for data analysis. The data used did not contain any personally identifiable information. In addition, efforts were made to ensure transparency throughout the design and conduct of this study.

Results

Research Data

We were able to judge sentiment through the sentiment analysis in 4,884,297 (97.74%) cases out of a total of 4,997,353 cases. In addition, the number of tweets was 1,374,025 (28.13%), and the number of retweets was 3,510,272 (71.87%). The number of tweets and retweets per day is shown in Table 3, and the daily trends for the data from March 23, 2020, to April 21, 2020, are shown in Multimedia Appendix 1.

Table 3. Daily tweet and retweet counts.

Date	Tweets (n=1,374,025), n (%)	Retweets (n=3,510,272), n (%)
2020/3/23	4666 (0.34)	13,643 (0.39)
2020/3/24	25,067 (1.82)	71,040 (2.02)
2020/3/25	33,759 (2.46)	87,476 (2.49)
2020/3/26	41,944 (3.05)	115,842 (3.30)
2020/3/27	39,433 (2.87)	103,798 (2.96)
2020/3/28	37,160 (2.70)	106,915 (3.05)
2020/3/29	37,804 (2.75)	108,868 (3.10)
2020/3/30	74,353 (5.41)	209,297 (5.96)
2020/3/31	51,765 (3.77)	144,594 (4.12)
2020/4/1	48,902 (3.56)	121,864 (3.47)
2020/4/2	48,127 (3.50)	119,259 (3.40)
2020/4/3	52,918 (3.85)	123,835 (3.53)
2020/4/4	48,470 (3.53)	113,346 (3.23)
2020/4/5	54,358 (3.96)	115,172 (3.28)
2020/4/6	75,831 (5.52)	175,918 (5.01)
2020/4/7	76,184 (5.54)	195,158 (5.56)
2020/4/8	60,645 (4.41)	179,707 (5.12)
2020/4/9	55,231 (4.02)	156,760 (4.47)
2020/4/10	51,078 (3.72)	134,393 (3.83)
2020/4/11	44,901 (3.27)	111,213 (3.17)
2020/4/12	42,403 (3.09)	96,575 (2.75)
2020/4/13	42,117 (3.07)	107,539 (3.06)
2020/4/14	42,800 (3.11)	105,344 (3)
2020/4/15	44,185 (3.22)	118,456 (3.37)
2020/4/16	48,618 (3.54)	122,458 (3.49)
2020/4/17	44,494 (3.24)	132,009 (3.76)
2020/4/18	38,270 (2.79)	111,351 (3.17)
2020/4/19	38,872 (2.83)	110,308 (3.14)
2020/4/20	39,611 (2.88)	116,187 (3.31)
2020/4/21	30,059 (2.19)	78,522 (2.24)

Percentage of Emotions in the Sentiment Analysis

The results of the sentiment analysis on the tweet and retweet groups for the period between midnight on March 23, 2020, to 23:59:59 on April 6, 2020 (before the declaration of the state of emergency) are shown in Figure 3. The results for the period between midnight on April 7, 2020, and 23:59:59 on April 21, 2020 (after the declaration of the state of emergency) are shown in Figure 4. In the tweet group, the positive emotion "joy" was highest both before and after the state of emergency declaration at 40.5% (n=272,879) and 31% (n=217,074), respectively, while in the retweet group, the negative sentiment of "worry" was 34% (n=587,540), and "disgust" was 18.6% (n=322,462) during the period before the state of emergency declaration. These percentages were higher than those for the other emotions.

Figure 3. Sentiment analysis value ratio in the tweet group.

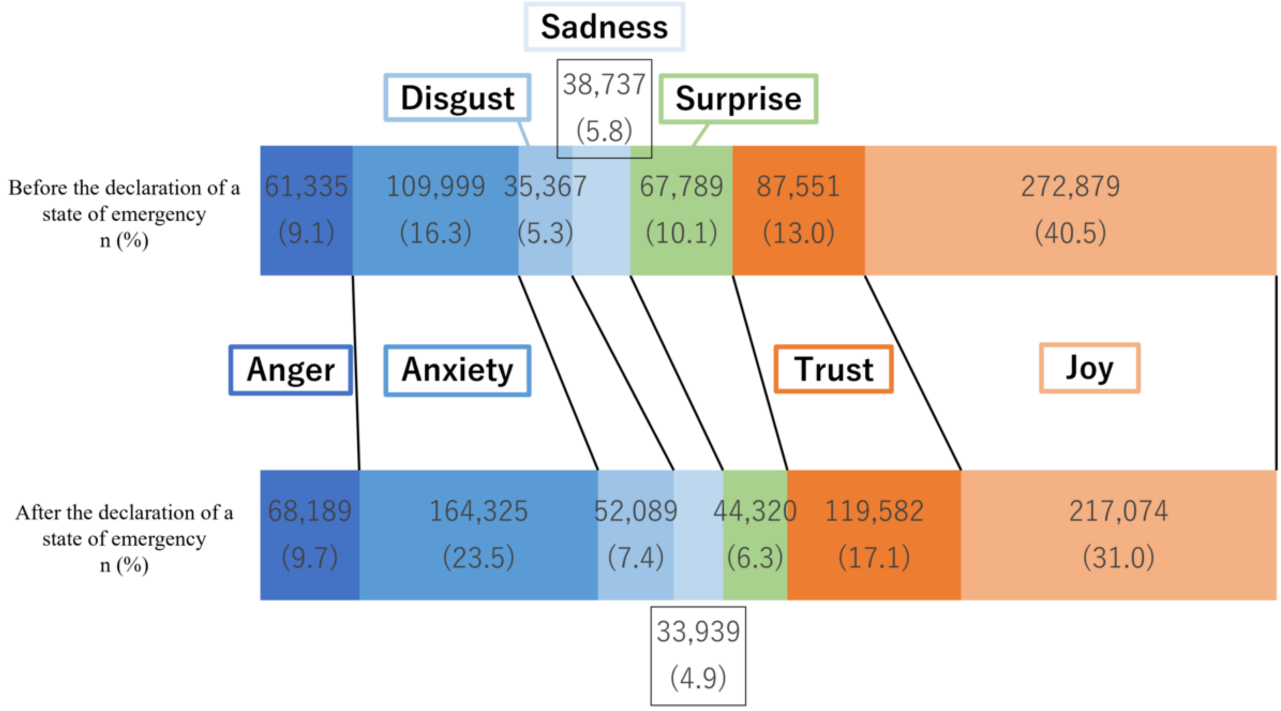
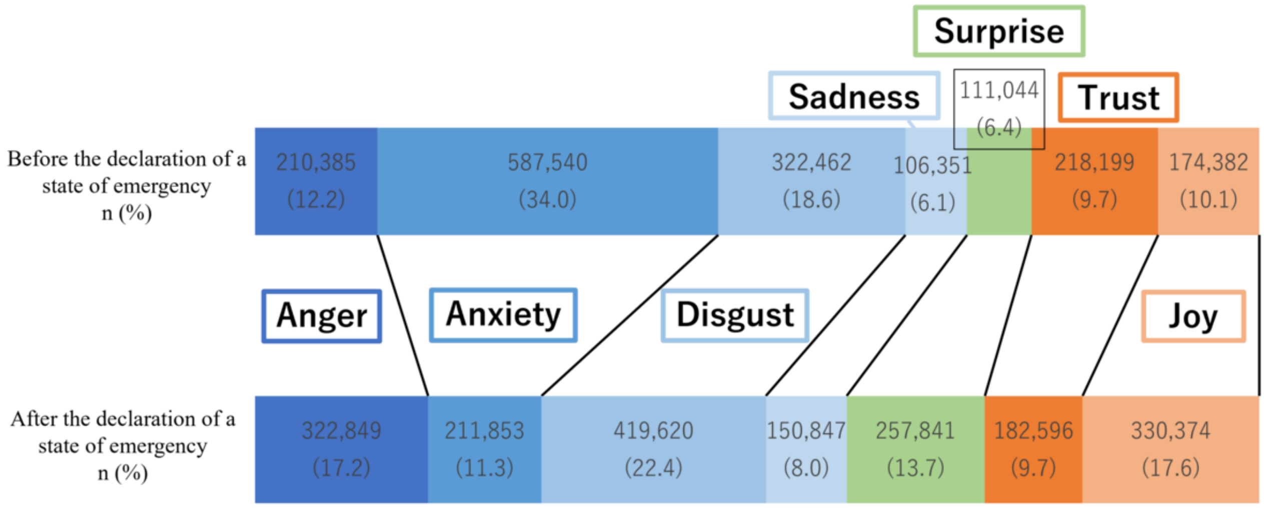


Figure 4. Sentiment analysis value ratio in the retweet group.



Changes in Sentiment Before and After the Declaration of a State of Emergency

Table 4 shows the results of the sentiment analysis yielding the proportions of the 7 emotion types before and after the declaration of the state of emergency. The Mann-Whitney U test comparison of differences in median values revealed that the sentiment of joy significantly increased in the retweet group (*P*<0.05). However, no significant differences were observed for the other emotions.

Table 5 and Figure 5 show the results of testing the change of positive and negative content between the tweet group and retweet groups. In the 2 weeks before and after the emergency declaration, the retweet group tended to post more negative content than the tweet group (before *r*=0.29, *P*=.02; after *r*=.0.40, *P*=.002). However, there was no difference between the tweet and retweet groups in the percentage of positive responses.

Table 4. Sentiment changes before and after the state of emergency declaration[a].

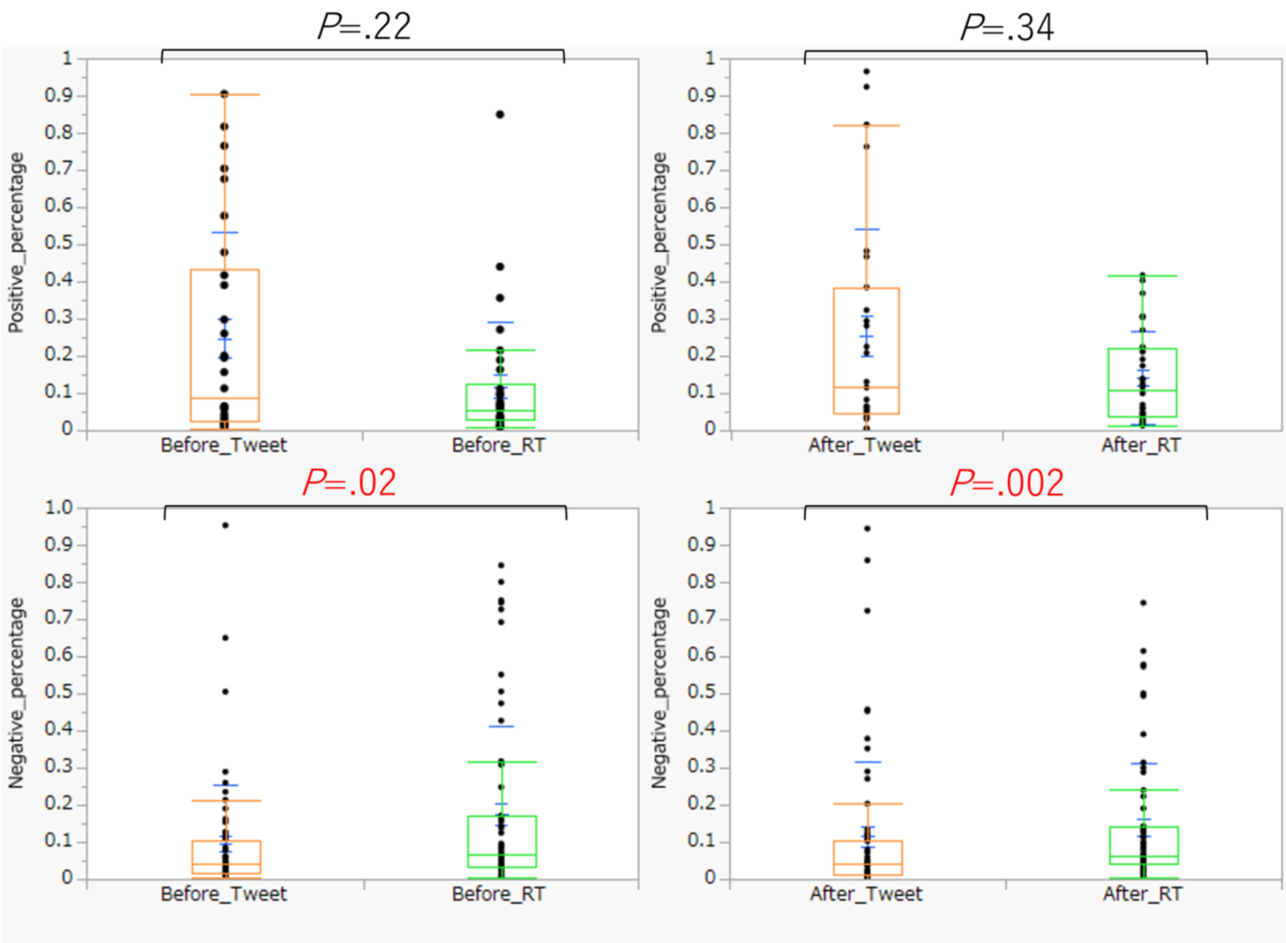
Sentiments	Before (n=15)		After (n=15)		*P* value
	Median	SD	Median	SD	
Anger tweet	0.042	0.061	0.024	0.082	.80
Anger retweet	0.042	0.051	0.063	0.035	.84
Anxiety tweet	0.063	0.078	0.050	0.135	.90
Anxiety retweet	0.210	0.307	0.054	0.052	.25
Disgust tweet	0.023	0.293	0.021	0.020	.59
Disgust retweet	0.073	0.136	0.127	0.100	.43
Sadness tweet	0.041	0.025	0.035	0.025	.51
Sadness retweet	0.041	0.045	0.055	0.025	.28
Surprise tweet	0.038	0.090	0.016	0.022	.16
Surprise retweet	0.051	0.023	0.035	0.104	.93
Trust tweet	0.035	0.033	0.038	0.056	.32
Trust retweet	0.059	0.032	0.061	0.050	.80
Joy tweet	0.390	0.258	0.281	0.191	.62
Joy retweet	0.041	0.057	0.191	0.097	.04

[a]Before refers to the period from midnight on March 23, 2020, until 11:59:59 PM on April 6, 2020, while after refers to the period from midnight on April 7, 2020, until 11:59:59 PM on April 21, 2020.

Table 5. Comparison results of positive and negative changes between the tweet and retweet groups.

Sentiment	Tweet		Retweet		*P* value
	Median	SD	Median	SD	
Positive (n=30)					
Before	0.089	0.286	0.054	0.174	.22
After	0.099	0.290	0.108	0.123	.34
Negative (n=60)					
Before	0.040	0.161	0.066	0.236	.02
After	0.038	0.202	0.063	0.173	.002

Figure 5. Graphs displaying the results from a comparative study illustrating changes in positive and negative sentiments between the tweet and retweet groups. RT: retweet.



Discussion

Principal Findings

The Japanese language sentiment analysis conducted during this study's target period, both before and after the declaration of the state of emergency, revealed that "joy," associated with a "positive" sentiment, accounted for high proportions within the tweet group at 40.5% (n=272,879) before and 31% (n=217,074) after. On the other hand, "anxiety" and "disgust," which express "negative" feelings, accounted for high percentages in both the tweet and retweet groups, especially in the retweet group, where "anxiety" accounted for 34% (n=587,540) and "disgust" accounted for 18.6% (n=322,464) of the total retweets before the state of emergency was declared. The self-restraint approach regulating behavior during the declaration of a state of emergency in Japan allowed movement across prefectures. This may have been a contributing factor to the widespread negative posts related to movements from the target area. This surge in negative sentiment was countered by a simultaneous rise in positive emotions, attributed to the anticipation of infection prevention following the state of emergency declaration. During the early stages of the COVID-19 spread in other countries, a previous study on English-speaking users indicated elevated levels of positive emotions linked to anticipations for potential policies [22]. A generally similar emotional response was apparent among the public in other countries. In the early stages of the spread of infection, when no vaccine or other countermeasures had been implemented, feelings of anxiety may have been expressed on social media, as well as expectations for strong countermeasures, such as behavioral restrictions. In contrast, the results of the sentiment analysis of English-language tweets corresponding to the same period showed that negative and positive emotions accounted for approximately the same proportions by late March, the end of the period covered in this study. Notably, the negative emotion "fear" occupied a higher percentage than other emotions around January and February [27]. In China and European countries, the first cases of infection were confirmed earlier than in Japan (where the initial expansion of the outbreak occurred in late March). Thus, the earlier spread of infection in those nations may have a significant impact on the sentiment analysis.

Comparative Study Between the Tweet and Retweet Groups

When comparing the tweet and retweet groups, the retweet group tended to post more negative sentiments. In this regard, a previous study revealed that in the early stage of the COVID-19 outbreak among English-speaking users, many tweets had a positive sentiment, while many retweets had a negative sentiment [28]. It is clear that much of the information users wished to disseminate was negative in nature. As for the difference between groups in this study, there is a research report that focuses on virality, one of the characteristics of sentiment analysis using social media [29] Virality is an explosive spread of attention and information through social media and word-of-mouth on the internet. Virality is derived from "viral"—as in a virus. Previous research indicates that negative posts increase virality, while positive posts decrease virality. Therefore, for topics that attract substantial public attention, such as COVID-19, the topic of this study, there is a tendency to spread negative content in retweets, consequently increasing virality. This suggests a noteworthy contrast between the tweet and retweet groups.

Limitations

There are a few key limitations of this study. First, the social media platform Twitter, which was used for the sentiment analysis in this study, had an age bias. According to a survey conducted by the Ministry of Internal Affairs and Communications in 2020, the Twitter usage rate is highest among teenagers (67.6%) and twentysomethings (79.8%) [30].

Additionally, data from the Ministry of Internal Affairs and Communications indicate that the usage rate declines with increasing age, especially among individuals aged 40 years and older. This suggests that the younger generation is the predominant user of Twitter as a whole. This suggests that the younger generation predominantly constitutes the main users of Twitter overall. Therefore, the results of the sentiment analysis in this study are not necessarily representative of the entire nation. In addition, the Twitter data used in this study were limited to Japanese-language content. We did not use location-based information or conduct analyses based on geographical data. As such, this data may originate from disproportionate samples depending on the prefecture. During Japan's initial state of emergency declaration in 2020, the target areas comprised 7 prefectures: Tokyo, Kanagawa, Chiba, Saitama, Osaka, Kobe, and Fukuoka. Subsequently, on April 16, 2020, the target area was expanded to the entire country [9]. Throughout the study period covered, only some of the target areas were declared as emergency areas; therefore, emotional variations in Twitter usage may exist depending on the location of the users.

Second, the sentiment analysis categorized each tweet into one of 7 predefined sentiment types, limiting its ability to capture multiple sentiments, such as "anger" and "surprise," within a single tweet or account for cases where the selected sentiments might not apply.

The Twitter data utilized in this study underwent random sampling for both tweets and retweets. Twitter incorporates a function known as "bot," which automatically generates tweets in response to specified times and keywords. Numerous accounts, commonly referred to as "bot accounts," are responsible for automatic posting. Shi et al [31] conducted a sentiment analysis on Twitter focusing on the #coronavirus hashtag from January 2020 to March 2020, including human and bot-generated tweets. Their findings revealed that bot-posted tweets had more negative sentiments compared to those posted by humans concerning the topic of COVID-19. This suggests that the bot feature intentionally promotes negative public opinion and sentiment. Consequently, it is plausible that the inclusion of a substantial amount of data posted by bot accounts in this study may have influenced the results of the sentiment analysis. Unfortunately, we were unable to preprocess the data to account for this aspect. For our future research, we anticipate that carrying out a network analysis using the results of this study will provide a deeper understanding of the specific subjects that capture public interest. In terms of social network analysis, Seungil [32] investigated how Twitter users in the United States accessed COVID-19–related information based on their posted data. The investigation revealed that during the initial outbreak period, users expressed significant concerns about the number of infections Additionally, the study highlighted that users were more likely to obtain COVID-19 information from news channel accounts and the official accounts of the president. Sakun et al [22] conducted a network analysis to uncover topics associated with different emotions based on the results of a sentiment analysis using Twitter text data They found that words like "pneumonia," "influenza," "infectious disease," and "quarantine" were frequently linked to the emotion of "fear." In addition, words like "pandemic," "disease," and "hospital" were associated with the emotion "sadness." These results suggest that Twitter data can be used to understand the public's awareness of and emotions toward pandemics, providing valuable insights for governmental responses. Hence, the results of the sentiment analysis should be used for further exploration in infodemiology, specifically by conducting a network analysis focusing on the topics associated with each sentiment identified in this study.

Conclusions

In this study, we conducted a sentiment analysis using Japanese tweet and retweet text data spanning approximately 2 weeks before and after the first state of emergency declaration in Japan to assess public sentiments toward the initial spread of COVID-19. We observed a combination of positive sentiments ("joy") and negative sentiments ("anxiety" and "disgust") during the target period. The results of the Mann-Whitney U test indicated that feelings of joy significantly increased in the retweet group before and after the state of emergency declaration. However, there was a significant tendency for the retweet group to post more negative content compared to the tweet group. After the first state of emergency declaration, the anticipation regarding infection prevention measures due to this declaration contributed to an increase in positive sentiments. Moreover, it appears that information, including negative content, was more likely to be disseminated on the topic of COVID-19. Based on the results of this study, we believe that further development through network analysis is possible.

Acknowledgments

We express our gratitude to all individuals who cooperated in the progression of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Total number of tweets and retweets per day.
[PNG File , 180 KB - infodemiology_v4i1e37881_app1.png]

References

1. Yamazaki T. Epidemiology of the early stages of the case outbreak in the Diamond Princess novel coronavirus infection case. IASR. 2020 Jul 31. URL: https://www.niid.go.jp/niid/images/idsc/iasr/41/485.pdf [accessed 2024-01-05]
2. Declaration of a state of emergency in response to the novel coronavirus disease. Prime Minister's Office of Japan. 2020 Apr 16. URL: https://japan.kantei.go.jp/ongoingtopics/_00020.html [accessed 2023-11-09]
3. Soubhik B, Stefano C, Antonio D, Paolo F, Thiemo F, Stefano F. Evaluating COVID-19 public health messaging in Italy: Self-reported compliance and growing mental health concerns. medRxiv. Preprint posted online on April 5, 2020 [FREE Full text] [doi: 10.1101/2020.03.27.20042820]
4. Outbreak of a case of pneumonia associated with novel coronavirus (case 1). Ministry of Health, Labour and Welfare. 2020 Jan 16. URL: https://www.mhlw.go.jp/stf/newpage_08906.html [accessed 2023-11-09]
5. Task force on new coronavirus infections (15th meeting). Prime Minister's Office of Japan. 2020 Feb 27. URL: https://www.kantei.go.jp/jp/98_abe/actions/202002/27corona.html [accessed 2023-11-09]
6. Regarding the "historical emergency". Cabinet Office Minister's Secretariat Document Management Division. 2020 Mar 10. URL: https://www8.cao.go.jp/chosei/koubun/rekiren/rekishitekikinkyujitai.pdf [accessed 2024-01-05]
7. Influenza countermeasures special measures law. Ministry of Health, Labour and Welfare. URL: https://www.mhlw.go.jp/web/t_doc?dataId=78ab2871&dataType=0 [accessed 2024-01-05]
8. Task force on new coronavirus infections (23th meeting). Prime Minister's Office of Japan. URL: https://www.kantei.go.jp/jp/98_abe/actions/202003/26corona.html [accessed 2023-11-09]
9. Facts about novel coronavirus infection. Cabinet Office. URL: https://www8.cao.go.jp/shougai/whitepaper/r03hakusho/zenbun/h2_01_02.html [accessed 2023-11-09]
10. Task force on new coronavirus infections (29th meeting). Prime Minister's Office of Japan. URL: https://www.kantei.go.jp/jp/98_abe/actions/202004/16corona.html [accessed 2023-11-09]
11. Zielinski C. Infodemics and infodemiology: a short history, a long future. Rev Panam Salud Publica 2021 Feb 23;45:e40. [doi: 10.26633/rpsp.2021.40]
12. 1st WHO Infodemiology Conference. World Health Organization. URL: https://www.who.int/news-room/events/detail/2020/06/30/default-calendar/1st-who-infodemiology-conference [accessed 2023-11-09]
13. Su Y, Xue J, Liu X, Wu P, Chen J, Chen C, et al. Examining the impact of COVID-19 lockdown in Wuhan and Lombardy: a psycholinguistic analysis on Weibo and Twitter. Int J Environ Res Public Health 2020 Jun 24;17(12) [FREE Full text] [doi: 10.3390/ijerph17124552] [Medline: 32599811]
14. de Las Heras-Pedrosa C, Sánchez-Núñez P, Peláez JI. Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. Int J Environ Res Public Health 2020 Jul 31;17(15):5542 [FREE Full text] [doi: 10.3390/ijerph17155542] [Medline: 32751866]
15. 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]
16. White paper 2021: Information and Communications in Japan. Ministry of Internal Affairs and Communications. URL: https://www.soumu.go.jp/johotsusintokei/whitepaper/eng/WP2021/2021-index.html [accessed 2023-11-09]
17. Nazuki no Oto. NTT Data Corporation. URL: https://nazuki-oto.com/main/ [accessed 2023-11-09]
18. Tomoya M, Mamoru K. Why is the Japanese language so difficult? Analysis of a realistic corpus of learners of Japanese and issues in language processing. IPSJ 2012 May:217-223.
19. Takuya K, Kaoru Y, Yuji M. Applying conditional random fields to Japanese morphological analysis. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing. 2004 Presented at: 2004 Conference on Empirical Methods in Natural Language Processing; July 25-26; Barcelona, Spain p. 230-237 URL: https://aclanthology.org/W04-3230
20. Sato T. Neologism dictionary based on language resources on the web for Mecab. GitHub. 2015. URL: https://github.com/neologd/mecab-ipadic-neologd [accessed 2023-11-09]
21. JIWC dictionary. GitHub. URL: https://github.com/sociocom/JIWC-Dictionary [accessed 2023-11-09]
22. 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 [FREE Full text] [doi: 10.2196/21978] [Medline: 33108310]
23. Tomas M, Kai C, Greg C, Jeffrey D. Efficient estimation of word representations in vector space. arXiv. Preprint posted online on January 16, 2013 [FREE Full text] [doi: 10.48550/arXiv.1301.3781]

24. Tomas M, Ilya S, Kai C, Greg C, Jeffrey D. Distributed representations of words and phrases and their compositionality. arXiv. Preprint posted online on October 16, 2013 [FREE Full text]
25. Zadeh L. Fuzzy sets. Information and Control 1965 Jun;8(3):338-353. [doi: 10.1016/s0019-9958(65)90241-x]
26. Bezdek JC. Pattern recognition with fuzzy objective function algorithms. In: Advanced Applications in Pattern Recognition. NY: Springer New York; 1981:43.
27. 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]
28. 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]
29. Jiménez-Zafra SM, Sáez-Castillo AJ, Conde-Sánchez A, Martín-Valdivia MT. How do sentiments affect virality on Twitter? R Soc Open Sci 2021 Apr 14;8(4):201756 [FREE Full text] [doi: 10.1098/rsos.201756] [Medline: 33996120]
30. 2020 Survey on information and communication media usage time and information behavior. Ministry of Internal Affairs and Communications. URL: https://www.soumu.go.jp/main_content/000765258.pdf [accessed 2023-11-10]
31. Shi W, Liu D, Yang J, Zhang J, Wen S, Su J. Social bots' sentiment engagement in health emergencies: a topic-based analysis of the COVID-19 pandemic discussions on Twitter. Int J Environ Res Public Health 2020 Nov 23;17(22):8701 [FREE Full text] [doi: 10.3390/ijerph17228701] [Medline: 33238567]
32. Yum S. Social network analysis for coronavirus (COVID-19) in the United States. Soc Sci Q 2020 Jul 28;101(4):1642-1647 [FREE Full text] [doi: 10.1111/ssqu.12808] [Medline: 32836475]

Abbreviations

CBOW: continuous bag-of-words
FCM: Fuzzy-C-Means
IPA: International Phonetic Alphabet
JIWC: Japanese Linguistic Inquiry and Word Count
SNS: social networking service
WHO: World Health Organization

Edited by T Mackey; submitted 11.03.22; peer-reviewed by J Chen, S Wei; comments to author 14.10.23; revised version received 20.11.23; accepted 17.12.23; published 06.02.24.

Please cite as:
Ueda R, Han F, Zhang H, Aoki T, Ogasawara K
Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users
JMIR Infodemiology 2024;4:e37881
URL: https://infodemiology.jmir.org/2024/1/e37881
doi:10.2196/37881
PMID:38127840

Original Paper

Evaluating the Disease-Related Experiences of TikTok Users With Lupus Erythematosus: Qualitative and Content Analysis

Lindsey J Wanberg[1], BA; David R Pearson[1,2], MD

[1]University of Minnesota Medical School, Minneapolis, MN, United States

[2]Department of Dermatology, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
David R Pearson, MD
University of Minnesota Medical School
516 Delaware St SE
Minneapolis, MN, MMC 98
United States
Phone: 1 612 626 8625
Email: pearsond@umn.edu

Abstract

Background: Lupus erythematosus (LE) is an autoimmune condition that is associated with significant detriments to quality of life and daily functioning. TikTok, a popular social networking platform for sharing short videos, provides a unique opportunity to understand experiences with LE within a nonclinical sample, a population that is understudied in LE research. This is the first qualitative study that explores LE experiences using the TikTok platform.

Objective: This study aims to evaluate the disease-related experiences of TikTok users with LE using qualitative and content analysis.

Methods: TikTok videos were included if the hashtags included #lupus, were downloadable, were in English, and involved the personal experience of an individual with LE. A codebook was developed using a standardized inductive approach of iterative coding until saturation was reached. NVivo (Lumivero), a qualitative analysis software platform, was used to code videos and perform content analysis. Inductive thematic analysis was used to derive themes from the data.

Results: A total of 153 TikTok videos met the inclusion criteria. The most common codes were *experiences with symptoms* (106/153, 69.3%), *mucocutaneous symptoms* (61/153, 39.9%), and *experiences with treatment* (59/153, 38.6%). *Experiences with symptoms* and *mucocutaneous symptoms* had the greatest cumulative views (25,381,074 and 14,879,109 views, respectively). Five thematic conclusions were derived from the data: (1) mucocutaneous symptoms had profound effects on the mental health and body image of TikTok users with LE; (2) TikTok users' negative experiences with health care workers were often derived from diagnostic delays and perceptions of "medical gaslighting"; (3) TikTok users tended to portray pharmacologic and nonpharmacologic interventions, such as diet and naturopathic remedies, positively, whereas pharmacologic treatments were portrayed negatively or referred to as "chemotherapy"; (4) LE symptoms, particularly musculoskeletal symptoms and fatigue, interfered with users' daily functioning; and (5) although TikTok users frequently had strong support systems, feelings of isolation were often attributed to battling an "invisible illness."

Conclusions: This study demonstrates that social media can provide important, clinically relevant information for health practitioners caring for patients with chronic conditions such as LE. As mucocutaneous symptoms were the predominant drivers of distress in our sample, the treatment of hair loss and rash is vital in this population. However, pharmacologic therapies were often depicted negatively, reinforcing the significance of discussions on the safety and effectiveness of these treatments. In addition, while TikTok users demonstrated robust support systems, feelings of having an "invisible illness" and "medical gaslighting" dominated negative interactions with others. This underscores the importance of providing validation in clinical interactions.

(JMIR Infodemiology 2024;4:e51211) doi:10.2196/51211

KEYWORDS

lupus; TikTok; autoimmune disease; qualitative research; quality of life

Introduction

Background

The term lupus erythematosus (LE) encompasses a group of autoimmune disorders that may have multiorgan involvement, as in systemic LE (SLE), which affects >3.4 million individuals worldwide, or have primarily cutaneous manifestations, as in discoid LE [1,2]. Patients with LE frequently experience detriments to quality of life and daily functioning [3,4]. Because of this, a number of qualitative studies were conducted over the last decade that attempted to better understand patient experiences with LE [5]. However, the large majority of these studies recruited participants from clinical settings; thus, clear gaps persist in understanding how experiences with LE can be improved in individuals outside the health care system [3,5].

Social media is underused in the qualitative research of individuals with LE. To our knowledge, only 2 thematic analyses have been conducted using LE-related content on social media forums, including an analysis of comments on an LE Facebook group and an analysis of LE-related Twitter (since rebranded as X) posts [6,7]. Qualitative research using social media is important because it captures individuals who are understudied in typical qualitative research because social media users represent a nonclinical sample and thus may have varied experiences with, and accessibility to, health care [8]. In addition, with estimates suggesting that 40% to 45% of individuals use social media to make medical decisions, it is clinically prudent to determine how diseases such as LE are being portrayed to patients seeking information about their condition [9-11].

TikTok (ByteDance) is an extremely popular social media platform, with >1 billion monthly users [11]. TikTok users post short—often <1 minute—videos and can add filters, music, and captions to their content within the app. TikTok videos provide an untapped, novel, and abundant source of patient experiences to examine; nevertheless, they have yet to be used in thematic analysis within the fields of dermatology or rheumatology. In fact, #lupus has 1.3 billion cumulative views on TikTok alone, indicating the popularity and prevalence of the topic on the app [12]. Furthermore, because TikTok is the fastest growing social media platform worldwide, it is imperative to study it because it is a rapidly expanding source of health information for patients seeking knowledge on the web [9,11].

Objectives

In this study, we used content analysis and thematic analysis to examine TikTok videos involving personal experiences with LE. By doing so, we hope to gain a better understanding of the disease-related experiences of TikTok users who have LE.

Methods

Data Collection

A new TikTok account was created for data collection to avoid TikTok algorithms that prioritize videos based on prior user activity [13,14]. The account was used to search for #lupus on the TikTok app on August 21, 2022; the app then displayed the most popular videos tagged with #lupus [13]. The links, captions, usernames, likes, comments, views, and shares were extracted from each video identified through this search. Videos were then evaluated for inclusion and exclusion criteria. Videos were included if they were downloadable, were in English, and involved the personal opinions or experiences of an individual with LE. They were excluded if they did not relate to LE or were primarily about a condition other than LE.

Codebook Development

A codebook was developed using a standardized inductive approach of iterative coding [13,15]. Study team members coded sets of 20 TikTok videos determined through random selection. After independently coding each set, the study team members met to reach consensus on the codes used for each video. Next, they collaboratively assigned labels, definitions, exclusions, and examples to new codes for the purpose of developing a preliminary codebook [15]. This process was repeated until saturation was reached, that is, no new major codes were created or adjusted, and the codebook could be finalized (refer to Multimedia Appendix 1 for the finalized codebook) [15].

Data Analysis

NVivo 2020 (Lumivero), a qualitative analysis software platform, was used for coding and analysis. Videos were imported into the NVivo software and transcribed. Line-by-line coding of each video's transcript and caption was accomplished within NVivo using the finalized codebook. For the quantitative content analysis aspect of this study, the prevalence and frequency of overlap of individual codes was obtained through NVivo. For the top codes, median views and median views per day were calculated. For the qualitative analysis aspect of this study, the study team members met to discuss the most predominant and rich codes, and they used inductive thematic analysis to derive major themes from the data [16].

Ethical Considerations

This research was determined by the University of Minnesota Institutional Review Board in July 2022 as not constituting human subjects research. Included videos had to be downloadable because this was seen as an indication that the user intended their content to be used and shared by others [13]. All identifying data, including usernames, were removed from the data before dissemination. The study team consulted with the University of Minnesota Medical School Office of Diversity, Equity, & Inclusion and elected not to document user demographics from individual videos because doing so would involve the assumption of identity through appearance on video, and objective demographic information of individual TikTok users is not publicly available.

Results

Overview

A total of 398 TikTok videos were identified through the #lupus search and underwent inclusion and exclusion criteria, after which 153 (38.4%) videos posted between December 19, 2019, and August 21, 2022, were included for analysis. Cumulatively, the videos had 29,446,765 views, with a median of 37,200 (IQR

146,936; range 163-2,300,000) views per video. A total of 76 distinct TikTok users were represented in the sample. Users were primarily female presenting. Of the 76 users, 3 (4%) contributed 35.9% (55/153) of the videos; the top user contributed 15.7% (24/153) of the videos, followed by a user with 11.8% (18/153) and a user with 8.5% (13/153). The most common codes were *experiences with symptoms* (107/153, 69.9%), *mucocutaneous symptoms* (62/153, 40.5%), and *experiences with treatment* (58/153, 37.9%). *Experiences with symptoms* and *mucocutaneous symptoms* had the greatest number of cumulative views (24,426,874 and 14,082,409 views, respectively; Table 1).

Five major thematic conclusions were derived from the data and are explored in the subsections that follow. Sample quotes for each theme can be found in Textbox 1.

Table 1. Top codes and their popularity metrics.

Codes	Videos coded (n=153), n (%)	Cumulative views	Views, median (IQR; range)	Views/d, median (IQR; range)
Experiences with symptoms	107 (69.9)	25,381,074	37,300 (201,100; 234-2,300,000)	582 (3732.5; 2-250,000)
Mucocutaneous symptoms	62 (40.5)	14,879,109	54,700 (280,700; 474-2,300,000)	699 (4507.3; 14-188,889)
Treatment experience	58 (37.9)	11,744,223	34,950 (170,334; 163-2,300,000)	1042 (6152.9; 1-188,889)
Health care experience	45 (29.4)	8,546,113	32,000 (123,322.5; 163-2,300,000)	582 (2315.8; 1-127,778)
Constitutional symptoms	43 (28.1)	9,993,593	31,100 (147,063; 234-2,300,000)	509 (4037.8; 3-250,000)
Mental health	42 (27.5)	7,703,071	32,150 (133,700; 234-2,300,000)	1748 (3391.5; 2-127,778)
Fatigue	41 (26.8)	9,835,548	31,100 (158,033; 234-2,300,000)	691 (4269.1; 3-250,000)
Hair loss	40 (26.1)	10,240,943	75,550 (311,900; 474-1,700,000)	733 (4526.5; 16-188,889)
Rash	39 (25.5)	8,767,466	43,800 (283,300; 474-1,700,000)	691 (4564.0; 14-188,889)
Humor	35 (22.9)	4,548,187	33,800 (57,683; 474-1,500,000)	545 (1263.6; 16-250,000)
Musculoskeletal symptoms	35 (22.9)	5,654,750	23,100 (83,400; 338-2,100,000)	165 (3627.2; 2-51,438)

Textbox 1. Summary of major themes with sample quotes, which have all been taken directly from users' spoken words, written captions, or written subtitles.

Themes and TikTok user quotes

- Mucocutaneous symptoms had profound effects on the mental health and body image of TikTok users with lupus erythematosus.
  - "In 2019, I had a full head of hair. Life can change in a blink of an eye. I was diagnosed with lupus in 2020. My hair was dropping like crazy. There was nothing I could do...[it was] painful and depressing. My scalp was filled with so many scabs. I couldn't even touch my head. [I] kept saying to myself 'this can't be it.'" [User 67]
  - "I had to cut my hair because of lupus (). I've been growing it for 7 and 1/2 years (). My hair started thinning out at the top because of the lupus (). So I knew soon I would have to cut it...I just found out I had lupus 4 months ago (). But at the end of the day (), my health is more important than my hair and the only way I could get it to grow back right is if I cut it (). This really broke my heart () but I'm still handsome [17] ()". [User 18]
- TikTok users' negative experiences with health care workers were often derived from diagnostic delays and perceptions of "medical gaslighting."
  - "Things doctors told me before I got diagnosed with lupus. Take some vitamins. Go outside more. You're just stressed, take these antidepressants. It's growing pains (I was 18). Have you tried yoga?" [User 33]
  - "How doctors would gaslight me until I was finally diagnosed:

[User portraying self] My body is aching so bad. I have really swollen lymph nodes and am losing weight fast.

[User portraying physician] Swollen lymph nodes are common. Maybe you're losing weight because you're depressed.

[User portraying self] I am so dizzy and have no energy. I am bruising everywhere and am always sick. Please, listen to me. I'm not making this up. Something is not right.

[User portraying physician] Are you exercising enough? You're too young to have something serious going on.

[User providing commentary] This was my experience for almost 3 years. Begging doctors to take me seriously. It traumatized me. You are your biggest advocate." [User 34]

- TikTok users tended to portray nonpharmacologic interventions, such as diet and naturopathic remedies, positively, whereas pharmacologic treatments were more commonly portrayed negatively or referred to as "chemotherapy."
  - "I did my part by reporting new symptoms to my rheumatologist and neurologist, which of course was downplayed. Five months later, I lost my mobility and couldn't do much for myself. I thought that I'd never bounce back. I researched natural practices/herbs and started a personal healing journey. I've regained my mobility and much more within a few months." [User 28]
  - "Nightshade vegetables will cause you extreme pain in the long run if you're someone who is dealing with autoimmune disease. I have lupus, but I put lupus in remission as quick as it came out of remission. So, my suggestion to you, unbeknownst to most people, stop eating nightshade vegetables. If you've got a garden, stop growing them. You'd be surprised, you could cure lupus immediately, just stop it." [User 21]
  - "Today is one full week on chemo. I don't want to keep doing this. [User 34]
  - "Sometimes you need things like chemotherapy...which sounds crazy, but [lupus] is that serious." [User 8]
  - [User portraying physician] Doctor: I'm sorry, but in order to control your flare we need to start steroids.

[User portraying self] My face: [user uses a special effect that causes their face to blow up to 3 times its size and resemble the front of a train. Audio of a train horn plays in the background.]

[Caption] All aboard the moon face express () [17]". [User 7]

- Lupus erythematosus symptoms, particularly musculoskeletal symptoms and fatigue, interfered with users' daily functioning.
  - "I was diagnosed with lupus almost four years ago. And lupus took a lot from me...lupus stole my social life. It took my freedom, it destroys my energy, it took my job. Do you know what it's like to be told you can't work? I've had to adjust to a new normal. This sucks but I can't let it win. That's why I can assure you I will not give up. I've come so far, I can't let it win." [User 16]
  - "A lot of days, my body hurts so badly that I don't possibly know how to get out of bed or survive for the next few hours." [User 70]
- Although TikTok users frequently had strong support systems, feelings of isolation or misunderstanding were often attributed to battling an "invisible illness."
  - "A true story about finding an *amazing* friend that actually gets it. She...is 100% understanding when I have to flake last minute because I don't feel well. No guilt. A rare gem indeed." [User 43]
  -

“[Footage plays of user getting their nails painted by their husband] I have lupus. My husband helps me through the ups and downs of this...He doesn’t care about doing things for me. Swollen ankles, messed up toes, and constant pain. This is where we’re at in life. 19 years together.” [User 71]

- “*Getting through life with lupus* Professors: Not understanding what I need. Most friends: Thinking I’m fine and insulting me unintentionally. Family: Making me feel like I’m doing it alone.” [User 12]

Theme 1: Mucocutaneous Symptoms Had Profound Effects on the Mental Health and Body Image of TikTok Users With LE

Mucocutaneous symptoms, defined as hair loss, rash, photosensitivity, ulcers, itch, or sicca symptoms, were the most frequently mentioned symptoms of LE in our sample; 40.5% (62/153) of the videos referred to a mucocutaneous manifestation of LE. Hair loss and rash were the most common and were each coded in approximately one-quarter of the videos (40/153, 26.1% and 39/153, 22.5%, respectively).

Mucocutaneous symptoms were highly distressing to users, affecting both body image and mental well-being. Of the 153 videos, 24 (15.7%) involved a user exhibiting negative body image, of which 79% (n=19) were related to hair loss, 46% (n=11) were related to rash, and 92% (n=22) were related to either hair loss or rash. Furthermore, *hair loss* was the second most common code to overlap with *mental health*; nearly one-fifth (29/153, 19%) of the *mental health* codes directly overlapped with *hair loss* codes.

Distress and body image concerns were apparent within TikTok videos that mentioned mucocutaneous symptoms. Users felt that their hair and skin changes led to a loss of identity:

> *Going through losing all my hair was really hard for me. Like, I didn’t realize how much of my self-worth I attached to my hair...I felt like I was losing a piece of myself...I felt so sad all the time. And it was so hard for me to just go outside because I felt so, you know, insecure.* [User 27]

Although hair loss and rash had significant impacts on users’ well-being and body image, these effects appeared to be mitigated by cosmetic measures and pharmacologic and nonpharmacologic treatments. Of the 24 videos, 13 (54%) involved users cosmetically modifying their hair with sew-ins, wigs, hair dye, hairstyles, haircuts, or scarves to hide hair loss or improve confidence. A user stated as follows:

> *In 2016 I lost my hair...this was before the shaved head hype. Back then my hair was my identity...I knew I needed to cover it up because I had to go into work, so I did a few scarf tutorials and ended up like this. I mean, I think it looks pretty dope.* [User 16]

Skin-directed treatments also helped users. A user showed old footage of large clumps of their hair that had fallen out in the sink. The user recalled how they felt at that time:

> *Super stressed. I’m going to be bald. My hair won’t go back to normal.* [User 68]

They then showed footage of dozens of boxes of prednisone and their scalp with hair growth, stating as follows:

> *OMG! It might be working...Now I can do the hair styles I want.* [User 68]

Theme 2: TikTok Users’ Negative Experiences With Health Care Workers Were Often Derived From Diagnostic Delays and Perceptions of “Medical Gaslighting”

Of the 153 videos, 45 (29.4%) involved a health care experience, of which 58% (n=26) depicted negative experiences, whereas 18% (n=8) depicted positive experiences. Of the 25 negative health care experiences, 22 (88%) could be attributed to an experience with a health care worker. Primarily, TikTok users expressed frustration due to diagnostic delays and “medical gaslighting,” which made up 64% (14/22) and 36% (8/22) of the negative interactions with health care workers, respectively (Textbox 2). Diagnostic delays described by users spanned from “months” to “years,” with 2 (14%) of the 14 users describing delays of ≥10 years.

In our sample, *diagnostic delays* and *medical gaslighting* frequently overlapped; 7 (4.6%) of all videos (n=153) described scenarios where users felt that their symptoms were belittled by medical professionals, leading to delayed diagnoses of LE. A TikTok user stated as follows:

> *[D]octors would gaslight me until I was finally diagnosed...this was my experience for almost 3 years. Begging doctors to take me seriously. It traumatized me.* [User 34]

In another video, a user made the following announcement:

> *I was diagnosed with lupus today after 10 years and 14 different sexist doctors, they finally found out I wasn’t just “overtired and dehydrated.”* [User 20]

Health care workers attributing users’ lupus symptoms to mental health causes seemed to be a common experience among those who experienced medical gaslighting (5/8, 63%). A TikTok video started with the following words:

> *Been sick since 2010. No doctor would listen.* [User 2]

The user then showed stock photos of 3 physicians, each depicted as saying, “It’s just anxiety, it’s just anxiety, it’s just anxiety.” The user finished the video by rolling their eyes at the camera and displaying the following words:

> *Finally diagnosed with lupus, rheumatoid arthritis, among other things, after over a decade...it became medical negligence a LONG time ago. They seriously need to stop telling people that.* [User 2]

Notably, in 4 (18%) of the 22 videos, the users’ negative experiences with health care workers directly caused them to pursue naturopathic practitioners to treat their lupus.

Textbox 2. Common hashtags explained.

Hashtag and description

- #medicalgaslighting
  - Describes medical providers minimizing symptoms or incorrectly attributing symptoms to a behavioral or psychological cause
  - Primarily developed to describe the experiences of women and Black, Indigenous, and patients from racial and ethnic minority groups [18,19]
  - Popularized in recent years by publications such as *The New York Times* and *The Atlantic* [18,19]
- #spoonie
  - An identity for an individual who experiences limited energy, often due to a chronic illness
  - Derived from "The Spoon Theory," written by Miserandino [20], a blogger with lupus; the theory describes spoons as units of energy that everyone starts their day with; however, people with a chronic illness only get a few spoons at the start of the day, whereas others get an excess of spoons [21]
- #chemotherapy
  - Used to describe immunosuppressive treatments for lupus, such as methotrexate
  - Possibly popularized by Selena Gomez, a singer and actor who has systemic lupus erythematosus; in a 2015 interview with *Billboard*, she shared that she received chemotherapy to treat her lupus [22,23]; this was met with backlash from people who thought that the term "chemotherapy" should be reserved for patients with cancer [23]
- #invisibleillness
  - An illness with symptoms that are predominantly "invisible" to others; this may lead to misunderstandings from others and diagnostic difficulties [24,25]

Theme 3: TikTok Users Tended to Portray Nonpharmacological Interventions, Such as Diet and Naturopathic Remedies, Positively, Whereas Pharmacological Treatments Were More Commonly Portrayed Negatively or Referred to as "Chemotherapy"

Of the 153 videos, 58 (37.9%) involved experiences with LE treatment, of which 41% (n=24) were on pharmacological treatments, and 28% (n=16) were on nonpharmacological treatments. Overall, 22% (13/58) involved positive experiences with treatment, and 36% (21/58) involved negative experiences with treatment.

Nonpharmacological treatments tended to be portrayed positively; 9 (56%) of the 16 videos that mentioned nonpharmacological treatments depicted a good experience. The majority of positive experiences with treatment involved nonpharmacological treatments (8/13, 62%), primarily diet (4/8, 50%) and naturopathic remedies (4/8, 50%). Furthermore, nonpharmacological treatments were often credited for disease remission; of the 9 videos that attributed LE remission to treatment, 6 (67%) cited nonpharmacological treatments, whereas only 3 (33%) cited pharmacological treatments. A user stated as follows:

> *Natural medicine saved my life...[I was] told that I would be on pharmaceuticals for life and that I would never be able to exercise again, I could barely walk, I wouldn't be able to have kids, I wouldn't be able to have a job...But luckily, I didn't listen. Because if I did, I don't know where I would be today. Instead, I run three different businesses, I found movement that works for me, I've completely reversed all fertility issues, I'm in remission from lupus, and most of my markers are completely normal. The secret to my healing, you may be asking? Well, it's the food that you eat, the herbs that you put into your body, and the habits that you practice on a daily basis that set the course for your entire life.* [User 26]

By contrast, pharmacological treatments tended to be portrayed negatively; of the 24 videos in which they were mentioned, they were depicted negatively in 10 (42%) and positively in only 3 (13%). Approximately three-fourths (16/21, 76%) of the videos about negative experiences with treatment involved pharmacologic treatments. Negative experiences included side effects (6/16, 38%); injection pain (2/16, 13%); and distress or difficulty with medication management (5/16, 31%) such as remembering to take pills, feeling as though they had too many prescriptions, having an emotional reaction to taking a medication, and relying on perishable and expensive prescriptions.

Furthermore, immunosuppression was repeatedly referred to negatively with the term "chemotherapy." In total, the term was used by 5 (7%) of the 76 users in 16 (10.5%) of the 153 videos and seemed to be used to portray the severity of disease or the gravity of treatment measures. A user relayed their experience with a flare:

> *I found out 3 years ago I had lupus. I had to have chemo. It's been manageable, but last week I had a bad flare up and ended up in the hospital to find out*

> *it's damaged my kidneys, and the doctors are talking about chemo again.* [User 23]

Another user stated as follows:

> *I am so physically exhausted from this disease and chemo but this is your reminder—don't give up.* [User 34]

Theme 4: LE Symptoms, Particularly Musculoskeletal Symptoms and Fatigue, Interfered With Users' Daily Functioning

Of the 153 videos, 29 (19%) referenced LE symptoms or treatments interfering with basic, instrumental, social, educational, or occupational functioning. Musculoskeletal symptoms and fatigue were the most common symptoms to directly overlap with codes related to interference with functioning.

Musculoskeletal symptoms tended to interfere with basic activities of daily living the most; 15 (83%) of the 18 instances of interference with basic functioning were directly attributable. Musculoskeletal symptoms, such as joint pain and stiffness, primarily affected ambulation (8/15, 53%). A user stated as follows:

> *A lot of days, my body hurts so badly that I don't possibly know how to get out of bed or survive for the next few hours.* [User 70]

Another user showed footage of themselves struggling to perform a variety of activities such as sit on a toilet, grip a marker, open a bottle of juice, and stand from a seated position. During this footage, the user displayed the following subtitles:

> *What it's like living with lupus. My joints get stiff. Doing normal things are a struggle now. Lupus affects my hands, wrists, and knees. I was just diagnosed and I hope to see improvement soon.* [User 48]

By contrast, fatigue resulting from LE seemed to be more likely to interfere with social, occupational, and educational functioning, contributing to 5 (56%) of the 9 references within these categories. A user talked about how fatigue affected their schooling:

> *You're worried about going back to school because you literally can't do anything without 15 hours of sleep, and you can't get your schoolwork done, and you can't study enough, and it's horrible.* [User 10]

Another user talked about feeling fatigued after driving to see a friend:

> *I had lunch with a friend today. I drove there and drove home, so naturally I am fatigued now. Good, bad, ugly, that a single activity can put me on the couch. I am not worried, but this is a reality.* [User 11]

Interestingly, 4 (5%) of the 76 users included the word "spoonie" in their TikTok videos, a term that has become an identity for individuals who experience fatigue from chronic illnesses.

Theme 5: Although TikTok Users Frequently Had Strong Support Systems, Feelings of Isolation or Misunderstanding Were Often Attributed to Battling an "Invisible Illness"

Of the 153 videos, 16 (10.5%) depicted users' support systems, which were composed of partners (n=8, 50%), family members (n=5, 31%), other TikTok users (n=3, 19%), and friends (n-3, 19%). TikTok users frequently expressed gratitude for the assistance they received from support people in navigating their LE symptoms and treatment. A user showed footage of their hospitalization for LE and displayed the following words:

> *It's been a rough few weeks. I couldn't express how grateful I am for my support system. My family. I couldn't have done it without you guys.* [User 73]

Another user, who similarly filmed their TikTok video when they were hospitalized for LE, wrote as follows:

> *I'm not recovering at the rate I hoped I would by now. I keep watching time go by as the pain gets worse as I lay here...You're left with the emptiness and questions what you did to deserve this and why you're here, and the only escape you have is those short few minutes you get a call from a friend or family member and can pretend it isn't happening.* [User 72]

However, TikTok users also reported discouraging interactions with others (14/153, 9.2%), including with people from school or work (n=3, 21%), friends (n=3, 21%), family (n=2, 14%), service industry workers (n=2, 14%), partners (n=1, 7%), other TikTok users (n=1, 7%), and neighbors (n=1, 7%). Most of these interactions arose from misunderstandings of LE. Users felt that because LE is primarily an "invisible illness," with many of its signs and symptoms not visibly apparent to others, others did not always recognize their needs; for example, a user's video provided a list of "things people with lupus are tired of hearing," which included comments such as "You can't be tired, you haven't done anything all day," "You don't look sick," "You're using it as an excuse to be lazy," and "It's not that bad" (User 41).

These misunderstandings led to feelings of isolation. A user, who filmed themselves lying in bed, commented as follows:

> *[Lupus is] so isolating because no one understands what you're going through. I just feel lame having to leave a group setting to have a flare up until your body goes back to normal after a few hours.* [User 44]

In all, of the 14 videos, 10 (71%) referred to LE as an "invisible illness" through hashtags, captions, or direct quotes, and 6 (43%) that were coded with "negative experiences with others" were also coded with "invisible illness."

Discussion

Principal Findings

This study represents the first qualitative and content analysis of TikTok videos involving personal experiences of users with LE. Patients are increasingly using social media to learn and share information about their health conditions [9-11,26]. Thus, social media provides a crucial fund of patient experiences that can be used to extract clinically relevant patient-centered information for clinicians that may ultimately improve patient care.

In this study, we found that TikTok videos on LE experiences have extensive audiences, garnering millions of views and high user engagement. Videos that mentioned mucocutaneous symptoms of LE, such as hair loss and rash, were pervasive among this sample, with *mucocutaneous symptoms* being the second most frequent code used. Consistent with findings of previous qualitative and survey studies, mucocutaneous symptoms seemed to be major drivers of poor mental health and negative body image among TikTok users with LE [27-33]. Our study suggests the need to regularly assess for mental health and body image concerns in patients with LE, especially among those with active dermatologic symptoms. It also underscores the importance of treating hair loss and rash to mitigate mental health burdens in this population (Textbox 3).

Textbox 3. Clinical applications of themes.

Themes and clinical applications

- Mucocutaneous symptoms had profound effects on the mental health and body image of TikTok users with lupus erythematosus (LE).
  - Treatment of hair loss and rash is important for quality of life and mental health of patients with LE.
  - Mental health should be regularly assessed at appointments, particularly for patients with rash or hair loss.
- TikTok users' negative experiences with health care workers were often derived from diagnostic delays and perceptions of "medical gaslighting."
  - Clinical strategies such as reflective listening and validation may enhance the clinician-patient relationship and prevent perceptions of medical gaslighting.
- TikTok users tended to portray nonpharmacologic interventions, such as diet and naturopathic remedies, positively, whereas pharmacologic treatments were more commonly portrayed negatively or referred to as "chemotherapy."
  - Clinicians should be aware of popular nonpharmacological treatments for LE.
  - Clinicians should engage in informed discussions of the safety and effectiveness of both pharmacological and nonpharmacological treatments with their patients.
- LE symptoms, particularly musculoskeletal symptoms and fatigue, interfered with users' daily functioning.
  - Treatment should focus on reducing musculoskeletal symptoms and fatigue for patients reporting interference in functioning.
  - Clinicians should assist patients in obtaining mobility devices, disability resources, and occupational and physical therapy that improve daily functioning.
- Although TikTok users frequently had strong support systems, feelings of isolation or misunderstanding were often attributed to battling an "invisible illness."
  - Involving support people in appointments could be a beneficial way to enhance existing support relationships and educate support people on LE morbidity and disability.

However, we found that pharmacologic therapies might be met with hesitancy by individuals with LE. Pharmacologic treatments were depicted negatively in our sample, with individuals citing side effects such as immunosuppression, weight gain, and fatigue. Notably, the term "chemotherapy" was used in several videos, which portrays the gravity that users associate with receiving immunosuppressive medications. By contrast, nonpharmacologic treatments such as diet and naturopathic remedies were depicted overwhelmingly positively in our sample, a finding that to our knowledge has only been reported once before, in a 2011 qualitative study on attitudes toward medications in South Asian patients with SLE [34]. The uniqueness of this finding could be because users can benefit monetarily from promoting diet or naturopathic remedies through promotion deals on TikTok. However, it is also possible that TikTok users, because they are a nonclinical sample, may have fewer or poorer experiences with clinical medicine and thus prefer nonpharmacologic treatments. Clinicians should be aware of common nonpharmacological options for patients with LE and should be prepared to counsel patients on the safety and effectiveness of these therapies (Textbox 3).

Concordantly, we did find that TikTok users with LE shared primarily negative interactions with the health care system and health care workers. Many of these experiences centered on instances of "medical gaslighting," which users felt resulted in diagnostic delays. Diagnostic delays are well documented in SLE qualitative and quantitative research and can have significant mental health ramifications for patients [3,35-39]. However, only a few qualitative studies have explored patients' perceptions of diagnostic delays resulting from physicians downplaying LE symptoms, with only 1 prior study capturing the term "gaslighting" in its analysis [3,38,39]. To maintain the therapeutic relationship, clinicians should combat perceptions of medical gaslighting through strategies such as validation and reflective listening (Textbox 3) [40].

Our sample had high symptom burden and frequently described how musculoskeletal symptoms and fatigue were interfering with daily functioning. TikTok users did not always feel that others understood these symptoms, leading to feelings of isolation and the thought that they have an "invisible illness." These ideas have been described in numerous qualitative analyses on LE [3,7,24,31,36,41-43]. Even so, overall, TikTok users demonstrated robust social support systems made up of friends, family, and partners. This is important because social support has been associated with improved mental health in patients with SLE, whereas a lack of substantial social support has been associated with increased disease activity [24,44-47]. These combined findings suggest the importance of educating patients' support people on LE morbidity and disability to facilitate successful support relationships (Textbox 3).

Strengths

This study has several notable strengths. First, because our qualitative data were derived from a nonclinical sample, we may have captured voices from individuals with diverse experiences with the health care system [8]. Second, because patient experiences were collected without interaction with the study team, patient experiences were unbiased by researcher presence or preset interview questions [48,49]. Third and last, in contrast to existing qualitative studies that often have low sample sizes, analyzing TikTok videos allowed us to gather the experiences of 76 distinct users.

Limitations

Although qualitative studies are inherently not designed to be generalizable because they provide rich, narrative data from the group being studied, it is important to note that this study only examined TikTok content and did not examine content from other web-based platforms [50]. Thus, these findings may not be representative of the entire LE web-based community. Furthermore, as the TikTok videos were sampled from a search revealing the most popular videos with #lupus, our findings may overrepresent ideas in popular videos, while underrepresenting ideas from users with fewer views. Furthermore, in comparison to qualitative studies in which interviewees are promised confidentiality when disclosing their experiences, TikTok videos in our sample were not confidential, and, in fact, were meant to be publicly shared. This means that patient experiences were subject to social desirability bias, and sensitive topics may have been avoided. Finally, demographic data of TikTok users are not publicly available, and thus detailed user demographics could not be characterized in this study.

Conclusions

TikTok provides a nonclinical, underused platform for qualitative and content analysis of patient experiences. This study summarizes key terminology and content in the LE TikTok community, which may be clinically relevant because a substantial number of patients use social media to obtain medical information [9,10]. Ultimately, this study presents 5 thematic conclusions paired with clinical applications, which offer an enhanced understanding of how the well-being of patients with LE is influenced by symptoms, treatments, support people, and health care experiences.

Authors' Contributions

Both LJW and DRP contributed to the design, execution, and reporting of this study.

Conflicts of Interest

DRP is a consultant for Biogen, Merck, and Pfizer, and a clinical trials investigator for EMD Serono, Daiichi Sankyo, and Priovant.

Multimedia Appendix 1
Final codebook.
[DOCX File , 34 KB - infodemiology_v4i1e51211_app1.docx]

References

1. Durcan L, O'Dwyer T, Petri M. Management strategies and future directions for systemic lupus erythematosus in adults. Lancet 2019 Jun 08;393(10188):2332-2343. [doi: 10.1016/S0140-6736(19)30237-5] [Medline: 31180030]
2. Tian J, Zhang D, Yao X, Huang Y, Lu Q. Global epidemiology of systemic lupus erythematosus: a comprehensive systematic analysis and modelling study. Ann Rheum Dis 2023 Mar;82(3):351-356 [FREE Full text] [doi: 10.1136/ard-2022-223035] [Medline: 36241363]
3. Sloan M, Bosley M, Blane M, Holloway L, Barrere C, D'Cruz D, et al. 'But you don't look sick': a qualitative analysis of the LUPUS UK online forum. Rheumatol Int 2021 Apr;41(4):721-732 [FREE Full text] [doi: 10.1007/s00296-020-04726-x] [Medline: 33104839]
4. Schmeding A, Schneider M. Fatigue, health-related quality of life and other patient-reported outcomes in systemic lupus erythematosus. Best Pract Res Clin Rheumatol 2013 Jun;27(3):363-375. [doi: 10.1016/j.berh.2013.07.009] [Medline: 24238693]
5. Petrocchi V, Visintini E, De Marchi G, Quartuccio L, Palese A. Patient experiences of systemic lupus erythematosus: findings from a systematic review, meta-summary, and meta-synthesis. Arthritis Care Res (Hoboken) 2022 Nov;74(11):1813-1821 [FREE Full text] [doi: 10.1002/acr.24639] [Medline: 34133081]
6. Venuturupalli S, Kumar A, Bunyan A, Davuluri N, Fortune N, Reuter K. Using patient-reported health data from social media to identify diverse lupus patients and assess their symptom and medication expressions: a feasibility study. Arthritis Care Res (Hoboken) 2023 Feb;75(2):365-372 [FREE Full text] [doi: 10.1002/acr.24868] [Medline: 35157364]

7. Colmenares-Roa T, Gastelum-Strozzi A, Crosley E, Fuentes-Silva Y, Reategui-Sokolova C, Elera-Fitzcarrald C, et al. Digital narratives of living with lupus: lived experiences and meanings for Latin American and Latino patients and their families. Arthritis Care Res (Hoboken) 2023 Mar;75(3):540-549. [doi: 10.1002/acr.24870] [Medline: 35188345]
8. Nelson PA. Getting under the skin: qualitative methods in dermatology research. Br J Dermatol 2015 Apr;172(4):841-843. [doi: 10.1111/bjd.13720] [Medline: 25827724]
9. Villa-Ruiz C, Kassamali B, Mazori DR, Min M, Cobos G, LaChance A. Overview of TikTok's most viewed dermatologic content and assessment of its reliability. J Am Acad Dermatol 2021 Jul;85(1):273-274. [doi: 10.1016/j.jaad.2020.12.028] [Medline: 33359080]
10. Social media "likes" healthcare: from marketing to social business. Health Research Institut. 2012 Apr. URL: https://adindex.ru/files2/access/2013_06/99606_tpc-health-care-social-media-report.pdf [accessed 2024-03-12]
11. Zheng DX, Mulligan KM, Scott JF. TikTok and dermatology: an opportunity for public health engagement. J Am Acad Dermatol 2021 Jul;85(1):e25-e26. [doi: 10.1016/j.jaad.2021.02.050] [Medline: 33639245]
12. #lupus. TikTok. URL: https://www.tiktok.com/tag/lupus?lang=en [accessed 2023-12-16]
13. Herrick SS, Hallward L, Duncan LR. "This is just how I cope": an inductive thematic analysis of eating disorder recovery content created and shared on TikTok using #EDrecovery. Int J Eat Disord 2021 Apr;54(4):516-526. [doi: 10.1002/eat.23463] [Medline: 33382136]
14. Ng R, Indran N. Hostility toward baby boomers on TikTok. Gerontologist 2022 Sep 07;62(8):1196-1206 [FREE Full text] [doi: 10.1093/geront/gnac020] [Medline: 35106587]
15. Roberts K, Dowell A, Nie JB. Attempting rigour and replicability in thematic analysis of qualitative research data; a case study of codebook development. BMC Med Res Methodol 2019 Mar 28;19(1):66 [FREE Full text] [doi: 10.1186/s12874-019-0707-y] [Medline: 30922220]
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
17. OpenMoji. GitHub. URL: https://openmoji.org/library/ [accessed 2024-03-12]
18. Moyer MW. Women are calling out 'medical gaslighting'. The New York Times. 2022 Mar 28. URL: https://www.nytimes.com/2022/03/28/well/live/gaslighting-doctors-patients-health.html [accessed 2024-03-12]
19. Fetters A. The doctor doesn't listen to her. But the media is starting to. The Atlantic. 2018 Aug 10. URL: https://www.theatlantic.com/family/archive/2018/08/womens-health-care-gaslighting/567149/ [accessed 2024-03-12]
20. Miserandino C. The spoon theory. In: Beginning with Disability. Milton Park, UK: Routledge; 2018.
21. Henderson W. 15 signs that confirm you're definitely a spoonie. Amyotrophic Lateral Sclerosis News Today. 2017 May 29. URL: https://tinyurl.com/4bnac8fw [accessed 2024-03-12]
22. Billboard cover sneak peek: 5 ways Selena Gomez is taking charge. Pinterest. URL: https://www.pinterest.com/pin/3799980912913174/ [accessed 2024-03-12]
23. Chemotherapy isn't only for cancer. Scleroderma & Raynaud's UK. URL: https://tinyurl.com/52y4xpxk [accessed 2023-01-04]
24. Brennan KA, Creaven AM. Living with invisible illness: social support experiences of individuals with systemic lupus erythematosus. Qual Life Res 2016 May;25(5):1227-1235. [doi: 10.1007/s11136-015-1151-z] [Medline: 26449351]
25. Donoghue PJ, Siegel ME. Sick and Tired of Feeling Sick and Tired: Living with Invisible Chronic Illness. Tempe, AZ: Norton; 1992.
26. Harter C, Ness M, Goldin A, Lee C, Merenda C, Riberdy A, et al. Exploring chronic pain and pain management perspectives: qualitative pilot analysis of web-based health community posts. JMIR Infodemiology 2023 May 30;3:e41672 [FREE Full text] [doi: 10.2196/41672] [Medline: 37252767]
27. Beckerman NL. Living with lupus: a qualitative report. Soc Work Health Care 2011;50(4):330-343. [doi: 10.1080/00981389.2011.554302] [Medline: 21512954]
28. Robinson D, Aguilar D, Schoenwetter M, Dubois R, Russak S, Ramsey‐Goldman R, et al. Impact of systemic lupus erythematosus on health, family, and work: the patient perspective. Arthritis Care Res 2010 Feb;62(2):266-273. [doi: 10.1002/acr.20077]
29. Farinha F, Freitas F, Águeda A, Cunha I, Barcelos A. Concerns of patients with systemic lupus erythematosus and adherence to therapy - a qualitative study. Patient Prefer Adherence 2017 Jul 14;11:1213-1219 [FREE Full text] [doi: 10.2147/PPA.S137544] [Medline: 28761334]
30. McElhone K, Abbott J, Gray J, Williams A, Teh LS. Patient perspective of systemic lupus erythematosus in relation to health-related quality of life concepts: a qualitative study. Lupus 2010 Dec;19(14):1640-1647. [doi: 10.1177/0961203310378668] [Medline: 20709719]
31. Phuti A, Schneider M, Makan K, Tikly M, Hodkinson B. Living with systemic lupus erythematosus in South Africa: a bitter pill to swallow. Health Qual Life Outcomes 2019 Apr 16;17:65. [doi: 10.1186/s12955-019-1132-y]
32. Tunnicliffe DJ, Singh-Grewal D, Craig JC, Howell M, Tugwell P, Mackie F, et al. Healthcare and research priorities of adolescents and young adults with systemic lupus erythematosus: a mixed-methods study. J Rheumatol 2017 Apr;44(4):444-451. [doi: 10.3899/jrheum.160720] [Medline: 28250139]
33. Jolly M, Pickard AS, Mikolaitis RA, Cornejo J, Sequeira W, Cash TF, et al. Body image in patients with systemic lupus erythematosus. Int J Behav Med 2012 Jun;19(2):157-164. [doi: 10.1007/s12529-011-9154-9] [Medline: 21380770]

34. Kumar K, Gordon C, Barry R, Shaw K, Horne R, Raza K. 'It's like taking poison to kill poison but I have to get better': a qualitative study of beliefs about medicines in Rheumatoid arthritis and Systemic lupus erythematosus patients of South Asian origin. Lupus 2011 Jul;20(8):837-844. [doi: 10.1177/0961203311398512] [Medline: 21511761]
35. Sloan M, Harwood R, Sutton S, D'Cruz D, Howard P, Wincup C, et al. Medically explained symptoms: a mixed methods study of diagnostic, symptom and support experiences of patients with lupus and related systemic autoimmune diseases. Rheumatol Adv Pract 2020 Feb 26;4(1):rkaa006 [FREE Full text] [doi: 10.1093/rap/rkaa006] [Medline: 32373774]
36. Hale ED, Treharne GJ, Lyons AC, Norton Y, Mole S, Mitton DL, et al. "Joining the dots" for patients with systemic lupus erythematosus: personal perspectives of health care from a qualitative study. Ann Rheum Dis 2006 May;65(5):585-589 [FREE Full text] [doi: 10.1136/ard.2005.037077] [Medline: 16219711]
37. Hale ED, Radvanski DC, Hassett AL. The man-in-the-moon face: a qualitative study of body image, self-image and medication use in systemic lupus erythematosus. Rheumatology (Oxford) 2015 Jul;54(7):1220-1225. [doi: 10.1093/rheumatology/keu448] [Medline: 25550393]
38. Mendelson C. Diagnosis: a liminal state for women living with lupus. Health Care Women Int 2009 May;30(5):390-407. [doi: 10.1080/07399330902785158] [Medline: 19350436]
39. Sloan M, Naughton F, Harwood R, Lever E, D'Cruz D, Sutton S, et al. Is it me? the impact of patient-physician interactions on lupus patients' psychological well-being, cognition and health-care-seeking behaviour. Rheumatol Adv Pract 2020 Jul 22;4(2):rkaa037 [FREE Full text] [doi: 10.1093/rap/rkaa037] [Medline: 32974426]
40. Kishor K, Singh Bisht D, Kalra S. Ghost of medical gaslighting. Indian J Clin Pract 2023;34(1):34-37.
41. Kier A, Midtgaard J, Hougaard KS, Berggreen A, Bukh G, Hansen RB, et al. How do women with lupus manage fatigue? a focus group study. Clin Rheumatol 2016 Aug;35(8):1957-1965. [doi: 10.1007/s10067-016-3307-9] [Medline: 27225245]
42. Mazzoni D, Cicognani E. Problematic social support from patients' perspective: the case of systemic lupus erythematosus. Soc Work Health Care 2014;53(5):435-445. [doi: 10.1080/00981389.2014.888124] [Medline: 24835088]
43. Booth S, Price E, Walker E. Fluctuation, invisibility, fatigue - the barriers to maintaining employment with systemic lupus erythematosus: results of an online survey. Lupus 2018 Dec;27(14):2284-2291 [FREE Full text] [doi: 10.1177/0961203318808593] [Medline: 30451638]
44. Bae SC, Hashimoto H, Karlson EW, Liang MH, Daltroy LH. Variable effects of social support by race, economic status, and disease activity in systemic lupus erythematosus. J Rheumatol 2001 Jun;28(6):1245-1251. [Medline: 11414264]
45. Karlson EW, Daltroy LH, Lew RA, Wright EA, Partridge AJ, Fossel AH, et al. The relationship of socioeconomic status, race, and modifiable risk factors to outcomes in patients with systemic lupus erythematosus. Arthritis Rheum 1997 Jan;40(1):47-56. [doi: 10.1002/art.1780400108] [Medline: 9008599]
46. Alarcón GS, Calvo-Alén J, McGwin G, Uribe AG, Toloza SM, Roseman JM, et al. Systemic lupus erythematosus in a multiethnic cohort: LUMINA XXXV. Predictive factors of high disease activity over time. Ann Rheum Dis 2006 Sep;65(9):1168-1174 [FREE Full text] [doi: 10.1136/ard.200X.046896] [Medline: 16905579]
47. Mazzoni D, Cicognani E. Social support and health in patients with systemic lupus erythematosus: a literature review. Lupus 2011 Oct;20(11):1117-1125. [doi: 10.1177/0961203311412994] [Medline: 21828159]
48. Hallberg LR. The "core category" of grounded theory: making constant comparisons. Int J Qual Stud Health Well-being 2009 Jul 12;1(3):141-148. [doi: 10.1080/17482620600858399]
49. Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualitative research methods in health technology assessment: a review of the literature. Health Technol Assess 1998;2(16):iii-ix, 1-iii-ix274 [FREE Full text] [Medline: 9919458]
50. Leung L. Validity, reliability, and generalizability in qualitative research. J Family Med Prim Care 2015;4(3):324-327 [FREE Full text] [doi: 10.4103/2249-4863.161306] [Medline: 26288766]

Abbreviations

LE: lupus erythematosus
SLE: systemic lupus erythematosus

Edited by T Mackey; submitted 24.07.23; peer-reviewed by T Phenwan, T Giardina; comments to author 29.11.23; revised version received 18.01.24; accepted 22.02.24; published 17.04.24.

Please cite as:
Wanberg LJ, Pearson DR
Evaluating the Disease-Related Experiences of TikTok Users With Lupus Erythematosus: Qualitative and Content Analysis
JMIR Infodemiology 2024;4:e51211
URL: https://infodemiology.jmir.org/2024/1/e51211
doi:10.2196/51211
PMID:38631030

Original Paper

Experiences of Women With Medical Abortion Care Reflected in Social Media (VEILLE Study): Noninterventional Retrospective Exploratory Infodemiology Study

Giulia Gouy[1*], MD; Luisa Attali[2*], PhD; Paméla Voillot[3*], MSc; Patrick Fournet[4*], MD; Aubert Agostini[5*], MD

[1]Service de Gynécologie-Obstétrique, Hôpital de la Croix-Rousse, Lyon, France

[2]Pôle de Gynécologie-Obstétrique et Fertilité, Hôpital de Hautepierre, Centre Hospitalier Universitaire de Strasbourg, Strasbourg, France

[3]Kap Code, Paris, France

[4]Centre de Santé SoMeD, Rouen, France

[5]Service de Gynécologie et d'Obstétrique, Hôpital de la Conception, Assistance Publique-Hôpitaux de Marseille, Marseille, France

[*]all authors contributed equally

Corresponding Author:
Giulia Gouy, MD
Service de Gynécologie-Obstétrique
Hôpital de la Croix-Rousse
103 Gd Rue de la Croix-Rousse
Lyon, 69004
France
Phone: 33 472 071 936
Email: giulia.gouy@chu-lyon.fr

Abstract

Background: Abortion (also known as termination of pregnancy) is an essential element of women's reproductive health care. Feedback from women who underwent medical termination of pregnancy about their experience is crucial to help practitioners identify women's needs and develop necessary tools to improve the abortion care process. However, the collection of this feedback is quite challenging. Social media offer anonymity for women who share their abortion experience.

Objective: This exploratory infodemiology study aimed to analyze, through French social media posts, personal medical symptoms and the different experiences and information dynamics associated with the medical abortion process.

Methods: A retrospective study was performed by analyzing posts geolocated in France and published from January 1, 2017, to November 30, 2021. Posts were extracted from all French-language general and specialized publicly available web forums using specific keywords. Extracted messages were cleaned and pseudonymized. Automatic natural language processing methods were used to identify posts from women having experienced medical abortion. Biterm topic modeling was used to identify the main discussion themes and the Medical Dictionary for Regulatory Activities was used to identify medical terms. Encountered difficulties were explored using qualitative research methods until the saturation of concepts was reached.

Results: Analysis of 5398 identified posts (3409 users) led to the identification of 9 major topics: personal experience (n=2413 posts, 44.7%), community support (n=1058, 19.6%), pain and bleeding (n=797, 14.8%), psychological experience (n=760, 14.1%), questioned efficacy (n=410, 7.6%), social pressure (n=373, 6.9%), positive experiences (n=257, 4.8%), menstrual cycle disorders (n=107, 2%), and reported inefficacy (n=104, 1.9%). Pain, which was mentioned in 1627 (30.1%) of the 5398 posts by 1024 (30.0%) of the 3409 users, was the most frequently reported medical term. Pain was considered severe to unbearable in 24.5% of the cases (399 of the 1627 posts). Lack of information was the most frequently reported difficulty during and after the process.

Conclusions: Our findings suggest that French women used social media to share their experiences, offer and find support, and provide and receive information regarding medical abortion. Infodemiology appears to be a useful tool to obtain women's feedback, therefore offering the opportunity to enhance care in women undergoing medical abortion.

(JMIR Infodemiology 2024;4:e49335) doi:10.2196/49335

KEYWORDS

infodemiology; medical abortion; patient experience; real-world evidence; social media; abortion; women's health; reproduction; reproductive; obstetric; obstetrics; gynecology; gynecological; text mining; topic model; topic modeling; natural language processing; NLP

Introduction

Background

Abortion is a common procedure. Worldwide from 2015 to 2019, there were 121.0 million unintended pregnancies annually leading to 73.0 million abortions (60%) [1]. In France, in 2022, there were 234,300 abortions, 78% of which were medical termination of pregnancy (MToP) that is usually performed early (<8 weeks of amenorrhea) [2].

Early MToP with an antiprogestin (ie, mifepristone) followed by a prostaglandin analog has revolutionized abortion. Since mifepristone was first approved in 1988, MToP has been authorized in numerous countries worldwide [3], and its practice has followed changes in health requirements and local regulations. For instance, the COVID-19 pandemic and associated mandatory lockdowns increased the overall rate of MToPs and the rate of MToPs performed at home [4]. Abortion was legalized in France on January 17, 1975 [5], and mifepristone was first approved for MToP in 1988. The current dose regimen for mifepristone and a prostaglandin analog was approved in 2007 [6]. The sequence of consultations includes a first visit to inform the woman, a second visit during which she signs a consent form and initiates the procedure, and a follow-up visit 14 to 21 days after mifepristone intake to ensure the success of the procedure [7].

Abortion is not perceived in the same way as other standard medical acts. In France, until 2001, a motivational interview was mandatory during the first visit. This allowed a survey showing that abortion was still considered to be too much of a taboo to be performed. Women indicated having elaborate defense strategies to protect themselves against this stigma, including keeping their abortion secret, leading to potential unsafe abortion practices [8]. Currently, women can generally express their abortion experience more freely. In particular, social media provide online anonymity and offer a safe opportunity for women to share and seek information.

With over 2.3 billion active users globally [9], social media have become a new data source for public health, as users can find, exchange, and discuss health information on the platforms. According to Médiamétrie, a company specializing in audience measurement and the study of the use of audiovisual and digital media, more than 85% of the French population are internet users [10]. Moreover, according to a recent report [11], 21.1% of French individuals frequently or very frequently use health-related social media such as Doctissimo, with 14.1% using Facebook and 9.8% using women's magazines and their associated websites to find health information.

Analyzing online information represents a developing alternative means to understand patients' health compared with self-administered questionnaires. These patient-generated health data are produced spontaneously (and are thus not limited to medical consultations, for instance), mostly anonymously. Therefore, these data may better correspond to patients' feelings compared with closed-ended questions. Moreover, text-mining techniques applied to analyze social media data can be used with relative ease [12], providing new opportunities to bridge the gap between qualitative and quantitative data analyses [13]. A new research discipline and methodology has thus emerged. This scientific discipline called infodemiology focuses on health-related content analysis published online [14].

Studies based on analysis of Instagram, Facebook, or Reddit posts have started to be published in peer-reviewed journals. These analyses have helped to characterize patients' experiences and related perceptions of an illness and its burden in many health fields (eg, in vitro fertilization [15], miscarriage [16], cesarean section [17], and breast cancer [18]), along with documenting the processes involved in abortion method decision-making [19]. However, to the best of our knowledge, few studies using social media data have been published to characterize the experiences and perceptions of women who underwent an abortion, in particular MToP [19]. Furthermore, none of these studies has been carried out in France, although women's experiences are likely influenced by clinical practices and cultural differences that preclude the generalization of data collected in other countries.

Objective

We conducted VEILLE, a 4-year retrospective infodemiology study, to analyze reported medical symptoms and the different experiences and information dynamics associated with the MToP process in France.

Abortion care is an essential element of women's reproductive health care [20]. Women's feedback about their experience is crucial to meeting women's needs during MToP. Collecting women's feedback about their MToP experiences and understanding these experiences could provide the necessary information for health care providers to respond to French women's needs during MToP.

Methods

Study Design

VEILLE is a noninterventional, retrospective study using a text-mining approach to retrieve and analyze medical abortion posts from social media posts.

All messages geolocated in France posted by women who had experienced MToP between January 1, 2017, and November 30, 2021, in French-speaking general and specialized web forums were considered. Only messages from publicly available sources were extracted.

The study name VEILLE is an allusion to the French translation of social media monitoring (*veille*) and a tribute to Simone Veil (same pronunciation, /v j/) who legalized abortion in France.

Data Extraction

Data (verbatim social media posts) were identified and pseudonymized by tokenization before being extracted. Irrelevant material was eliminated.

All public posts available on the web containing at least one of the relevant keywords related to MToP were identified using the Brandwatch social media data extractor [21]. This tool is based on queries that include selected keywords evocative of the subject of interest. Using the query, the Brandwatch extractor searches through available public data sources and identifies keywords within posts matching those in the query.

Posts were downloaded along with their associated metadata: URL/domain, publication date, forum, language used, hashtags, authors, and engagement type such as retweet. Posts and associated metadata constituted the corpus.

Keywords in French (eg, IVG for *interruption volontaire de grossesse*, voluntary pregnancy termination) and their synonyms were defined by the authors (see Multimedia Appendix 1).

Data Preprocessing and Modeling

Extracted posts were cleaned before being stored in the study-specific database. Posts from irrelevant sources such as potential advertising sites or forums related to pets and animals were removed using regular expression rules. Duplicates were managed by merging posts with either the same username on different platforms or the same post with another username. A machine-learning algorithm (extreme gradient boosting classifier [XGBoost]) was used to identify posts reporting personal experiences [22]. These posts constituted the study data set.

The algorithm was implemented based on message-level calculation of the user's probability of being a woman having experienced MToP according to specific features (lexical fields and regular forms, such as "I have [EXTRACTION TERM]") and coupled with pronoun variables. Filters (in French) were used to narrow down the search to only MToP experience, excluding surgical abortion (Multimedia Appendix 2).

Data Analysis

Descriptive analysis was performed for posts (number and source) and social media users (number, age, gender). A social media user's age was determined through the identification of regular expressions such as "*j'ai 45 ans*" ("I am 45 years old"), "*ayant 45 ans*" ("being 45") (Regex method) over all posts. Each pseudonym was associated with one gender (man, woman, or unidentified) and one age category (20 years or younger, 21-30 years, 31-40 years, and so on, or unidentified). Gender was confirmed using the Regex method and with the support vector machine algorithm (XGBoost method) through the identification of regular expressions in the content of each post: gendered participles, adjectives, and names (eg, Miss, pregnant) or grammatical features [23].

A *topic model* was applied to identify the topics addressed in the posts constituting the study data set [24]. Topic models consist of text-mining approaches that aim to automatically identify the abstract topics addressed in a collection of documents. Such models are based on the hypothesis that each document corresponds to a distribution of several topics.

A biterm topic model (BTM) was used to identify the topics without prior knowledge. A topic is defined as a subject of discussion, which amounts to tokens that frequently appear together in the posts from the data set. The BTM considers the whole data set as a mixture of topics, where each co-occurring pair in tokens (the biterm) is drawn from a specific topic independently and modeled topics are probability distributions over the biterms [24]. As topics are probability distributions over tokens of the study data set, they can be characterized by the highest per-topic probability tokens. Weighting these probabilities through term-frequency inverse document frequency (TF-IDF) weighting allows topic-specific tokens to be allocated with higher importance. In this case, the per-topic probability of a token is weighted by the inverse of the probabilities of this token in other topics. Therefore, for each topic, tokens were ranked from the highest to the lowest weighted probabilities TF-IDF value in this topic. The first 9 tokens were designated as the set of characteristic tokens and used to name the topic manually.

A specific list of symptoms related to MToP was established based on the Medical Dictionary for Regulatory Activities (version 23) terms [23]. The lexical field was enriched to consider verbal forms found on social networks. A single post could contain several medical terms.

For *difficulties* encountered by social media users, posts were randomly allocated to create a sample representing 30% of the extracted posts. A qualitative manual search was performed on this sample using a generic annotation grid, which helped to categorize each difficulty. A single post could contain several encountered difficulties. Given the diversity of encountered difficulties, data saturation was used to obtain a representative sample of expressed difficulties [25].

Saturation was checked by taking 5% samples of the total number of social media users (N=1964) and analyzing the number of new types of difficulties or unmet needs per 5% sample (n=98). Saturation was considered to be achieved when two consecutive samples no longer yielded more than one newly identified difficulty category. Two additional batches of 5% each were analyzed after saturation was first reached for further validation of the findings [25].

Ethical Considerations

Data collection and treatment followed the European Union General Data Protection Regulation. The study was conducted within the frame of legitimate interest. The study involved data issued from publicly available sources. Consent was not required as the study involved publicly available posts and as users automatically grant their consent for the reuse of their data when they post on public platforms. Following this and as this falls under the R1121-1 Article of the French Public Health Code [26] (in effect since July 1, 2021), we did not seek ethics board review or approval for this study. Private groups or web pages were excluded from our data extraction process. The results of the study do not contain any identifiable information and are presented taken together. A privacy-by-design approach was adopted as all usernames, web forum names, geographic locations, URLs, or any other sensitive information was substituted by identifiers before being stored in the analysis corpus.

Results

Population and Posts

After cleaning/filtering of the corpus, 8326 posts published by 6223 users were extracted from social media platforms to be preprocessed/modeled. Therefore, 5398 posts published by 3409 social media users were identified and constituted the study data set (Figure 1). The top 6 keywords are presented in Table 1.

As only posts reported by individuals having experienced MToP could be included in the data set, all posts were deemed to be written by women. Gender was confirmed for 2898 of the 3409 (85%) social media users. Age was found in the posts of 8.1% of the social media users (n=275): 1.5% (n=52) were ≤20 years, 4.2% (n=142) were between 20 and 30 years, 1.5% (n=51) were between 30 and 40 years, and 0.9% (n=30) were >40 years. The median age was 26 years.

The 5398 posts were retrieved from a total of 22 web forums (Table 1); 78% of the posts were issued from two specialized forums (Doctissimo and aufeminin.com) and one general forum (Facebook). Doctissimo, which was the top-ranked source is a French specialized medical site, whereas aufeminin.com, ranking second, is an online women's magazine. The remaining sources (22%) were specialized (ie, women, patient, or disease-driven) web forums, except for Twitter, Reddit, and YouTube.

Figure 1. Study framework and flowchart of data extraction and analysis. BTM: biterm topic model; MedDRA; Medical Dictionary for Regulatory Activities; Xgboost: extreme gradient boosting.

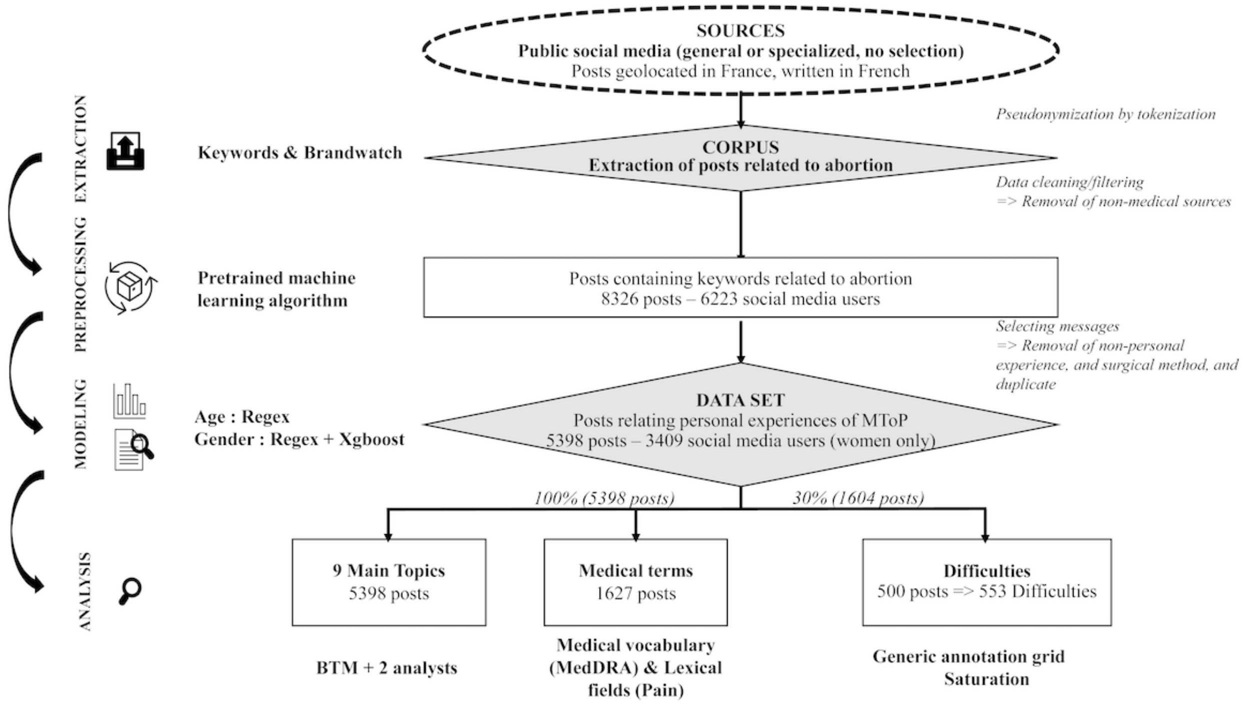


Table 1. List of forums reporting women's medical termination of pregnancy experiences and the top 6 extraction keywords.

Feature	Posts (N=5398), n (%)	Social media users (N=3409), n (%)
Keyword extraction (top 6)		
IVG[a]/IVG medicament[b]	3642 (67.5)	N/A[c]
Avort[d]	761 (14.1)	N/A
GYMISO	350 (6.5)	N/A
MIFEGYNE/MIFEGINE	211 (3.9)	N/A
MISOPROSTOL	188 (3.5)	N/A
MISOONE	111 (2.1)	N/A
Forums		
Specialized forums		
Doctissimo	2218 (41.1)	1018 (29.9)
AuFeminin	1001 (18.5)	768 (22.5)
Journal des femmes	353 (6.5)	241 (7.1)
Babycenter.fr	336 (6.2)	254 (7.5)
Journaldesfemmes.com	86 (1.6)	58 (1.7)
Mamanandco	57 (1.1)	38 (1.1)
Mademoizelle.com	16 (0.3)	15 (0.4)
Magicmaman	15 (0.3)	15 (0.4)
Enceinte.com	14 (0.3)	14 (0.4)
Psychologies	13 (0.2)	13 (0.4)
Fiv.fr	10 (0.2)	10 (0.3)
Beauté test	9 (0.2)	9 (0.26)
Parents.fr	5 (<0.1)	5 (0.15)
Thyroide	2 (<0.1)	2 (<0.1)
Lymphome espoir	2 (<0.1)	2 (<0.1)
Notrefamille	2 (<0.1)	2 (<0.1)
Alexia.fr	2 (<0.1)	2 (<0.1)
Entrepatients.net	1 (<0.1)	1 (<0.1)
General forums		
Facebook	993 (18.4)	743 (21.8)
Twitter	174 (3.2)	136 (4.0)
Reddit	55 (1.0)	34 (1.0)
YouTube	34 (0.6)	29 (0.9)

[a]IVG: French abbreviation for voluntary termination of pregnancy.

[b]medicament: French word for medicine.

[c]N/A: not applicable.

[d]Avort: first letters of the French word *avortement*, which means abortion.

Discussion Topics

From the 5398 posts, 9 topics of interest were identified (Table 2). Personal experience and community support were the most prominent topics. The 7 other topics were as follows (in decreasing order): pain and bleeding, psychological experience, questioned efficacy, social pressure, positive experiences, menstrual cycle disorders, and reported inefficacy.

Table 2. Topics and topic description ranked by frequency.[a]

Rank	Topic	Posts (N=5398), n (%)	Description
1	Personal experience	2413 (44.7)	Users shared personal experiences. They described what they experienced during their medical abortion, the details of the procedure, and what they felt at that moment.
2	Community support	1058 (19.6)	Looking for community support. Some users looked for experiences shared by other users about the procedure to increase their knowledge and to be prepared for it, as well as to feel reassured.
3	Pain and bleeding	797 (14.8)	Seeking for testimonies about pain and bleeding. Highlighted a lack of information on these drug-related adverse events. Users were concerned about what they were about to experience, and they found nonreassuring testimonies on social media.
4	Psychological experience	760 (14.1)	Users expressed regrets and mental outcomes such as depression and emotional distress with short- and long-term consequences. They also reported that medical abortion was "traumatic" and that if they had known they would have chosen surgical abortion.
5	Questioned efficacy	410 (7.6)	Efficacy was questioned.
6	Social pressure	373 (6.9)	The pressure was from the family and mostly from the partner.
7	Positive experience	257 (4.8)	Shared positive experiences with medical abortion.
8	Menstrual cycle disorders	107 (2.0)	Some users reported menstrual cycle disorders following abortion.
9	Reported inefficacy	104 (1.9)	Some users reported inefficacy of the procedure (medical abortion).

[a]A single post may contain several topics.

Medical Terms

Pain was the most frequently reported medical term related to difficulties (Table 3). Bleeding was the second most frequent medical term. Pain and bleeding were reported both during and after medical abortion. Other medical terms reported during and after medical abortion were nausea or vomiting (475/5398, 8.8%) and fatigue. Stress and anxiety were directly associated with the medical abortion procedure, including the fear of abortion inefficacy. Emotional distress, echography, and delayed menstruation were reported both before (when pregnancy was confirmed) and after medical abortion. After the procedure, emotional distress was associated with the feeling of regret and grief.

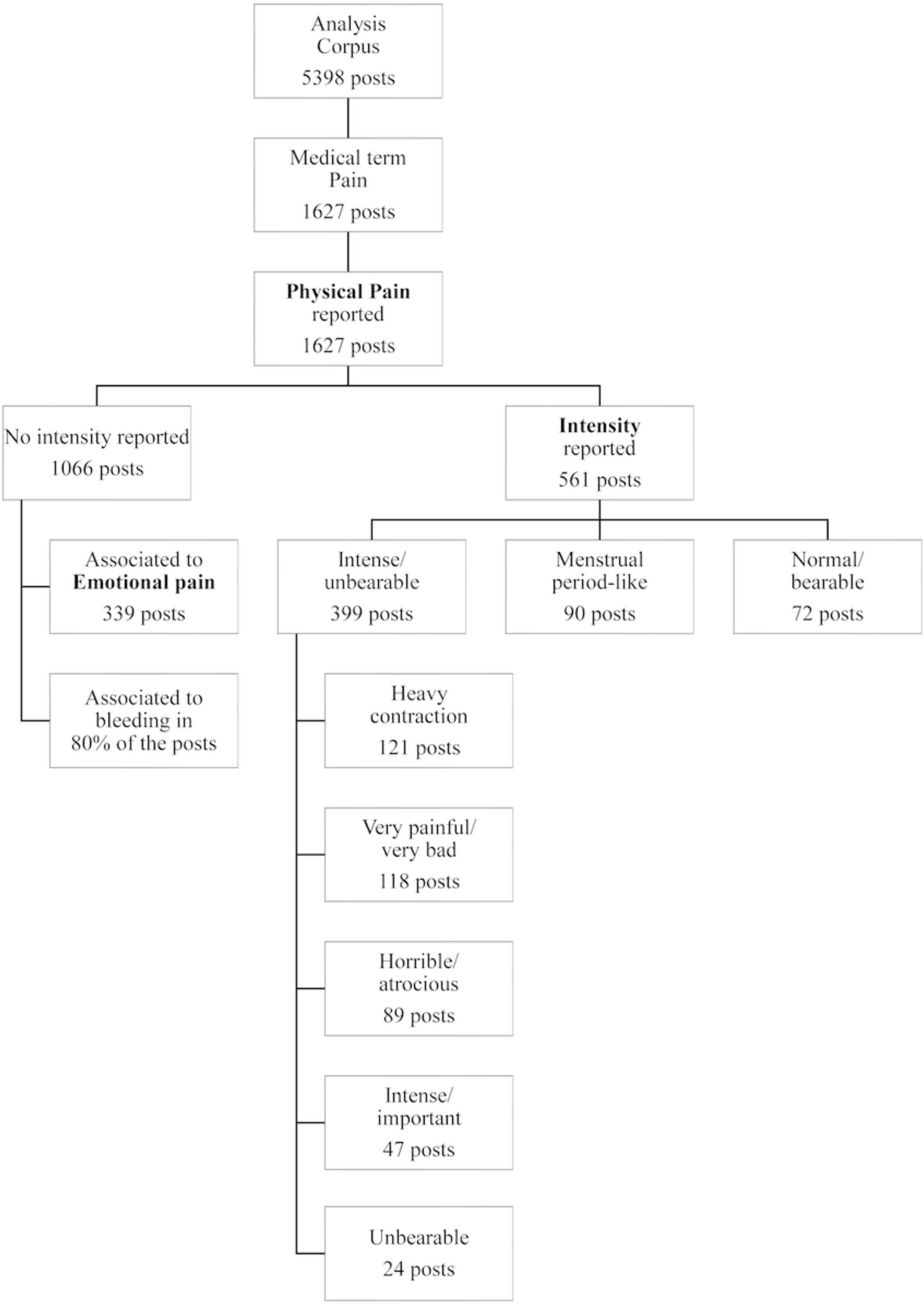
Pain is a multimodal concept with subjectivity, which was reported by 1024 of the 3409 (30.3%) users in 1627 of the 5398 posts (30.1%). Pain usually occurred after the second drug intake (prostaglandin analogs). Using topic modeling, different types of pain were identified, providing details to characterize each type. As a result, two main types of pain were identified: physical and emotional pain. Of the 1627 posts regarding physical pain, the pain intensity was described in 561 posts (34.5%) and was considered severe to unbearable in 399 posts (24.5%) (Figure 2). These rates are not associated with new safety signals in MToP.

Table 3. Most frequently reported medical terms related to difficulties after medical termination of pregancy.[a]

Rank	Medical term	Posts mentioning term (N=5398), n (%)
1	Pain	1627 (30.1)
2	Bleeding	1112 (20.6)
3	Emotional distress	997 (18.5)
4	Echography	492 (9.1)
5	Stress/anxiety	460 (8.5)
6	Fatigue	360 (6.7)
7	Nausea	252 (4.7)
8	Vomiting	223 (4.1)
9	Delayed menstruation	134 (2.5)
10	Grief	114 (2.1)

[a]A single post may contain several medical terms.

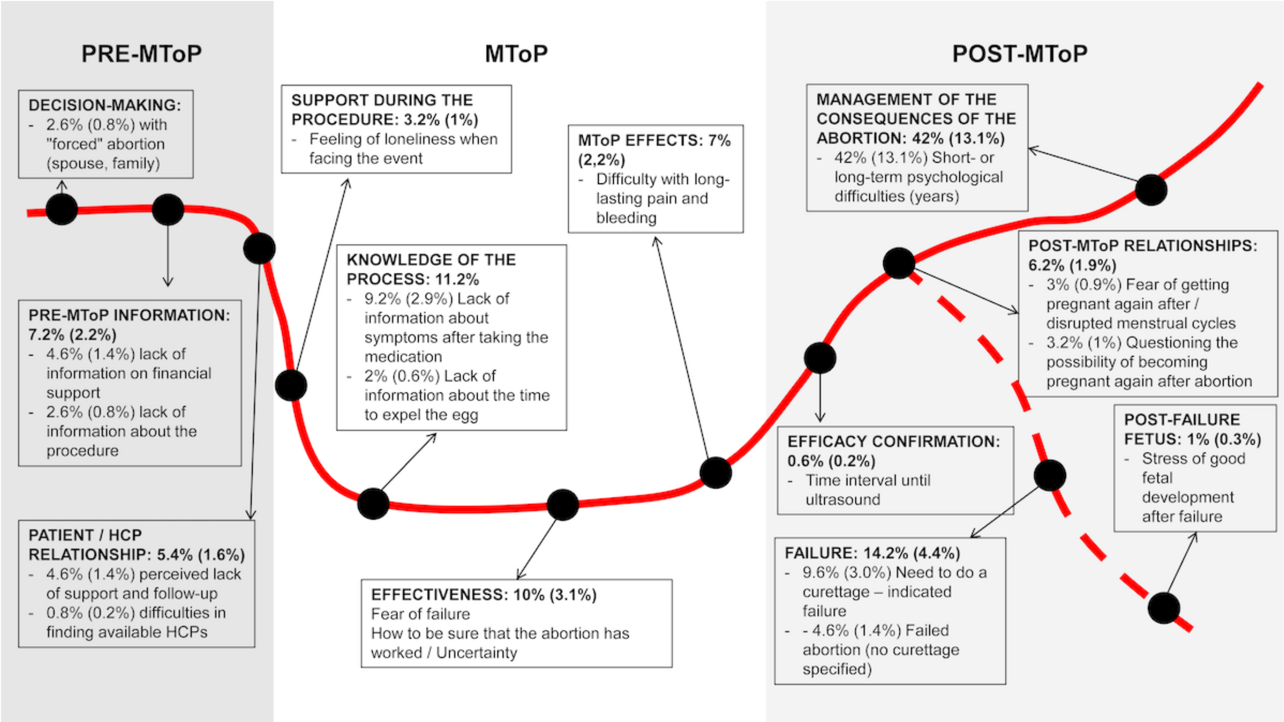
Figure 2. Focus on pain medical terms.



Encountered Difficulties

A total of 553 difficulties were identified from 500 posts derived from a randomized sample containing 1604 posts (30%). Reported difficulties were encountered along the MToP process (Figure 3).

Figure 3. Encountered difficulties. The red line represents the pathway from premedical MToP to post-MToP via MToP. The dashed red line represents the pathway at the end of the procedure in case of failure. Encountered difficulties are summarized in boxes distributed along the whole pathway. HCP: health care professional; MToP: medical termination of pregnancy; x% (y%): number of difficulties among the 553 identified difficulties (number of difficulties among the 1604 posts of the randomized sample).



Lack of information was reported within each stage of the MToP process (before, during, and after medical abortion). Before the procedure, users enquired about reimbursement and the MToP process in detail (ie, drug dosage, route of administration). Details about the MToP process were still misunderstood during the process as users were asking questions about the potential side effects, delay of action, and monitoring. After taking the medications, users were concerned about pain and bleeding, especially for the duration of symptoms, time of occurrence, level of intensity, and types of symptoms. For some women, the pain and bleeding lasted longer than expected (several weeks or even months), and they were looking for community advice. MToP efficacy was a fundamental question during and after the MToP process. Users were wondering how to be sure that the procedure worked, and they expressed fear of failure. They also enquired about embryo expulsion: the appearance, at what moment it was supposed to be expelled, and how to know whether or not it had occurred. They also expressed concerns about surgical management of incomplete abortion. Postabortion sexual life was also of cardinal importance. Future fertility was also a concern: users were wondering when they could get pregnant again and when their menstrual periods would return to normal.

The need for psychological support was quite common after medical abortion. During the procedure, users reported feelings of loneliness. After the procedure, some users mentioned difficulties in overcoming the event, sometimes for a long time.

Overall, 5.4% of the 553 encountered difficulties (n=30) highlighted a lack of health care support during and after the procedure. Users reported that they were looking for health care providers who did not judge them and who could provide them with an environment of mutual trust.

Discussion

Principal Findings

The results of this exploratory study found that infodemiology could help to collect French women's feedback about their MToP experiences. The results showed that women used social media to share their experiences, offer and find support, and provide and receive information regarding medical abortion. The extent of the need for information during and after the MToP procedure suggested that there is still room for improvement.

In the context of shared decision-making in medicine, where both the patient and physician contribute to the process and agree on treatment decisions, the relationship between women and health professionals is crucial and needs to be built up, including beyond the MToP act. To extend this connection, health care professionals can already rely on tools such as leaflets, institutional websites, or mobile apps (eg, chatbots) to answer additional questions from women. This ensures high-quality, standardized, and reachable information. In this study, the need for information did not necessarily mean a lack of information. The discovery of an unintended pregnancy, the idea of terminating it, the fear of stigmatization, and facing medical terms for an unprecedented situation can generate anxiety and ultimately difficulties for the woman in integrating the information provided by the health care professional. Independent of the counseling method, unbiased nondirective information should remain accessible (19% of posts sought advice) [27].

"Pain and bleeding" and "psychological experience" were also among the main topics. The psychological experience linked to the procedure was evoked at all stages of the process, from the pregnancy discovery to the MToP follow-up, showing that some women became apprehensive about this experience. Pain and anxiety were tightly associated; pain (physical and psychological) was mentioned in more than 1 out of 4 posts. This confirmed the previous evidence that some women needed timely counseling and education through this experience. Studies emphasize that listening to and accompanying women is essential [28]. The possibility of verbalizing physical pain could allow women to better bear the pain [29]. To satisfy the need for psychological support, the integration of the contact information of volunteer psychologists in the directories of health care professionals involved in the abortion process could facilitate access to psychological follow-up for women who wish to do so, instead of having to navigating the experience alone through the testimonies of online community members. This nonpharmacological individualized anxiety management could advantageously complete a pharmacological pain relief strategy.

Other encountered difficulties were reported (Figure 2). Fear of failure and its fallouts were mainly mentioned during and after abortion. Once the decision was made, there was an apparent need for reassurance about the success of their action. A timely counsel and education through a health care professional (eg, via telemedicine) or a community could meet this need. In the absence of a patient organization, the online community can offer adequate support. Indeed, pressure from the entourage and loneliness were mentioned in a small percentage of the posts. Abortion is not a neutral topic, and it can be either strongly encouraged or discouraged by the environment. This underlines the importance of meeting the woman alone to ensure the freedom of her choice [28]. When there is a language barrier, it is important to be able to call upon a professional interpreter [28]. In France, the law of March 20, 2017, protects women against disruption of access to abortion as a medical act and misinformation on the abortion procedure, particularly on the internet and social media [30]. Despite this law, no proceedings have led to a conviction so far. The persistent stigmatization of women who have recourse to abortion and the fear of the possible consequences of public exposure to a private and intimate situation may explain this, especially since misinformation is difficult to assess clearly and the law is still often misunderstood [31]. Nevertheless, a study conducted in 2019 in the planning center of a French hospital center (108 women) showed that 36% of women made their decision alone and 68% of women made their decision without difficulty (decision-making was assessed using the Decisional Conflict Scale) [32].

Concerning the psychological effects of MToP, the messages reporting regrets (including in the long term) underlined the importance of providing a caring listening ear (without bias or judgment) to women's requests, and the importance of being able to offer women, when they feel the need, the possibility of psychological support.

Study Strengths

The present results were obtained using data from social media. The use of social media to collect information has several advantages.

First, with, 2.3 billion users voluntarily sharing their data, experience, and outcomes, social media represent the new El Dorado to gather patient feedback [9,18]. Furthermore, the broad variety of social media and the long-term storage of public posts offer access to a large-scale data set allowing focus on specific topics, time periods, and locations. Indeed, the reactivity of social media facilitates carrying out analyses at a given moment and then over time. As such, these data make it possible to quickly measure the impact and acceptance of the implementation of a new health care procedure.

Second, the analysis of social media posts makes an important contribution by generating patient-centered perspectives from an underutilized data source. Our goal was to identify the direct experiences of MToP. Anonymity likely allowed women to express themselves without fear of recognition or judgment in this context. This alternative to in-hospital interviews helps to circumvent any form of white-coat bias [8,9].

Moreover, obtaining data from social media is facilitated by the low acquisition cost. This makes infodemiology an affordable methodology complementing standard clinical methods (ie, clinical studies or surveys), as it enables accessing a large data set while avoiding some of the intrinsic biases of standard methods.

Third, the median age of social media users reported in the study was consistent with the age for abortion in France [2].

Finally, the combined methodology of quantitative analysis and qualitative examination enabled robust characterization of topics, as previously described in peer-reviewed papers. This proven study type helps to give a voice to women experiencing MToP with limited background noise on this topic.

Study Limitations

First, our study is subject to the inherent limitations of all infodemiology studies. One of these limitations results from the fact that, despite an abundant amount of data available, worldwide regulation prevented us from extracting posts from private forums/groups or those that are exchanged directly between users. Moreover, not all social media users are active. van Mierlo et al [33] estimated that approximately 90% of social media users are observers and do not actively participate in content creation; only 9% contribute sparingly and 1% create most of the content.

Another limitation is the variability of the level of contribution according to age (young people express themselves more than other age groups), gender (women express themselves more than men), country, socioprofessional class, and other factors [9,34]. In our study, as the data collected via social media were not representative of the population, there is a limitation in generalizing the findings to the whole French population of women who have experienced MToP [35].

Since the data were issued from the internet, our study could also present recall bias. Social media users tend to more frequently verbalize negative rather than positive experiences (ie, recall bias). This could lead to an overrepresentation of negative observations related to MToP in our study. It should also be noted that data published on social media could be deleted or modified, limiting the reproducibility of the results. The quality of the data collected was very heterogeneous and varied among social media users. Verifying the accuracy of published data is challenging due to the anonymity offered by social media. Content bots or users pretending to be others could have created some of the analyzed content.

Moreover, our analysis was based on the spontaneous testimonies of social media users on a single topic of interest and according to their feelings (subjectivity). The media or influential people could direct the discussions and encourage a peak of comments at a given moment (eg, the change of legislation around abortion in the United States).

Finally, due to variations among clinical practices and cultural differences, the conclusions of our study are not reproducible in different countries and regions.

Conclusion

This exploratory study showed the added value of infodemiology. Applied to medical abortion, the results indicate that French women who underwent an MToP used social media to document their experiences, offer and find support, and provide and receive information regarding the procedure. This suggests that there is still room for improvement during and after the process, particularly in providing women with the opportunity to be properly informed, be listened to, and express themselves.

Acknowledgments

The authors thank Anne-Lise Rossi and Laura Levitte (Nordic Pharma SAS) for their contributions in the study conception, study design, and manuscript review. They also thank the team of Kap Code, a contract research organization startup specializing in medical social network analysis, and in particular Adel Mebarki, for their contributions in the study conception and conduct (data extraction and analysis). Thanks also go to Camille Foottit and Fabienne Péretz (Abelia Science) for their help in writing the manuscript. The article was written using a collaborative writing tool (MyPubli.online).

Data Availability

According to French law, the data sets can only be shared through controlled access. Please contact Kap Code for any questions.

Conflicts of Interest

This study, including editorial assistance, was funded by Nordic Pharma France. The authors were not paid for the publication. GG, LA, PF, and AA report fees for board membership from Nordic Pharma, consultant and speaker fees from Nordic Pharma, and participation in congresses on invitation by Nordic Pharma. PV is employed by the contract research organization Kap Code.

Multimedia Appendix 1
Keywords for data extraction.
[PNG File , 301 KB - infodemiology_v4i1e49335_app1.png]

Multimedia Appendix 2
Data extraction filters.
[PNG File , 33 KB - infodemiology_v4i1e49335_app2.png]

References

1. Bearak J, Popinchalk A, Ganatra B, Moller A, Tunçalp Ö, Beavin C, et al. Unintended pregnancy and abortion by income, region, and the legal status of abortion: estimates from a comprehensive model for 1990-2019. Lancet Glob Health 2020 Sep;8(9):e1152-e1161 [FREE Full text] [doi: 10.1016/S2214-109X(20)30315-6] [Medline: 32710833]
2. Le nombre des interruptions volontaires de grossesse augmente en 2022. Direction de la Recherche, des Études, de l'Évaluation et des Statistiques. 2023. URL: https://drees.solidarites-sante.gouv.fr/sites/default/files/2023-09/ER1281MAJ.pdf [accessed 2024-02-15]
3. Map of Mifepristone Approvals. Gynuity Health Projects. 2017. URL: https://gynuity.org/resources/map-of-mifepristone-approvals [accessed 2023-04-21]
4. Mallaury V, Steve N, Geneviève V, Olivier B, Florence B, Aubert A. Impact of the COVID-19 pandemic and the emergency measures on abortion care taken during this period in a French region (Provence Alpes Côte d'Azur). J Gynecol Obstet Hum Reprod 2022 Dec;51(10):102478 [FREE Full text] [doi: 10.1016/j.jogoh.2022.102478] [Medline: 36108936]
5. 40ème anniversaire de la loi sur l'IVG. Assemblée Nationale. 2015. URL: https://www2.assemblee-nationale.fr/14/evenements/2015/anniversaire-loi-veil#node_9805 [accessed 2023-04-21]

6. Rapport public d'évaluation Mifégyne 200 mg, comprimé (mifépristone). Agence Française de Sécurité Sanitaire des Produits de Santé. 2007. URL: https://archive.ansm.sante.fr/var/ansm_site/storage/original/application/776b24988710e52725e3584edb4cdc2a.pdf [accessed 2023-04-21]
7. IVG médicamenteuse jusqu'à 9 semaines: pérenniser ce droit pour les femmes. Haute Autorité de Santé (HAS). 2021 Apr 12. URL: https://tinyurl.com/mr58pvnz [accessed 2024-01-30]
8. Divay S. L'avortement: une déviance légale. Deviance Société 2004;28(2):195-209. [doi: 10.3917/ds.282.0195]
9. Donegan K, Ovelgonne H, Flores G, Fuglerud P, Georgescu A. Social media and m-Health data. Subgroup report. European Medicines Agency. URL: https://www.ema.europa.eu/en/documents/report/social-media-m-health-data-subgroup-report_en.pdf [accessed 2023-04-21]
10. Audience Internet Global en Février 2024. Médiamétrie. 2024 Mar 26. URL: https://www.mediametrie.fr/fr/audience-internet-global-en-fevrier-2024 [accessed 2024-04-05]
11. Cordonier L. Information et santé. Analyse des croyances et comportements d'information des Français liés à leur niveau de connaissances en santé, au refus vaccinal et au renoncement médical. Fondation Descartes. 2023. URL: https://www.fondationdescartes.org/2023/10/information-et-sante/ [accessed 2024-01-30]
12. McDonald L, Malcolm B, Ramagopalan S, Syrad H. Real-world data and the patient perspective: the PROmise of social media? BMC Med 2019 Jan 16;17(1):11 [FREE Full text] [doi: 10.1186/s12916-018-1247-8] [Medline: 30646913]
13. Wiedemann G. Opening up to big data: computer-assisted analysis of textual data in social sciences. Forum Qualitative Sozialforschung 2013 May 25;14(2):23. [doi: 10.17169/fqs-14.2.1949]
14. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med 2002 Dec;113(9):763-765. [doi: 10.1016/S0002-9343(02)01473-0]
15. Perone H, Herweck A, Stump H, Levine H, Wong A, Carugno J. The virtual infertility community: a qualitative analysis of patient experiences shared on Instagram. J Assist Reprod Genet 2021 Mar;38(3):613-620 [FREE Full text] [doi: 10.1007/s10815-020-02028-6] [Medline: 33411326]
16. Mercier R, Senter K, Webster R, Henderson Riley A. Instagram users' experiences of miscarriage. Obstet Gynecol 2020 Jan;135(1):166-173. [doi: 10.1097/AOG.0000000000003621] [Medline: 31809440]
17. Vazquez Corona M, Betrán AP, Bohren M. The portrayal and perceptions of cesarean section in Mexican media Facebook pages: a mixed-methods study. Reprod Health 2022 Feb 22;19(1):49 [FREE Full text] [doi: 10.1186/s12978-022-01351-8] [Medline: 35193590]
18. Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What patients can tell us: topic analysis for social media on breast cancer. JMIR Med Inform 2017 Jul 31;5(3):e23 [FREE Full text] [doi: 10.2196/medinform.7779] [Medline: 28760725]
19. 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]
20. Kapp N, Lohr PA. Modern methods to induce abortion: safety, efficacy and choice. Best Pract Res Clin Obstet Gynaecol 2020 Feb;63:37-44. [doi: 10.1016/j.bpobgyn.2019.11.008] [Medline: 32029379]
21. Industry leading social media management solution. Brandwatch. URL: https://www.brandwatch.com/suite/social-media-management/ [accessed 2023-04-21]
22. Talmatkadi M, Foulquié P, Déguilhem A, Renner S, Châteauneuf L, Mebarki A, et al. Création d'un algorithme d'identification d'expériences vécues par des patients ou leurs proches à partir de messages issus des réseaux sociaux : un cas d'usage sur le COVID long. Rev Épidémiol Santé Publique 2022 Nov;70:S265. [doi: 10.1016/j.respe.2022.09.045]
23. Schäfer F, Faviez C, Voillot P, Foulquié P, Najm M, Jeanne J, et al. Mapping and modeling of discussions related to gastrointestinal discomfort in French-speaking online forums: results of a 15-year retrospective infodemiology study. J Med Internet Res 2020 Nov 03;22(11):e17247 [FREE Full text] [doi: 10.2196/17247] [Medline: 33141087]
24. Voillot P, Riche B, Portafax M, Foulquié P, Gedik A, Barbarot S, et al. Social media platforms listening study on atopic dermatitis: quantitative and qualitative findings. J Med Internet Res 2022 Jan 28;24(1):e31140 [FREE Full text] [doi: 10.2196/31140] [Medline: 35089160]
25. Renner S, Loussikian P, Foulquié P, Arnould B, Marrel A, Barbier V, 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 [FREE Full text] [doi: 10.2196/37518] [Medline: 36125861]
26. Article R1121-1. French Public Health Code. 2021. URL: https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000034696952/
27. Baum SE, Ramirez AM, Larrea S, Filippa S, Egwuatu I, Wydrzynska J, et al. "It's not a seven-headed beast": abortion experience among women that received support from helplines for medication abortion in restrictive settings. Health Care Women Int 2020 Oct;41(10):1128-1146. [doi: 10.1080/07399332.2020.1823981] [Medline: 33156737]
28. Attali L. Psychological aspects of abortion. J Gynecol Obstet Biol Reprod 2016 Dec;45(10):1552-1567. [doi: 10.1016/j.jgyn.2016.09.030] [Medline: 28029465]
29. Attali L, Guceve E, Schoch F, Bettahar K, Warynski F. Histoires d'IVG, Histoires de femmes. In: Recherches Familiales. Paris: Vuibert; Jan 2021:144.

30. Loi du 20 mars 2017 relative à l'extension du délit d'entrave à l'interruption volontaire de grossesse. Vie publique. 2017. URL: https://www.vie-publique.fr/loi/20989-ivg-extension-du-delit-dentrave-linterruption-volontaire-de-grosses [accessed 2023-04-21]
31. Mathieu M. L'avortement en France: du droit formel aux limites concrètes à l'autonomie des femmes. Droit Société 2022;2(111):335-355. [doi: 10.3917/drs1.111.0335]
32. Troccard-Fourcade AC. Évaluation de la certitude décisionnelle des patientes se présentant pour la réalisation d'une interruption volontaire de grossesse au centre de planification du centre hospitalier de la Côte Basque. Médecine Humaine et Pathologie. 2019. URL: https://dumas.ccsd.cnrs.fr/dumas-02487399/document [accessed 2023-04-21]
33. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014 Feb 04;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966] [Medline: 24496109]
34. Vigoureux S. Epidemiology of induced abortion in France. J Gynecol Obstet Biol Reprod 2016 Dec;45(10):1462-1476. [doi: 10.1016/j.jgyn.2016.09.024] [Medline: 27793490]
35. Leng HK. Methodological issues in using data from social networking sites. Cyberpsychol Behav Soc Netw 2013 Sep;16(9):686-689. [doi: 10.1089/cyber.2012.0355] [Medline: 23679568]

Abbreviations

BTM: biterm topic model
MToP: medical termination of pregnancy
TF-IDF: term-frequency inverse document frequency
XGBoost: extreme gradient boosting

Edited by T Mackey; submitted 30.05.23; peer-reviewed by G Li, C Li; comments to author 09.11.23; revised version received 20.02.24; accepted 06.03.24; published 02.05.24.

Please cite as:
Gouy G, Attali L, Voillot P, Fournet P, Agostini A
Experiences of Women With Medical Abortion Care Reflected in Social Media (VEILLE Study): Noninterventional Retrospective Exploratory Infodemiology Study
JMIR Infodemiology 2024;4:e49335
URL: https://infodemiology.jmir.org/2024/1/e49335
doi:10.2196/49335
PMID:38696232

Original Paper

The Role of Scientific Research in Human Papillomavirus Vaccine Discussions on Twitter: Social Network Analysis

Geneviève Jessiman-Perreault[1], BA, MSc; Jean-Christophe Boucher[2], PhD; So Youn Kim[2], MDSA; Nicole Frenette[1], MSW; Abbas Badami[2], MEng; Henry M Smith[2], MA; Lisa K Allen Scott[1,3,4], PhD

[1]Alberta Health Services, Calgary, AB, Canada

[2]School of Public Policy, University of Calgary, Calgary, AB, Canada

[3]Department of Oncology, University of Calgary, Calgary, AB, Canada

[4]Department of Community Health Sciences, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Geneviève Jessiman-Perreault, BA, MSc
Alberta Health Services
2210 2 St SW
Calgary, AB, T2S 3C3
Canada
Phone: 1 587 794 4232
Email: genevieve.jessiman-perreault@albertahealthservices.ca

Abstract

Background: Attitudes toward the human papillomavirus (HPV) vaccine and accuracy of information shared about this topic in web-based settings vary widely. As real-time, global exposure to web-based discourse about HPV immunization shapes the attitudes of people toward vaccination, the spread of misinformation and misrepresentation of scientific knowledge contribute to vaccine hesitancy.

Objective: In this study, we aimed to better understand the type and quality of scientific research shared on Twitter (recently rebranded as *X*) by vaccine-hesitant and vaccine-confident communities.

Methods: To analyze the use of scientific research on social media, we collected tweets and retweets using a list of keywords associated with HPV and HPV vaccines using the Academic Research Product Track application programming interface from January 2019 to May 2021. From this data set, we identified tweets referring to or sharing scientific literature through a Boolean search for any tweets with embedded links, hashtags, or keywords associated with scientific papers. First, we used social network analysis to build a retweet or reply network to identify the clusters of users belonging to either the vaccine-confident or vaccine-hesitant communities. Second, we thematically assessed all shared papers based on typology of evidence. Finally, we compared the quality of research evidence and bibliometrics between the shared papers in the vaccine-confident and vaccine-hesitant communities.

Results: We extracted 250 unique scientific papers (including peer-reviewed papers, preprints, and gray literature) from approximately 1 million English-language tweets. Social network maps were generated for the vaccine-confident and vaccine-hesitant communities sharing scientific research on Twitter. Vaccine-hesitant communities share fewer scientific papers; yet, these are more broadly disseminated despite being published in less prestigious journals compared to those shared by the vaccine-confident community.

Conclusions: Vaccine-hesitant communities have adopted communication tools traditionally wielded by health promotion communities. Vaccine-confident communities would benefit from a more cohesive communication strategy to communicate their messages more widely and effectively.

(JMIR Infodemiology 2024;4:e50551) doi:10.2196/50551

KEYWORDS

human papillomavirus; HPV; vaccine; immunization; social media; misinformation; social network analysis

Introduction

Background

Cervical cancer is one of the most preventable types of cancer in the world. Almost all cases are attributable to human papillomavirus (HPV), for which an effective vaccine exists [1]. Part of the global strategy to eliminate cervical cancer includes fully vaccinating 90% of girls with the HPV vaccine by the age of 15 years [2]. However, the global HPV immunization coverage currently remains suboptimal [3]. While many countries are experiencing vaccine supply issues, even high-income countries with reliable vaccine supply and comprehensive school-based programs are still failing to meet vaccine targets, largely due to vaccine hesitancy [4].

Studies show that people now search the web for health information more often than they talk to health professionals about these matters [5]. The popularity of social media platforms has also created a phenomenon wherein people not only use the web to access health information but also play an active role in cocreating the information and ideas (in the form of opinions, anecdotes, and links to other sources of information) that they encounter in these web-based spaces [6]. Social media spaces create an important setting for people to interact and for communities to emerge, as they are not geographically bound but rather reflect patterns of shared interests, purpose, or identities [7]. As such, vaccine-confident and vaccine-hesitant groups represent distinctive ideologies and create distinctive web-based communities. The distinction between these 2 groups lies in their attitudes, beliefs, and behaviors associated with vaccine decision-making, in that vaccine-confident groups reflect public trust in vaccines and the evidence supporting their efficacy, effectiveness, and safety, which leads to their uptake of recommended vaccines. Vaccine-hesitant groups, for their part, tend to doubt this information, demonstrated by their reluctance or refusal to receive recommended vaccines [8,9].

Despite a large body of evidence demonstrating the safety and efficacy of the HPV vaccine [10,11], attitudes toward the vaccine and the accuracy of information shared about this topic in web-based settings vary markedly from extremely negative and erroneous to supportive and factually accurate [12]. In addition, in recent years, there has been a rapid increase in the accessibility of scientific journals and subsequent dissemination of scientific findings through social media [13]. Simultaneously, there has been a decline in the role of unbiased science journalists and other communication experts as mediators between scientists and the public [14]. While these changes have had a democratizing effect on scientific knowledge and allowed for better communication between scientific communities and the public, this unfiltered access to scientific research also creates an environment where individuals may have difficulty in differentiating valid and credible information from biased and unreliable information or may misinterpret legitimate findings [15]. In contrast, researchers have also noted that the growth of open science can create opportunities for people to discuss novel research across polarized boundaries [16], but the type and quality of scientific research about HPV vaccination that is being shared in web-based discussions is unknown. Finally, with a wealth of open-access scientific research available, there are concerns about how ideologically motivated communities, such as vaccine-hesitant groups, integrate scientific knowledge into their social media communication strategies to amplify uncertainty around vaccines [17]. It is prudent to investigate how scientific research is integrated into web-based HPV vaccine discussions, given that web-based information is typically considered to be more credible, reliable, and authoritative if supported by scientific citation, notwithstanding the source of journal, authorship, or other features [18].

Twitter (recently rebranded as *X*; as data collection occurred before the rebrand, we will be using its former name throughout this paper) is one of the largest, most popular, and most influential social media platforms in the world. Twitter has also traditionally been a preferred source of public opinion data for applied public health research [19-22]. This is because social media feeds such as Twitter offer an avenue for continuous, near-real–time collection of unsolicited information generated by many individuals regarding a variety of topics of interest [23,24]. Several studies have recently demonstrated the benefits of leveraging social media over traditional methods such as surveys as a source of primary data for health promotion interventions, including those aimed at increased participation in HPV immunization programs [25].

Objectives

Exposure to web-based discussions about HPV immunization on Twitter, regardless of geographic location, may influence peoples' attitudes toward the vaccine [22,26,27]. Thus, there is significant interest among public health professionals to better understand how scientific knowledge about HPV immunization is wielded on Twitter, both to understand the impact of scientific knowledge on vaccine hesitancy and to identify opportunities for novel interventions aimed at countering or debunking misinformation and supporting increased uptake of the HPV vaccine [6,28]. Therefore, in this study, we aimed to do the following:

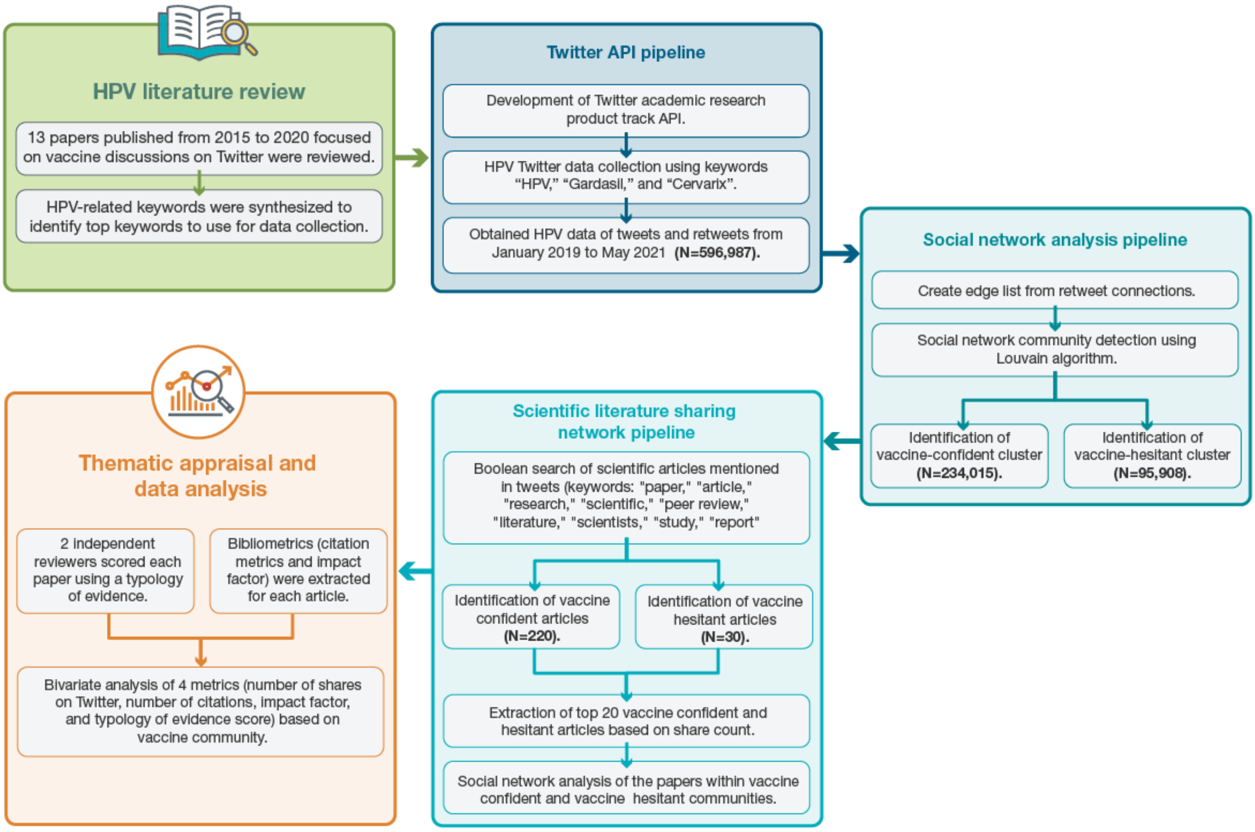
1. Describe and visualize the vaccine-hesitant and vaccine-confident communities' patterns of sharing HPV vaccination–related scientific literature on Twitter
2. Thematically analyze the scientific literature shared by both vaccine-hesitant and vaccine-confident communities using a typology of research evidence
3. Determine whether there are differences in shares, quality of evidence, and other bibliometric indicators of the scientific literature shared by each community

Methods

Overview

Our methods followed a multistep process. First, we conducted a rapid review to inform HPV and HPV vaccine keywords. Second, we used these keywords to filter tweets and create a data set. Third, we detected vaccine-confident and vaccine-hesitant communities and generated social network maps of each community based on tweets and retweet. Fourth, we detected the mentions of scientific literature in each community and extracted those papers for future statistical and social network analysis. A summary of these methods is presented in Figure 1 (adapted from the paper by Elyashar et al [29]), and further details are presented in the following sections.

Figure 1. Summary of the study methods. API: application programming interface; HPV: human papillomavirus.



Literature Review to Inform Data Collection

To determine the most applicable keywords to guide this study, a rapid review was first conducted to determine the most frequently used keywords in literature focused on HPV and HPV immunization discourse on Twitter. The rapid review methodology was selected due to its efficiency in synthesizing a large volume of information in a timely yet systematic manner [30]. This review yielded 13 papers published between 2015 and 2020 about the topic of HPV immunization discussions on social media, with 11 (85%) focusing on HPV immunization discussions on Twitter specifically. We extracted the keywords used in each paper to filter content on social media (Textbox 1). Then, we synthesized these keywords to compile a list of the most used keywords to represent HPV and HPV vaccine discussions on social media, and the top 3 keywords were used to generate the data set.

Textbox 1. Papers yielded from the rapid review and the associated human papillomavirus (HPV)–related and HPV vaccine–related keywords.

Papers and keywords

- Shapiro et al [31]
  - “Gardasil,” “Cervarix,” “HPV AND vaccin*,” and “cervical AND vaccin*”
- Massey et al [32]
  - “HPV,” “HPV vaccine,” “HPV shot,” “Gardasil,” and “Cervarix” (and hashtag equivalents)
- Keim-Malpass et al [33]
  - “#HPV” and “#Gardasil”
- Du et al [21]
  - “HPV,” “human papillomavirus,” “Gardasil,” and “Cervarix”
- Nelon et al [34]
  - “#vaccines,” “#vaccine,” “#vaccinations,” and “#vaccination”
- Surian et al [35]
  - “HPV AND vaccine,” “HPV AND vaccination,” “Gardasil,” “cervical AND vaccination,” “cervical AND vaccine,” and “Cervarix”
- Zhou et al [36]
  - “HPV,” “vaccine,” “Gardasil,” “Cervarix,” “vaccination,” “cervical,” and “cancer”
- Becker et al [37]
  - “Pentavalent OR pentavac OR quinvaxem”
- Dyda et al [38]
  - “Cervical,” “Cervarix,” “HPV,” “human papillomavirus,” “vaccine,” “vaccination,” and “Gardasil”
- Chakraborty et al [20]
  - “HPV,” “papilloma,” “pappiloma,” “papiolma,” “papillomavirus,” “Gardasil,” “Gardisil,” “Guardisil,” “Guardasil,” “Cervarix,” “cervical shot,” “cervical shots,” “cervical vaccine,” “cervical vaccines,” “cervical vax,” “cervical vaxine,” “cervical vaxines,” “cervical vaxx,” “cervical vaxxine,” “cervical vaxxines,” “cervical vaccination,” and “cervical vaccinations”
- Dunn et al [39]
  - “Gardasil,” “Cervarix,” “HPV AND vaccine,” and “cervical AND vaccin”
- Budenz et al [40]
  - “HPV,” “HPV vaccine,” “HPV shot,” “Gardasi,” and “Cervarix” (and hashtag equivalents)
- Zhang et al [41]
  - “Cervarix,” “Gardasil,” “HPV,” “human papillomavirus,” “Gardasil,” “HPV AND vaccin*,” and “cervical AND vaccin*”

Data Collection

Using 3 of the most common keywords that emerged from the initial rapid review (“HPV” OR “Gardasil” OR “Cervarix”), a data set of tweets and retweets was created (N=596,987). Then, tweets were collected using the Academic Research Product Track application programming interface (API) from January 2019 to May 2021 [42]. Data were collected using the Twitter API Python wrapper (Python Software Foundation, version 3.8.5) [43]. The construction of the API, data collection, and data processing (ie, importing, exporting, and filtering of data) were performed in Python [44].

Ethical Considerations

This study received an exemption from ethics approval as determined by The Conjoint Faculties Research Ethics Board at the University of Calgary. This was due to its use of only publicly available information from an existing data set. Furthermore, the published results have omitted all identifiable information and are only presented in aggregate form.

Social Network Analysis

First, we created a social network of accounts by creating an edge list using retweets. The retweet edge list consisted of nodes representing individual Twitter accounts and edges representing accounts that are being retweeted. The individual Twitter accounts were identified using the "username" information from the API, and the source of the retweet account information was extracted using the account mentions beside the "RT" in the tweets' text in our data set. Our data set consisted of 57,109 retweets and 25,898 original or quoted tweets. Retweet networks were analyzed as they are found on aggregate to better reflect agreement among users and thus represent an ideological community on issues such as vaccination [45]. Second, we used a Louvain modularity method to classify subclusters of web-based communities in the resulting social network [46]. This method was chosen because the algorithm was designed to accurately detect subcommunities within large networks and operate fast computationally. Third, the social network analysis map also illustrated a strong polarization of the subclusters. Through this polarization and the identification of primary influencers within a subcommunity, the vaccine-confident (n=234,015) and vaccine-hesitant (n=95,908) web-based communities were identified. The primary influencers were detected by measuring the degree centrality, which is the measure of the number of connections each user has within the network. Thus, the accounts with the highest measure of degree centrality were categorized as primary influencers, as a high degree centrality demonstrates a high number of connections an account has within the network. These primary influencers, along with the content of the account's bio descriptions and tweets, were qualitatively studied to examine their expressed positions regarding HPV vaccination. Edge list was constructed using Python, and the retweet social network analysis was conducted using Gephi- (Gelphi, version 0.9.2) [47].

Scientific Literature Sharing Network Analysis

From the vaccine-confident and vaccine-hesitant data sets, we identified tweets that either mentioned or shared scientific literature through a Boolean search for tweets with an embedded http secure link or any of the select list of words ("paper," "article," "research," "scientific," "peer review," "literature," "scientists," "study," and "report") [48]. This filter identified 220 papers from the vaccine-confident community and 30 papers from the vaccine-hesitant community. The titles of or links to these papers were extracted from the data set along with associated metrics such as number of shares for further analysis (as described in the *Data Analysis* section). We identified the top 20 most shared scientific publications in these respective communities. We chose to identify the top 20 most shared scientific publications due to the proportion of shares that these papers had—accounting for >97% of shares in the vaccine-hesitant community and approximately 61% in the vaccine-confident community. Then, we repeated the social network analysis steps by creating a retweet network of accounts sharing the top 20 prominent scientific publications within the vaccine-confident and vaccine-hesitant communities. The edge list for the vaccine-confident community comprised 989 nodes and 1013 edges, whereas the vaccine-hesitant group had 355 nodes and 422 edges. The primary influencers in this network were again identified using degree centrality measures, and we qualitatively analyzed these accounts on Twitter through their Twitter bio descriptions. The social network analysis of the scientific papers was conducted using Gephi (version 0.9.2) [47].

Typology of Evidence for Thematic and Critical Appraisal

Overall, 2 members of the research team (GJP and NF) with subject area expertise in HPV immunization independently reviewed all scientific papers from each network using a typology of evidence, proposed by Gray [49], based on the suitability of the study design for the research question posed. This typology was determined to be the most appropriate and feasible approach to critically appraise the scientific papers because it allowed for the ability to schematically differentiate between diverse study designs (from in vivo to clinical trials and reviews). First, we classified the objective, research question, or aim of the study based on 9 categories that were used to classify research papers based on the typology by Gray [49] (presented in the first column of Table 1). Next, we classified each paper according to the study design. On the basis of these 2 metrics, a score ranging from 0 to 2 was assigned to each paper, where 0 indicates the least appropriate study design for the research question posed and 2 indicates the most appropriate design for the research question posed (refer to Table 1 for details about the scoring of the typology of evidence). The same 2 members of the research team compared their classifications and scoring, and if consensus could not be reached, a third member of the research team (LKAS) made the final decision. In addition, we extracted information about the characteristics of the paper (study design, research question, or objective), journal (journal name and year published), and author (names, affiliations, and conflicts of interest; refer to Multimedia Appendices 1 [50-70] and 2 [52,71-89] for results of the top 20 most shared papers obtained from the vaccine-confident and vaccine-hesitant communities). These data were used to conduct bibliometric analyses of the journal and descriptive analysis of the research content shared by each community, which are further described in the following sections.

Table 1. A typology of evidence (example questions in columns refer to human papillomavirus [HPV] vaccination for the prevention of cancer) based on appropriateness of study design for the research question posed (adapted from the papers by Gray [49] and Petticrew and Roberts [90]).

	In vivo and in vitro studies	Qualitative research	Cross-sectional survey	Case-control studies	Cohort studies	RCTs[a]	Quasi-experimental studies	Nonexperimental evaluations	Scoping reviews and narrative reviews
Effectiveness (does this work? does doing this work better than doing that?)	0	0	0	0	1	2	1	0	2
Process of service delivery (how does it work?)	0	2	1	0	0	0	0	1	2
Salience (does it matter?)	0	2	2	0	0	0	0	0	2
Safety (will it do more harm than good?)	0	1	0	1	1	2	1	1	2
Acceptability (will the focus population be willing to or want to take up the HPV vaccine?)	0	2	1	0	0	1	1	1	2
Cost-effectiveness (is it worth delivering this service?)	0	0	0	0	0	2	0	0	2
Appropriateness (is this the right service for this population?)	0	2	2	0	0	0	0	0	1
Satisfaction with the service (is this population satisfied with the service?)	0	2	2	1	1	0	0	0	0
Basic science (what is the cellular mechanism of action?)	1	0	0	0	0	0	0	0	0

[a]RCT: randomized controlled trial.

Bibliometric Indicators

Traditionally, the prestige and quality of a journal was evaluated using citation metrics such as impact factor [91]. In the past few years, as assessment of scientific information has grown exponentially, new tools have been developed to capture the visibility and reach of web-based scientific information. Examples of these alternative metrics or altmetrics include likes, shared tweets, and retweets [92]. To compare traditional scholarly measures of quality to altmetrics, we collected data about the number of times the paper was shared by each vaccine community and the impact factor of the journal the paper was published in. We also collected data about the number of citations each shared paper had received through Google Scholar. Given that citations are impacted by the length of time since publication, we used the SCImago Journal Ranking (SJR) indicator, which provides a weighted average score that remains consistent each year and accounts for the prestige of the citing journal and the differences across subject fields, allowing for more equal comparisons across subject fields [93]. Each paper was assigned an SJR indicator, where a lower score indicates lower-ranking journals and higher scores indicate higher-ranking journals [94]. Journals that were not indexed in the Scopus database were not assigned an SJR score and were marked as missing in our database. These metrics were used to assess the influence of the shared papers in scientific research and the prestige of the journal the shared papers were published in.

Data Analysis

Once these bibliometrics and typology-of-evidence scores were collected in a data set, basic descriptive results of these 4 metrics (number of shares on Twitter, number of citations, impact factor, and typology of evidence score) were calculated using median and IQR, given their skewed distributions. We also performed the Mann-Whitney *U* test, given the nonnormal distribution of these data [95], to determine whether there were statistically significant differences in the 4 indicators between the papers shared in the vaccine-hesitant and vaccine-confident communities. The four indicators examined were (1) the number of shares that the original tweet sharing the publication on Twitter received, (2) the SJR score of the journal the paper was published in, (3) the number of citations the paper received, and

(4) the typology of evidence score that the paper received. Statistical significance was determined using *P* value <.05. Effect size was calculated using Cohen *d*, where a standardized difference of 0.2 indicates a small difference, difference of 0.5 indicates a medium difference, and difference of 0.8 indicates a large difference [96]. All data analyses were conducted using SAS Studio (SAS Institute, version 3.6).

Results

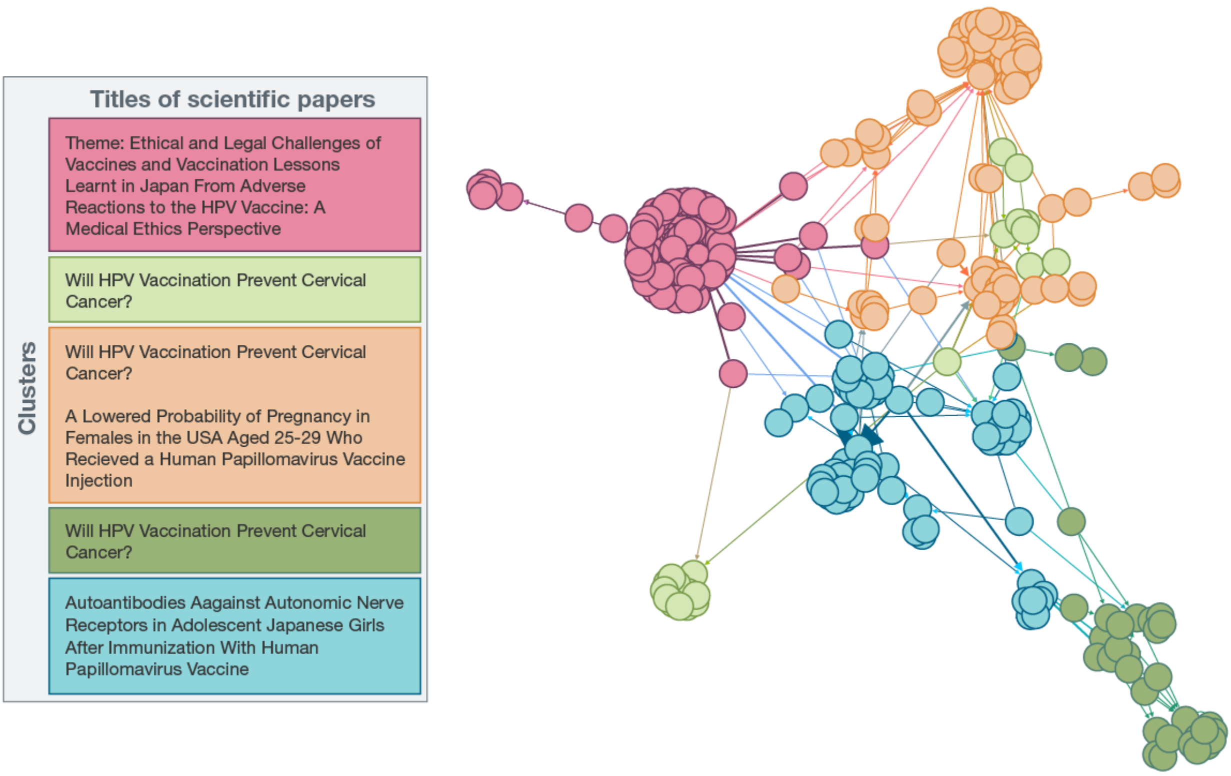
Overview

In total, 250 scientific papers (n=30, 12% in the vaccine-hesitant community and n=220, 88% in the vaccine-confident community) shared between January 2019 and May 2021 were identified. These papers received a combined total of 2247 shares on Twitter, with 562 (25.01%) shares for vaccine-hesitant papers and 1685 (74.99%) shares for vaccine-confident papers. On average, vaccine-hesitant papers received approximately 19.2 (SD 35.6) shares, whereas vaccine-confident papers received approximately 7.7 (SD 30.5) shares. Of these 250 scientific papers, the top 20 most shared papers from each vaccine community were used to produce a social network map of all tweets interacting with or sharing scientific papers about the HPV vaccine on Twitter (Multimedia Appendix 3).

Vaccine-Hesitant Social Network

Figure 2 presents the social network of all tweets sharing or interacting with tweets discussing scientific papers among the vaccine-hesitant community. As can be seen in Figure 2, the retweet network of scientific literature in the vaccine-hesitant community can be categorized into 5 distinct subclusters. Accounts associated with the red cluster shared papers focusing on the safety and ethical considerations around vaccination, with a journalist from a conservative news network emerging as the most influential account holder in this cluster. The most commonly shared paper in this cluster was a case study about the safety of the HPV vaccine in the context of alleged adverse reactions to the HPV vaccine in Japan [50]. In the light green cluster, 1 particular influencer, whose account was later suspended by Twitter, was similarly influential by sharing a paper focused on the effectiveness of HPV vaccination in the prevention of cervical cancer, namely, a widely circulated review paper about this topic [51]. Leading accounts linked to the orange cluster and the dark green cluster were personal user accounts, and both shared the same paper as the light green cluster, calling into question the efficacy of the HPV vaccine in the prevention of cervical cancer.

Figure 2. Network analysis of the vaccine-hesitant community sharing scientific research on Twitter. HPV: human papillomavirus.

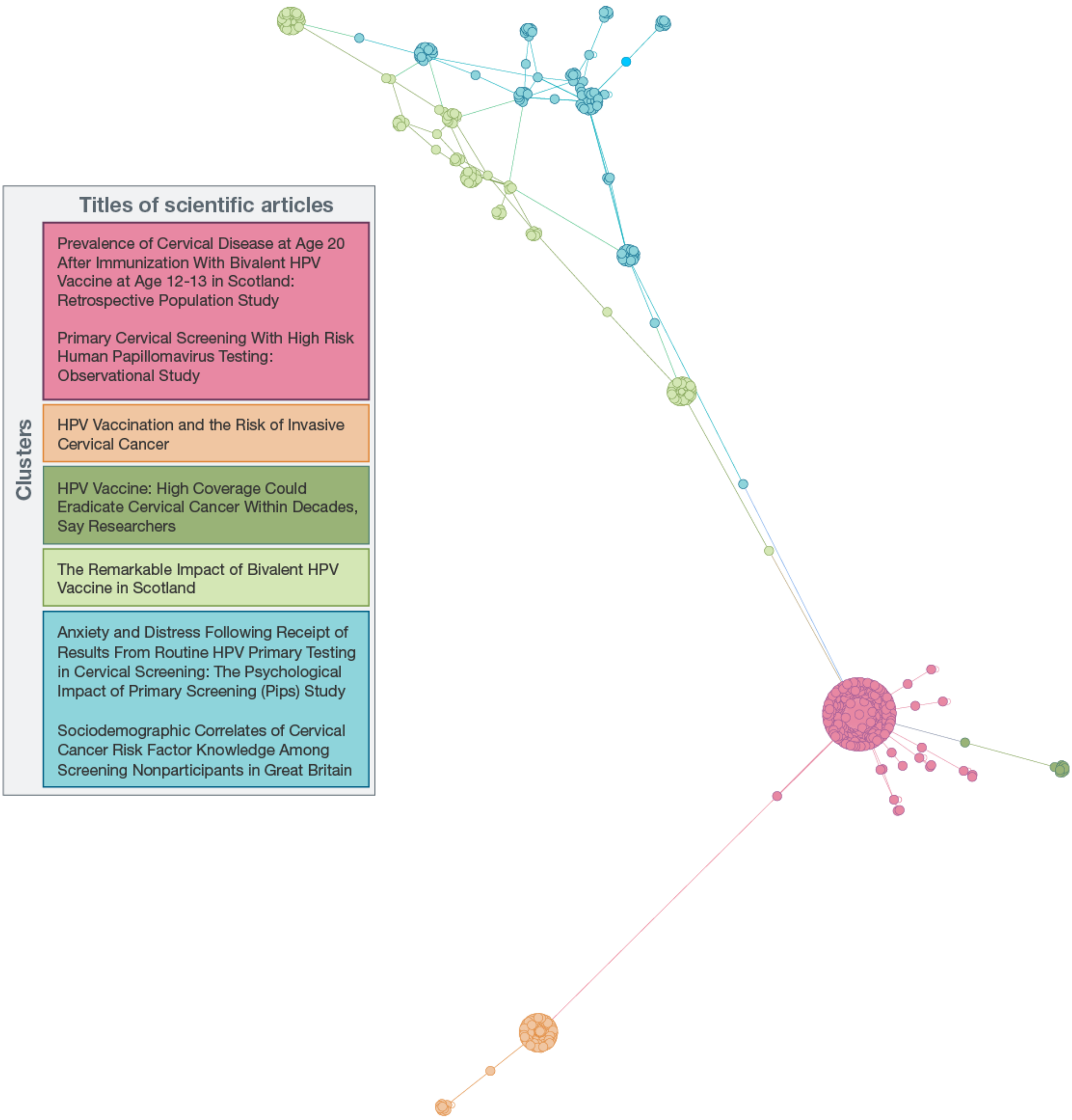


The orange cluster of the vaccine-hesitant community circulated a retracted paper, which alleged that HPV vaccines affected the vaccine recipients' fertility and focused on safety [52]. Furthermore, the orange cluster's location in the network (ie, adjacent to the light green cluster) suggests social influence and connection between the 2 clusters. In contrast, there was little interaction between the accounts in the light green cluster and the dark green cluster, suggesting that the influential accounts in these clusters independently found the same scientific literature and circulated it among a relatively isolated cohort of users. Finally, in the blue cluster, a European support group for those who had experienced vaccine injuries was the leading influential account, whereas a medical society's account that published a widely shared paper in this cluster [51] was an account of secondary influence. Again, the influential accounts in this cluster shared scientific papers, which were retweeted by accounts that are more peripheral to the central clusters of influential accounts. The primary scientific paper circulated among users in this cluster focused on the theme of safety of the HPV vaccine by measuring the serum levels of autoantibodies in a cohort of girls who had possible adverse reactions following the receipt of the HPV vaccine [53].

Vaccine-Confident Social Network

The retweet network of scientific research shared among the web-based vaccine-confident community can similarly be divided into 5 distinct subclusters, as shown in Figure 3. The red cluster primarily included users retweeting literature from the *British Medical Journal*. There were 2 main papers circulated in this cluster, both of which focused on the effectiveness of the HPV vaccine. The first was a retrospective population study about the efficacy of the HPV vaccine in the prevention of cervical cancer in Scotland, focusing on the theme of satisfaction with service [71], whereas the second was an observational study about the outcomes of HPV screening in high-risk populations in England [72]. In the orange cluster, we observed a similar influence exerted by a government-funded public health agency, which shared a popular paper about effectiveness, focusing on the potential of the HPV vaccine to lower the risk of cervical cancer in a cohort population [73].

Figure 3. Network analysis of the vaccine-confident community sharing scientific research on Twitter. HPV: human papillomavirus.



In the red, orange, and dark green clusters, there were physicians and health care workers among the users who retweeted influential tweets. For example, in the orange cluster, 1 particularly influential physician circulated an editorial paper about the effectiveness of the HPV vaccine, which indicated that high HPV vaccine coverage could eradicate cervical cancer within a few decades [74]. A science correspondent for a pre-eminent American newspaper was the leading influencer in the light green cluster wherein the primary paper circulated was an editorial, also focused on effectiveness, related to the positive impacts of HPV vaccination in Scotland [75]. Finally, in the blue cluster, a leading cancer prevention researcher from a British research institute was the leading influencer and author of the scientific papers circulated. In this cluster, papers about the psychological impacts of HPV screening [76] and the sociodemographic correlates of cervical cancer risk among those who did not participant in cervical screening programs in the United Kingdom [77] were recirculated by the accounts influenced by the leading researcher. Unlike the other clusters, health care workers were not overrepresented in the light green and blue clusters.

Overall, results from the vaccine-confident community suggest that health care, scientific, and news media communities are operating in closed systems. As we can see in Figure 3, there are relatively few bridging connections among the different communities discussing influential HPV vaccination literature in the vaccine-confident space. In contrast, the vaccine-hesitant space (Figure 2) is a more cohesive and tightly connected community, suggesting that there are stronger knowledge flows between subclusters in this group. Twitter accounts in the vaccine-hesitant community appear to be more efficient in sharing information than the more fragmented vaccine-confident community (Multimedia Appendix 3). Furthermore, the vaccine-hesitant Twitter accounts are more effective in communicating the results and research of interest to one another, whereas those in the vaccine-confident space appear to struggle to disseminate the research of interest beyond their personal and professional communities. These findings are supported by the descriptive statistics presented later in the paper, which indicated that while the vaccine-confident community shares far more scientific papers than the vaccine-hesitant community, the scientific literature shared by the vaccine-confident community received far fewer shares per paper despite being published in higher-ranked journals.

Typology of Evidence and Bibliometric Analysis

Table 2 presents the distribution of typology of evidence categorized by vaccine community type. Most of the scientific papers shared by the vaccine-hesitant community focused on safety (16/30, 55%) or effectiveness (8/30, 28%), exemplifying the key concerns legitimizing vaccine hesitancy. The vaccine-confident community shared papers related to a wider range of research themes, the most common being papers that focused on basic science (56/220, 25.7%), effectiveness (55/220, 25.2%), acceptability (49/220, 22.5%), and salience (38/220, 17.4%). While the level of focus on effectiveness was similar between the 2 communities, there was very little overlap in the specific papers selected for sharing.

Table 2. Description of the typology of evidence of all papers shared on Twitter categorized based on vaccine-confident and vaccine-hesitant communities.

	Vaccine-confident community (N=220), n (%)	Vaccine-hesitant community (N=30), n (%)
Effectiveness	55 (25)	8 (26.7)
Safety	12 (5.5)	16 (53.3)
Process of service delivery	4 (1.8)	1 (3.3)
Satisfaction with the service	1 (0.5)	0 (0)
Salience	38 (17.3)	1 (3.3)
Acceptability	49 (22.3)	1 (3.3)
Cost-effectiveness	5 (2.3)	0 (0)
Basic science	56 (25.5)	3 (10)

Table 3 presents the descriptive statistics about the 4 metrics for all the scientific papers shared by vaccine-confident (220/250, 88% papers) and vaccine-hesitant (30/250, 12% papers) communities. The 4 metrics described in Table 3 are the median shares per paper, the median number of citations each shared paper received, the median SJR score of the journal that published each shared paper, and the median typology of evidence score. Table 3 also presents the results from the Mann-Whitney *U* test. Tweets containing scientific papers shared by the vaccine-confident community received a median of 3 shares, compared to a median of 4 shares by the vaccine-hesitant community. Results from the Mann-Whitney *U* test indicate that there are statistically significant differences (*P*=.01) in shares of tweets containing papers about HPV vaccination between the vaccine-hesitant and vaccine-confident communities and that this difference is small (Cohen *d*=0.37). Scientific papers shared by the vaccine-confident community received a median of 13 citations compared to a median of 17 citations for the scientific papers shared by the vaccine-hesitant community. We did not find evidence of statistically significant differences in the number of citations received by papers shared between the vaccine-confident and vaccine-hesitant communities. Scientific papers shared by the vaccine-confident community received a median SJR score of 1.83 compared to a median score of 0.84 for the papers shared by the vaccine-hesitant community. Results from the effect size calculation found this to be a medium standardized difference (Cohen *d*=61). The Mann-Whitney *U* test also found evidence of statistically significant (*P*<.001) differences in SJR scores of the HPV-related papers shared between the vaccine-confident and vaccine-hesitant communities. Finally, scientific papers shared by both the vaccine-confident and the vaccine-hesitant communities received a median typology of evidence score of 1, and results from the Mann-Whitney *U* test did not find evidence of a statistically significant difference.

Table 3. Results from the Mann-Whitney U test for shares, number of citations, SCImago Journal Ranking (SJR), and typology of evidence score categorized based on human papillomavirus vaccine-confident and vaccine-hesitant communities.

	Vaccine-confident community (n=220), median (IQR)	Vaccine-hesitant community (n=30), median (median)	*P* value	Effect size (Cohen *d*)
Shares	3 (1.0-6.5)	4 (2.0-15.0)	.007	0.37
Number of citations	13 (5.0-75.0)	17 (9.0-44)	.28	0.19
SJR	1.83 (1.25-3.44)	0.84 (0.68-1.30)	<.001	0.61
Typology of evidence	1 (0.0-10)	1 (1.0-1.0)	.22	0.14

Discussion

Principal Findings

The increase in the volume of scientific publications shared on the web [13] and the growth of open-access scientific publishing [16] have created an environment of greater access to scientific literature among lay audiences. However, little is known about how scientific literature is being incorporated into web-based communication strategies of vaccine-confident and vaccine-hesitant communities. Our study examined how scientific literature focusing on the HPV vaccine is being shared by vaccine-hesitant and vaccine-confident networks on Twitter. We found that despite the increased quantity of scientific literature being shared, such literature is often used by the vaccine-hesitant community to proliferate misinformation about vaccination, which is amplified in a web-based environment such as Twitter. Therefore, Kata [97] has described four key tactics that are used by the antivaccination movement to spread their messages on the web: (1) skewing the science, (2) shifting the hypotheses, (3) censorship, and (4) attacking the critics. A study conducted by van Schalkwyk et al [17] demonstrated that vaccine-hesitant groups are strategic in their use of scientific literature on social media to amplify uncertainty about vaccine safety and that vaccine-hesitant accounts who use large arsenals of scientific literature play important roles in dissemination of information across multiple communication networks. Findings from our thematic analysis of the papers shared by the vaccine-hesitant networks confirm this. Our study also found that the vaccine-hesitant community was much more likely to share scientific publications that questioned the safety and effectiveness of the HPV vaccine, whereas the vaccine-confident community shared scientific publications on a wider range of topics. This aligns with the tactic of skewing the science (identified by Kata [97]), which focuses on criticizing scientific studies while simultaneously calling for more studies, particularly focusing on the need for randomized controlled trials that compare vaccinated children and unvaccinated children. Moreover, most of the papers shared by the vaccine-confident community focused on basic science (ie, in vitro or in vivo studies), and this focus lowered the typology of evidence score of the vaccine-confident community, while failing to contribute to a unified message in the vaccine-confident community.

Furthermore, the quality of journals that published the papers shared in these communities varied markedly. The scientific publications shared by the vaccine-confident community were significantly more likely to be published in higher-ranked journals and therefore obtained higher SJR scores, compared with those shared by the vaccine-hesitant community. Other researchers have found that critical appraisal is often absent when vaccine-hesitant individuals share “scientific evidence” on the web, which often includes citations that blur the line between legitimate scientific publications and fraudulent studies [98]. However, there is little evidence of communication across networks, despite repeated calls from public health communication experts to prebunk and debunk vaccine misinformation on the web [99,100]. Notably, both communities share a retracted paper, but their framing of the paper varies. The vaccine-confident community mocks the paper for its outlandish claims, whereas the vaccine-hesitant community highlights the findings as if they were accurate. This highlights 2 issues. First, despite not supporting the findings of the retracted paper, the vaccine-confident community still shared the paper, thus amplifying its reach. Second, the vaccine-hesitant communities’ definition of “scientific evidence” does not align with accepted norms, as retracted papers can no longer be considered part of the scientific evidence base.

Vaccine-hesitant groups have been shown to co-opt the perceived authority of professional sources (eg, WebMD and the American Medical Association) to bolster their claims, even when the associated evidence does not support their arguments [101]. Interestingly, past studies have shown that while both groups point out knowledge deficits in their counterparts and attempt to correct misinformation by offering alternate sources of evidence, vaccine-confident groups have been shown to infrequently cite scientific evidence to correct misinformation or present counterarguments in web-based forums [102]. However, our analysis shows that the vaccine-confident community often shares scientific literature on the web as a form of self-promotion or knowledge translation, rather than as a tool to counter misinformation or correct misinterpretations.

Consequently, consistent with others in this field, we suggest that vaccine researchers should take a more active role in the HPV-related conversations that are occurring on the web, beyond simply promoting their own studies and instead countering misinformation and disinformation on the web [103]. Researchers and practitioners hoping to meaningfully contribute to the conversation about HPV vaccination on the web should explore training in science communication and social media engagement strategies, including the monitoring and correcting of public misinterpretation of their studies on various social media platforms [103,104]. Studies show that the way in which health information is communicated affects recipients’ perception of it, with transparent communication fostering trust in health authorities and reducing the proliferation of conspiratorial beliefs [105].

Limitations

While Twitter provides us with a large body of unfiltered discussions to examine, the use of Twitter is not universal, and younger individuals (aged 18-29 years) and minority groups tend to be overrepresented on Twitter [20,24]. Therefore, while this analysis is not universal for all demographics, such as those who do not use Twitter as a social media platform, it provides opportunities to collect information about the health opinions held by members of several priority populations. While this study provides a way of studying web-based social interaction, further studies are needed to understand vaccine hesitancy among the general population who may not use Twitter.

The creation of the data set of HPV-related and HPV vaccine–related tweets was based on 3 commonly used hashtags derived from a rapid review of published papers; therefore, there is the potential that we missed some tweets that also discussed HPV and HPV vaccine but were not captured by these hashtags. In addition, we extracted a variety of metrics about the papers and journals included in our data set, but given the wide variation in study design among the extracted papers, conducting a formal critical appraisal of quality was unfeasible for this project and is an area for future study. Furthermore, this study did not measure the engagement rate of tweets, which is a new analytic metric offered by Twitter and is calculated by dividing the number of engagements (ie, total number of times a user interacted with a post including retweets, replies, likes, and follows) by the number of impressions (ie, number of times a user is shown a particular post in their timeline or search results). It should be reinforced that the number of shares of a tweet is not equivalent to the impact of the content shared.

Another limitation is that one of the metrics collected in our study was the number of citations each paper had received, for which we chose to use the "cited by" count provided by Google Scholar. While there has been criticism about the *cited by* metrics provided by Google Scholar due to double counting of citations from published journals and other sources [51], Google Scholar covers a larger breadth of sources (eg, conference papers and book chapters) than alternative platforms such as Web of Science [106]. Finally, the time frame we selected to collect tweets for this study, that is, January 2019 to May 2021, presents a limitation. We chose to expand our data collection to 2021 to allow us to acquire a sufficiently large data set, because the COVID-19 pandemic began shortly after the start of our data collection period. With the emergence of the COVID-19 pandemic, health discussions on Twitter became heavily focused on COVID-19 instead of other topics, including HPV vaccination. We ultimately extended our data collection time frame beyond our original timeline to provide us with a sufficiently large corpus of tweets to analyze. Given the unique period of data collection (ie, before and during the COVID-19 pandemic), which influenced the quantity of discussion about non–COVID-19 topics, the generalizability of these findings is reduced. Our experience in collecting these data over the course of the COVID-19 pandemic has been explored further in another publication, where we examined the attitudes and sentiment on Twitter toward HPV vaccination amidst the context of the pandemic [107].

Strengths

This study contributes to the growing body of knowledge about the discussions about HPV immunization in web-based settings by using novel mixed methods to identify what papers about HPV and HPV vaccine are being shared on the web and how vaccine-confident and vaccine-hesitant communities are using this knowledge in their web-based communication strategies. Our study demonstrates that vaccine-hesitant communities are using strategies of scientific authority by presenting them as "scientific evidence" on Twitter, regardless of the quality of the papers themselves. Vaccine-confident communities do not appear to be sharing papers to build consensus, rather they share their scientific studies. These findings are relevant to health communication experts who aim to combat vaccine misinformation and disinformation on the web by providing them with concrete examples of papers used to create distrust in HPV vaccines. Moreover, HPV researchers and health promotion organizations that use Twitter might find these results helpful in crafting a more deliberative knowledge translation strategy.

Our study has several strengths. First, we used a large body of data from Twitter to track near–real-time conversations about HPV vaccination on the web. Twitter, in its previous iteration, was one of the largest and most popular social media platforms and was seen as a preferred source of public opinion data for applied public health research due to the following features: (1) quick processes for collecting data sets, (2) low costs for data collection, (3) ability to monitor trends over time, and (4) ability to avoid researcher biases that are inherent to the design and delivery of traditional research tools such as surveys [21,24]. Therefore, this data set provided us access to a large number of unfiltered discussions from populations that are traditionally difficult to access through conventional data collection methods.

Next, our use of social network analysis allowed us to examine how scientific literature is shared and its connection within wider networks representing communities of interest. Thus, we were also able to identify key influencers within networks who potentially act as leverage points to amplify future health communication campaigns, while also shedding light on the density of vaccine-hesitant influencers compared to vaccine-confident influencers within the respective social networks. Finally, while the vaccine-hesitant community has attempted to use or distort scientific literature to support their viewpoints for a long time, to the best of our knowledge, this is the first study to examine how scientific evidence has been used and shared on the web by comparing both vaccine-hesitant and vaccine-confident web-based communities in discussions specifically related to the HPV vaccine.

Conclusions

Many of the communication strategies initially used by health promotion communities, including the use of the logical fallacy such as appealing to scientific authority and scientific knowledge, appear to have been co-opted by the vaccine-hesitant community and are being used to create controversy by focusing on questions about the effectiveness and safety of the HPV vaccine. While the scientific literature shared within these vaccine-hesitant communities is often published in lower-ranked journals, they deliver a substantially more successful, coordinated strategy when it comes to communicating about HPV vaccine on Twitter, compared to the vaccine-confident communities. By widely sharing a curated selection of scientific publications among like-minded individuals, the vaccine-hesitant community members' communication around the HPV vaccine yields much more interaction (ie, shares and retweets) than is observed in the vaccine-confident community's efforts to disseminate research findings. While the scientific literature shared by members of the vaccine-confident community is published in higher-ranked journals, these papers receive far fewer interactions and have lesser reach on Twitter.

While the vaccine-hesitant community has successfully incorporated communication tools that were traditionally wielded by health promotion communities to advance their agenda, the web-based vaccine-confident community could benefit from paying attention to their dissemination techniques for using web-based platforms such as Twitter to amplify their messaging. However, it is crucial that the vaccine-confident community's messages ultimately be transmitted in a manner that fosters long-term trust and credibility, which stems from accurate and transparent communication.

Acknowledgments

Funding was provided, in whole or in part, by Alberta Health. Strategic direction and applied research support were provided by the Alberta Health Services Cancer Prevention and Screening Innovation team. Provision of funding by Alberta Health does not indicate that this project represents the policies or views of Alberta Health.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of the top 20 most shared scientific papers on Twitter by the vaccine-hesitant community.
[DOCX File , 21 KB - infodemiology_v4i1e50551_app1.docx]

Multimedia Appendix 2
Summary of the top 20 most shared scientific papers on Twitter by the vaccine-confident community.
[DOCX File , 23 KB - infodemiology_v4i1e50551_app2.docx]

Multimedia Appendix 3
Retweet network map of human papillomavirus immunization conversations (N=596,987).
[DOCX File , 457 KB - infodemiology_v4i1e50551_app3.docx]

References

1. Volesky KD, El-Zein M, Franco EL, Brenner DR, Friedenreich CM, Ruan Y. Corrigendum to "Cancers attributable to infections in Canada" [Prev. Med. 122 (2019) 109-117]. Prev Med 2019 Dec;129:105730. [doi: 10.1016/j.ypmed.2019.05.018] [Medline: 31151672]
2. Global strategy to accelerate the elimination of cervical cancer as a public health problem. World Health Organization. 2020 Nov 17. URL: https://www.who.int/publications/i/item/9789240014107 [accessed 2024-02-05]
3. Hopkins TG, Wood N. Female human papillomavirus (HPV) vaccination: global uptake and the impact of attitudes. Vaccine 2013 Mar 25;31(13):1673-1679. [doi: 10.1016/j.vaccine.2013.01.028] [Medline: 23375978]
4. Dubé E, Gagnon D, MacDonald N, Bocquier A, Peretti-Watel P, Verger P. Underlying factors impacting vaccine hesitancy in high income countries: a review of qualitative studies. Expert Rev Vaccines 2018 Nov;17(11):989-1004. [doi: 10.1080/14760584.2018.1541406] [Medline: 30359151]
5. Tonsaker T, Bartlett G, Trpkov C. Health information on the internet: gold mine or minefield? Can Fam Physician 2014 May;60(5):407-408 [FREE Full text] [Medline: 24828994]
6. Lutkenhaus RO, Jansz J, Bouman MP. Mapping the Dutch vaccination debate on Twitter: identifying communities, narratives, and interactions. Vaccine X 2019 Mar 21;1:100019 [FREE Full text] [doi: 10.1016/j.jvacx.2019.100019] [Medline: 31485580]
7. Castells M. Networks of Outrage and Hope: Social Movements in the Internet Age, 2nd Edition. Hoboken, NJ: Wiley; Apr 2015.
8. Baldwin AS, Tiro JA, Zimet GD. Broad perspectives in understanding vaccine hesitancy and vaccine confidence: an introduction to the special issue. J Behav Med 2023 Apr;46(1-2):1-8 [FREE Full text] [doi: 10.1007/s10865-023-00397-8] [Medline: 36802315]
9. Larson HJ, Gakidou E, Murray CJ. The vaccine-hesitant moment. N Engl J Med 2022 Jul 07;387(1):58-65 [FREE Full text] [doi: 10.1056/NEJMra2106441] [Medline: 35767527]

10. Huh WK, Joura EA, Giuliano AR, Iversen OE, de Andrade RP, Ault KA, et al. Final efficacy, immunogenicity, and safety analyses of a nine-valent human papillomavirus vaccine in women aged 16-26 years: a randomised, double-blind trial. Lancet 2017 Nov 11;390(10108):2143-2159. [doi: 10.1016/S0140-6736(17)31821-4] [Medline: 28886907]
11. Arbyn M, Xu L. Efficacy and safety of prophylactic HPV vaccines. A Cochrane review of randomized trials. Expert Rev Vaccines 2018 Dec;17(12):1085-1091. [doi: 10.1080/14760584.2018.1548282] [Medline: 30495978]
12. Madden K, Nan X, Briones R, Waks L. Sorting through search results: a content analysis of HPV vaccine information online. Vaccine 2012 May 28;30(25):3741-3746. [doi: 10.1016/j.vaccine.2011.10.025] [Medline: 22019758]
13. Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, et al. The state of OA: a large-scale analysis of the prevalence and impact of open access articles. PeerJ 2018 Feb 13;6:e4375 [FREE Full text] [doi: 10.7717/peerj.4375] [Medline: 29456894]
14. Bucchi M. Credibility, expertise and the challenges of science communication 2.0. Public Underst Sci 2017 Nov;26(8):890-893. [doi: 10.1177/0963662517733368] [Medline: 29025369]
15. Steffens MS, Dunn AG, Leask J. Meeting the challenges of reporting on public health in the new media landscape. Aust J Rev 2017 Dec;39(2):119-132.
16. Barberá P. How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S. In: Proceedings of the American Political Science Association Annual Meeting 2015. 2015 Presented at: APSA 2015; September 3-6, 2015; San Francisco, CA URL: http://pablobarbera.com/static/barbera_polarization_APSA.pdf
17. van Schalkwyk F, Dudek J, Costas R. Communities of shared interests and cognitive bridges: the case of the anti-vaccination movement on Twitter. Scientometrics 2020 Jun 13;125:1499-1416. [doi: 10.1007/s11192-020-03551-0]
18. Boothby C, Murray D, Waggy AP, Tsou A, Sugimoto CR. Credibility of scientific information on social media: variation by platform, genre and presence of formal credibility cues. Quant Sci Stud 2021 Nov 05;2(3):845-863. [doi: 10.1162/qss_a_00151]
19. Bello-Orgaz G, Hernandez-Castro J, Camacho D. Detecting discussion communities on vaccination in Twitter. Future Gener Comput Syst 2017 Jan;66:125-136. [doi: 10.1016/j.future.2016.06.032]
20. Chakraborty P, Colditz JB, Silvestre AJ, Friedman MR, Bogen KW, Primack BA. Observation of public sentiment toward human papillomavirus vaccination on Twitter. Cogent Med 2017 Dec 12;4(1). [doi: 10.1080/2331205x.2017.1390853]
21. Du J, Xu J, Song H, Liu X, Tao C. Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. J Biomed Semantics 2017 Mar 03;8(1):9 [FREE Full text] [doi: 10.1186/s13326-017-0120-6] [Medline: 28253919]
22. Du J, Xu J, Song HY, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):69 [FREE Full text] [doi: 10.1186/s12911-017-0469-6] [Medline: 28699569]
23. Kunneman F, Lambooij M, Wong A, van den Bosch A, Mollema L. Monitoring stance towards vaccination in Twitter messages. BMC Med Inform Decis Mak 2020 Feb 18;20(1):33 [FREE Full text] [doi: 10.1186/s12911-020-1046-y] [Medline: 32070334]
24. 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]
25. Du J, Luo C, Shegog R, Bian J, Chen Y, Tao C. Deep learning and behavioral theory: an improved analytic method to understand HPV vaccination intentions from Twitter discussion. SSRN J 2019 Oct 03:1-30 [FREE Full text] [doi: 10.2139/ssrn.3460688]
26. Shah Z, Surian D, Dyda A, Coiera E, Mandl KD, Dunn AG. Automatically appraising the credibility of vaccine-related web pages shared on social media: a Twitter surveillance study. J Med Internet Res 2019 Nov 04;21(11):e14007 [FREE Full text] [doi: 10.2196/14007] [Medline: 31682571]
27. Sanawi JB, Samani MC, Taibi M. #vaccination: identifying influencers in the vaccination discussion on Twitter through social network visualisation. Int J Bus Soc 2017 Oct;18(S4):718-726.
28. Caulfield T, Marcon AR, Murdoch B, Brown JM, Perrault ST, Jarry J, et al. Health misinformation and the power of narrative messaging in the public sphere. Can J Bioeth 2019 Mar 20;2(2):52-60. [doi: 10.7202/1060911ar]
29. Elyashar A, Plochotnikov I, Cohen IC, Puzis R, Cohen O. The state of mind of health care professionals in light of the COVID-19 pandemic: text analysis study of Twitter discourses. J Med Internet Res 2021 Oct 22;23(10):e30217 [FREE Full text] [doi: 10.2196/30217] [Medline: 34550899]
30. Dobbins M. Rapid Review Guidebook. Hamilton, ON: National Collaborating Centre for Methods and Tools; 2017.
31. Shapiro GK, Surian D, Dunn AG, Perry R, Kelaher M. Comparing human papillomavirus vaccine concerns on Twitter: a cross-sectional study of users in Australia, Canada and the UK. BMJ Open 2017 Oct 05;7(10):e016869 [FREE Full text] [doi: 10.1136/bmjopen-2017-016869] [Medline: 28982821]
32. Massey PM, Leader A, Yom-Tov E, Budenz A, Fisher K, Klassen AC. Applying multiple data collection tools to quantify human papillomavirus vaccine communication on Twitter. J Med Internet Res 2016 Dec 05;18(12):e318 [FREE Full text] [doi: 10.2196/jmir.6670] [Medline: 27919863]

33. Keim-Malpass J, Mitchell EM, Sun E, Kennedy C. Using Twitter to understand public perceptions regarding the #HPV vaccine: opportunities for public health nurses to engage in social marketing. Public Health Nurs 2017 Jul;34(4):316-323. [doi: 10.1111/phn.12318] [Medline: 28261846]
34. Nelon JL, Moscarelli M, Stupka P, Sumners C, Uselton T, Patterson MS. Does scientific publication inform public discourse? A case study observing social media engagement around vaccinations. Health Promot Pract 2021 May;22(3):377-384. [doi: 10.1177/1524839919899925] [Medline: 32009444]
35. Surian D, Nguyen DQ, Kennedy G, Johnson M, Coiera E, Dunn AG. Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. J Med Internet Res 2016 Aug 29;18(8):e232 [FREE Full text] [doi: 10.2196/jmir.6045] [Medline: 27573910]
36. Zhou X, Coiera E, Tsafnat G, Arachi D, Ong MS, Dunn AG. Using social connection information to improve opinion mining: identifying negative sentiment about HPV vaccines on Twitter. Stud Health Technol Inform 2015;216:761-765. [Medline: 26262154]
37. Becker BF, Larson HJ, Bonhoeffer J, van Mulligen EM, Kors JA, Sturkenboom MC. Evaluation of a multinational, multilingual vaccine debate on Twitter. Vaccine 2016 Dec 07;34(50):6166-6171. [doi: 10.1016/j.vaccine.2016.11.007] [Medline: 27840012]
38. Dyda A, Shah Z, Surian D, Martin P, Coiera E, Dey A, et al. HPV vaccine coverage in Australia and associations with HPV vaccine information exposure among Australian Twitter users. Hum Vaccin Immunother 2019;15(7-8):1488-1495. [doi: 10.1080/21645515.2019.1596712] [Medline: 30978147]
39. Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. Vaccine 2017 May 25;35(23):3033-3040 [FREE Full text] [doi: 10.1016/j.vaccine.2017.04.060] [Medline: 28461067]
40. Budenz A, Massey P, Leader A, Klassen A, Fisher K, Yom-Tov E. Let's talk about HPV: a focus on gender- and MSM-specific Twitter communication. In: Proceedings of the American Public Health Association Annual Meeting 2017. 2017 Presented at: APHA 2017; November 4-8, 2017; Atlanta, GA.
41. Zhang H, Wheldon C, Dunn AG, Tao C, Huo J, Zhang R, et al. Mining Twitter to assess the determinants of health behavior toward human papillomavirus vaccination in the United States. J Am Med Inform Assoc 2020 Feb 01;27(2):225-235 [FREE Full text] [doi: 10.1093/jamia/ocz191] [Medline: 31711186]
42. X data for academic research. X Developer Platform. URL: https://developer.twitter.com/en/use-cases/do-research/academic-research [accessed 2024-02-05]
43. X API documentation. X Developer Platform. URL: https://developer.twitter.com/en/docs/twitter-api [accessed 2024-02-05]
44. Van Rossum G, Drake FL. Python 3 Reference Manual. North Charleston, SC: CreateSpace; Mar 2009.
45. Yan Y, Toriumi F, Sugawara T. Understanding how retweets influence the behaviors of social networking service users via agent-based simulation. Comput Soc Netw 2021 Sep 13;8:18. [doi: 10.1186/S40649-021-00099-8]
46. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech 2008 Oct 09;2008:P10008. [doi: 10.1088/1742-5468/2008/10/P10008]
47. 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]
48. pandas.Series.str.contains. Pandas. URL: https://pandas.pydata.org/docs/reference/api/pandas.Series.str.contains.html [accessed 2024-02-05]
49. Gray JA. Evidence-Based Healthcare. London, UK: Churchill Livingstone; 1997.
50. Beppu H, Minaguchi M, Uchide K, Kumamoto K, Sekiguchi M, Yaju Y. Lessons learnt in Japan from adverse reactions to the HPV vaccine: a medical ethics perspective. Indian J Med Ethics 2017 Apr 1:82 [FREE Full text] [doi: 10.20529/IJME.2017.021]
51. Rees CP, Brhlikova P, Pollock AM. Will HPV vaccination prevent cervical cancer? J R Soc Med 2020 Feb 21;113(2):64-78 [FREE Full text] [doi: 10.1177/0141076819899308] [Medline: 31962050]
52. DeLong G. A lowered probability of pregnancy in females in the USA aged 25-29 who received a human papillomavirus vaccine injection. Retracted in: J Clin Hypertens (Greenwich) 2018 Jun 11;81(14):661-674. [doi: 10.1080/15287394.2018.1477640] [Medline: 29889622]
53. Hineno A, Ikeda S, Scheibenbogen C, Heidecke H, Forster K, Junker J, et al. Autoantibodies against autonomic nerve receptors in adolescent Japanese girls after immunization with human papillomavirus vaccine. Ann Arthritis Clin Rheumatol 2019;2:1014 [FREE Full text]
54. Atkinson AE, Mandujano CA, Bejarano S, Kennedy LS, Tsongalis GJ. Screening for human papillomavirus in a low- and middle-income country. J Glob Oncol 2019 May;5:JGO1800233 [FREE Full text] [doi: 10.1200/JGO.18.00233] [Medline: 31050922]
55. Bizjak M, Bruck O, Praprotnik S, Dahan S, Shoenfeld Y. Pancreatitis after human papillomavirus vaccination: a matter of molecular mimicry. Immunol Res 2017 Mar;65(1):164-167. [doi: 10.1007/s12026-016-8823-9] [Medline: 27421720]
56. Blitshteyn S, Brinth L, Hendrickson JE, Martinez-Lavin M. Autonomic dysfunction and HPV immunization: an overview. Immunol Res 2018 Dec;66(6):744-754. [doi: 10.1007/s12026-018-9036-1] [Medline: 30478703]

57. Chang J, Campagnolo D, Vollmer TL, Bomprezzi R. Demyelinating disease and polyvalent human papilloma virus vaccination. J Neurol Neurosurg Psychiatry 2011 Nov;82(11):1296-1298. [doi: 10.1136/jnnp.2010.214924] [Medline: 20935322]
58. Debeer P, De Munter P, Bruyninckx F, Devlieger R. Brachial plexus neuritis following HPV vaccination. Vaccine 2008 Aug 18;26(35):4417-4419. [doi: 10.1016/j.vaccine.2008.06.074] [Medline: 18602437]
59. Fischer S, Bettstetter M, Becher A, Lessel M, Bank C, Krams M, et al. Shift in prevalence of HPV types in cervical cytology specimens in the era of HPV vaccination. Oncol Lett 2016 Jul;12(1):601-610 [FREE Full text] [doi: 10.3892/ol.2016.4668] [Medline: 27347187]
60. Guo F, Hirth JM, Berenson AB. Comparison of HPV prevalence between HPV-vaccinated and non-vaccinated young adult women (20-26 years). Hum Vaccin Immunother 2015;11(10):2337-2344 [FREE Full text] [doi: 10.1080/21645515.2015.1066948] [Medline: 26376014]
61. Inbar R, Weiss R, Tomljenovic L, Arango MT, Deri Y, Shaw CA, et al. Behavioral abnormalities in female mice following administration of aluminum adjuvants and the human papillomavirus (HPV) vaccine Gardasil. Immunol Res 2017 Mar;65(1):136-149. [doi: 10.1007/s12026-016-8826-6] [Medline: 27421722]
62. Jørgensen L, Gøtzsche PC, Jefferson T. Benefits and harms of the human papillomavirus (HPV) vaccines: systematic review with meta-analyses of trial data from clinical study reports. Syst Rev 2020 Mar 28;9(1):43 [FREE Full text] [doi: 10.1186/s13643-019-0983-y] [Medline: 32106879]
63. Little DT, Ward HR. Adolescent premature ovarian insufficiency following human papillomavirus vaccination: a case series seen in general practice. J Investig Med High Impact Case Rep 2014;2(4):2324709614556129 [FREE Full text] [doi: 10.1177/2324709614556129] [Medline: 26425627]
64. Masson JD, Crépeaux G, Authier FJ, Exley C, Gherardi RK. Critical analysis of reference studies on the toxicokinetics of aluminum-based adjuvants. J Inorg Biochem 2018 Apr;181:87-95. [doi: 10.1016/j.jinorgbio.2017.12.015] [Medline: 29307441]
65. Castle PE, Xie X, Xue X, Poitras NE, Lorey TS, Kinney WK, et al. Impact of human papillomavirus vaccination on the clinical meaning of cervical screening results. Prev Med 2019 Jan;118:44-50. [doi: 10.1016/j.ypmed.2018.10.001] [Medline: 30316878]
66. Ran J, Yang JY, Lee JH, Kim HJ, Choi JY, Shin JY. Signal detection of human papillomavirus vaccines using the Korea Adverse Events Reporting System database, between 2005 and 2016. Int J Clin Pharm 2019 Oct;41(5):1365-1372. [doi: 10.1007/s11096-019-00881-9] [Medline: 31313003]
67. Riva C, Spinosa JP. Has the HPV vaccine approval ushered in an era of over-prevention? J Sci Pract Integr 2020 Dec;2(1). [doi: 10.35122/001c.18180]
68. Ryan M, Marlow L, Waller J. Socio-demographic correlates of cervical cancer risk factor knowledge among screening non-participants in Great Britain. Prev Med 2019 Aug;125:1-4 [FREE Full text] [doi: 10.1016/j.ypmed.2019.04.026] [Medline: 31085204]
69. Álvarez-Soria MJ, Hernández-González A, Carrasco-García de León S, del Real-Francia M, Gallardo-Alcañiz MJ, López-Gómez JL. [Demyelinating disease and vaccination of the human papillomavirus]. Rev Neurol 2011 Apr 16;52(8):472-476 [FREE Full text] [Medline: 21425100]
70. Tomljenovic L, Shaw CA. No autoimmune safety signal after vaccination with quadrivalent HPV vaccine Gardasil? J Intern Med 2012 Nov;272(5):514-5; author reply 516 [FREE Full text] [doi: 10.1111/j.1365-2796.2012.02551.x] [Medline: 22540172]
71. Palmer T, Wallace L, Pollock KG, Cuschieri K, Robertson C, Kavanagh K, et al. Prevalence of cervical disease at age 20 after immunisation with bivalent HPV vaccine at age 12-13 in Scotland: retrospective population study. BMJ 2019 Apr 03;365:l1161 [FREE Full text] [doi: 10.1136/bmj.l1161] [Medline: 30944092]
72. Rebolj M, Rimmer J, Denton K, Tidy J, Mathews C, Ellis K, et al. Primary cervical screening with high risk human papillomavirus testing: observational study. BMJ 2019 Feb 06;364:l240. [doi: 10.1136/bmj.l240]
73. Lei J, Ploner A, Elfström KM, Wang J, Roth A, Fang F, et al. HPV vaccination and the risk of invasive cervical cancer. N Engl J Med 2020 Oct;383(14):1340-1348. [doi: 10.1056/nejmoa1917338]
74. Torjesen I. HPV vaccine: high coverage could eradicate cervical cancer within decades, say researchers. BMJ 2019 Jun 27;365:l4450. [doi: 10.1136/bmj.l4450] [Medline: 31248937]
75. Brotherton JM. The remarkable impact of bivalent HPV vaccine in Scotland. BMJ 2019 Apr 03;365:l1375. [doi: 10.1136/bmj.l1375] [Medline: 30944088]
76. McBride E, Marlow LA, Forster AS, Ridout D, Kitchener H, Patnick J, et al. Anxiety and distress following receipt of results from routine HPV primary testing in cervical screening: the psychological impact of primary screening (PIPS) study. Int J Cancer 2020 Apr 15;146(8):2113-2121 [FREE Full text] [doi: 10.1002/ijc.32540] [Medline: 31251820]
77. Ryan M, Marlow L, Waller J. Socio-demographic correlates of cervical cancer risk factor knowledge among screening non-participants in Great Britain. Prev Med 2019 Aug;125:1-4 [FREE Full text] [doi: 10.1016/j.ypmed.2019.04.026] [Medline: 31085204]

78. Artemchuk H, Eriksson T, Poljak M, Surcel HM, Dillner J, Lehtinen M, et al. Long-term antibody response to human papillomavirus vaccines: up to 12 years of follow-up in the Finnish maternity cohort. J Infect Dis 2019 Jan 29;219(4):582-589. [doi: 10.1093/infdis/jiy545] [Medline: 30239832]
79. Baker P, Kelly D, Medeiros R, Morrissey M, Price R. Eliminating HPV-caused cancers in Europe: achieving the possible. J Cancer Policy 2021 Jun;28:100280. [doi: 10.1016/j.jcpo.2021.100280] [Medline: 35559909]
80. Colafrancesco S, Perricone C, Tomljenovic L, Shoenfeld Y. Human papilloma virus vaccine and primary ovarian failure: another facet of the autoimmune/inflammatory syndrome induced by adjuvants. Am J Reprod Immunol 2013 Oct;70(4):309-316. [doi: 10.1111/aji.12151] [Medline: 23902317]
81. Cromwell I, Smith LW, van der Hoek K, Hedden L, Coldman AJ, Cook D, et al. Cost-effectiveness analysis of primary human papillomavirus testing in cervical cancer screening: results from the HPV FOCAL Trial. Cancer Med 2021 May;10(9):2996-3003 [FREE Full text] [doi: 10.1002/cam4.3864] [Medline: 33811457]
82. Drolet M, Bénard É, Pérez N, Brisson M, HPV Vaccination Impact Study Group. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. Lancet 2019 Aug 10;394(10197):497-509 [FREE Full text] [doi: 10.1016/S0140-6736(19)30298-3] [Medline: 31255301]
83. Gee J, Naleway A, Shui I, Baggs J, Yin R, Li R, et al. Monitoring the safety of quadrivalent human papillomavirus vaccine: findings from the Vaccine Safety Datalink. Vaccine 2011 Oct 26;29(46):8279-8284. [doi: 10.1016/j.vaccine.2011.08.106] [Medline: 21907257]
84. Gleber-Netto FO, Rao X, Guo T, Xi Y, Gao M, Shen L, et al. Variations in HPV function are associated with survival in squamous cell carcinoma. JCI Insight 2019 Jan 10;4(1):e124762 [FREE Full text] [doi: 10.1172/jci.insight.124762] [Medline: 30626753]
85. Kardas-Nelson M. Vaccine uptake and prevalence of HPV related cancers in US men. BMJ 2019 Mar 18;364:l1210. [doi: 10.1136/bmj.l1210] [Medline: 30885904]
86. Klein NP, Hansen J, Chao C, Velicer C, Emery M, Slezak J, et al. Safety of quadrivalent human papillomavirus vaccine administered routinely to females. Arch Pediatr Adolesc Med 2012 Dec;166(12):1140-1148. [doi: 10.1001/archpediatrics.2012.1451] [Medline: 23027469]
87. Nagarsheth NB, Norberg SM, Sinkoe AL, Adhikary S, Meyer TJ, Lack JB, et al. TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers. Nat Med 2021 Mar;27(3):419-425 [FREE Full text] [doi: 10.1038/s41591-020-01225-1] [Medline: 33558725]
88. Simms KT, Hanley SJ, Smith MA, Keane A, Canfell K. Impact of HPV vaccine hesitancy on cervical cancer in Japan: a modelling study. Lancet Public Health 2020 Apr;5(4):e223-e234 [FREE Full text] [doi: 10.1016/S2468-2667(20)30010-4] [Medline: 32057317]
89. Hoss A, Meyerson BE, Zimet GD. State statutes and regulations related to human papillomavirus vaccination. Hum Vaccin Immunother 2019;15(7-8):1519-1526 [FREE Full text] [doi: 10.1080/21645515.2019.1627817] [Medline: 31241406]
90. Petticrew M, Roberts H. Evidence, hierarchies, and typologies: horses for courses. J Epidemiol Community Health 2003 Jul;57(7):527-529 [FREE Full text] [doi: 10.1136/jech.57.7.527] [Medline: 12821702]
91. Callaham M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. JAMA 2002 Jun 05;287(21):2847-2850. [doi: 10.1001/jama.287.21.2847] [Medline: 12038930]
92. Scarlat MM, Mavrogenis AF, Pećina M, Niculescu M. Impact and alternative metrics for medical publishing: our experience with International Orthopaedics. Int Orthop 2015 Aug;39(8):1459-1464. [doi: 10.1007/s00264-015-2766-y] [Medline: 25947897]
93. Guerrero-Bote VP, Moya-Anegón F. A further step forward in measuring journals' scientific prestige: the SJR2 indicator. J Informetr 2012 Oct;6(4):674-688. [doi: 10.1016/j.joi.2012.07.001]
94. Scimago journal and country rank homepage. Scimago Journal & Country Rank. URL: https://www.scimagojr.com/ [accessed 2023-07-04]
95. McKnight PE, Najab J. Mann-Whitney U Test. In: The Corsini Encyclopedia of Psychology. Hoboken, NJ: John Wiley & Sons; 2010.
96. Cohen J. Statistical Power Analysis for the Behavioral Sciences Second Edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
97. Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm--an overview of tactics and tropes used online by the anti-vaccination movement. Vaccine 2012 May 28;30(25):3778-3789. [doi: 10.1016/j.vaccine.2011.11.112] [Medline: 22172504]
98. Martini C. Ad hominem arguments, rhetoric, and science communication. Stud Log Gramm Rhetor 2018;55(1):66. [doi: 10.2478/slgr-2018-0033]
99. Vivion M, Anassour Laouan Sidi E, Betsch C, Dionne M, Dubé E, Driedger SM, et al. Prebunking messaging to inoculate against COVID-19 vaccine misinformation: an effective strategy for public health. J Commun Healthc 2022 Mar 04;15(3):232-242. [doi: 10.1080/17538068.2022.2044606]
100. van der Linden S, Roozenbeek J, Maertens R, Basol M, Kácha O, Rathje S, et al. How can psychological science help counter the spread of fake news? Span J Psychol 2021 Apr 12;24:e25. [doi: 10.1017/sjp.2021.23]

101. Love B, Himelboim I, Holton A, Stewart K. Twitter as a source of vaccination information: content drivers and what they are saying. Am J Infect Control 2013 Jun;41(6):568-570. [doi: 10.1016/j.ajic.2012.10.016] [Medline: 23726548]
102. Meyer SB, Violette R, Aggarwal R, Simeoni M, MacDougall H, Waite N. Vaccine hesitancy and Web 2.0: exploring how attitudes and beliefs about influenza vaccination are exchanged in online threaded user comments. Vaccine 2019 Mar 22;37(13):1769-1774. [doi: 10.1016/j.vaccine.2019.02.028] [Medline: 30826142]
103. Caulfield T, Bubela T, Kimmelman J, Ravitsky V. Let's do better: public representations of COVID-19 science. FACETS 2021 Jan 01;6(1):403-423. [doi: 10.1139/facets-2021-0018]
104. Wang JT, Power CJ, Kahler CM, Lyras D, Young PR, Iredell J, et al. Communication ambassadors-an Australian social media initiative to develop communication skills in early career scientists. J Microbiol Biol Educ 2018 Mar;19(1):19.1.25 [FREE Full text] [doi: 10.1128/jmbe.v19i1.1428] [Medline: 29904520]
105. Petersen MB, Bor A, Jørgensen F, Lindholt MF. Transparent communication about negative features of COVID-19 vaccines decreases acceptance but increases trust. Proc Natl Acad Sci U S A 2021 Jul 20;118(29):e2024597118 [FREE Full text] [doi: 10.1073/pnas.2024597118] [Medline: 34292869]
106. Harzing AW, Alakangas S. Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. Scientometrics 2015 Nov 26;106(2):787-804. [doi: 10.1007/s11192-015-1798-9]
107. Boucher JC, Kim SY, Jessiman-Perreault G, Edwards J, Smith H, Frenette N, et al. HPV vaccine narratives on Twitter during the COVID-19 pandemic: a social network, thematic, and sentiment analysis. BMC Public Health 2023 Apr 14;23(1):694 [FREE Full text] [doi: 10.1186/s12889-023-15615-w] [Medline: 37060069]

Abbreviations

API: application programming interface
HPV: human papillomavirus
SJR: SCImago Journal Ranking

Edited by T Mackey; submitted 05.07.23; peer-reviewed by A Elyashar, WG Woodall, I Ballalai; comments to author 28.11.23; revised version received 13.02.24; accepted 20.03.24; published 09.05.24.

Please cite as:
Jessiman-Perreault G, Boucher JC, Kim SY, Frenette N, Badami A, Smith HM, Allen Scott LK
The Role of Scientific Research in Human Papillomavirus Vaccine Discussions on Twitter: Social Network Analysis
JMIR Infodemiology 2024;4:e50551
URL: https://infodemiology.jmir.org/2024/1/e50551
doi:10.2196/50551
PMID:38722678

Original Paper

Analyzing Questions About Alcohol in Pregnancy Using Web-Based Forum Topics: Qualitative Content Analysis

Nessie Felicia Frennesson[1], MSc; Julie Barnett[2], PhD; Youssouf Merouani[3], MSc; Angela Attwood[1], PhD; Luisa Zuccolo[4], PhD; Cheryl McQuire[5], PhD

[1]Tobacco and Alcohol Research Group, School of Psychological Science, University of Bristol, Bristol, United Kingdom

[2]Department of Psychology, University of Bath, Bath, United Kingdom

[3]Department of Economic History, School of Economics and Management, Lund University, Lund, Sweden

[4]Health Data Science Centre, Human Technopole, Milan, Italy

[5]Centre for Public Health, Bristol Medical School, University of Bristol, Bristol, United Kingdom

Corresponding Author:
Nessie Felicia Frennesson, MSc
Tobacco and Alcohol Research Group
School of Psychological Science
University of Bristol
12a Priory Road
Bristol, BS8 1TU
United Kingdom
Phone: 44 1173746633
Email: felicia.frennesson@bristol.ac.uk

Abstract

Background: Prenatal alcohol exposure represents a substantial public health concern as it may lead to detrimental outcomes, including pregnancy complications and fetal alcohol spectrum disorder. Although UK national guidance recommends abstaining from alcohol if pregnant or planning a pregnancy, evidence suggests that confusion remains on this topic among members of the public, and little is known about what questions people have about consumption of alcohol in pregnancy outside of health care settings.

Objective: This study aims to assess what questions and topics are raised on alcohol in pregnancy on a web-based UK-based parenting forum and how these correspond to official public health guidelines with respect to 2 critical events: the implementation of the revised UK Chief Medical Officers' (CMO) low-risk drinking guidelines (2016) and the first COVID-19 pandemic lockdown (2020).

Methods: All thread starts mentioning alcohol in the "Pregnancy" forum were collected from Mumsnet for the period 2002 to 2022 and analyzed using qualitative content analysis. Descriptive statistics were used to characterize the number and proportion of thread starts for each topic over the whole study period and for the periods corresponding to the change in CMO guidance and the COVID-19 pandemic.

Results: A total of 395 thread starts were analyzed, and key topics included "Asking for advice on whether it is safe to consume alcohol" or on "safe limits" and concerns about having consumed alcohol before being aware of a pregnancy. In addition, the Mumsnet thread starts included discussions and information seeking on "Research, guidelines, and official information about alcohol in pregnancy." Topics discussed on Mumsnet regarding alcohol in pregnancy remained broadly similar between 2002 and 2022, although thread starts disclosing prenatal alcohol use were more common before the introduction of the revised CMO guidance than in later periods.

Conclusions: Web-based discussions within a UK parenting forum indicated that users were often unclear on guidance and risks associated with prenatal alcohol use and that they used this platform to seek information and reassurance from peers.

(JMIR Infodemiology 2024;4:e58056) doi:10.2196/58056

KEYWORDS

social media; web-based forum; alcohol; pregnancy; prenatal health; prenatal alcohol exposure

Introduction

Background

Prenatal alcohol exposure (PAE) can lead to several detrimental outcomes, such as fetal alcohol spectrum disorder (FASD) [1] and developmental effects on both physical [2] and mental health [3]. In addition, those with FASD have a higher risk of experiencing problems in school, getting into trouble with the law, and having problems with alcohol and illicit drug use [4]. FASD has been mentioned in research since the early 1970s [5], and it is a complex diagnosis with a high rate of comorbidity [6], usually requiring a multidisciplinary team to diagnose it [7]. Estimates have shown that, on average, 9.8% of women worldwide consume alcohol in pregnancy, and in the United Kingdom, 40.1% of women report consuming alcohol in pregnancy, ranking the United Kingdom as having the fourth highest prevalence of PAE in the world [8]. It has been shown that approximately 8 in 1000 individuals have FASD in the general population globally, with the European region having a prevalence of approximately 20 in 1000, with approximately 1 in 13 children being born with FASD after PAE [9]. Estimates show that 1 in 67 pregnant women consuming alcohol gives birth to a child with fetal alcohol syndrome (the dysmorphic subtype of FASD) [8]. While there is strong evidence that high levels of PAE can be harmful, evidence on the effects of low to moderate PAE has been less conclusive. Uncertainties surrounding the risk of harm at low levels of PAE, coupled with conflicting messages from health professionals on low-moderate PAE, have been cited as reasons why some people choose not to abstain and for ongoing confusion about the risks of PAE [10]. Nevertheless, studies that have used robust methods to support causal inference have found that low-moderate levels of PAE can lead to adverse perinatal, physical [11], and developmental outcomes [12,13] and that apparent null and protective effects of PAE are likely due to residual confounding. Consequently, a recent review concluded that “any amount of prenatal alcohol exposure appeared to risk healthy child development” [14].

Previously, there was a consensus that UK guidelines on alcohol in pregnancy could be confusing because there was no clear recommendation [15]. In earlier guidelines, the National Institute for Health and Care Excellence stated that pregnant women should avoid alcohol in the first 3 months of pregnancy, and if they choose to drink, they should not drink more than 1 to 2 units twice per week [16]. In 2016, the UK Chief Medical Officers (CMOs) changed the guidelines to advise “if you are pregnant or planning a pregnancy, the safest approach is not to drink alcohol at all” [17]. This change of guidelines caused much debate. While some welcomed the introduction of the revised guidelines [18], others claimed that the new guidelines would lead to women having feelings such as guilt and anxiety if they were consuming alcohol during pregnancy [19]. Pregnant women feel there are too many guidelines to follow, arguably leading to increased stress [20]. It has also been expressed that pregnancy can lead to a perceived lack of agency and control [21] in this context, and the abstinence guideline can be perceived as “policing women” [22]. A study conducted in Denmark, a country with similar estimates of PAE as the United Kingdom, evaluated both knowledge and attitudes toward alcohol in pregnancy before and after their guidelines changed to advising women not to drink during pregnancy and showed no changes in either knowledge or attitudes [23].

Furthermore, the concept of abstaining from alcohol during pregnancy is not always clear, with some women experiencing confusion about, for example, whether food containing alcohol is safe or whether it is acceptable to consume so-called no- or low-alcohol (NoLo) products [20].

There is little research on the experiences and attitudes toward alcohol in pregnancy among the UK general population. This is particularly true for research on conversations in more informal and “naturalistic” settings, such as those on social media platforms. Within the context of the changed guidelines on alcohol in pregnancy and ongoing debates, a gap emerged in understanding informal dialogues, notably during events such as the COVID-19 pandemic. As concerns heightened regarding increased alcohol intake during pregnancy amid the lockdown, the need for research on the correlation between COVID-19 pandemic–related anxieties and evolving attitudes toward alcohol during and after the lockdown became apparent. In March 2020, the United Kingdom experienced its first lockdown [24]. This led to millions of people having to change how they lived, with everything from how to visit the midwife to having to give birth without their partners and loved ones present [25]. There was also the worry that people with alcohol dependence would not get the help they needed while the society was shut down [26]. In addition, there was a concern surrounding a potential increase in the level of alcohol consumption during pregnancy due to the stress and anxiety that the lockdown brought to many people [27]. Research has suggested that the shifts in drinking habits during the COVID-19 pandemic could have lasting effects on alcohol-related harm in the future for the general population in England [28]. However, to this date, there has been limited evidence supporting any relationship between COVID-19 concerns and increased alcohol consumption during pregnancy [29]. Therefore, more research is needed to examine the potential change in attitudes toward alcohol in pregnancy after the COVID-19 pandemic lockdown.

Many pregnant women use the internet to search for information related to their pregnancy [30]. A recent study showed that as many as 44% of new mothers used social media to keep in contact and communicate with others in the same situation [31]. Mumsnet [32] was founded in early 2000 and is one of the most prominent web-based forums for parents in the United Kingdom, with approximately 7 million monthly visitors [33]. Mumsnet was initially created as a web-based space where people could ask for and give advice and share knowledge to make parents’ lives easier [33]. While Mumsnet is open to anyone, it has previously been described as having a majority of middle-class and university-educated users [34]. Because many people use social media and the internet to seek health information [35] and to find support during pregnancy [30], Mumsnet presents an excellent opportunity for researchers to capture the unmediated opinions and thoughts about alcohol during pregnancy. The forum has previously been used to address topics such as breastfeeding [36], regretting motherhood [37], and maternal feelings [38]. Therefore, with the use of Mumsnet thread starts, this study will explore what topics related to alcohol during pregnancy are discussed and if their nature has changed since the start of Mumsnet in 2002, with the change of CMO guidelines in 2016 and the COVID-19 pandemic as key time points for comparison.

Aims and Objectives

Given that social media and the internet can be used by those seeking to gain real-time insight into people's behaviors and attitudes as well as to identify how people perceive public health messages [39], this study aimed to explore what issues and topics are raised with regard to alcohol use in pregnancy in web-based parenting forums. In addition, it aimed to explore if there has been a change in the different issues and topics with respect to 3 time points: before the implementation of the current CMO low-risk drinking guidelines, after the implementation of the CMO guidelines, and after the first COVID-19 pandemic lockdown.

The specific research questions in this study were as follows:

- What topics relating to alcohol in pregnancy are raised on web-based parenting forums?
- Have these topics changed in content or volume with respect to 3 time points: before the implementation of the current CMO guidelines, after the implementation of the CMO guidelines, and after the first COVID-19 pandemic lockdown?

Methods

Data Source

Mumsnet allows members to post anonymously on the forum called "Mumsnet Talk." Talk consists of different subforums (eg, the "Pregnancy" subforum). Mumsnet users can post thread starts, usually by asking a question, and other users can reply to these thread starts by adding a comment. Users are identified by unique usernames. The forum is open for everyone to view; however, users must be registered to post content.

Ethical Considerations

This study followed ethical guidelines for internet-mediated research. Ethical considerations for research using social media data differ from traditional research [40], and it is essential to distinguish between public and private data [41]. Because Mumsnet does not require users to log in to read the forum, and is available to the public, it was considered public data and informed consent was not required [42]. To support anonymity in our study, we have not included usernames and endeavored to remove personally identifiable information from the data during the cleaning process. We excluded thread starts in which the user stated that they were aged <18 years. In line with the British Psychological Society guidance [43], direct quotes were not reproduced in this study, and all quotes have been paraphrased.

A favorable ethical opinion was obtained from the School of Psychological Science Research Ethics Committee at the University of Bristol in August 2023 (ethics approval code 14455).

Search Strategy

Data were collected through web scraping code by authors NFF and YM. The data were collected from the "Pregnancy" topic in the "Talk" part of Mumsnet and included original thread starts that mentioned alcohol in the title or text. Duplicate thread starts and threads unrelated to alcohol use during pregnancy were removed manually. After web scraping, thread starts were stored in an Excel (Microsoft Corp) file with information on the username, date and time of the thread start, and the thread start itself.

Data Analysis

Thread starts were analyzed in 3 groups according to the time and date in which they were posted: "Pre-CMO recommendation update," before the introduction of revised low-risk drinking guidelines in 2016 (August 24, 2002, to January 7, 2016); "Post-CMO recommendation update," after the change of guidelines (January 8, 2016, to March 22, 2020); and "Post–COVID-19 Pandemic Lockdown," from the first lockdown in the United Kingdom up until the last date of the data collection (March 23, 2020, to November 12, 2022). A content analysis was conducted following the steps described in the study by Elo and Kyngäs [44]. This approach was appropriate as the aim of the study was to map the landscape of discussions on alcohol in pregnancy on Mumsnet, including patterns and time trends in people's views and experiences.

NFF read the thread starts, became familiarized with the data, and applied preliminary code labels to organize the data. An inductive approach was used because little is known about how mothers use web-based forums to discuss alcohol and pregnancy. Categories were generated after finding patterns among the codes to better understand the data [44]. Throughout the process, the categories were reviewed and refined. During each step of the analysis, both codes and categories were discussed among the researchers NFF, JB, AA, LZ, and CM and refined accordingly. We used descriptive statistics to describe the proportion of thread starts for each category for each of the prespecified periods. The analysis workflow is based on the process outlined in the study by Elo and Kyngäs [44].

Results

Overview

The web scraping resulted in 803 thread starts, which, after eliminating duplicates and irrelevant thread starts, resulted in 395 thread starts included in the analysis. Multimedia Appendix 1 provides an overview of the 9 categories and each code within them, together with the number of times the categories appear in each period. Although the categories are presented separately, some overlap does exist.

The results show that while the categories of Mumsnet thread starts relevant to alcohol use in pregnancy remained broadly similar over time, there were some changes in the relative prevalence of different topics over time. Category headings, frequencies over time, and illustrative quotes are presented in Multimedia Appendix 1. The categories are presented in detail in subsequent sections.

Asking for Advice on Whether It Is Safe to Consume Alcohol or on Safe Limits

Looking at the questions raised within this category, it became apparent that the people posting on Mumsnet felt insecure about whether it is safe to consume alcohol during pregnancy or if it is safe to have a glass or 2 on a special occasion such as weddings or birthday celebrations:

> *I am 30 weeks pregnant and haven't had a single drink but it is my friend's wedding and I really want to have a glass of champagne, is this ok?*

Thread starts in this category also addressed the issue of not knowing whether it is safe to eat certain foods or desserts as they contain alcohol, for example, a tiramisu or red wine sauce:

> *Is it ok to eat dessert that has alcohol in it? I am in my third trimester.*

There was uncertainty around NoLo options, with questions raised about whether a 0.5% level of alcohol is safe to consume during pregnancy:

> *Is it okay to drink none alcoholic ciders? This might be a stupid question but it does say 0.5% so is there still alcohol in there that can hurt my baby?*

Overall, 17.5% (69/395) of all the included thread starts appeared in this category. The category saw a slight decrease in the percentage of thread starts asking about safe limits to drink over time. Many of the thread starts in the first period, before the CMO recommendation update, mentioned the timing of the pregnancy, which could be a result of the change in guidelines.

Consumed Alcohol Before Knowing About Pregnancy

Most of the thread starts within this category showed some expression of worry or anxiety that the thread starters had consumed alcohol before they found out about their pregnancy. They also sought reassurance from others who have been in a similar situation:

> *I've just found out I'm pregnant after weeks of unknowingly consuming alcohol and indulging in partying. I'm feeling guilty and concerned about any potential harm to my baby. Can anyone share their experiences if they have been in a similar situation?*

Moreover, many thread starters mentioned that they were usually not heavy drinkers. Still, due to situations such as birthday parties or Christmas celebrations, their alcohol intake had been higher than usual:

> *I am freaking out please help. Just found out that I am 6 weeks pregnant and have been drinking so much, especially because of Christmas celebrations, I promise I am usually not a heavy drinker. What should I do? Could not live with myself if something happens to the baby!*

Some thread starters were asking if they should consider having an abortion following an unintended PAE, even if the baby is wanted:

> *Need advice since I am worrying myself sick! I am pregnant and have been drinking because I was on holiday (usually only have a glass of wine once a month). Has anyone else experienced this and their baby turned out fine? Should I just have an abortion even if I really want this baby? How could I be this stupid!*

Overall, 28.1% (111/395) of all the included thread starts appeared in this category. Throughout the different periods, there is a notable increase in the percentage of thread starts regarding the worry that they may have harmed their baby; this worry is expressed more frequently after the revised CMO low-risk guidelines were introduced in 2016.

Research, Guidelines, and Official Information About Alcohol in Pregnancy

The threads started in this category were all related to research, guidelines, and information from official sources (eg, National Health Service) about alcohol in pregnancy. Thread starters throughout all periods expressed that they found this information confusing, conflicting, or untrustworthy:

> *It's a bit puzzling to me. The NHS advises against it, and I've come across articles saying the same, yet in my real-life circle, many pregnant women I know enjoy the occasional drink, even if it's just a glass. It got me thinking if there's a significant gap between official recommendations and what's happening in practice?*

In the first period, thread starters were asking what others think about the guidelines and also sharing information on how it is acceptable to consume small amounts during pregnancy:

> *To be completely honest, there is no research showing that it is really bad for the baby!*

Some skepticism toward the guidelines can be seen in how thread starters expressed that there is no evidence that small amounts of alcohol in pregnancy have an adverse effect. Those who posted in the later periods also expressed that the guidelines were not feasible for "real people" and that they were too strict while contending that most people do not follow them:

> *The internet just gives you information about that you shouldn't drink and that no amount is safe but surely this is not how real people see it. I think they are just trying to scare us with all of these rules!*

Of the 395 included thread starts, 46 (11.6%) of these appeared in this category. Closer to 17.5% (37/213) of the thread starts in the first period (pre-CMO updated guidance) were related to research, guidelines, and information about alcohol use in pregnancy, but this was less frequent in the later periods with 4.9% (5/103) of the thread starts in the post-CMO recommendation update and 5.1% (4/79) in the post–COVID-19 pandemic lockdown.

NoLo Products

This category covered conversations about the consumption of NoLo products during pregnancy. Reasons for seeking NoLo alternatives included not feeling left out in a group that is consuming alcohol and also because some missed the taste of alcohol. The need for this type of product appeared greater during celebrations such as Christmas and weddings:

> *I don't want to feel left out at the wedding so do you have any recommendations of what I can drink instead of champagne?*

There were also thread starters who had been consuming NoLo options during pregnancy, assuming that these were completely alcohol free (ie, 0.0% alcohol by volume [ABV]) and later realized that they contained some alcohol. This was associated with a concern that this alcohol level might have harmed their babies, and they were seeking reassurance from others on Mumsnet:

> *I am crying so much, had a couple of none-alcohol beers and now I realised that they are 0.05%! What if I have hurt my baby?*

Of the 395 included thread starts, 52 (13.2%) of these appeared in this category. All of the 3 periods included conversations about alternative products to alcohol to drink during pregnancy, and the proportions of these remained relatively stable over time.

How to Hide Not Consuming Alcohol to Conceal Pregnancy?

This category reflected the worry thread starters felt regarding how to conceal them not consuming alcohol. This was of particular concern when they were invited to social situations where alcohol would be available:

> *How can I hide that I am not drinking when we go to the pub? With table service it's going to be so much harder!*

Overall, 10.4% (41/395) of all the included thread starts appeared in this category. This category saw a decrease in thread starts discussing how to hide not consuming alcohol in pregnancy during the post–COVID-19 pandemic lockdown period.

Have Been Consuming Alcohol During Pregnancy But Now Worried

The thread starts within this category mainly address the issue of wanting reassurance that the baby will be fine although they have been consuming alcohol during pregnancy:

> *I am in my third trimester and had two drinks yesterday and now I am thinking that I might have hurt my baby, can someone tell me that this is fine?*

Some thread starters mentioned that they have been consuming alcohol but later on read information about how alcohol can affect the baby, and therefore regret the decision to drink:

> *I have just had a few glasses of wine here and there and haven't really thought of it but now I started reading about FAS and I am terrified. I cannot have an abortion since it is too late, but what should I do?*

Overall, 4.3% (17/395) of all the included thread starts appeared in this category. For this category, the proportion of thread starts for the first period was 6.6% (14/213), the second period had a percentage of 1.9% (2/103), and the last period had a proportion of 1.3% (1/79), falling in this category.

Are Consuming Alcohol During Pregnancy, Not Worried About PAE

Some of the thread starters in this category were seeking reassurance from others on Mumsnet in which they were seeking affirmation that consuming alcohol does not make them a bad person or mother:

> *I am going for a nice meal to celebrate, but I am worried that people will judge me if I have a drink.*

Of the 395 included thread starts, 18 (4.6%) of these appeared in this category. All the thread starts in this category appeared in the first and second periods.

Consumed Alcohol by Mistake

This category covers those who have consumed alcohol by mistake while eating a dessert, consuming nonalcoholic beers, or being served alcohol without realizing and being worried about that:

> *I went to the pub with my friends and ordered a non-alcohol option but after drinking most of it I realised that it was alcohol! Will I be ok? Freaking out!*

There was also a concern that using other products containing alcohol, for example, a mouthwash or hand sanitizer, could have also harmed their baby:

> *I have been using mouthwash throughout my whole pregnancy but it has alcohol in it, what have I done? So scared right now!*

Overall, 6.3% (25/395) of all the included thread starts appeared in this category. The highest percentage of thread starts in this category occurred in the post–COVID-19 pandemic lockdown period with 15% (12/79) of thread starts belonging to this category. The first period had 1.9% (4/213) of thread starts belonging to this category, and the second period had 8.7% (9/103) of the thread starts belonging to this category.

It Is Hard Not to Consume Alcohol During Pregnancy

Although a smaller category, this category expressed a sadness or emptiness about having to give up alcohol during the pregnancy. Some users suggest that this could be a reflection of the life changes that come with having a child and also feeling left out from social situations:

> *I feel sad about not drinking, or maybe it's about my life changing so much with this new baby on its way, am I the only one with these thoughts?*

Overall, 4.1% (16/395) of all the included thread starts appeared in this category and the proportion of this category remained stable during the different periods.

Discussion

Principal Findings

This study sought to explore the topics relating to alcohol in pregnancy, which are raised on a web-based parenting forum.

Moreover, it sought to explore if these topics had changed in content or volume, both after the change of CMO guidelines in 2016 (changing from advising women to abstain for the first trimester and not drinking more than 1 to 2 units per week to advising complete abstinence if pregnant or planning a pregnancy) and after the first COVID-19 pandemic lockdown in 2020. Through our content analysis of thread starts on Mumsnet, the United Kingdom's leading parenting web-based forum, it was possible to evaluate what topics were commonly raised with regard to alcohol consumption in pregnancy and if and how these topics had changed over time. It is important to note that the 3 periods span 16, 4, and 2 years.

Much of the discussion on Mumsnet was around seeking reassurance and wanting to know if others had been in the same situation. For example, thread starters seeking reassurance that they have not harmed their babies by consuming alcohol before knowing about the pregnancy and asking if others had experienced something similar. It was evident that thread starters were not only seeking reassurance but also wanting to obtain information about alcohol in pregnancy both by asking for guidelines or wanting to know if such a thing as a safe limit exists. This brings up the risk of inaccurate information being shared among the users on Mumsnet. According to the World Health Organization [45], too much information and false information could lead to worsening outcomes in terms of health. Further research is required to investigate if this is true for the information shared on Mumsnet. In many categories, the results showed how thread starters were confused or worried about safe limits, including whether it is safe or not to consume products that may contain traces of alcohol. This is in line with previous research, showing that the concept of abstinence is not always clear, with confusion about, for example, if food containing alcohol is safe or if it is acceptable to consume nonalcohol options during pregnancy [20].

It was evident that there was a concern among the thread starters about having consumed alcohol before they found out that they were pregnant and that this caused stress and anxiety for some thread starters. There is no known safe limit for alcohol consumption during pregnancy. Some studies have shown that there was no relationship between consuming alcohol during the early days of pregnancy and outcomes such as low birth weight and spontaneous preterm birth [46], while others have suggested that alcohol consumption during the first trimester of pregnancy can increase the risk of spontaneous abortion [47,48]. Many of the thread starters had discussed the consumption of alcohol with their midwives or other professionals, but the concern remained. Previous research has shown that there is a lack of a standardized approach to how midwives approach the topic of alcohol consumption during pregnancy [16]. This could indicate that there is a need for professionals to give accurate information and at the same time being able to reduce any anxiety and stress that alcohol consumption could have caused. It has been suggested that midwives should be offered training in communication skills and in delivering alcohol interventions [16]. Reducing stress during pregnancy is especially important as it has been reported that stress can lead to outcomes such as low birth weight [49] and obesity in the offspring [50]. Research has shown that there exists a social pressure to consume alcohol, which causes a challenge when someone wishes to conceal their pregnancy [51]. This was also prevalent in the discussions on Mumsnet, where thread starters wanted advice on hiding that they were not consuming alcohol due to their pregnancy.

Our study demonstrated that the proportion of thread starts being brought up regarding alcohol consumption has changed over time. These temporal changes were most evident in the category "Research, guidelines, and official information about alcohol in pregnancy" as well as the 2 categories addressing consuming alcohol during pregnancy. The former contains topics such as not believing in the research carried out about alcohol in pregnancy or sharing information on how small amounts of alcohol are not harmful, with these all disappearing in the later periods. One topic that appeared in all periods was how confusing or conflicting the research or guidelines on alcohol consumption in pregnancy were. This is in line with previous research, showing how conflicting advice can cause stress in pregnancy and the need for reliable information [52]. Those who posted thread starts on Mumsnet were also expressing how the guidelines and research were not clear, and quotes illustrated that some felt that no "real people" could follow all the rules. This is in line with previous research about how women feel like there are too many guidelines [20] and how the abstinence message can be perceived as policing women [22]. Interestingly, thread starts asking for the guidelines or asking for more information were only observed before the introduction of the updated CMOs' low-risk drinking guidance on alcohol in pregnancy [17]. This may indicate that the updated CMO guidance has made the recommendation to avoid alcohol during pregnancy clearer and easier to understand. The latter categories, addressing consuming alcohol during pregnancy, show that expressing that one is consuming alcohol during pregnancy on the internet was more common before the introduction of the revised CMOs' guidance than in the later periods. This could indicate that the actual prevalence of alcohol consumption during pregnancy has gone down. However, research suggests that the prevalence still remains high [8]. It could also indicate that it has become less socially acceptable to disclose alcohol use during pregnancy. Furthermore, no threads started during the COVID-19 pandemic period expressed any alcohol consumption due to the lockdown. This is in line with previous research showing that the reported rates of alcohol consumption during pregnancy were lower after the pandemic than before the pandemic [29].

Moreover, many of the thread starts in the category "Asking for advice on whether it is safe to consume alcohol or on safe limits" during the first period mentioned the timing of the pregnancy, which could have been a result of the change in guidelines. The change in guidelines was that the previous National Institute for Health and Care Excellence guidelines suggested that pregnant women should avoid alcohol in the first 3 months of pregnancy, and if they chose to drink, they should not consume more than 1 to 2 units twice per week [16]. However, a 2020 survey by the National Organisation for FASD showed that awareness of the current CMOs guidance that the safest approach is not to consume alcohol at all if you are pregnant or if you could become pregnant remains low among some population subgroups, particularly young people (aged 18-25 years) [53]. This shows how important it is to communicate research in a way that is acceptable, understandable, and accessible for all. Our study showed that some Mumsnet users missed drinking alcohol while they were pregnant and wanted to know if others agreed that a glass of alcohol was acceptable, particularly on special occasions such as weddings or birthday celebrations. It was also evident that some of the forum users were trying to minimize the potential risks of consuming alcohol by referring to how previous generations had been consuming alcohol without clear adverse outcomes or how other countries have less strict guidelines. This was most apparent in content posted before the introduction of updated CMO low-risk drinking guidance in 2016.

The categories identified in this study highlight the importance of providing reliable and trustworthy information about alcohol consumption in pregnancy. This is relevant for scholars, professionals, and organizations, such as midwives and the Public Health of England. The study highlights the interactive nature of web-based forums, demonstrating an endeavor to establish social connections and seek peer reassurance. For future research, it is essential to investigate how these thread starts are replied to and how people manage their worries through interactions with others. It will also be important to investigate how conversations on social media can be used to identify knowledge gaps and preferences for the nature and format of prenatal health messaging and to explore the measurable impact of key public health and policy events on outcomes related to PAE.

Strengths and Limitations

This study retrieved all available data related to alcohol in pregnancy from the largest dedicated web-based parenting forum in the United Kingdom. Moreover, to our knowledge, it is the first study to provide insight into the nature of web-based conversations on alcohol use in the United Kingdom and how the trends in these have changed over time in relation to key policy and public health events. Moreover, these findings are relevant to policy makers. This includes the current 2023 consultation [54] on NoLo products. The consultation intended to set out if the ABV that can be deemed “NoLo products” should be increased. The findings in this study suggest that increasing the threshold for ABV from 0.05% to 0.5% could exacerbate concerns among pregnant people who report having mistakenly consumed alcohol and also increase the uncertainties about the safe limits of these products. In addition, time trends in conversations are significant as they reveal uncertainties among pregnant people regarding topics such as the current CMO guidance and can offer valuable priorities to inform improved communication, reach, and preference for prenatal health information.

The choice to specifically search for mentions of “alcohol” within the “Pregnancy” topic on Mumsnet was made to ensure that only thread starts relevant to alcohol consumption during pregnancy were included, thereby excluding discussions unrelated to this specific context (such as threads discussing alcohol consumption while breastfeeding). This may have resulted in some critical thread starts being excluded. In 2021, it was reported that approximately 20,000 posts were created daily on Mumsnet [55]. Given the number of daily posts, collecting all of them and manually going through them was not feasible. For future research, natural language processing or topic modeling could be used to analyze a larger data set. Furthermore, because not everyone has the same access to the digital space, the voices heard on Mumsnet might not be representative of the United Kingdom, which could have led to some potential bias in the data. This is especially true as the demographics of Mumsnet have previously been described as middle class and university educated [34], thus omitting other socioeconomic groups from this analysis. It is important to note that previous research has shown that one predictor of alcohol consumption during pregnancy is higher education [56], which could limit the relevance of the findings to some subgroups of the general population. Previous research has also shown that social media use is more common among those in a higher socioeconomic group [57].

Conclusions

This study provides insight into how mothers and expecting mothers use Mumsnet to raise topics that are important to them regarding alcohol consumption in pregnancy and illustrates how these topics have changed since the start of Mumsnet. The findings suggest that mothers and expecting mothers use Mumsnet primarily to seek reassurance and information from others in similar situations. Our findings also suggest that the topics and the proportion of thread starts relating to each topic have changed over time, with results indicating less confusion about the current guidelines and research about alcohol in pregnancy in more recent times. The study also provides insight into the worries and anxiety that pregnant women report experiencing if they had consumed alcohol before finding out about the pregnancy and the importance of seeking advice and reassurance from peers on how to manage that worry. These findings suggest that innovative interventions, such as peer support initiatives, may offer a promising approach to prenatal alcohol prevention, warranting further investigation.

Acknowledgments

This work was supported in part by grant MR/N0137941/1 for the GW4 BioMed Medical Research Council Doctoral Training Partnership, awarded to the Universities of Bath, Bristol, Cardiff, and Exeter from the Medical Research Council and United Kingdom Research and Innovation for NFF. During this work, CM was supported by the National Institute for Health and Care Research School for Public Health Research (grant PD-SPH-2015). The views expressed are those of the authors and not necessarily those of the National Institute for Health and Care Research or the Department of Health and Social Care.

Data Availability

Because the site contents are copyright of Mumsnet, publishing a data set collecting posts or threads is not possible (refer to "Copyright" [58]). Instead, the summary statistics or analysis results will be published via data.bris.

Authors' Contributions

NFF became familiarized with Mumsnet to decide what data should be downloaded and analyzed. NFF and YM used a web scraping technique to download the data. NFF read and reread the data multiple times, made notes about initial impressions, and removed data irrelevant to the study. JB supported in reading the data to make any notes of insights. NFF coded the data and discussed the codes and their meaning with JB. NFF started grouping the codes if they shared the same meaning. NFF started forming categories from the coded data, which involved JB reviewing the codes and categories. NFF, JB, AA, LZ, and CM reviewed the categories and their related codes and discussed their definitions. NFF wrote the draft for the paper with JB, AA, LZ, and CM, who all provided critical feedback and contributed to the review and editing. NFF designed the study with JB, AA, LZ, and CM contributing to its development. NFF and YM web scraped the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Included categories.
[DOCX File , 24 KB - infodemiology_v4i1e58056_app1.docx]

References

1. Fetal alcohol spectrum disorders. Nat Rev Dis Primers 2023 Feb 23;9(1):10. [doi: 10.1038/s41572-023-00426-5] [Medline: 36823194]
2. Hamułka J, Zielińska MA, Chądzyńska K. The combined effects of alcohol and tobacco use during pregnancy on birth outcomes. Rocz Panstw Zakl Hig 2018;69(1):45-54 [FREE Full text] [Medline: 29517191]
3. Burd L. FASD and ADHD: are they related and how? BMC Psychiatry 2016 Sep 22;16(1):325 [FREE Full text] [doi: 10.1186/s12888-016-1028-x] [Medline: 27655173]
4. Streissguth AP, Bookstein FL, Barr HM, Sampson PD, O'Malley K, Young JK. Risk factors for adverse life outcomes in fetal alcohol syndrome and fetal alcohol effects. J Dev Behav Pediatr 2004 Aug;25(4):228-238. [doi: 10.1097/00004703-200408000-00002] [Medline: 15308923]
5. Jones KL, Smith DW. Recognition of the fetal alcohol syndrome in early infancy. Lancet 1973 Nov 03;302(7836):999-1001. [doi: 10.1016/s0140-6736(73)91092-1] [Medline: 4127281]
6. Popova S, Lange S, Shield K, Mihic A, Chudley AE, Mukherjee RA, et al. Comorbidity of fetal alcohol spectrum disorder: a systematic review and meta-analysis. Lancet 2016 Mar 05;387(10022):978-987. [doi: 10.1016/S0140-6736(15)01345-8] [Medline: 26777270]
7. Cook JL, Green CR, Lilley CM, Anderson SM, Baldwin ME, Chudley AE, et al. Fetal alcohol spectrum disorder: a guideline for diagnosis across the lifespan. CMAJ 2016 Feb 16;188(3):191-197 [FREE Full text] [doi: 10.1503/cmaj.141593] [Medline: 26668194]
8. Popova S, Lange S, Probst C, Gmel G, Rehm J. Estimation of national, regional, and global prevalence of alcohol use during pregnancy and fetal alcohol syndrome: a systematic review and meta-analysis. Lancet Glob Health 2017 Mar;5(3):e290-e299 [FREE Full text] [doi: 10.1016/S2214-109X(17)30021-9] [Medline: 28089487]
9. Lange S, Probst C, Gmel G, Rehm J, Burd L, Popova S. Global prevalence of fetal alcohol spectrum disorder among children and youth: a systematic review and meta-analysis. JAMA Pediatr 2017 Oct 01;171(10):948-956 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.1919] [Medline: 28828483]
10. Muggli E, Halliday J, Hearps S, Nguyen TN, Penington A, Thompson DK, et al. Low to moderate prenatal alcohol exposure and neurodevelopment in a prospective cohort of early school aged children. Sci Rep 2024 Mar 27;14(1):7302 [FREE Full text] [doi: 10.1038/s41598-024-57938-7] [Medline: 38538856]
11. Muggli E, Matthews H, Penington A, Claes P, O'Leary C, Forster D, et al. Association between prenatal alcohol exposure and craniofacial shape of children at 12 months of age. JAMA Pediatr 2017 Aug 01;171(8):771-780 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.0778] [Medline: 28586842]
12. Mamluk L, Jones T, Ijaz S, Edwards HB, Savović J, Leach V, et al. Evidence of detrimental effects of prenatal alcohol exposure on offspring birthweight and neurodevelopment from a systematic review of quasi-experimental studies. Int J Epidemiol 2021 Jan 23;49(6):1972-1995 [FREE Full text] [doi: 10.1093/ije/dyz272] [Medline: 31993631]
13. Zuccolo L, Lewis SJ, Smith GD, Sayal K, Draper ES, Fraser R, et al. Prenatal alcohol exposure and offspring cognition and school performance. A 'Mendelian randomization' natural experiment. Int J Epidemiol 2013 Oct;42(5):1358-1370 [FREE Full text] [doi: 10.1093/ije/dyt172] [Medline: 24065783]

14. Römer P, Mathes B, Reinelt T, Stoyanova P, Petermann F, Zierul C. Systematic review showed that low and moderate prenatal alcohol and nicotine exposure affected early child development. Acta Paediatr 2020 Dec;109(12):2491-2501. [doi: 10.1111/apa.15453] [Medline: 32603488]
15. Mather M, Wiles K, O'Brien P. Should women abstain from alcohol throughout pregnancy? BMJ 2015 Oct 06;351:h5232. [doi: 10.1136/bmj.h5232] [Medline: 26442904]
16. Schölin L, Watson J, Dyson J, Smith L. Alcohol guidelines for pregnant women: barriers and enablers for midwives to deliver advice. The Institute of Alcohol Studies. 2019 Aug. URL: https://www.ias.org.uk/uploads/pdf/IAS%20reports/rp37092019.pdf [accessed 2024-06-08]
17. Alcohol guidelines review – report from the guidelines development group to the UK Chief Medical Officers. Department of Health, United Kingdom Government. 2016 Jan. URL: https://assets.publishing.service.gov.uk/media/5a7507d8ed915d3c7d529d6f/GDG_report-Jan2016.pdf [accessed 2024-06-08]
18. Hear our voices: FASD stakeholders share their experiences with policy makers. National Organisation for Foetal Alcohol Syndrome-UK. 2018 May. URL: https://www.sendreviewportal.net/ws/media-library/f89f082d1d294f0285c44692895153a3/hear-our-voices-fasd.pdf [accessed 2024-06-08]
19. Alcohol in pregnancy – what are the issues? British Pregnancy Advisory Service. URL: https://www.bpas.org/our-cause/campaigns/briefings/alcohol-in-pregnancy/ [accessed 2024-06-08]
20. Pehlke-Milde J, Radu I, Gouilhers S, Hammer R, Meyer Y. Women's views on moderate and low alcohol consumption: stages of the subjective transition from pregnancy to postpartum. BMC Pregnancy Childbirth 2022 Dec 05;22(1):902 [FREE Full text] [doi: 10.1186/s12884-022-05247-0] [Medline: 36464711]
21. Haaker M. The power of the pregnant body: perspectives of agency and autonomy in pregnancy. Genealogy 2021 Jan 31;5(1):13. [doi: 10.3390/genealogy5010013]
22. Holland K, McCallum K, Walton A. 'I'm not clear on what the risk is': women's reflexive negotiations of uncertainty about alcohol during pregnancy. Health Risk Soc 2016 Mar 28;18(1-2):38-58. [doi: 10.1080/13698575.2016.1166186]
23. Kesmodel US, Urbute A. Changes in drinking patterns, and attitudes toward and knowledge about alcohol consumption during pregnancy in a population of pregnant Danish women. Alcohol Clin Exp Res 2019 Jun;43(6):1213-1219. [doi: 10.1111/acer.14031] [Medline: 30908663]
24. Brown J, Kirk-Wade E, Baker C, Barber S. Coronavirus: a history of English lockdown laws. House of Commons Library. 2021 Dec 22. URL: https://commonslibrary.parliament.uk/research-briefings/cbp-9068/ [accessed 2024-06-04]
25. Gray A, Barnett J. Welcoming new life under lockdown: exploring the experiences of first-time mothers who gave birth during the COVID-19 pandemic. Br J Health Psychol 2022 May;27(2):534-552 [FREE Full text] [doi: 10.1111/bjhp.12561] [Medline: 34633132]
26. Finlay I, Gilmore I. Covid-19 and alcohol-a dangerous cocktail. BMJ 2020 May 20;369:m1987. [doi: 10.1136/bmj.m1987] [Medline: 32434792]
27. Sher J. Fetal alcohol spectrum disorders: preventing collateral damage from COVID-19. Lancet Public Health 2020 Aug;5(8):e424 [FREE Full text] [doi: 10.1016/S2468-2667(20)30159-6] [Medline: 32681819]
28. Angus C, Henney M, Pryce R. Modelling the longer-term health and health inequality impacts of changes in alcohol consumption during the COVID-19 pandemic in England. J Public Health (Oxf) 2024 May 29;46(2):286-293 [FREE Full text] [doi: 10.1093/pubmed/fdae010] [Medline: 38304989]
29. Kar P, Tomfohr-Madsen L, Giesbrecht G, Bagshawe M, Lebel C. Alcohol and substance use in pregnancy during the COVID-19 pandemic. Drug Alcohol Depend 2021 Aug 01;225:108760 [FREE Full text] [doi: 10.1016/j.drugalcdep.2021.108760] [Medline: 34102507]
30. Sayakhot P, Carolan-Olah M. Internet use by pregnant women seeking pregnancy-related information: a systematic review. BMC Pregnancy Childbirth 2016 Mar 28;16:65. [doi: 10.1186/s12884-016-0856-5] [Medline: 27021727]
31. Baker B, Yang I. Social media as social support in pregnancy and the postpartum. Sex Reprod Healthc 2018 Oct;17:31-34. [doi: 10.1016/j.srhc.2018.05.003] [Medline: 30193717]
32. Mumsnet. URL: https://www.mumsnet.com/ [accessed 2024-06-11]
33. About Mumsnet. Mumsnet. URL: https://www.mumsnet.com/i/about-us [accessed 2024-06-08]
34. Pedersen S. The good, the bad and the 'good enough' mother on the UK parenting forum Mumsnet. Womens Stud Int Forum 2016;59:32-38. [doi: 10.1016/j.wsif.2016.09.004]
35. Alamer F, Al-Edreese T. Pregnant women utilization of internet and social media for health education in Saudi Arabia: a thematic analysis. J Health Inform Dev Ctries 2021 Mar 14;15(1):1-16 [FREE Full text]
36. Caes L, Abbott K, Currie S. Exploring women's perceptions of pain when breastfeeding using online forums. Int Breastfeed J 2021 Oct 18;16(1):84 [FREE Full text] [doi: 10.1186/s13006-021-00426-9] [Medline: 34663383]
37. Matley D. "I miss my old life": regretting motherhood on Mumsnet. Discourse Context Media 2020 Oct;37:100417. [doi: 10.1016/j.dcm.2020.100417]
38. Pedersen S, Lupton D. 'What are you feeling right now?' Communities of maternal feeling on Mumsnet. Emot Space Soc 2018 Feb;26:57-63. [doi: 10.1016/j.emospa.2016.05.001]

39. Social media strategy development: a guide to using social media for public health communication. European Centre for Disease Prevention and Control. 2016. URL: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/social-media-strategy-guide-for-public-health-communication.pdf [accessed 2024-06-08]
40. Using social media for social research: an introduction. Social Media Research Group, United Kingdom Government. 2016 May. URL: https://assets.publishing.service.gov.uk/media/5a75c2e740f0b6488c78ec0a/GSR_Social_Media_Research_Guidance_-_Using_social_media_for_social_research.pdf [accessed 2024-06-08]
41. Townsend L, Wallace C. Social media research: a guide to ethics. The University of Aberdeen. URL: https://www.gla.ac.uk/media/Media_487729_smxx.pdf [accessed 2024-06-08]
42. Privacy policy. Mumsnet. URL: https://www.mumsnet.com/i/privacy-policy [accessed 2024-06-08]
43. Ethics guidelines for internet-mediated research. The British Psychological Society. 2021 Jun 07. URL: https://www.bps.org.uk/guideline/ethics-guidelines-internet-mediated-research [accessed 2024-06-08]
44. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
45. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic#tab=tab_1 [accessed 2024-06-08]
46. McCarthy F, O'Keeffe LM, Khashan AS, North RA, Poston L, McCowan LM, et al. Association between maternal alcohol consumption in early pregnancy and pregnancy outcomes. Obstet Gynecol 2013 Oct;122(4):830-837. [doi: 10.1097/AOG.0b013e3182a6b226] [Medline: 24084541]
47. Sundermann AC, Velez Edwards DR, Slaughter JC, Wu P, Jones SH, Torstenson ES, et al. Week-by-week alcohol consumption in early pregnancy and spontaneous abortion risk: a prospective cohort study. Am J Obstet Gynecol 2021 Jan;224(1):97.e1-97e16 [FREE Full text] [doi: 10.1016/j.ajog.2020.07.012] [Medline: 32673615]
48. Mamluk L, Edwards HB, Savović J, Leach V, Jones T, Moore TH, et al. Low alcohol consumption and pregnancy and childhood outcomes: time to change guidelines indicating apparently 'safe' levels of alcohol during pregnancy? A systematic review and meta-analyses. BMJ Open 2017 Aug 03;7(7):e015410 [FREE Full text] [doi: 10.1136/bmjopen-2016-015410] [Medline: 28775124]
49. Lima SA, El Dib RP, Rodrigues MR, Ferraz GA, Molina AC, Neto CA, et al. Is the risk of low birth weight or preterm labor greater when maternal stress is experienced during pregnancy? A systematic review and meta-analysis of cohort studies. PLoS One 2018 Jul 26;13(7):e0200594 [FREE Full text] [doi: 10.1371/journal.pone.0200594] [Medline: 30048456]
50. Eberle C, Fasig T, Brüseke F, Stichling S. Impact of maternal prenatal stress by glucocorticoids on metabolic and cardiovascular outcomes in their offspring: a systematic scoping review. PLoS One 2021 Jan 22;16(1):e0245386 [FREE Full text] [doi: 10.1371/journal.pone.0245386] [Medline: 33481865]
51. Jones SC, Telenta J. What influences Australian women to not drink alcohol during pregnancy? Aust J Prim Health 2012;18(1):68-73. [doi: 10.1071/PY10077] [Medline: 22394665]
52. Lyall V, Wolfson L, Reid N, Poole N, Moritz KM, Egert S, et al. "The problem is that we hear a bit of everything…": a qualitative systematic review of factors associated with alcohol use, reduction, and abstinence in pregnancy. Int J Environ Res Public Health 2021 Mar 26;18(7):3445 [FREE Full text] [doi: 10.3390/ijerph18073445] [Medline: 33810338]
53. National poll of young adults (April-May 2020). National Organisation for Foetal Alcohol Spectrum Disorder. URL: https://nationalfasd.org.uk/learn-more/resources/research/ [accessed 2024-06-08]
54. Updating labelling guidance for no and low-alcohol alternatives: consultation. Office for Health Improvement & Disparities, United Kingdom Government. 2023 Sep 28. URL: https://tinyurl.com/mr37j58d [accessed 2024-06-08]
55. Written evidence submitted by Mumsnet. UK Parliament. URL: https://committees.parliament.uk/writtenevidence/116697/pdf/ [accessed 2024-06-08]
56. Mårdby AC, Lupattelli A, Hensing G, Nordeng H. Consumption of alcohol during pregnancy-a multinational European study. Women Birth 2017 Aug;30(4):e207-e213. [doi: 10.1016/j.wombi.2017.01.003] [Medline: 28111037]
57. Yates S, Lockley E. Social media and social class. Am Behav Sci 2018 May 04;62(9):1291-1316. [doi: 10.1177/0002764218773821]
58. Terms of use. Mumsnet. URL: https://www.mumsnet.com/i/terms-of-use [accessed 2024-06-11]

Abbreviations

ABV: alcohol by volume
CMO: Chief Medical Officer
FASD: fetal alcohol spectrum disorder
NoLo: no- or low-alcohol
PAE: prenatal alcohol exposure

Edited by T Mackey; submitted 04.03.24; peer-reviewed by R Mccarthy, D Gilberg; comments to author 26.03.24; revised version received 15.04.24; accepted 13.05.24; published 20.06.24.

Please cite as:
Frennesson NF, Barnett J, Merouani Y, Attwood A, Zuccolo L, McQuire C
Analyzing Questions About Alcohol in Pregnancy Using Web-Based Forum Topics: Qualitative Content Analysis
JMIR Infodemiology 2024;4:e58056
URL: https://infodemiology.jmir.org/2024/1/e58056
doi:10.2196/58056
PMID:38900536

Original Paper

Exploring How Youth Use TikTok for Mental Health Information in British Columbia: Semistructured Interview Study With Youth

Roxanne Turuba[1,2], MPH; Willow Cormier[1], BA; Rae Zimmerman[1], BA; Nikki Ow[1], PhD; Marco Zenone[1,3], PhD; Yuri Quintana[4,5,6], PhD; Emily Jenkins[7], PhD; Shelly Ben-David[8], PhD; Alicia Raimundo[2], BA; Alessandro R Marcon[3], MA; Steve Mathias[1,2,9], MD; Jo Henderson[10,11,12], PhD; Skye Barbic[1,2,9], PhD

[1]Department of Occupational Science and Occupational Therapy, University of British Columbia, Vancouver, BC, Canada
[2]Foundry, Providence Health Care, Vancouver, BC, Canada
[3]Health Law Institute, University of Alberta, Edmonton, AB, Canada
[4]Division of Clinical Informatics, Beth Israel Deaconess Medical Centre, Boston, MA, United States
[5]Harvard Medical School, Boston, MA, United States
[6]Homewood Research Institute, Guelph, ON, Canada
[7]School of Nursing, University of British Columbia, Vancouver, BC, Canada
[8]School of Social Work, University of British Columbia, Vancouver, BC, Canada
[9]Centre for Advancing Health Outcomes, Vancouver, BC, Canada
[10]Centre for Addiction and Mental Health, Toronto, ON, Canada
[11]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[12]Youth Wellness Hubs Ontario, Toronto, ON, Canada

Corresponding Author:
Roxanne Turuba, MPH
Department of Occupational Science and Occupational Therapy
University of British Columbia
T325-2211 Wesbrook Mall
Vancouver, BC, V6T 2B5
Canada
Phone: 1 6725155337
Email: roxanne.turuba@ubc.ca

Abstract

Background: TikTok (ByteDance) experienced a surge in popularity during the COVID-19 pandemic as a way for people to interact with others, share experiences and thoughts related to the pandemic, and cope with ongoing mental health challenges. However, few studies have explored how youth use TikTok to learn about mental health.

Objective: This study aims to understand how youth used TikTok during the COVID-19 pandemic to learn about mental health and mental health support.

Methods: Semistructured interviews were conducted with 21 youths (aged 12-24 years) living in British Columbia, Canada, who had accessed TikTok for mental health information during the COVID-19 pandemic. Interviews were audio-recorded, transcribed verbatim, coded, and analyzed using an inductive, data-driven approach.

Results: A total of 3 overarching themes were identified describing youth's experiences. The first theme centered on how TikTok gave youth easy access to mental health information and support, which was particularly helpful during the COVID-19 pandemic to curb the effects of social isolation and the additional challenges of accessing mental health services. The second theme described how the platform provided youth with connection, as it gave youth a safe space to talk about mental health and allowed them to feel seen by others going through similar experiences. This helped normalize and destigmatize conversations about mental health and brought awareness to various mental health conditions. Finally, the last theme focused on how this information led to action, such as trying different coping strategies, discussing mental health with peers and family, accessing mental health services, and advocating for themselves during medical appointments. Across the 3 themes, youth expressed having to be mindful of bias and misinformation, highlighting the barriers to identifying and reporting misinformation and providing individualized advice on the platform.

Conclusions: Findings suggest that TikTok can be a useful tool to increase mental health awareness, reduce stigma, and encourage youth to learn and address their mental health challenges while providing a source of peer connection and support. Simultaneously, TikTok can adversely impact mental health through repetitive exposure to mentally distressing content and misleading diagnosis and treatment information. Regulations against harmful content are needed to mitigate these risks and make TikTok safer for youth. Efforts should also be made to increase media and health literacy among youth so that they can better assess the information they consume online.

(JMIR Infodemiology 2024;4:e53233) doi:10.2196/53233

KEYWORDS

youth; adolescents; young adults; mental health; TikTok; social media; qualitative research

Introduction

TikTok is a social media platform that is mainly used by youth and young adults younger than 30 years of age to create and share short videos [1]. TikTok was created as a music-based entertainment app that involved lip-syncing and dancing [2] but has since evolved as a way to share information about diverse topics, ranging from news coverage [3,4] and politics [4,5] to public health [6] and social justice issues [3,4,7]. TikTok uses a complex algorithm to recommend content based on viewers' interactions with previous posts, such as viewing time, likes, shares, and comments, resulting in content catered to each individual's interests [8]. This content floods users' "For You" page, which runs through an automated loop that users can scroll through.

TikTok experienced a surge in popularity during the COVID-19 pandemic, providing youth with a platform to interact with each other and share their lived experiences and thoughts related to the pandemic, including social isolation, public health measures, and ways to cope with mental health challenges [9-11]. TikTok has also been used to incorporate humor when sharing experiences in the face of psychological distress, which can facilitate coping and social connection [12]. For example, the "stupid walk challenge," where TikTok users would refer to "going on a stupid walk for my stupid mental health," became a popular way to share the common struggle of maintaining mental well-being during social isolation [13]. The hashtag #MentalHealth also became popular, with over 100 billion views on TikTok as of August 2023, with users sharing their mental health journeys and encouraging conversations about mental health [13]. Although this proliferation of social media content has increased mental health awareness [14], there have been growing concerns about the quality of health advice being shared on the platform [15] and the impact of persistent exposure to mentally distressing content [6]. A recent content analysis [6] found that while most TikTok videos using the hashtag #MentalHealth were filled with supportive and validating comments, almost half of these videos reported or expressed symptoms of mental distress, which could have concerning effects on viewers' mental health.

With youth spending a significant amount of time on the platform—on average, TikTok users spend 95 minutes per day on the app [16]—there is a need to better understand the types of information youth are accessing and how it impacts their mental health. In British Columbia, 26% of youths (ages 12-19 years) reported having a mental health condition in 2018, with anxiety (19%) and depression (15%) being the most common [17]. The COVID-19 pandemic greatly exacerbated youth's mental health issues [18-20], which was reflected in increased emergency department visits and hospitalizations across Canada for suicidal ideation, self-poisoning, and self-harm [21]. Many youth have reported that their mental health has worsened during the COVID-19 pandemic, mainly due to social isolation, missing out on life experiences, fear of getting sick, and challenges accessing mental health services [18,20].

Although TikTok has become a popular platform for mental health information, very little research has explored how youth use TikTok to learn about mental health and how the COVID-19 pandemic influenced this behavior. As such, this study aimed to address the following research question: How did youth use TikTok during the COVID-19 pandemic to access information and support about mental health? This is one of the first studies to explore the experiences and perceptions of youth accessing mental health information on TikTok through qualitative inquiry.

Methods

Study Design and Setting

We conducted a qualitative interview study with youth across British Columbia, Canada. This paper follows the Standards for Reporting Qualitative Research (SRQR) [22]. British Columbia is a province on Canada's west coast, with approximately 5.4 million people. A total of 88% of the population resides in a metropolitan area, while the remaining 12% live in rural and remote communities across a large mass of land, which limits the availability of mental health services in rural areas.

Ethical Considerations

This study was reviewed and approved by the University of British Columbia Behavioural Research Ethics Board (REB #H21-02948). Eligible youth were sent a copy of the study information sheet and consent form. All participants met with the research coordinator (author RT) and a research assistant to go over the consent form provided verbal informed consent over Zoom (Zoom Video Communications). The research team member filled out the consent form with each participant over Zoom and sent them a final copy for their records. To protect the participants' identity, each youth was assigned a participant ID number which was used to deidentify the collected data. The consent forms containing participants' names were stored separately on an encrypted USB key stored in a locked cabinet at the principal investigator's office at the University of British

Columbia (UBC). All participants received a CAD $30 (US $23.27) honorarium to acknowledge their contributions.

Sampling and Recruitment

We recruited youth between the ages of 12-24 years who lived in British Columbia, Canada, spoke English, and had accessed TikTok for mental health information during the COVID-19 pandemic. This age range was chosen to reflect the largest population using TikTok (13-30 years of age) [1] and the age criteria for youth services in British Columbia, which typically ends at 25 years of age. Youth were recruited through social media channels belonging to Foundry, a province-wide network of integrated youth services that provides mental health care, substance use services, physical and sexual health, peer support, and social services to youth aged 12-24 years. Foundry offers on-site and app-based services and has had over 250,000 youths visit since 2018. Foundry centers across the province and other partnering mental health organizations reshared the social media posts for the study. Interested youth emailed the research coordinator (author RT), which was followed by a brief screening call to confirm their eligibility.

Data Collection

Data were collected between June and August 2022. Semistructured interviews were held over Zoom and lasted between 30 and 60 minutes. Participants were given the option to use a different name on Zoom and keep their cameras off during the interview to further protect their identity. The interviews were audio-recorded and transcribed via Zoom and saved on a secure UBC-licensed Zoom cloud server. The research team revised and anonymized the interview transcripts and saved them to a secure UBC server and folder, while the audio recordings were deleted. A research assistant facilitated the interviews, while another research team member took general notes during the interview. The research assistants (authors WC and RZ) were undergraduate students close in age to most of the study participants, which helped to build rapport with the interviewees. They also had their own experience with using the TikTok platform and accessing mental health content on the app, which helped stimulate dialogue during the interviews. Before beginning the interview, participants were asked to complete a short demographic survey distributed through Qualtrics (Qualtrics Developer Platform). Participants were not asked to share their survey responses verbally and were only identified by their unique participant ID number. Throughout data collection, the research team debriefed frequently and discussed whether new topics arose in order to determine whether additional questions should be included in the interview guide to further our understanding of youth's overall experiences. Interview questions focused on exploring why youth use TikTok for mental health information, the type of content youth access, and the benefits and barriers to using TikTok for mental health information. Participants were also asked how the COVID-19 pandemic impacted their use of TikTok, their accessed content, and their ability to access mental health services.

Data Analysis

All interviews were audio-recorded, transcribed verbatim, and uploaded to NVivo (version 12; Lumivero) to facilitate analysis. The research coordinator led the analysis using Braun and Clarke's reflexive thematic analysis method [23,24]. This process began by reading the transcripts and interview notes multiple times while taking additional memos and reflections to support analysis. A data-driven approach was used to generate verbatim codes, which were categorized thematically using a thematic map to visualize and refine. Peer debriefing meetings were held between the research coordinator and 2 youth research assistants (authors WC and RZ) who cofacilitated the interviews to discuss the relationships between the codes and identify potential themes and subthemes. This involved reflecting on our biases stemming from our experience accessing mental health content on social media. The research assistants also supported the selection of quotes that best represented the overarching themes and subthemes.

Results

Overview

A total of 21 youths participated in this study. Participants' median age was 18 (IQR 16-21) years and they primarily identified as women (12/21, 57.1%), bisexual or pansexual (9/21, 42.9%), and White (12/21, 57.1%). Most youth had accessed mental health services before (16/21, 76.2%), including counseling (15/21, 71.4%), prescription medicine (10/21, 47.6%), psychiatry (6/21, 23.8%), peer support (4/21, 19.0%), and case management (3/21, 14.3%; see Table 1).

Table 1. Participant characteristics.

Characteristics	Values
Sociodemographics	
Age (years)	
Median (IQR)	18 (16-21)
Range	13-24
Gender identity, n (%)	
Woman	12 (57.1)
Man	5 (23.8)
Nonbinary	3 (14.3)
Transgender man	1 (4.8)
Sexual orientation[a], n (%)	
Bisexual or pansexual	9 (42.9)
Heterosexual	6 (28.6)
Gay or lesbian	1 (4.8)
Queer	1 (4.8)
Homoromantic asexual	1 (4.8)
Ethnicity[b], n (%)	
White	12 (57.1)
First Nation, Métis, or Inuit	4 (19.0)
South Asian (eg, East Indian, Pakistani, and Sri Lankan)	3 (14.3)
Chinese	2 (6.7)
Black or African	2 (9.5)
Latin American	2 (9.5)
Filipino	1 (4.8)
Middle Eastern or North African	1 (4.8)
West Asian (eg, Vietnamese, Cambodian, Laotian, and Thai)	1 (4.8)
Living location, n (%)	
Fraser Health	7 (33.3)
Vancouver Coastal Health	7 (33.3)
Vancouver Island Health	4 (19.0)
Interior Health	3 (14.3)
School or employed, n (%)	
Both	11 (52.4)
School	7 (33.3)
Employed	2 (9.5)
Neither	1 (4.8)
Highest level of education[a], n (%)	
Some high school	8 (38.1)
High school diploma	5 (23.8)
Some college or technical school education	1 (4.8)
Some university education	2 (9.5)
Bachelor’s degree	4 (19.0)

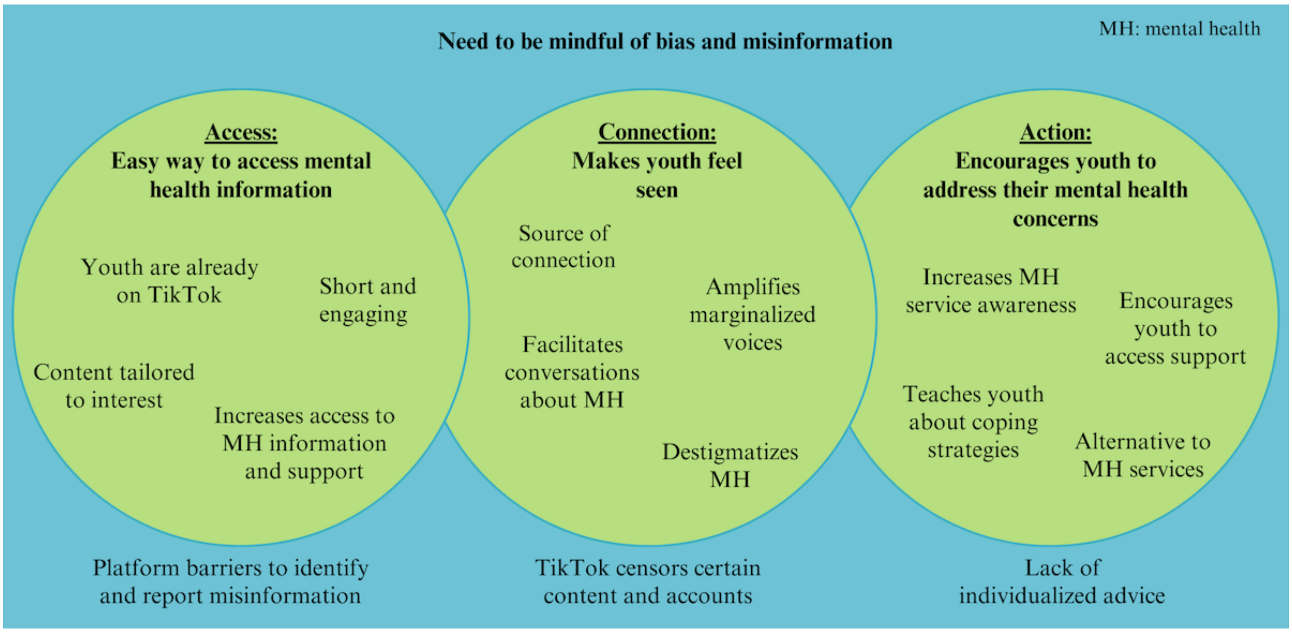
Characteristics	Values
Current living situation, n (%)	
I live with my parents or guardians	16 (76.2)
I live in an apartment or house independently or with roommates	4 (19.0)
I live with my partner	1 (4.8)
Services accessed for mental health (past 12 months)[b], n (%)	
None (I have never gotten treatment for mental health)	5 (23.8)
Counseling	15 (71.4)
Prescription medicine	10 (47.6)
Psychiatry	6 (28.6)
Peer support	4 (19.0)
Case management	3 (14.3)
Type of health service environment accessed for mental health (past 12 months)[b], n (%)	
Family doctor's office	10 (47.6)
School counseling services	10 (47.6)
Foundry center	7 (33.3)
Private office or clinic	5 (23.8)
Community health center	4 (19.0)
Emergency room or department	1 (4.8)

[a]Prefer not to answer (n=1).

[b]Participants could select more than 1 response. Therefore, the number of responses may be greater than the total number of participants who completed the survey.

A total of 3 overarching and interconnected themes were identified to describe youth's experiences using TikTok during the COVID-19 pandemic for mental health information and support (see Figure 1). These themes described how TikTok led to increased access to mental health information (theme 1), which in turn increased their connection with others going through similar experiences, making youth feel seen (theme 2). This content led to action, encouraging youth to address their mental health concerns (theme 3). Across these 3 themes, youth highlighted the need to be mindful of bias and misinformation when consuming mental health content on the platform in different ways, which are represented as subthemes under each overarching theme in Figure 1. The following sections summarize each main theme in greater detail, supported by participant quotes.

Figure 1. Overarching themes and subthemes describing youth's experiences using TikTok for mental health information. Across all 3 themes, youth described the need to be mindful of bias and misinformation when consuming mental health information on TikTok. MH: mental health.



Access: TikTok is an Easy Way to Access Mental Health Information

Most youth had joined TikTok for entertainment; however, they began seeing an increase in mental health information on their "For You" page during the COVID-19 pandemic. Most of the youth described how the pandemic exacerbated their mental health challenges, leading them to seek mental health information online. As 1 youth shared:

> *I guess I started kind of looking at [mental health information] and TikTok ended up having a lot of it during the pandemic because I think a lot of people started searching what they might have when the pandemic kind of exacerbated their issues.* [P01]

Consequently, youth described how the platform was an important source of mental health information and support during the pandemic. As 1 participant expressed, it was "the perfect way to connect with young people and spread information" (P11), while another participant noted that it made "something that's quite inaccessible and expensive (i.e., mental health services), free" (P03). Youth were drawn to the content's short video format and the various ways of presenting information (eg, sketches and using humor) on the platform, which made content engaging, digestible, and relatable. Although many of the youth appreciated informational content on TikTok, they described that they mainly used the platform to disengage from reality and distract themselves from external stressors (eg, social isolation, boredom, and mental health issues). Thus, they were more likely to engage with humorous and trendy content. As 1 youth outlined:

> *I think I would be more prone to follow somebody if they had a mix of like funny content, where it's like almost following the, you know how there's like waves of trends through TikTok. Like different audios are more popular at certain times. I'd be more inclined to follow somebody if they had a mix of that plus, like, informational stuff and stuff like that, just because when I go on TikTok I'm not necessarily looking to learn as much as I can, I'm just kind of looking to turn my brain off for a little bit.* [P19]

The use of such engaging strategies was often lacking among the content created by health professionals on TikTok, who were sometimes seen as "kind of boring, like just people saying facts and stuff" (P11). As such, youth tended to gravitate toward content creators with similar experiences and used engaging ways to connect with their audience. Youth also emphasized the benefits of TikTok's algorithm, as it would tailor mental health information based on what they engaged with and were interested in; however, this also meant that youth who did not engage with mental health content would likely not have access to it. As 1 participant described:

> *So even if there's someone who needs mental health support, if it's not an interest of theirs or it's not something that they're engaging with, it may not show up for them.* [P09]

Further, youth expressed barriers to finding information specific to their location, given the amount of information shared on TikTok and the algorithm's inability to filter based on region. As 1 youth described:

> *when TikTok came around, it was like, yeah, there are services. I would get the idea that there are mental health services, but I just didn't know what they were because these services were from everywhere around the world...I still don't know to this day, what are the mental health services provided in my area and what can someone do, near me, to help other than my school counsellor.* [P20]

Some youth described additional challenges accessing mental health services during the COVID-19 pandemic, particularly those who did not enjoy connecting to service providers virtually. As 1 youth explained:

> *I feel like having that sort of disconnect online is really challenging because it's hard to trust someone if you only know them through a screen. It made it harder to access personalized services.* [P05]

Thus, youth appreciated the availability of mental health content on TikTok, as it provided them with alternative mental health support that could be accessed immediately and free of charge. This was also seen as "less intimidating" (P08) than accessing professional services:

> *It's like these 30-second clips that—there's not a whole lot of pressure to engage with. If you want to learn more or if you want to engage, you can, but there's not necessarily a lot of pressure, like this pressure to go out of your way to follow-up or make an appointment or anything like that.* [P09]

Access to such content was especially helpful for youth who described having negative mental health care experiences, had been denied mental health support by their parents or caregivers, or simply did not feel comfortable or ready to access mental health support. As 1 participant shared:

> *It was interesting and helpful hearing people talk about those things...like experiences of autism. That's not something that I'm really comfortable seeking mental health support for so I am finding it really interesting and helpful, like informally listening to people talk about it and kind of learning more about it and how it might apply to me in that way.* [P03]

Although youth appreciated having easy access to mental health information via TikTok, youth discussed the importance of being mindful of misinformation because "it's ultimately all from strangers on the internet" (P06). Youth expressed concerns about content creators boasting professional credentials, particularly counselor influencers whose perceived goal was to gain followers for paid sponsorships. As 1 youth described:

> *listing their credentials at the beginning of every video that they make, I think it sort of implies to an audience, particularly a not very media critical audience, that this is professional advice, even if they give a little disclaimer like "I'm not trying to give advice, this is just me speaking as a PhD in psychiatry," like, it kind of offsets their disclaimer when they pump their credentials like that.* [P03]

Youth highlighted how numerous aspects of the platform exacerbated the spread of misinformation, such as the inability to embed evidence links, short caption sizes, a lack of verified information, and barriers to reporting misinformation, "which made it difficult to tell whether or not something is true" (P15).

Connection: Mental Health Content on TikTok Makes Youth Feel Seen

TikTok was described as a safe space for people to talk about their mental health experiences without fear of judgment. Many youth appreciated watching TikTok videos about content creators' mental health experiences (eg, symptoms, coping strategies, and treatments), which helped youth "feel validated" (P21), "know that [they're] not alone" (P09), and gave "hope that [they] could move forward and get past it" (P04). This was amplified by the wide range of voices (eg, youth, adults, celebrities, and mental health professionals) and mental health topics discussed on the platform. As 1 participant offered:

> *because of how many different people there are and how wide a range of different mental health issues and topics they cover, and how many different people are affected by mental health differently, it kind of helped me figure out where I can personally work on my mental health and helped me kind of recognize that in myself.* [P017]

The platform was also described as supporting youth to engage with "various facets of TikTok that are so niche about me and my identity" (P19) and find online communities to which they could relate. This was described as a unique feature found on TikTok compared to other social media platforms. As 1 participant shared:

> *I liked how it also gave platforms to people who were marginalized and kind of overlooked on other platforms. I saw a lot more of those kinds of people on TikTok than I had on like Instagram, for example, and it kind of made me and a lot of people feel seen.* [P01]

Youth discussed how the content they accessed on TikTok portrayed mental health concerns as a "common struggle" (P12), which helped normalize and destigmatize conversations related to mental health. Youth described how this was particularly helpful in increasing awareness of certain mental health conditions and experiences that were not often discussed (eg, psychiatric hospitalizations, schizophrenia, and how autism presents in women):

> *I also like the people who just try and talk about issues that aren't really talked about much in person. For example, hospitalizations. I don't see a lot of people talking about their experiences with that because it's still I think very heavily stigmatized, going to a psychiatric hospital...When people share videos of their stay or what care they received and stuff like that, not only does it destigmatize going to a hospital in general, it also makes people less scared to access that help because a lot of people are scared to go into those sorts of places.* [P01]

Youth also discussed the "other side" of TikTok, which tended to portray mental health challenges as a personal responsibility rather than something to seek support for. However, this content generally did not reach their For You page as it did not align with the type of information they resonated and interacted with, and thus, was not pushed by the platform's algorithm. As 1 youth described:

> *I think it depends on what side of TikTok you're on, right? So if you're on mental health TikTok, then it's very positive, it's very "you're not alone"...but if you're on the quote unquote "wrong side" of TikTok, then you're going to get a lot of neoliberal "pull yourself up by your bootstraps" attitude.* [P12]

Youth raised concerns about aspects of the platform that could further contribute to misinformation and stigmatization. For instance, the censoring of specific accounts and mental health content by TikTok moderators' was described as adding,

> *to the stigma around mental health...the fact that it can't be talked about and people can't access those resources.* [P05]

For example, 1 youth discussed how content about suicide was often taken down and had to find ways to circumvent the censorship:

> *when people caption their videos, for example, if they're trying to talk about suicide...they have to spell it like "seweslid" or something like that. And it's just like, every time I see that it kind of makes me cringe a little bit because it's like they're censoring something that is important to talk about.* [P19]

A few youth also discussed how TikTok's algorithm could be "biased against certain demographics and certain topics" (P19) and would often suppress creators who were less "conventionally attractive" (P01) and had disabilities, limiting their access to relevant mental health information.

Youth also expressed concerns about the romanticization of certain mental health conditions and the overgeneralization of symptoms. This could lead to youth pathologizing "normal" human experiences and making inaccurate self-diagnosis. As 1 participant articulated:

> *There's a lot of room for self-diagnosis on TikTok. Especially if people do like, put a finger down challenge for ADHD [attention-deficit/hyperactivity disorder], and if you get over five, good chance you have ADHD and I'm like, well I have six and I know I don't have ADHD. So there's a lot of room for assumptions and taking in a lot of information that you don't know if it's true or not. Because anybody can post on TikTok, anybody can make a video and post it and you don't know what's true. So that's definitely a negative and you have to be really careful about what you just even subconsciously take in and believe.* [P13]

Youth underscored how this phenomenon could negatively impact people who are living with a medical diagnosis, as it further stigmatizes and trivializes their lived experiences.

Youth also highlighted aspects of the platform that could negatively impact their mental health. For instance, being regularly exposed to negative experiences and "trauma dumping" (P20) was described as discouraging and took a toll on youth's mental health. As 1 participant expressed:

> *it's kind of an endless cycle...when I just see someone talking about their own experience it kind of takes me into a rabbit hole of just all these things people have gone through, and it's like now I feel that weight is on my shoulders.* [P20]

Although many youth described how mental health videos often led to others sharing their experiences and further discussion in the comments section, a few highlighted the damaging impact negative comments could have (ie, trolling). As 1 youth described:

> *the comments were like really attacking the creators for absolutely no reason, and I just thought it was really vicious to look at every single day.* [P05]

Youth expressed concerns about the addictive nature of the platform, which led them to spend too much time on TikTok and negatively impact their mental health. As 1 youth described:

> *it's really easy to endlessly scroll on the app and I find myself sometimes just not doing the things that are on my to-do list and going on TikTok instead...after I get off of it, I feel really terrible about myself because I wasted so much time.* [P10]

This led to some youth deleting the app altogether or finding strategies to limit the amount of time they spent on the platform, which became easier as public health restrictions were eased and youth became busier with other things (eg, school, work, socializing with friends, and going outside).

Action: Mental Health Content on TikTok Encourages Youth to Address Their Mental Health Concerns

Youth described how accessing mental health information on TikTok encouraged them to "be more mindful" (P11) about their mental health challenges and identify ways to mitigate negative symptoms they were experiencing in real time. Most youth described using TikTok as a way to learn coping strategies to reduce stress, anxiety, anger, procrastination, and negative self-talk, including managing panic attacks. Strategies they considered helpful included meditation, breathing exercises, grounding techniques, inner-child work, journaling, and reaching out to friends. Although the youth acknowledged that some strategies might not work for them, they appreciated having the ability to learn and try new things. As 1 participant shared:

> *I've come out with a couple breathing techniques I hadn't heard of before, or like my partner has ADHD and we've tried a couple things and they actually work so...the couple of hack videos that they have on there, like everyone's different, but if you even come out with one thing that's new that works for you, I think it's pretty awesome.* [P12]

Notably, this required youth to be mindful that "everybody has different experiences, everybody reacts to certain things in different ways" (P05) and that not everything they consumed would apply to their situation since mental health advice provided on the platform cannot be tailored to each individual. This required a certain level of media and health literacy, as 1 youth describes:

> *I also try not to fall victim to applying everything to myself because I understand that like, you know, a lot of it doesn't apply to me and that some of the stuff that I can't—that is said on TikTok is very generalizable and broad definitions of these actual mental illnesses or effects of mental illnesses—so I do try to be aware of that.* [P17]

As such, youth appreciated content from people with lived experience, as they were "just speaking from experience and what's worked for them" (P14), and content creators who avoided "single solution-based videos" (P12). As 1 youth described:

> *There's a lot of, like, fat overweight women on TikTok and that was really nice because a lot of the time, you know, doctors are like "Oh you're depressed and you're fat, so the answer is exercise," and so seeing like those peer support and those other creators kind of offering alternatives and things like mindful eating and all these other things like not necessarily these crash course single-based solutions was really nice as well.* [P12]

While youth shared how TikTok could lead to the misdiagnosis of mental health conditions, they also saw the benefit in helping people identify certain behaviors within themselves that they "might not have found otherwise" (P03). This particularly helped those lacking the ability to get properly diagnosed or where sufficient research on a specific mental health condition was lacking (eg, how symptoms of autism present in women), as long as they were mindful that this was not equivalent to a professional diagnosis.

Mental health content shared on TikTok made youth more aware of the types of resources and services available to them and what they should be looking for to address their individual mental health challenges. As 1 participant offered:

> *if I were to have looked for a counsellor before I went on TikTok and learned more about my own mental health, I definitely would just look for someone who is generally just a good counsellor...but after looking into that, I now know that it's very important to look for someone specifically, who specializes more in autism and someone who specializes in youth.* [P17]

This also normalized getting help, made services feel more approachable, and encouraged some youth to access mental health services:

> *...people on TikTok talking about going to therapy, that persuaded me to go to therapy and then also made me feel a bit more comfortable going to therapy too...I feel like TikTok helped normalize that a lot.* [P06]

Youth felt that TikTok helped them stay connected with peers, which helped mitigate the mental health impacts of social isolation during the COVID-19 pandemic. Youth described using the platform to share relatable content and initiate conversations about mental health with their peers, parents, and service providers that they may not have had otherwise. For instance, youth described how they would often share TikTok videos with friends, which made it "a lot easier for us to talk about [mental health]" (P20):

> *It can be a good way to send something to your friend and say like "hey, this is really funny, but at the same time like, yeah this is something that we're probably both dealing with" and it can be a good tool to facilitate conversations that way I think.* [P09]

Youth also described accessing content that helped them advocate for themselves during medical appointments and receive appropriate care. As 1 participant stated:

> *there's a lot of videos on self-advocacy as well, so I was able to kind of take some tips to my doctor and get a psychiatrist for myself which was really great.* [P12]

While most youth acknowledged that TikTok did not replace the need for professional services when dealing with "serious" (P20) mental health concerns, they considered it a helpful tool to identify mental health challenges and potential coping strategies. As 1 youth described:

> *I don't think TikTok is the best way to deal with a mental health crisis or to deal with...if you're trying to diagnose yourself. I think it can be almost harmful in those ways, but I do also think that for people who are trying to, you know, figure themselves out more in relation to mental health, that it can be very beneficial in finding coping mechanisms.* [P17]

As such, TikTok allowed youth to access support in real time by accessing mental health content on the platform, enabling conversations with friends, family, and service providers, and encouraging them to access appropriate services.

Discussion

Principal Findings

The COVID-19 pandemic had profound and deleterious mental health impacts, with youth uniquely affected. Our findings describe how the interviewed youth used TikTok as a means of mental health information and support. TikTok was perceived as an easy way to access mental health information, connect with others going through similar experiences, and learn how to address mental health challenges. Such content helped facilitate conversations about mental health with family, friends, and service providers and encouraged youth to access support. This source of mental health support and connection was significant for confronting the negative impacts of the COVID-19 pandemic, such as social isolation, boredom, exacerbated mental health issues, and challenges accessing mental health services. While TikTok was not seen as a replacement for mental health services, it was viewed as a tool to increase awareness, reduce stigma, and encourage people to address their mental health issues. Conversely, many youth expressed concerns over the lack of safeguards to deal with misinformation and the negative mental health effects that can occur when spending too much time on the platform. As such, youth discussed having to be mindful of bias and misinformation when consuming mental health content on TikTok while limiting the time spent on the platform.

Increasing Youth's Access to Information and Community

The findings from this study are particularly timely given that youth increasingly rely on online sources (eg, websites, social media, apps, and online communities) for health information, given the ease and anonymity it provides [25,26]. However, prior studies have found that youth have been unsatisfied by the lack of youth-specific information available online [25], which may partly explain why TikTok is replacing Google as a search engine among younger generations as its content is specifically designed for younger audiences [27]. Indeed, a systematic review exploring the barriers and facilitators to youth mental health-seeking behavior [26] found that societal views and attitudes toward mental health were the biggest obstacles preventing youth from accessing mental health services, contributing to their lack of knowledge about mental health and available mental health services. Youth from this study described how TikTok helped them identify and normalize their mental health challenges and encouraged them to access mental health support, demonstrating the platform's potential for destigmatizing mental health among youth. This also suggests that TikTok may be useful for mental health professionals and organizations to spread accurate mental health information specifically aimed at youth.

Several studies have found that people turn to online communities to reduce anxiety, depression, and feelings of loneliness [28-30]. This was a common coping mechanism for youth during the COVID-19 pandemic when face-to-face interactions were not possible [31,32]. For instance, a qualitative study exploring youth's perspectives using online sources for health information [25] found that many youth benefited from seeing what others were doing to support their health, which was a source of motivation and inspiration. These findings align with our study results, as youth described feeling less alone when viewing others going through similar experiences. Moreover, TikTok was described as enabling them to connect with communities and try coping strategies that they would not otherwise have accessed.

Online communities can particularly benefit those who experience greater barriers to accessing supports and services (eg, rural and remote communities and marginalized groups), experience social anxiety, or are concerned about being stigmatized [33,34]. For instance, research suggests that people who identify as LGBTQ2S+ (lesbian, gay, bisexual, transgender, queer, two-spirit, plus) experience greater mental health benefits from active social media use compared to cisgender people as it provides them an outlet for self-expression and access to social support they may not have access to offline [35,36]. Participants from our study mainly identified as LGBTQ2S+ (12/21, 57.1%) and also expressed how TikTok helped them find communities they could relate to. While these sentiments were consistent across our sample, these findings support the idea that social media platforms can be particularly beneficial for marginalized individuals. The anonymity provided online can also help youth feel more comfortable sharing personal stories and increase their access to comprehensive information, which may not always be accessible through caregivers or health care providers [25]. For example, another qualitative study exploring youth's perspectives on using technology for health information found that youth relied on online sources for comprehensive sexual education, as their school education was solely abstinence-based. This resonates with what we heard from youth in this study, who described how TikTok increased their ability to access mental health information unavailable to them when their parents were involved in their care. As such, TikTok presents an opportunity to promote online support services to increase youth's access to mental health support; however, data privacy remains a concern [37]. Albeit all social media platforms mine personal data for profit, a study exploring data sharing behaviors across social media platforms found that TikTok shared more personal data to third-party entities compared to all other social media platforms and it is unclear what that information is being used for [38]. As such, it is essential that governments develop legislation and regulations to limit TikTok's ability to collect and share personal information with third-party services, given the amount of sensitive information being collected and shared on the platform. While TikTok implemented new measures in 2023 to protect younger youth on the platform (eg, limiting screen time, direct messages, and comments from strangers) [39], having a caregiver who understands how the platform works and can ensure these measures are in place is vital.

Mitigating the Negative Effects of TikTok

Although mental health information on TikTok can be helpful, participants from this study raised concerns about the negative mental health effects of spending too much time on the platform, suggesting that youth may benefit from time limits on social media platforms to limit daily usage. Youth also highlighted the lack of safeguards to report and identify misinformation and having to rely on their own judgment to assess the accuracy of the information presented and read through the comments section to corroborate their assumptions. Misinformation is rampant on social media—this became evident during the COVID-19 pandemic, which led to a boom of misinformation about the disease, its transmission, treatment, and prevention [40]. Yeung et al [41] conducted a cross-sectional study that examined the quality of 100 TikTok videos about attention-deficit/hyperactivity disorder and found that 51% of videos (posted mainly by non–health care professionals) were misleading. Similarly, a systematic review exploring online health information-seeking behavior [42] identified a lack of mechanisms to verify misinformation, making it difficult for consumers to determine credibility. Our findings extend this prior research and suggest that these missing safeguards can lead youth to pathologize everyday experiences, contributing to unnecessary stress, anxiety, and self-diagnosis. This can further reinforce existing stereotypes and lead youth to try inappropriate or ineffective treatments without consulting a health care provider which can have detrimental consequences [43].

This study brings into question TikTok's responsibility as a platform to prevent the spread of misinformation. While incorporating mechanisms to report misinformation on TikTok can help, this responsibility cannot solely lie on users who may not be able to differentiate between credible and misleading information. Currently, social media platforms have opaque processes for reviewing posts for potential harms [44]. As such, regulations are needed on disclosing policies and methods used to identify and remove harmful content [45].

Other ways to help mitigate the spread of misinformation may include more content from health care professionals. While the number of health care professionals sharing mental health content on the platform is increasing, McCashin and Murphy [46] reported that they are less likely to use the full range of features on TikTok and tend to share more serious content. Thus, health care professionals often have lower engagement levels than peers speaking from lived experience [46], which resonates with our study findings. This highlights the need for health care professionals to create more engaging content to maximize the spread of accurate mental health information and educate viewers about media and health literacy and how to identify misinformation. While working collaboratively with youth to create relatable content on TikTok may be a helpful solution for reaching more youth on the platform, this requires significant time, effort, and investment, which may be difficult for many. More research is needed to understand how to create engaging content that gets sufficient rank and visibility with TikTok's algorithm.

Alternatively, TikTok could provide mental health professionals and organizations with opportunities to highlight their services through sponsored advertisements and location-specific hashtags. This would help address the lack of location-specific information available about mental health services, a noted barrier to acting on what youth learned on the platform. Incorporating media and health literacy education in schools should also be a priority to ensure youth have opportunities to build critical thinking skills to discern the difference between expert advice and anecdotal experiences, given the time youth spend online. Although anecdotal experiences can foster hope and connection and provide practical suggestions, it is important for youth to assess the relevance of these experiences in relation to their unique situation. Another possible approach would be to mandate public service announcements for mental health support on TikTok from legitimate and trusted health care providers and nonprofits so that all viewers have access to reliable mental health information.

While sharing personal stories on TikTok can be therapeutic, these stories can also be traumatic or triggering to those who have experienced similar situations and may not be ready to review this type of content. Little is known about the impact of consuming repetitive traumatic or triggering content online, a common trend on TikTok [6,47], exacerbated by the repetitive nature of TikTok's algorithm [48]. When users interact with negative content (eg, self-harm and eating disorders), the algorithm will continuously promote this type of content on their For You page. A study by the Centre for Countering Digital Hate [48] analyzed recommended TikTok videos within the first 30 minutes of creating standard youth accounts and found that content on eating disorders came up every 4.1 minutes, while content on self-harm and suicide was shown every 20 minutes. Youth accounts containing the phrase "lose weight" in their usernames were 3 times more likely to see this type of content on their For You page, particularly self-harm and suicide content, which was 12 times more likely to be shown. This can seriously impact youth's mental health, especially those who are already struggling. For instance, the parents of a 16-year-old from the United States who died by suicide in 2022 found his For You page flooded with distressing content, including the glorification of suicide [49]. These concerns were raised by participants in our study, highlighting the need for further research on the impacts of mentally distressing content online and to identify ways to mitigate these risks, such as developing legislation that regulates harmful content and TikTok's algorithm.

Limitations and Future Directions

While the findings from this study provide important implications for future mental health education efforts, there are limitations to consider. Since TikTok's algorithm tailors content based on users' interests and interactions, youth from our study all desired access to mental health information. This suggests that we may be missing the perspectives of youth less likely to discuss mental health and access mental health support. This also highlights the potential pitfalls of TikTok's algorithm, which prevents mental health information from reaching individuals who may benefit from it the most, and the need to incorporate disclaimers linking to evidence-based mental health information to mitigate the spread of misinformation. Further efforts to reach this population are warranted to ensure everyone has access to accurate mental health information and support. Additionally, our findings represent the experiences of youth who mainly identified as White and women. Having a more ethnic and gender-diverse sample could have provided further insights into youth's use of TikTok for mental health information as minority populations experience different access to mental health care and support compared to White, cisgender individuals [35,36]. While our sample did include many youth who identified as LGBTQ2S+, further research is warranted to identify new opportunities that social media platforms can provide in terms of mental health support for minority groups. While these limitations impact the generalizability of our findings, this is one of the first studies to explore the experiences of youth with TikTok, specifically their experiences accessing mental health information on the platform. The findings provide important direction for future research on the mental health impacts of the platform and its algorithms. They also have important considerations for policy makers shaping legislation regulating social media content to prevent harm and relay accurate mental health information.

Conclusions

Our findings shed light on how TikTok was an important source of mental health information and support for youth during the COVID-19 pandemic and how it has increased mental health awareness, encouraged youth to address their mental health issues, and reduced the stigma of help-seeking. This study highlights how TikTok can be a valuable tool for youth to access relatable mental health information, peer connection, and support while helping facilitate conversations about mental health between youth and friends, family, and service providers. Conversely, the lack of safeguards to deal with misinformation and harmful content on the platform is concerning and suggests the need for better regulations against harmful content and increased media and health literacy education among youth. Further research on the impact of TikTok's algorithm and repetitive exposure to mentally distressing content online is needed to mitigate the potential risks associated with the platform.

Acknowledgments

The TikTok Study team is grateful for the study to have taken place on the ancestral lands of many different Indigenous nations and peoples across what we now call British Columbia, Canada. We are also incredibly grateful to the participants who shared their experiences and perspectives. The TikTok Study has been made possible through the financial contributions of the Canadian Institutes of Health Research (CIHR). The views herein do not necessarily represent the views of CIHR.

Authors' Contributions

SB, SM, JH, ARM, AR, SBD, EJ, YQ, MZ, and NO contributed to the study design. RT, WC, and RZ collected the data. RT led the analysis with support from WC and RZ. RT took the lead in writing the first draft of the paper. All authors revised the paper and approved the final version.

Conflicts of Interest

None declared.

References

1. Curry D. TikTok report: revenue, users, dempgraphics and data points. Business of Apps. 2023. URL: https://www.businessofapps.com/data/tiktok-report/?utm_source=tiktok&utm_medium=click&utm_campaign=Hyperlink+report#stripe-edd-window [accessed 2024-06-10]
2. Geyser W. What is TikTok? - The fastest growing social media app uncovered. Internet Influencer Marketing Hub; 2021. 2022. URL: https://influencermarketinghub.com/what-is-tiktok/ [accessed 2023-08-24]
3. Abbas L, Fahmy S, Ayad S, Ibrahim M, Ali A. TikTok intifada: analyzing social media activism among youth. Online Media Glob Commun 2022;1(2):287-314 [FREE Full text] [doi: 10.1515/omgc-2022-0014]
4. Literat I, Boxman-Shabtai L, Kligler-Vilenchik N. Protesting the protest paradigm: TikTok as a space for media criticism. Int J Press Polit 2022 Aug 01;28(2):362-383. [doi: 10.1177/19401612221117481]
5. Medina Serrano JC, Papakyriakopoulos O, Hegelich S. Dancing to the partisan beat: a first analysis of political communication on TikTok. 2020 Presented at: 12th ACM Conference on Web Science; October 7, 2020; United States p. 07-10 URL: https://dl.acm.org/doi/10.1145/3394231.3397916 [doi: 10.1145/3394231.3397916]
6. 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 [FREE Full text] [doi: 10.2196/38340] [Medline: 35588057]
7. Le CD, Klug D. "It's viral!" A study of the behaviors, practices, and motivations of TikTok users and social activism. 2021 Presented at: Companion Publication of the 2021 Conference on Computer Supported Cooperative WorkSocial Computing; October, 23 2021; United States p. 108-111 URL: https://dl.acm.org/doi/10.1145/3462204.3481741 [doi: 10.1145/3462204.3481741]
8. Klug D, Qin Y, Evans M, Kaufman G. Trick and please. A mixed-method study on user assumptions about the TikTok algorithm. 2021 Presented at: WebSci '21: Proceedings of the 13th ACM Web Science Conference 2021; June 2021; United States p. 84-92 URL: https://dl.acm.org/doi/10.1145/3447535.3462512 [doi: 10.1145/3447535.3462512]
9. Kendall T. From binge-watching to binge-scrolling: TikTok and the rhythms of #LockdownLife. Film Q 2021;75(1):41-46 [FREE Full text] [doi: 10.1525/fq.2021.75.1.41]
10. Krug D, Evans M, Kaufman G. How TikTok served as a platform for young people to share and cope with lived COVID-19 experiences. MedieKulture 2023;73:152-170 [FREE Full text] [doi: 10.33767/osf.io/ytnmu]
11. Feldkamp J. The rise of TikTok: the evolution of a social media platform during COVID-19. In: Hovestadt C, Recker J, Richter J, Werder K, editors. Digital Responses to Covid-19: Digital Innovation, Transformation, and Entrepreneurship During Pandemic Outbreaks. Switzerland: springer nature Switzerland AG. Cham: Springer; 2021:73-85.
12. McGraw AP, Williams LE, Warren C. The rise and fall of humor. Soc Psychol Personal Sci 2013 Dec 11;5(5):566-572. [doi: 10.1177/1948550613515006]
13. TikTok. 2023. URL: https://www.tiktok.com/search/video?lang=en&q=%23stupidwalkchallenge&t=1686694480211 [accessed 2023-06-13]
14. Milton A, Ajmani L, DeVito MA, Chancellor S. "I see me here": mental health content, community, and algorithmic curation on TikTok. 2023 Presented at: CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023;480; April 19, 2023; United States p. 1-17. [doi: 10.1145/3544548.3581489]
15. 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]
16. Doyle B. TikTok statistics–Updated May 2024. WALLAROO. 2023. URL: https://wallaroomedia.com/blog/social-media/tiktok-statistics/ [accessed 2023-08-25]
17. Smith A, Peled M, Reinhart S, Thawer Z, McCreary Centre Society. Doing OK? Checking in on the mental health of BC youth. McCreary Centre Society. URL: https://mcs.bc.ca/pdf/doing_ok.pdf [accessed 2024-06-10]

18. Smith A, Peled M, McGowan E, Horton K, Reinhard S, McCreary Centre Society. Youth mental health and the COVID-19 pandemic: conversations with young people. Vancouver Coastal Health. 2022. URL: https://mcs.bc.ca/pdf/VCH_Youth%20mental%20health%20and%20COVID_final%20report.pdf [accessed 2024-06-10]
19. Samji H, Wu J, Ladak A, Vossen C, Stewart E, Dove N, et al. Review: mental health impacts of the COVID-19 pandemic on children and youth - a systematic review. Child Adolesc Ment Health 2022 May;27(2):173-189 [FREE Full text] [doi: 10.1111/camh.12501] [Medline: 34455683]
20. Youth mental health in the spotlight again, as pandemic drags on. Statistics Canada. 2022. URL: https://www.statcan.gc.ca/o1/en/plus/907-youth-mental-health-spotlight-again-pandemic-drags [accessed 2023-10-15]
21. Poonai N, Freedman SB, Newton AS, Sawyer S, Gaucher N, Ali S, Pediatric Emergency Research Canada (PERC) Network. Emergency department visits and hospital admissions for suicidal ideation, self-poisoning and self-harm among adolescents in Canada during the COVID-19 pandemic. CMAJ 2023 Sep 18;195(36):E1221-E1230 [FREE Full text] [doi: 10.1503/cmaj.220507] [Medline: 37722746]
22. 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 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
25. Radovic A, McCarty CA, Katzman K, Richardson LP. Adolescents' perspectives on using technology for health: qualitative study. JMIR Pediatr Parent 2018;1(1):e2 [FREE Full text] [doi: 10.2196/pediatrics.8677] [Medline: 30740590]
26. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. Eur Child Adolesc Psychiatry 2021 Feb;30(2):183-211 [FREE Full text] [doi: 10.1007/s00787-019-01469-4] [Medline: 31965309]
27. Huang K. For Gen Z, TikTok is the new search engine. The New York Times. 2022. URL: https://www.nytimes.com/2022/09/16/technology/gen-z-tiktok-search-engine.html [accessed 2023-10-10]
28. Highton-Williamson E, Priebe S, Giacco D. Online social networking in people with psychosis: a systematic review. Int J Soc Psychiatry 2015 Feb;61(1):92-101. [doi: 10.1177/0020764014556392] [Medline: 25381145]
29. Stuart J, O'Donnell K, O'Donnell A, Scott R, Barber B. Online social connection as a buffer of health anxiety and isolation during COVID-19. Cyberpsychol Behav Soc Netw 2021 Aug;24(8):521-525. [doi: 10.1089/cyber.2020.0645] [Medline: 33601944]
30. Morahan-Martin J, Schumacher P. Loneliness and social uses of the internet. Comput Human Behav 2003;19(6):659-671 [FREE Full text] [doi: 10.1016/s0747-5632(03)00040-2]
31. Pahayahay A, Khalili-Mahani N. What media helps, what media hurts: a mixed methods survey study of coping with COVID-19 using the media repertoire framework and the appraisal theory of stress. J Med Internet Res 2020 Aug 06;22(8):e20186 [FREE Full text] [doi: 10.2196/20186] [Medline: 32701459]
32. Klug D, Evans M, Kaufman G. How TikTok served as a platform for young people to share and cope with lived COVID-19 experiences. J Media Commun 2023;38(73):152-170 [FREE Full text] [doi: 10.7146/mk.v38i73.128463]
33. Changrani J, Lieberman M, Golant M, Rios P. Online cancer support groups: experiences with underserved immigrant Latinas. Prim Psychiatry 2008;15:55-62.
34. Wright KB, Bell SB, Wright KB, Bell SB. Health-related support groups on the internet: linking empirical findings to social support and computer-mediated communication theory. J Health Psychol 2003 Jan;8(1):39-54. [doi: 10.1177/1359105303008001429] [Medline: 22113899]
35. Coyne SM, Weinstein E, Sheppard JA, James S, Gale M, Van Alfen M, et al. Analysis of social media use, mental health, and gender identity among US youths. JAMA Netw Open 2023 Jul 03;6(7):e2324389 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.24389] [Medline: 37486631]
36. Charmaraman L, Zhang A, Wang K, Chen B. Sexual minorities and loneliness: exploring sexuality through social media and gender-sexuality alliance (GSA) supports. Int J Environ Res Public Health 2024 Mar 04;21(3):300 [FREE Full text] [doi: 10.3390/ijerph21030300] [Medline: 38541299]
37. Iwaya LH, Babar MA, Rashid A, Wijayarathna C. On the privacy of mental health apps: an empirical investigation and its implications for app development. Empir Softw Eng 2023;28(1):2 [FREE Full text] [doi: 10.1007/s10664-022-10236-0] [Medline: 36407814]
38. Brown M. TikTok says it has safety measures in place, but just how safe is it for kids? Parents. 2024. URL: https://www.parents.com/kids/safety/internet/is-tiktok-safe-for-kids/#citation-5 [accessed 2024-04-26]
39. Klais B. New research across 200 IOS apps hints that surveillance marketing is still growing strong. URL GENIUS. 2022. URL: https://app.urlgeni.us/blog/new-research-across-200-ios-apps-hints-surveillance-marketing-may-still-be-going-strong [accessed 2024-04-26]

40. O'Sullivan NJ, Nason G, Manecksha RP, O'Kelly F. The unintentional spread of misinformation on 'TikTok'; a paediatric urological perspective. J Pediatr Urol 2022 Jun;18(3):371-375. [doi: 10.1016/j.jpurol.2022.03.001] [Medline: 35331640]
41. Yeung A, Ng E, Abi-Jaoude E. TikTok and attention-feficit/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]
42. Jia X, Pang Y, Liu LS. Online health information seeking behavior: a systematic review. Healthcare (Basel) 2021 Dec 16;9(12):1740 [FREE Full text] [doi: 10.3390/healthcare9121740] [Medline: 34946466]
43. TikTok's wellness trends breed misinformation. Coda. 2021. URL: https://www.codastory.com/waronscience/bad-science-on-tiktok/ [accessed 2023-10-10]
44. Gourd C. Critics hit out at social media platforms? Disinformation reports. Internet Politico. 2023. URL: https://www.politico.eu/article/critics-social-media-platforms-disinformation-report-european-union-meta-youtube-twitter-tiktok/ [accessed 2023-10-10]
45. Perakslis E, Quintana Y. Social media is addictive and influences behavior: should it be regulated as a digital therapeutic? J Med Internet Res 2023 Jan 26;25:e43174 [FREE Full text] [doi: 10.2196/43174] [Medline: 36701180]
46. 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 [FREE Full text] [doi: 10.1177/13591045221106608] [Medline: 35689365]
47. Hung YH, Miles A, Trevino Z, DAniello C, Wood H, Bishop A, et al. BIPOC experiences of racial trauma on TikTok: a qualitative content analysis. Contemp Fam Ther 2023 May 09:1-11 [FREE Full text] [doi: 10.1007/s10591-023-09669-6] [Medline: 37361259]
48. Deadly by design: TikTok pushes harmful content promoting eating disorders and self-harm into users' feeds. Centre for Countering Digital Hate. 2022. URL: https://counterhate.com/wp-content/uploads/2022/12/CCDH-Deadly-by-Design_120922.pdf [accessed 2023-10-15]
49. Carville O. TikTok?s algorithm keeps pushing suicide to vulnerable kids. Bloomberg. 2023. URL: https://www.bloomberg.com/news/features/2023-04-20/tiktok-effects-on-mental-health-in-focus-after-teen-suicide [accessed 2023-04-19]

Abbreviations

LGBTQ2S+: lesbian, gay, bisexual, transgender, queer, two-spirit, plus
SRQR: Standards for Reporting Qualitative Research
UBC: University of British Columbia

Edited by E Lee; submitted 26.10.23; peer-reviewed by M Shoukri, R Zhuang; comments to author 01.04.24; revised version received 27.04.24; accepted 26.05.24; published 05.07.24.

Please cite as:
Turuba R, Cormier W, Zimmerman R, Ow N, Zenone M, Quintana Y, Jenkins E, Ben-David S, Raimundo A, Marcon AR, Mathias S, Henderson J, Barbic S
Exploring How Youth Use TikTok for Mental Health Information in British Columbia: Semistructured Interview Study With Youth
JMIR Infodemiology 2024;4:e53233
URL: https://infodemiology.jmir.org/2024/1/e53233
doi:10.2196/53233
PMID:38967966

Original Paper

Public Perception of the Tobacco 21 Amendment on Twitter in the United States: Observational Study

Liane M Schneller-Najm[1*], MS, PhD; Zidian Xie[2*], PhD; Jiarui Chen[3*]; Sarah Lee[3]; Emily Xu[3]; Dongmei Li[2], PhD

[1]Department of Health Behavior, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States

[2]Department of Clinical and Translational Research, University of Rochester Medical Center, Rochester, NY, United States

[3]Goergen Institute for Data Science, University of Rochester, Rochester, NY, United States

[*]these authors contributed equally

Corresponding Author:
Liane M Schneller-Najm, MS, PhD
Department of Health Behavior
Roswell Park Comprehensive Cancer Center
Elm and Carlton Streets
Buffalo, NY, 14203
United States
Phone: 1 716 845 5881
Email: Liane.Najm@RoswellPark.org

Abstract

Background: Following the signing of the Tobacco 21 Amendment (T21) in December 2019 to raise the minimum legal age for the sale of tobacco products from 18 to 21 years in the United States, there is a need to monitor public responses and potential unintended consequences. Social media platforms, such as Twitter (subsequently rebranded as X), can provide rich data on public perceptions.

Objective: This study contributes to the literature using Twitter data to assess the knowledge and beliefs of T21.

Methods: Twitter data were collected from November 2019 to February 2021 using the Twitter streaming application programming interface with keywords related to vaping or e-cigarettes, such as "vape," "ecig," etc. The temporal trend of the T21 discussion on Twitter was examined using the mean number of daily T21-related tweets. Inductive methods were used to manually code the tweets into different sentiment groups (positive, neutral, and negative) based on the attitude expressed toward the policy by 3 coders with high interrater reliability. Topics discussed were examined within each sentiment group through theme analyses.

Results: Among the collected 3197 tweets, 2169 tweets were related to T21, of which 444 tweets (20.5%) showed a positive attitude, 736 (33.9%) showed a negative attitude, and 989 (45.6%) showed a neutral attitude. The temporal trend showed a clear peak in the number of tweets around January 2020, following the enactment of this legislation. For positive tweets, the most frequent topics were "avoidance of further regulation" (120/444, 27%), "Enforce T21" (110/444, 24.8%), and "health benefits" (81/444, 18.2%). For negative tweets, the most frequent topics were "general disagreement or frustration" (207/736, 28.1%) and "will still use tobacco" (188/736, 25.5%). Neutral tweets were primarily "public service announcements (PSA) or news posts" (782/989, 79.1%).

Conclusions: Overall, we find that one-third of tweets displayed a negative attitude toward T21 during the study period. Many were frustrated with T21 and reported that underage consumers could still obtain products. Social media data provide a timely opportunity to monitor public perceptions and responses to regulatory actions. Continued monitoring can inform enforcement efforts and potential unintended consequences of T21.

(JMIR Infodemiol 2024;4:e53899) doi:10.2196/53899

KEYWORDS

tobacco policy; tobacco regulation; social media; tobacco use; tobacco; health belief; sentiment analysis; smoking; cigarettes; social media analysis; vaping; e-cigarettes; health behavior; public opinion

Introduction

Nearly all current daily adult cigarette users reported first trying a cigarette before the age of 18 years. While cigarette use has decreased among adults, as well as youth, other products, such as electronic nicotine delivery systems, have gained popularity. In 2023, 12.6% of high schoolers reported current use of tobacco (in the past 30 days), including cigarettes, cigars, smokeless tobacco, hookah, heated tobacco products, and electronic nicotine delivery systems [1]. Furthermore, 3.9% of high schoolers reported current use of 2 or more products in the past 30 days [1] and are therefore more likely to continue nicotine use into adulthood [2-4]. In order to prevent tobacco initiation, the Tobacco 21 Amendment (T21) of the Federal Food, Drug, and Cosmetic Act was signed in December 2019 [5]. T21 raised the federal minimum age to purchase tobacco products from 18 to 21 years [5].

T21 was vital to interrupting the US youth and early young adult tobacco epidemic as tobacco product use is the leading cause of preventable diseases and death [2], and most current tobacco users started by the age of 18 [2,3]. Youth may obtain tobacco products from social sources, including peers and classmates, as well as commercial sources [6,7]. Before the Federal T21 legislation, 19 states and many more localities passed legislation to increase the minimum age of tobacco sales as early as January 2016 [8]. Simulations and many other evaluations of local legislation found that raising the minimum age to purchase tobacco reduced product sales [9-11], product use prevalence [12-20], initiation of nicotine use [13,21], and tobacco-related health disparities among youth [20]. Therefore, one would expect to see similar changes at a national level. However, underage tobacco sales and use have been an issue for decades and require retailer compliance, with more youth and early young adults reporting purchasing their products directly from the retailer than any other source [6,7]. In April 2018, the Truth Initiative surveyed 12- to 17-year-old, past 30-day JUUL (a vape system developed by Juul Labs, Inc) users nationally and found that 74% (n/N) of them purchased their product from a retailer, 52% (n/N) from a social source such as a friend or family member, and 6% (n/N) web-based retailers [6]. Therefore, it is important to prevent youth smokers from acting as suppliers and to make it more difficult for 16- and 17-year-olds to pass as legal purchasers. Policy enforcement and retailer compliance are crucial to reducing tobacco use among youth [11,22]. Furthermore, social determinants have been shown to impact retailer compliance, including previous state and local tobacco control policies, neighborhood demographics, retail signage, and scanners for ID checks [23-25]. A study assessing the impact of the Federal T21 after 1 year of signing found that middle- and high-school students perceived it to be more difficult to buy tobacco products from a store, but this was not the case for purchasing products from web-based retailers [25].

Following T21, there is an urgent need to monitor the knowledge and beliefs of T21. Before the Federal T21 policy, previous studies have assessed attitudes toward raising the minimum age to purchase tobacco products. In general, the studies found that the majority of adults were in favor of increasing the age to purchase tobacco to 21 years, regardless of smoking status and demographics [26-28]. Social media platforms, such as Twitter (subsequently rebranded as X), can provide rich data on public perceptions. In the past, Twitter has been used to examine discussions focused on government policies, such as e-cigarette flavor policies and T21 [29-34]. Many tweets in these previous studies that used Twitter data to assess sentiment toward T21 identified a neutral tone toward T21; however, an alternative to previous survey findings, more tweets (over one-third) portrayed opposition to the policy rather than support [32-34]. Common themes identified in these studies included unfairness to youth who were already addicted to nicotine and skepticism toward the policy efficacy [32], a disjunction for other age restrictions (eg, military, alcohol, and voting) [33], and incorrectly describing the policy as a purchase law [34]. However, these Twitter studies used data from the months leading up to the signing of the amendment. This study uses Twitter data a month before and over a year following the signing of the amendment to assess the public attitude toward T21 and its potential impact on tobacco use behavior, such as policy avoidance or seeking out cessation advice.

Methods

Data Collection

Publicly available Twitter data from November 12, 2019, to February 26, 2021, was previously collected using the Twitter streaming application programming interface with keywords related to vaping or e-cigarettes for a previous study [35]. The methodology and vaping or e-cigarette keywords can be found published elsewhere [35]. Next, we filtered out a subset of the data using keywords related to the T21 policy, such as "tobacco21," "tobacco 21," "t21," "tobacco age," "vaping age," "tobacco purchase law," "buy tobacco," "tobacco age restriction," "tobacco 21 laws", "minimum age to buy tobacco," "vape age," "smoking age," "legal age," and "tobacco age." The tweets without any of the keywords were eliminated. Then, we applied 2 filters to remove tweets related to the commercial promotion of smoking and vaping products [35]. The first filter was applied to the Twitter username, and the keyword list contained "dealer," "store," "promo," etc. If a tweet is associated with a username containing any of these keywords, it is removed from the data set. The second filter targeted the tweet content. The keywords contained, but were not limited to, "discount," "sale," "percent off," and "store." Finally, repetitive tweets and retweets were removed, and we obtained a data set with 6489 tweets. In addition, any additional commercial tweets were manually excluded from the data set, resulting in 49.3% (3197/6489) of tweets being hand-coded.

Content Analysis

Content analyses were conducted from April 2022 to May 2022 using inductive methods. From the processed data set, we randomly selected 10% (320/3197) of the tweets to manually review the tweets, identify key attitudes and themes, and develop the codebook for the content analysis of the entire data set. First, we determined whether each tweet was related to the T21 policy, and then we grouped the related tweets into 3 categories (positive, negative, and neutral) based on the attitude expressed toward the policy. Next, we identified the topics of supportive and antagonistic reasoning and labeled each positive and negative tweet with the identified topic. Neutral tweets were labeled in a slightly different way due to the lack of apparent reasoning and were divided into "public service announcements (PSA) or news," "dialogue or discussion," and "extend to other regulations." A list of all topics and their definitions are shown in Table 1.

When developing the codebook, 2 coders reviewed the tweet independently and respectively summarized the topics. Then the commons and differences between the 2 versions were collectively discussed among all 6 authors to reach a consensus on the final codebook. Next, the 2 coders reviewed sample tweets again to adjust their original results according to the revised codebook. The 2 coders achieved an overall kappa agreement of 0.80 on attitudes and an overall rate of 0.74 on the specific topic labels. For the disagreement, a third coder was introduced to resolve the discrepancy during the development of the codebook. After the codebook was developed, the remaining tweets were divided into 3 parts and were labeled by 3 coders separately. Each tweet was identified into 1 attitude group and then one of the topics within the attitude group. We chose not to collectively label all the remaining tweets due to the large number of tweets. Finally, the results for sample tweets and the remaining tweets were combined for analysis.

Table 1. Codebook for hand-coding Tobacco 21–related tweets.

Attitude and topic	Description	Examples
Positive		
Avoidance of further regulation	The tweet claims that the governors should let T21[a] do its job before implementing additional regulations.	“as an adult, I will not have my rights stripped away, including flavors. Tobacco 21 will be the best solution. we have respiratory therapists who recommend vaping over smoking, no use of thc!”
Enforce T21	The tweet calls for thorough execution and enforcement of the policy.	“it's already illegal for children to buy tobacco and vaping products. how about this: how about you release who funds your little group? and yes, protects the childrens. protect them by enforcing the laws already on the books. stop the hysteria!”
Health benefit	The tweet states that T21 reduces the harm to the health of the youth.	“congress should act to ban characterizing flavors in all tobacco products, give fda user fees to regulate e-cigs, and raise tobacco age to 21. these steps can combat youth tobacco use, preserve harm reduction, and further reduce tobacco death and disease.”
General support	The tweet offers no specific reason for supporting T21 but generally expresses a positive attitude.	“tobacco 21 legislation is critical! learn more about youth tobacco use in the granite state here.”
Reduce usage	The tweet states that T21 will be effective in reducing the usage of tobacco in the youth.	“for me flavours play a critical role in disassociating vaping from smoking. lots of others feel the same, it's a matter of making the product better than smoking, and thus more enjoyable than smokes. t21 keeps the product out of the hands of kids.”
Other	Other tweets with positive attitudes that do not belong to any of the other categories.	“wow, believe your listening to us vapers! children should not vape or smoke, educating parents is very important! t21 is the law and should be followed! if i had this when i was 16, i would of never started smoking! #vaping is for adults who want to quit smoking!”
Negative		
Parenting or blaming	The tweet argues that the essence of the problem is not tobacco, but other factors such as advertisement, parenting, etc.	“i agree that the shops should be held accountable and that t-21 should have been voted on not put forcefully in place that’s not how our government works!!! there has been a decline in youth vaping so now there will be a rise in youth smoking because it’s easy access”
Inconsistency in adult age definition or limits personal choice	The tweet argues that the policy limits adults’ rights and personal choices.	“yup. not to mention, adults can make their own damn decisions. here in michigan, you now have to be 21 to buy tobacco or vape. i no longer smoke, but if i’m old enough to die for my country, i’m old enough to decide if i buy this shit.”
Will still use tobacco	The tweet believes that T21 is ineffective in restricting the usage of tobacco among youth, and the underage group will still find a way to purchase tobacco products.	“so, vaping is illegal for <21 in mass and <19 in nh. sounds a lot like alcohol, or cigarettes. kids have been working around these sorts of barriers for generations. vaping regulations are no different than alcohol regulations afa kids are concerned.”
Not the priority	The tweet argues that T21 should not be the priority of the agenda, and government should focus on policies in other areas.	““the govt acted swiftly when vaping deaths became more frequent. but still stay silent on mass shootings… congress passes bill raising minimum tobacco and vape smoking age to 21.”
Mocking	The tweet mocks or contains irony and satire about the policy.	“i voted to send the majority leader to washington to repeal and replace obamacare. and all i got was a minimum vaping age. let’s be sure to name a lot buildings and highway after this guy.”
General disagreement or frustration	The tweet offers no specific reasons for opposing T21 but expresses a negative attitude or frustration in general.	“i knew that also.. and i know i dont like everthing ..like the vape or smoking age to 21 .. but will still vote for him again”
Other	Other tweets with negative attitudes that do not belong to any of the other categories.	“i almost bought a pack yesterday – i think because it's always in the ether. I'm only a year in so i’'m one of the vulnerable ones – being told by these policies that my health isn't as valuable as these young people's.”
Neutral		
PSA[b] or news	The content of the tweet is a news headline, which typically states the policy change.	“#us expected to raise #vaping age to 21”

Attitude and topic	Description	Examples
Dialogue or discussion	The tweet is a part of the conversation or discussion among users with no obvious attitudes.	"sorry to disappoint you! it is illegal for youth to obtain vaping products, t18 or t21 depending on the state,Äôs law. it was not even targeted to non smokers. it was and always has been meant for smokers!"
Extend to other regulations	The tweet discusses and compares the age restrictions across smoking, voting, drinking, etc.	"repealing the 26th amendment to raise the voting age back up to 21 and make the draft only apply to those 21 and older. justify by pointing out the drinking age=21, smoking/vape age going up to 21, buying guns/ammo going up to 21 means voting should go up to 21 also."

[a]T21: Tobacco 21 Amendment.

[b]PSA: public service announcement.

Data Analysis

To understand the prevalence of T21-related discussion, we examined the temporal trend of daily counts of T21-related posts over the studied period. In addition, we summarized the distribution of the attitude and the specific topics within each attitude group by calculating the frequencies and proportions of tweets in each topic within each attitude group.

Ethical Considerations

This study only analyzed publicly available data, and the results do not contain any identifiable information and are presented in aggregate.

Results

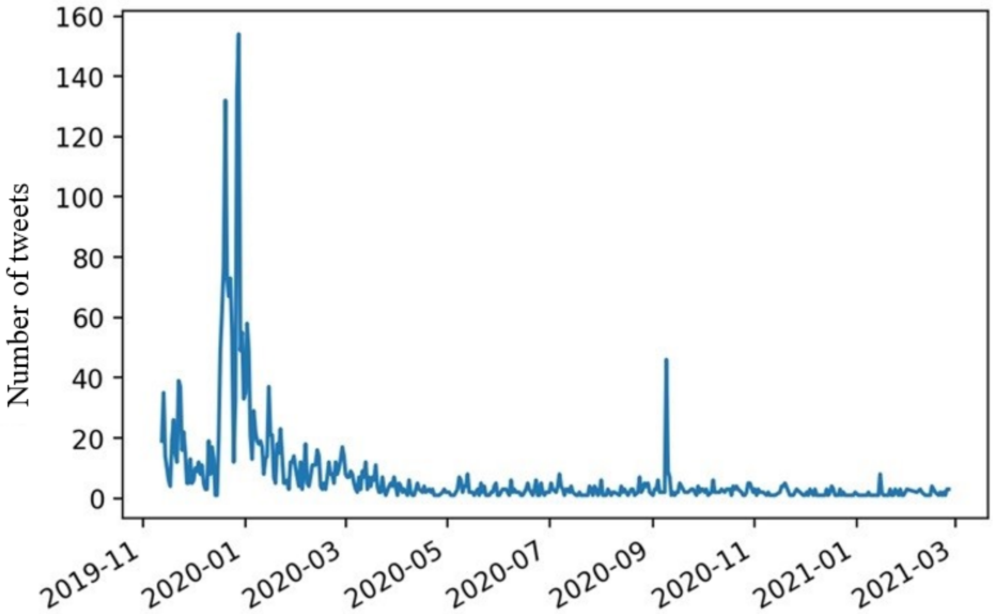
Codebook Development for Thematic Analysis of T21 Tweets

From our data set of 3197 tweets, we determined that 2169 (67.8%) tweets were relevant with regard to T21-related discussion. Of these 2169 tweets, we hand-coded each to their respective attitude and topic based on our defined codebook, which were mutually exclusive groups (Table 1). Among tweets that portrayed a positive attitude, themes identified included (1) avoidance of further regulation, where the tweet claims that T21 should do its job before implementing additional regulations such as an e-cigarette flavor restriction; (2) Enforce T21, where the tweet calls for thorough execution and enforcement of the policy; (3) health benefits, where the tweet states that T21 can reduce the harm to the health of youth; (4) general support tweets offered no specific reasons for supporting T21 but expresses a positive attitude in general; (5) reduced usage, where the tweet states that T21 will be effective to reduce the usage of tobacco among youth; and (6) other tweets that portrayed a positive attitude but did not fit into any of the other categories. Among tweets that showed a negative attitude, themes included (1) parenting or blaming, where the tweet argues that the problem is not tobacco but other factors; (2) inconsistency in adult age definition or limits personal choice, where the tweet argues that the T21 amendment is limiting adult rights and personal choice; (3) will still use tobacco, where the Twitter user believes that the policy will be ineffective in restricting youth use of tobacco and underage youth will still find a way to obtain tobacco products; (4) mocking, where the tweet mocks or contains irony and satires about the policy; (5) general disagreement or frustration, where the tweet offers no specific reasons for opposing T21 but expresses a negative attitude or frustration in general; and (6) other tweets that showed a negative attitude but did not fit into any of the other categories. Finally, tweets that were more neutral in nature contained themes such as (1) PSA or news, where the tweet contained a news headline that stated the policy change; (2) dialogue or discussion, where the tweet was part of a conversation or discussion among users with no obvious attitudes; and (3) extend to other regulations where the tweet discussed and compares age restrictions across different activities (eg, smoking, drinking, gambling, and voting).

Temporal Trends of Attitudes Toward T21

It was determined that out of 2169 tweets, 444 (20.47%) showed a positive attitude, 736 (33.93%) showed a negative attitude, and 989 (45.60%) showed a neutral attitude. Figure 1 shows the temporal trend of T21-related mentions on Twitter. There is a clear peak in the number of tweets around January 2020, 2 months following the enactment of this legislation. A secondary peak occurs during September 2020. This secondary peak appears to be associated with a discussion of Florida Governor Ron DeSantis vetoing Florida's T21 policy and e-cigarette flavor restriction.

Figure 1. Temporal trend of mentioning Tobacco 21 on Twitter.



Themes Associated With the Mentioning of T21

The temporal trend of themes associated with the mentioning of T21 was not clearly distinctive from one another and, therefore, is not depicted. Table 2 displays each topic and its associated attitude, amount, and proportion of tweets. Proportions were calculated by attitude. For positive tweets, the most frequent topic was “avoidance of further regulation” (120/444 tweets, 27.0%), followed closely by “Enforce T21” (110/444 tweets, 24.8%), then “health benefit” (81/444 tweets, 18.2%), “other” (78/444, 17.6%), “general support” (34/444 tweets, 7.7%), and “reduce usage” (21/444 tweets, 4.7%). For negative tweets, the most frequent topic was “general disagreement or frustration” (207/736 tweets, 28.1%), followed closely by “will still use tobacco” (188/736 tweets, 25.5%), then “other” (142/736 tweets, 19.3%), “inconsistency in adult age definition or limits personal choice” (70/736 tweets, 15.8%), “mocking” (66/736 tweets, 9.0%), “parenting or blaming” (40/736 tweets, 5.4%), and “not the priority” (23/736 tweets, 3.1%). For neutral tweets, the most frequent topic was “PSA or news” (782/989 tweets, 79.1%), then “dialogue or discussion” (198/989 tweets, 20.0%), and “extended to other regulations” (9/989 tweets, 0.9%).

Table 2. Main topics in Tobacco 21–related tweets.

Attitude and topic	N=2169, n (%)
Positive	444 (20.5)
Avoidance of further regulation	120 (27)
Enforce T21[a]	110 (24.8)
Health benefit	81 (18.2)
General support	34 (7.7)
Reduce usage	21 (4.7)
Other	78 (17.6)
Negative	736 (33.9)
General disagreement or frustration	207 (28.1)
Will still use tobacco	188 (25.5)
Inconsistency in adult age definition or limits personal choice	70 (15.8)
Mocking	66 (9)
Parenting or blaming	40 (5.4)
Not the priority	23 (3.1)
Other	142 (19.3)
Neutral	989 (45.6)
PSA[b] or news	782 (79.1)
Dialogue or discussion	198 (20)
Extend to other regulations	9 (0.9)

[a]T21: Tobacco 21 Amendment.

[b]PSA: public service announcement.

Discussion

Principal Findings

This analysis used an existing data set of Twitter posts related to vaping and e-cigarettes to assess discussions associated with the signing of T21 in December 2019. Most discussions associated with T21 occurred in the month following the signing of the amendment, with some discussion in the preceding month. A spike in discussions was observed in late 2020. Many tweets at this time discussed Florida Governor Ron DeSantis vetoing Florida's T21 policy and e-cigarette flavor restriction on September 8, 2020, stating that it was unnecessary due to the Federal T21 policy [36].

The majority of tweets had a neutral attitude toward the T21 amendment. However, more showed a negative attitude than positive. Most negative posts discussed general disagreement and frustration, followed by a discussion that those younger than 21 years will still use tobacco. Studies have assessed the penalty structure for T21 violations and have suggested that monetary fines, no matter the severity, are not as effective as license suspension, revocation, or criminal penalties [37,38]. In addition, local policies with procedures to conduct inspections and impose penalties can ensure the effectiveness and enforcement of T21 [38-40]. Another common topic of discussion was the age definition of an adult; people feel that if one can vote and enlist in the military, then one should be able to purchase tobacco and alcohol. The age of 21 was chosen since most current smokers report trying their first cigarette before turning 21 years old [3]. Furthermore, young adults, 18-25 years of age, are highly influenced by their peers and environment as their brains are still developing [3,41]. However, the Institute of Medicine report on the public health implications of increasing the minimum age to purchase tobacco products found that increasing the age to 25 years as opposed to 21 years would result in considerably smaller effects [42]. Topics of discussion among tweets that showed a positive attitude toward T21 overall had a general theme that enforcement of T21 would work to reduce the use of tobacco and benefit public health. Previous research has also shown the positive impact and great support for increasing the minimum legal age to purchase tobacco to 21 years among various populations and users of tobacco [9,10,12-18,21,40,43].

Limitations

Social media data can provide rich data on the public's perception of a policy, such as T21. In addition, the tweets for developing the codebook were randomly sampled and, therefore, can fully represent the whole Twitter data set in our study. However, there are some limitations to this study. First, the demographics of the Twitter users are not available. Furthermore, the geolocation of the users was not available, nor was there a clear distinction if the discussion was referring to the federal policy or a state or local policy, which could be useful to evaluate policies implemented at local and state levels before the federal amendment. Second, less than a quarter of the US population has a Twitter account. In addition, some users have private accounts, so their posts are unavailable. Third, this data set was not collected for the purpose of this analysis, and it focused on vaping and e-cigarettes. Therefore, our findings are only generalizable to users who show an interest in discussing vaping and e-cigarettes on Twitter. Fourth, we only assessed the content of the tweets and no other form of data or interaction between Twitter users (eg, follow, retweet, and favorites). Fifth, there may have been T21 content that was missed due to keyword filtering, and accounts were not assessed to see if they were bots. In addition, about 90% (1952/2169) of tweets in this study were single-coded, which could lead to potential bias even though the intercoder reliability for 10% (217/2169) of tweets was high. Finally, there were a lot of events (eg, COVID-19 and cartridge-based e-cigarette flavor restriction) that may have masked the discussion associated with T21. Therefore, our analysis is not representative of the general population, and an assessment of posts that were not restricted to vaping and e-cigarettes could provide additional information on the public's perception of the T21 amendment.

Conclusion

We observed an overall negative public attitude toward the T21 policy on Twitter, with major discussions around the frustration about the T21 policy due to continued underage use of tobacco products. While the attitudes and themes found in this assessment of tweets are consistent with previous studies assessing sentiment toward T21 using Twitter data, this study provided a more comprehensive understanding of reasons either supporting or against T21 policy [32-34]. Greater enforcement and penalties within communities would likely minimize the continued underage use of tobacco products. This analysis of Twitter posts provided a comprehensive look at the public's perception of the US T21 amendment. Continued monitoring can inform enforcement efforts and potential unintended consequences of T21. Considering more tweets with a negative attitude toward T21 than those with a positive attitude based on our results, it is important to enhance health communication about the underage use of tobacco products, for example, launching health communication campaigns on social media, which can reach more underage population. Furthermore, machine learning models could allow for the assessment of a greater number of tweets or posts found on other platforms that may be lengthier.

Acknowledgments

This research was supported by the National Cancer Institute of the National Institutes of Health and the US Food and Drug Administration Center for Tobacco Products (award U54CA228110). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the US Food and Drug Administration.

Data Availability

The data and scripts used for the analysis and creation of figures are available upon request from the corresponding author LMSN.

Authors' Contributions

LMSN, ZX, and DL conceived and designed the study. JC, SL, and EX analyzed the data. LMSN wrote the paper. LMSN, ZX, and DL assisted with the interpretation of analyses and edited the paper.

Conflicts of Interest

None declared.

References

1. Birdsey J, Cornelius M, Jamal A, Park-Lee E, Cooper MR, Wang J, et al. Tobacco product use among U.S. middle and high school students - National Youth Tobacco Survey, 2023. MMWR Morb Mortal Wkly Rep 2023;72(44):1173-1182 [FREE Full text] [doi: 10.15585/mmwr.mm7244a1] [Medline: 37917558]
2. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. The health consequences of smoking—50 years of progress: a Report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; 2014.
3. Preventing tobacco use among youth and young adults: a Report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services; 2012.
4. Gomez Y, Creamer M, Trivers KF, Anic G, Morse AL, Reissig C, et al. Patterns of tobacco use and nicotine dependence among youth, United States, 2017-2018. Prev Med 2020;141:106284. [doi: 10.1016/j.ypmed.2020.106284] [Medline: 33068604]
5. Tobacco 21. US Food and Drug Administration. URL: https://www.fda.gov/tobacco-products/retail-sales-tobacco-products/tobacco-21 [accessed 2024-07-16]
6. Where are kids getting JUUL? Truth Initiative. 2018. URL: https://truthinitiative.org/research-resources/emerging-tobacco-products/where-are-kids-getting-juul [accessed 2021-04-16]

7. Where do youth smokers get their cigarettes. Campaign for Tobacco-Free Kids. 2020. URL: https://www.tobaccofreekids.org/assets/factsheets/0073.pdf [accessed 2021-04-16]
8. Tobacco 21. Preventing Tobacco Addiction Foundation. URL: https://tobacco21.org/ [accessed 2021-04-08]
9. Ali FRM, Rice K, Fang X, Xu X. Tobacco 21 policies in California and Hawaii and sales of cigarette packs: a difference-in-differences analysis. Tob Control 2020;29(5):588-592 [FREE Full text] [doi: 10.1136/tobaccocontrol-2019-055031] [Medline: 31645377]
10. Glover-Kudon R, Gammon DG, Rogers T, Coats EM, Loomis B, Johnson L, et al. Cigarette and cigar sales in Hawaii before and after implementation of a Tobacco 21 law. Tob Control 2021;30(1):98-102 [FREE Full text] [doi: 10.1136/tobaccocontrol-2019-055248] [Medline: 31932332]
11. Zhang X, Vuong TD, Andersen-Rodgers E, Roeseler A. Evaluation of California's 'tobacco 21' law. Tob Control 2018;27(6):656-662 [FREE Full text] [doi: 10.1136/tobaccocontrol-2017-054088] [Medline: 29440328]
12. Kessel Schneider S, Buka SL, Dash K, Winickoff JP, O'Donnell L. Community reductions in youth smoking after raising the minimum tobacco sales age to 21. Tob Control 2016;25(3):355-359 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-052207] [Medline: 26071428]
13. Bonnie RJ, Stratton K, Kwan LY, editors. Public Health Implications of Raising the Minimum Age of Legal Access to Tobacco Products. Washington, DC: Institute of Medicine; 2015.
14. Ahmad S. Closing the youth access gap: the projected health benefits and cost savings of a national policy to raise the legal smoking age to 21 in the United States. Health Policy 2005;75(1):74-84 [FREE Full text] [doi: 10.1016/j.healthpol.2005.02.004] [Medline: 16298230]
15. Friedman AS, Buckell J, Sindelar JL. Tobacco-21 laws and young adult smoking: quasi-experimental evidence. Addiction 2019;114(10):1816-1823 [FREE Full text] [doi: 10.1111/add.14653] [Medline: 31342591]
16. Friedman A, Wu RJ. Do local tobacco-21 laws reduce smoking among 18 to 20 year-olds? Nicotine Tob Res 2020;22(7):1195-1201. [doi: 10.1093/ntr/ntz123] [Medline: 31348515]
17. Boettiger DC, White JS. Cigarette pack prices and sales following policy changes in California, 2011-2018. Am J Public Health 2020;110(7):1002-1005 [FREE Full text] [doi: 10.2105/AJPH.2020.305647] [Medline: 32437272]
18. Dove MS, Stewart SL, Tong EK. Smoking behavior in 18-20 year-olds after tobacco 21 policy implementation in California: a difference-in-differences analysis with other states. Prev Med 2021;148:106553 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106553] [Medline: 33862032]
19. Hansen B, Sabia JJ, McNichols D, Bryan C. Do tobacco 21 laws work? J Health Econ 2023;92:102818 [FREE Full text] [doi: 10.1016/j.jhealeco.2023.102818] [Medline: 37950948]
20. Colston DC, Xie Y, Patrick ME, Thrasher JF, Titus AR, Elliott MR, et al. Tobacco 21 laws may reduce smoking and tobacco-related health disparities among youth in the U.S. Prev Med Rep 2022;27:101762 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101762] [Medline: 35340271]
21. Ahmad S, Billimek J. Limiting youth access to tobacco: comparing the long-term health impacts of increasing cigarette excise taxes and raising the legal smoking age to 21 in the United States. Health Policy 2007;80(3):378-391 [FREE Full text] [doi: 10.1016/j.healthpol.2006.04.001] [Medline: 16698112]
22. Macinko J, Silver D. Impact of New York City's 2014 increased minimum legal purchase age on youth tobacco use. Am J Public Health 2018;108(5):669-675. [doi: 10.2105/AJPH.2018.304340] [Medline: 29565664]
23. Roberts M, Klein EG, Ferketich AK, Keller-Hamilton B, Berman ML, Chacko M, et al. Beyond strong enforcement: understanding the factors related to retailer compliance with tobacco 21. Nicotine Tob Res 2021;23(12):2084-2090 [FREE Full text] [doi: 10.1093/ntr/ntab093] [Medline: 33982115]
24. Dai H, Tamrakar N, Rathnayake N, Samson K. Geographical distribution and social determinants of tobacco 21 policy adoption and retail inspections in the United States, 2015-2019. Tob Induc Dis 2021;19:55 [FREE Full text] [doi: 10.18332/tid/140148] [Medline: 34602933]
25. Agaku IT, Nkosi L, Agaku QD, Gwar J, Tsafa T. A rapid evaluation of the US federal tobacco 21 (T21) law and lessons from statewide T21 policies: findings from population-level surveys. Prev Chronic Dis 2022;19:E29 [FREE Full text] [doi: 10.5888/pcd19.210430] [Medline: 35653765]
26. King BA, Jama AO, Marynak KL, Promoff GR. Attitudes toward raising the minimum age of sale for tobacco among U.S. adults. Am J Prev Med 2015;49(4):583-588 [FREE Full text] [doi: 10.1016/j.amepre.2015.05.012] [Medline: 26163165]
27. Gentzke AS, Glover-Kudon R, Tynan M, Jamal A. Adults' attitudes toward raising the minimum age of sale for tobacco products to 21 years, United States, 2014-2017. Prev Med 2020;133:106012 [FREE Full text] [doi: 10.1016/j.ypmed.2020.106012] [Medline: 32027916]
28. Winickoff JP, McMillen R, Tanski S, Wilson K, Gottlieb M, Crane R. Public support for raising the age of sale for tobacco to 21 in the United States. Tob Control 2016;25(3):284-288. [doi: 10.1136/tobaccocontrol-2014-052126] [Medline: 25701856]
29. Gao Y, Xie Z, Li D. Investigating the impact of the New York State flavor ban on e-cigarette-related discussions on twitter: observational study. JMIR Public Health Surveill 2022;8(7):e34114 [FREE Full text] [doi: 10.2196/34114] [Medline: 35802417]

30. 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;8(3):e25697 [FREE Full text] [doi: 10.2196/25697] [Medline: 35348461]
31. Sun L, Lu X, Xie Z, Li D. Public reactions to the New York State policy on flavored electronic cigarettes on twitter: observational study. JMIR Public Health Surveill 2022;8(2):e25216 [FREE Full text] [doi: 10.2196/25216] [Medline: 35113035]
32. Dobbs PD, Boykin AA, Ezike N, Myers AJ, Colditz JB, Primack BA. Twitter sentiment about the US federal tobacco 21 law: mixed methods analysis. JMIR Form Res 2023;7:e50346 [FREE Full text] [doi: 10.2196/50346] [Medline: 37651169]
33. Dobbs PD, Colditz JB, Shields S, Meadows A, Primack BA. Policy and behavior: comparisons between twitter discussions about the US tobacco 21 law and other age-related behaviors. Int J Environ Res Public Health 2022;19(5):2613 [FREE Full text] [doi: 10.3390/ijerph19052613] [Medline: 35270306]
34. Dobbs PD, Schisler E, Colditz JB, Primack BA. Miscommunication about the US federal tobacco 21 law: a content analysis of twitter discussions. Tob Control 2023;32(6):696-700. [doi: 10.1136/tobaccocontrol-2021-057099] [Medline: 35173067]
35. Lu X, Chen L, Yuan J, Luo J, Luo J, Xie Z, et al. User perceptions of different electronic cigarette flavors on social media: observational study. J Med Internet Res 2020;22(6):e17280 [FREE Full text] [doi: 10.2196/17280] [Medline: 32579123]
36. Florida governor vetoes bill because of liquid nicotine ban. 2020. URL: https://apnews.com/general-news-953d409f3f246ba8e42a20f00f999228 [accessed 2024-02-26]
37. D Dobbs P, Chadwick G, W Ungar K, M Dunlap C, White KA, Kelly MC, et al. Development of a tobacco 21 policy assessment tool and state-level analysis in the USA, 2015-2019. Tob Control 2020;29(5):487-495 [FREE Full text] [doi: 10.1136/tobaccocontrol-2019-055102] [Medline: 31611425]
38. DiFranza JR. Best practices for enforcing state laws prohibiting the sale of tobacco to minors. J Public Health Manag Pract 2005;11(6):559-565. [doi: 10.1097/00124784-200511000-00014] [Medline: 16224293]
39. Dobbs PD, Chadwick G, Dunlap CM, White KA, Cheney MK. Tobacco 21 policies in the U.S.: the importance of local control with federal policy. Am J Prev Med 2021;60(5):639-647. [doi: 10.1016/j.amepre.2020.12.009] [Medline: 33602533]
40. Juarez R, Haidar S, Brookins-Fisher J, Hancher-Rauch H, Ohneck M, Thompson A, et al. Practices and perceptions of local health officers/commissioners regarding tobacco 21 policy advocacy. Health Promot Pract 2022;23(3):463-472. [doi: 10.1177/1524839921989271] [Medline: 33533278]
41. Winters KC, Arria A. Adolescent Brain Development and Drugs. Prev Res 2011;18(2):21-24 [FREE Full text] [Medline: 22822298]
42. Committee on the Public Health Implications of Raising the Minimum Age for Purchasing Tobacco Products. Public Health Implications of Raising the Minimum Age of Legal Access to Tobacco Products. Washington, DC: The National Academies Press; 2015.
43. Ickes MJ, Butler K, Wiggins AT, Rayens MK, Hahn EJ. Support for tobacco 21 in a tobacco-growing state. West J Nurs Res 2019;41(8):1203-1215. [doi: 10.1177/0193945918822523] [Medline: 30608019]

Abbreviations

PSA: public service announcement
T21: Tobacco 21 Amendment

Edited by T Mackey; submitted 23.10.23; peer-reviewed by H Dashtian, M Navarro; comments to author 07.02.24; revised version received 26.02.24; accepted 20.03.24; published 25.09.24.

Please cite as:
Schneller-Najm LM, Xie Z, Chen J, Lee S, Xu E, Li D
Public Perception of the Tobacco 21 Amendment on Twitter in the United States: Observational Study
JMIR Infodemiology 2024;4:e53899
URL: https://infodemiology.jmir.org/2024/1/e53899
doi:10.2196/53899
PMID:

Original Paper

Detection and Characterization of Online Substance Use Discussions Among Gamers: Qualitative Retrospective Analysis of Reddit r/StopGaming Data

Nicolette Le[1], MA; Tiana McMann[1,2,3], MA; Luning Yang[2], BS; Zhuoran Li[2,3], MS; Raphael E Cuomo[4], MPH, PhD; Tim K Mackey[1,2,3], MAS, PhD

[1]Global Health Program, Department of Anthropology, University of California San Diego, La Jolla, CA, United States

[2]Global Health Policy and Data Institute, San Diego, CA, United States

[3]S-3 Research, San Diego, CA, United States

[4]Department of Anesthesiology, School of Medicine, University of California San Diego, San Diego, CA, United States

Corresponding Author:
Tim K 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 9514914161
Email: tmackey@ucsd.edu

Abstract

Background: Video games have rapidly become mainstream in recent decades, with over half of the US population involved in some form of digital gaming. However, concerns regarding the potential harms of excessive, disordered gaming have also risen. Internet gaming disorder (IGD) has been proposed as a tentative psychiatric disorder that requires further study by the American Psychological Association (APA) and is recognized as a behavioral addiction by the World Health Organization. Substance use among gamers has also become a concern, with caffeinated or energy drinks and prescription stimulants commonly used for performance enhancement.

Objective: This study aimed to identify substance use patterns and health-related concerns among gamers among a population of Reddit users.

Methods: We used the public streaming Reddit application programming interface to collect and analyze all posts from the popular subreddit, r/StopGaming. From this corpus of posts, we filtered the dataset for keywords associated with common substances that may be used to enhance gaming performance. We then applied an inductive coding approach to characterize substance use behaviors, gaming genres, and physical and mental health concerns. Potential disordered gaming behavior was also identified using the tentative IGD guidelines proposed by the APA. A chi-square test of independence was used to assess the association between gaming disorder and substance use characteristics, and multivariable logistic regression was used to analyze whether mental health discussion or the mention of any substance with sufficient sample size was significantly associated with IGD.

Results: In total, 10,551 posts were collected from Reddit from June 2017 to December 2022. After filtering the dataset for substance-related keywords, 1057 were included for further analysis, of which 286 mentioned both gaming and the use of ≥1 substances. Among the 286 posts that discussed both gaming and substance use, the most mentioned substances were alcohol (n=132), cannabis (n=104), and nicotine (n=48), while the most mentioned genres were role-playing games (n=120), shooters (n=90), and multiplayer online battle arenas (n=43). Self-reported behavior that aligned with the tentative guidelines for IGD was identified in 66.8% (191/286) posts. More than half, 62.9% (180/286) of the posts, discussed a health issue, with the majority (n=144) cited mental health concerns. Common mental health concerns discussed were depression and anxiety. There was a significant association between IGD and substance use (P<.001; chi-square test), and there were significantly increased odds of

IGD among those who self-reported substance use (odds ratio 2.29, *P*<.001) and those who discussed mental health (odds ratio 1.64, *P*<.03).

Conclusions: As gaming increasingly becomes highly prevalent among various age groups and demographics, a better understanding of the interplay and convergence among disordered gaming, substance use, and negative health impacts can inform the development of interventions to mitigate risks and promote healthier gaming habits.

(JMIR Infodemiology 2024;4:e58201) doi:10.2196/58201

KEYWORDS

internet gaming disorder; gaming disorder; substance use; alcohol use; nicotine use; stimulants; gaming; internet gaming; video games; addiction; addiction medicine; digital mental health; reddit

Introduction

In recent decades, video games have become mainstream, with US gamers spending an average of 13 hours playing each week and over half of the US population involved in some form of digital or internet gaming [1]. Furthermore, more than 90% of children, teenagers, and young adults in the United States play video games, and these demographics spend between 4.5 and 7 hours per day [2,3]. Though the effects of video games on mental health have been examined over the years, with periodic media sensationalization bringing the discussion to the forefront of popular discussion, public debate remains ongoing over the effect of video games on individuals and broader society.

As video games have grown in popularity, concerns about the potential harmful effects of excessive and disordered gaming have also arisen and become an important topic of research. This has led to the concept of online gaming as an addiction in the 1990s and, more recently, the recognition of gaming disorder as a behavioral addiction by the World Health Organization (WHO) and the proposal of internet gaming disorder (IGD) in the latest edition of the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]) as a tentative psychiatric disorder that requires further study by the American Psychological Association (APA) [2,4]. The proposed disorder has been defined as a condition of “persistent and repeated internet use for gaming, often with other players, that can lead to significant distress or impairment” and includes 9 criteria, of which 5 must be self-reported by an individual to meet the criteria for IGD [5,6]. Criteria are available in Multimedia Appendix 1. Notably, symptoms can be exhibited in either online or offline settings [3]. Since its proposal, prevalence rates of IGD have been estimated to be between 1% and 9% based on various demographics. However, among American youth, a recent study found that 8.5% of gamers met the disorder criteria [3,7].

Concerningly, mental health concerns and IGD can co-occur together and can share similar symptomology, including problems in an individual’s personal, social, and occupational life [8]. A 2018 review of the literature identified significant correlations between IGD and a number of psychological and other health problems, including bipolar disorder, anxiety, depression, and social phobia [9]. Similar findings were reported in a 2019 survey-based study, which also identified associations between problematic video game use and lower self-esteem, life satisfaction, perceived social support, self-efficacy, and academic performance [10]. Anxiety has also been identified as a risk factor for IGD [11].

IGD also shares many core features with other behavioral and addiction disorders, including substance use disorder [12,13]. Both are thought to share underlying factors, including impulsivity and escaping reality [13,14]. Hence, their intersection has become a growing area of research. Specifically, certain substances have significant use among gamers, with caffeine, energy drinks, and prescription stimulants commonly used for performance enhancement [15,16]. Recent studies have found 40% of adult gamers use performance enhancing drugs [17]. Furthermore, substance use and excessive gaming can share a bidirectional relationship with excessive gamers at an increased risk of alcohol, tobacco, and cannabis use and an earlier onset and increased consumption of certain substances correlated with increased gaming [13].

However, there remains limited research examining social media discussions of substance use and disordered gaming co-occurrence, despite that gamers commonly use online discussion forums (such as Reddit [Advance Publications], Discord, and dedicated online gaming community boards and groups) to talk to others both during online gameplay, and in offline settings. Online discussion forums represent access to important communities, which provide rich contextual information regarding gaming, as well as increased transparency among those who post due to the ability to remain anonymous [18]. Hence, this study aimed to identify substance use patterns and health-related concerns among gamers who have engaged in online discussions on Reddit regarding compulsive gaming or video game addiction, adding to the growing body of literature on video games, disordered gaming, and health. Furthermore, we aimed to assess the potential relationship of individuals who self-reported IGD characteristics, substance use co-occurrence, and health concerns among a specific Reddit gaming community and potential influences by game genre.

Methods

Participants

The subreddit r/StopGaming is a community on Reddit that is described as a space to “to help those who struggle with or have struggled with compulsive gaming or video game addiction,” with an estimated 52,000 members and ranked in the top 3% of communities on Reddit by size, at the time of data collection.

Procedure

For this study, the public streaming Reddit application programming interface was used to collect and analyze all publicly available posts from r/StopGaming. Reddit was chosen based on previous studies that have used the platform to characterize substance use and gaming behaviors separately [19,20]. First, to ensure this subreddit is appropriate for examination, the dataset was filtered for keywords related to disordered gaming, including "internet gaming disorder" and "gaming addiction." Then, to identify, characterize, and elucidate substance use behavior among habitual gamers, the dataset was filtered for keywords associated with substances commonly used to enhance gaming performance and experience (eg, stimulants, depressants, and other substances) [15]. The full list of study keywords is available in Textbox 1. An inductive coding approach was then applied to all initial posts that contained selected substance use-related keywords of interest to characterize self-reported substance use, game genres, and physical and mental health concerns. Potential disordered gaming behavior was also identified through self-reports using the tentative guidelines for IGD, as proposed by the APA. Coding criteria are available in Multimedia Appendix 1. To facilitate analysis, individual games within a series were grouped into a single property. Genre coding was determined based on the commercial market's classification of the game [21,22]. However, each game was assigned a single genre, chosen to most comprehensively describe the gameplay. Genre categorizations included the following: role-playing games, shooters, multiplayer online battle arenas, action-adventure, simulation and sports, strategy, survival horror, platformer, puzzler and party, sandbox, and other (ie, encompassing all other gaming genre categories not fitting the genres described). First through fourth authors coded the data for study inclusion and inductively coded themes and achieved a high intercoder reliability (Cohen κ=0.95). Discrepancies in coding were discussed among authors, and a consensus on the correct classification was reached.

Data Analysis

Mean imputation was used to provide a complete dataset for continuous variables, and mode imputation was used for categorical variables. A chi-square test of independence was used to assess the association between IGD and substance use characteristics, and multivariable logistic regression was used to analyze whether mental health discussion or the mention of any substance with sufficient sample size (ie, nicotine, caffeine, alcohol, or cannabis) was significantly associated with IGD after accounting for the association with substance use.

Textbox 1. Keywords related to disordered gaming and substance use were used for dataset filtering of posts collected from r/StopGaming.

Disordered gaming-related keywords:

- IGD
- internet gaming disorder
- gaming addiction
- game addiction
- video game addiction

Stimulant, depressants, and other substance-related keywords:

- Nicotine
- Vape
- Vaping
- Cigarette
- Cig
- Adderall
- Addy
- Vyvanse
- Ritalin
- Alcohol
- Booze
- Drinking
- Energy drinks
- Soda
- Caffeine
- Weed
- Cannabis
- Smoke
- Smoking
- Drugs
- Stimulants
- Marijuana
- Cocaine
- Meth

Ethical Considerations

This study involves the analysis of publicly available data from Reddit, specifically posts from the r/StopGaming subreddit. The data used in this research are entirely anonymous and deidentified. No personally identifiable information was collected or analyzed.

Results

A total of 10,551 posts dated from June 2017 to December 2022 were collected from r/StopGaming. After filtering the dataset for disordered gaming-related keywords, 736 posts were identified, averaging 367 words per post, and authored by 649 unique Reddit usernames. After filtering the dataset for substance-related keywords, 1057 posts, averaging 443 words per post and authored by 937 unique usernames, were included for further analysis. Out of these, 27.1% (286/1057) posts contained self-reports of both gaming and the use of 1 or more substances. The overwhelming majority (277/286, 96.9%) were written in first person about the user themselves, while the rest (9/286, 3.1%) were written in third person about someone in the user's personal life. The age of the individual engaged in gaming and substance use was self-reported in 30.4% (87/286) posts, with a median age of 27 (range 14-45). The summary of detected health topics, substances, and gaming genres and their frequency is available in Table 1.

Table 1. Qualitative characteristics of discussions found within r/StopGaming posts.

Characteristics	Count
r/StopGaming output, n	
Posts collected from r/StopGaming	10,551
Posts with relevant keywords	1057
Discussion of gaming and substance use (signal posts)	286
Discussion topics among signal posts (n=286) , n (%)	
Discussion of specific game genres	132 (46.2)
Discussion of health	180 (62.9)
Discussion of mental health	144 (50.3)
Potential IGD[a]	191 (66.8)
Discussion of game genre (n=350) , n (%)	
Role-playing game	120 (34.3)
Shooter	90 (25.7)
Multiplayer online battle arena	43 (12.3)
Action adventure	27 (7.7)
Simulation and sports	26 (7.4)
Strategy	16 (4.6)
Survival horror	10 (2.9)
Platformer	6 (1.7)
Other	6 (1.7)
Puzzler and party	4 (1.1)
Sandbox	2 (0.6)
Total mentions of genres	350 (100)
Discussion of specific substance in posts (n=286 posts; n=372 total mentions) , n (%)	
Alcohol	132 (46.2)
Cannabis	104 (36.4)
Nicotine	48 (16.8)
Other	40 (14)
Caffeine	31 (10.8)
Prescription stimulants	10 (3.5)
Illicit stimulants	3 (1)
Hallucinogens	3 (1)
Total mentions of substance	372 (100)

[a]IGD: internet gaming disorder.

Self-reported behavior that aligned with the tentative guidelines for IGD was identified in 66.8% (191/286) posts. More than half, 62.9% (180/286), of posts, discussed a health issue, with the majority of those posts (n=144) citing mental health concerns. Common mental health concerns discussed were depression and anxiety. Among posts that discussed both gaming and substance use, some mentioned more than 1 substance, with a total of 372 mentions of a substance in 286 posts. The majority (159/286, 55.6%) of substance use discussions indicated current use at the time of posting, while the remainder (127/286, 44.4%) indicated past use.

Discussion within the posts included a variety of topics, such as reflections on the impact of gaming disorder on authors' lives, progress updates, and expressions of encouragement. Some authors made comparisons between their relationship with gaming and their relationship with substance use. The tone of the posts varied widely, ranging from expressions of deep struggle and continued efforts to overcome the disorder to feelings of embarrassment or shame for how disordered gaming has negatively affected areas of their lives (eg, school and academic performance, personal relationships, job performance), and also sincere appeals for support and guidance from others in the community. The selected examples are available in Table 2, which have been truncated for brevity and anonymity.

The most mentioned substances in all posts collected were alcohol (132/372, 35.5%), cannabis (104/372, 28.0%), and nicotine (48/372, 12.9%). These numbers likely underrepresent the use of cannabis and nicotine due to ambiguous discussions related to "smoking" without specifying a substance. These instances were instead coded as "Other."

There were 133 posts that identified at least 1 specific game name, title, or game genre. Within these, there were 350 mentions of a specific name or game genre. The genres most mentioned were role-playing games (120/350, 34.3%), shooters (90/350, 25.7%), and multiplayer online battle arenas (43/350, 12.3%). Overall, 120 distinct gaming properties were identified across all posts. The top 3 properties, by number of times mentioned, accounted for over two-thirds of all mentions. These were League of Legends (37/120, 30.8%; Riot Games), World of Warcraft (25/120, 20.8%; Blizzard Entertainment), and Call of Duty (19/120, 15.8%; Activision).

The chi-square test for independence between IGD and substance use disorder revealed a statistically significant association (χ^2_1=29.83, *P*<.001). The expected percentage for individuals without substance use and without IGD was approximately 31.2% (330/1057), compared with the observed percentage of 36.6% (387/1057). For individuals with substance use and without IGD, the expected percentage was about 11.6% (123/1057), while the observed percentage was 6.3% (67/1057). For those without substance use and with IGD, the expected percentage was around 49.3% (521/1057), with the observed percentage being 44% (465/1057). For individuals with both substance use and IGD, the expected percentage was approximately 18.3% (193/1057), whereas the observed percentage was 23.6% (249/1057). These observed percentages significantly deviated from the expected percentages, underscoring a meaningful relationship between IGD and substance use.

In the multivariable logistic regression model where IGD was the dependent variable and substance use and discussions of mental health impacts were covariates, the model demonstrated successful convergence with a log-likelihood of –624.03 after 5 iterations. For the substance use variable, the odds ratio (OR) was 2.29 (95% CI 1.69-3.11; *z*=5.36; *P*<.001), with the odds of having IGD as 69% for those with substance use and 30% for those without substance use. For the impact on physical and mental health covariate, the OR was 1.64 (95% CI 1.05-2.56; *z*=2.17; *P*=.03), with the odds of having IGD as 62% for those with physical and mental health impacts, compared with 38% for those without such impacts. Nonsignificant associations included nicotine use (28% for individuals using nicotine compared to 29% for those not using nicotine; OR 0.96, 95% CI 0.85-1.08; *z*=–0.594; *P*=.55), caffeine use (36% for individuals not using caffeine compared with 37% for those not using caffeine; OR 0.99, 95% CI 0.89-1.11; *z*=–0.276; *P*=.78), alcohol use (40% for individuals using alcohol compared to 39% for those not using alcohol; OR 1.01, 95% CI 0.90-1.12; *z*=0.239; *P*=.81), and cannabis use (31% for individuals using cannabis compared with 32% for those not using cannabis; OR 0.98, 95% CI 0.88-1.10; *z*=–0.377; *P*=.71).

Table 2. Selected example posts from r/StopGaming (truncated for brevity and anonymity).

Themes	Examples
Author reflects on money spent on games and missed life opportunities.	"This year I've spent $2000 on games I don't play. I don't enjoy video games anymore. I hate them but I play because of habit and boredom. I'd rather have that money now. I want to sell my PlayStation and never turn back. I lost so many opportunities because I would smoke weed and play video games all day like a loser. It's time I take accountability for my actions and turn away from what makes me miserable. Instead of praying to God for a change I'm going to make a change myself. Will I have the sudden urge to play games? Yes, but I can do something else like reading or drawing. I'm sorry video games, but I'm done with you. I'm 25 now, I have shit I have to do. I ignored a lot of people who are no longer in my life because I put YOU first."
Author compares gaming addiction to drug addiction.	"[…] Gaming is barely seen as addictive. I only recognized I was addicted after feeling the extreme emotional high that reminded me of doing a lot of cocaine. I recognized "… I have a problem". Addiction is rooted in denial, and the very disease of addiction makes you think that you have no problem. If people aren't aware of the fact that gaming can be addictive at all, then there is no problem, right? Look at the sub count in this bar. There are 32,000 subs here. Marijuana is barely seen as addictive in mainstream culture, and yet /r/leaves has over 4 times the number of subs. […]"
Author discusses reasons for quitting and offers insights from alcohol recovery.	"[…] I'm sure I have a similar story to most of you, but I'm likely older than most (age redacted). Like many addicts, I'm chronically depressed and have a pretty negative internal dialogue. Quitting games will give me the opportunity to be more "in the moment" and spend time on activities that make me happy. IDK how often I'll check in here because, for me, Reddit and gaming are intertwined to a high degree. But I'm here to learn and offer any advice I can from my own recovery from alcohol."
Author reflects on how time spent playing games impacting academic progress.	"Online gaming with friends can be fun, but I am depressed by the amount of my life that I wasted in a virtual world. I have over 400 hours in games. I was looking at the numbers today and became quite depressed. I have barely played any games for the past 3 months because I have been so burnt out with school. I am a year and a half behind in college. I withdrew from multiple semesters because I was so addicted to videogames. I would skip class and not do homework so that I could play games all day. Then I'd take Adderall so I wouldn't sleep and play even more. […]"

Discussion

As video games become widely mainstream across diverse demographics of users (including console gaming, PC gaming, mobile gaming, and through immersive technologies such as virtual reality gaming), the concepts of disordered gaming and video game addiction have gained considerable attention in the public eye, leading to several countries having established protocol for regulating gaming behavior and diagnosing and treating IGD [3,23]. Previous research has found reason for concern with similarities drawn between disordered gaming and behavioral and addiction disorders, including substance use disorder, and has highlighted concerns for the implications disordered gaming has on social behavior and mental health [12,14].

This study focused on a subset of gamers who already self-identify as experiencing problematic gaming behaviors evident from their posts in a dedicated gaming addiction community, r/StopGaming, on the social networking website Reddit. In these posts, individuals report various disordered gaming behaviors and the impacts disordered gaming has had on their lives, the vast majority from a first-hand standpoint. Among the 1057 posts we analyzed, more than a quarter (286/1057, 27.1%) reported the use of at least 1 substance and active gaming. Furthermore, findings reveal that a significant portion, over 3-quarters, of those who self-report the use of at least 1 substance also meet the proposed criteria for IGD based on self-reported behavior. A smaller percentage of those who self-reported IGD-related behavior also self-reported the use of at least 1 substance. These findings suggest that substance use could contribute to the development of IGD, which supports previous research that increased consumption or earlier onset of use of certain substances may correlate to increased gaming [13].

Furthermore, role-playing, shooter, and multiplayer online battle arena games emerged as the most frequently mentioned genres. These findings are similar to those from a previous questionnaire-based study of 613 participants, which found the risk of gaming disorder to be approximately 2 times higher in individuals who preferred first-person shooter games and massively online role-playing games [2]. These observations may indicate that certain genres and games may be more prevalent among gamers who engage in problematic gaming behavior or may have features that promote addictive behaviors inherent to their design or gameplay. However, just as with the identified gaming properties, it important to note that these genre preferences could also reflect broader trends in gaming choices and may simply be representative of the distribution of popular games and genres among the general gaming population.

The study found significant associations between IGD and adverse mental health impacts, underscoring the complex and potentially harmful interplay of gaming, substance use, and mental health. These findings align with a previous review, which also found significant correlations between IGD and mental health conditions, depression, and anxiety [9]. While this study did not seek to elucidate the authors' gender from their posts, the same review found higher video game usage and higher levels of IGD among males [9]. Previous research focused on adolescent males has found associations between playing video games and poorer social and mental health [24]. These findings highlight the importance of recognizing disordered gaming as a significant psychological concern and, potentially, an emerging public health and men's mental health issue, though recent news reports also suggest that the number of female gamers is also increasing as a whole and requires further attention in the context of unique risk factors potentially associated with IGD [25].

Yet, it is crucial to acknowledge that this study's findings are specific to a subset of gamers, those who see themselves as having some degree of gaming addiction due to their participation in the online r/StopGaming community. As such, it may not capture the diversity of gaming behaviors and attitudes among all gamers and lacks generalizability to the general population of all those who play video games (eg, there are also other online forums where users discuss gaming disorders, addiction, and relapse, such as the forum "Game Quitters" and its associated Discord channel). In addition, though Reddit offers users a significant degree of anonymity through features like customizable usernames, participation in topic-specific subreddits without revealing personal information, 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. Furthermore, factors such as age, gender, and specific gaming preferences may have influenced both gaming behaviors and substance use patterns, acting as potential confounding variables. Finally, the categorization of video games by one 1 single genre that best captures overall gameplay is imperfect, as video games may be multifaceted, offer a variety of gameplay options, and have features that encompass more than 1 genre.

Future studies should focus on using validated measures to gain deeper insights into the motivation behind problematic gaming behavior and track longitudinal behavioral changes over time to explore the cross-cultural influences on gaming behaviors and substance use attitudes. A better understanding of the influence of particular games and genres, as well as online versus offline play, on disordered gaming behavior will also require a more robust and standardized framework for defining different elements of gameplay. However, the associations identified do suggest a need for the development of more accessible interventions tailored to individuals engaging in problematic gaming behaviors. As gaming continues to grow as a US $347 billion global market with widespread popularity and increasingly diverse options to play and compete (eg, esports), a better understanding of the interplay and convergence between disordered gaming, substance use, and negative health impacts can inform the development of targeted interventions aimed at mitigating risks and promoting healthier gaming habits needed for a population of multigenerational gamers worldwide [26].

Acknowledgments

This manuscript has been seen by all authors, who have approved of its content.

Data Availability

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

Conflicts of Interest

TM, ZL, and TKM are also employees of the startup company S-3 Research LLC. TKM also holds equity in the company S-3 Research LLC. S-3 Research is a startup funded with previous and current funding from the National Institutes of Health - National Institute on Drug Abuse through a Small Business Innovation and Research program for social media research and technology commercialization. TKM is also the Editor-in-Chief of *JMIR Infodemiology*.

Multimedia Appendix 1
Coding Criteria for Disordered Gaming on r/StopGaming Based on *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]*) Proposed Criteria for Internet Gaming Disorder.
[DOCX File , 15 KB - infodemiology_v4i1e58201_app1.docx]

References

1. Totilo S. Number of gamers in U.S. dips slightly to 216 million. URL: https://www.axios.com/2022/06/07/video-game-players-america-222 [accessed 2023-08-02]
2. Ünal E, Gökler ME, Turan Ş. An evaluation of the factors related to internet gaming disorder in young adults. Addict Health 2022;14(4):279-287 [FREE Full text] [doi: 10.34172/ahj.2022.1381] [Medline: 37559796]
3. Gentile DA, Bailey K, Bavelier D, Brockmyer JF, Cash H, Coyne SM, et al. Internet gaming disorder in children and adolescents. Pediatrics 2017;140(Suppl 2):S81-S85 [FREE Full text] [doi: 10.1542/peds.2016-1758H] [Medline: 29093038]
4. Darvesh N, Radhakrishnan A, Lachance CC, Nincic V, Sharpe JP, Ghassemi M, et al. Exploring the prevalence of gaming disorder and Internet gaming disorder: a rapid scoping review. Syst Rev 2020;9(1):68 [FREE Full text] [doi: 10.1186/s13643-020-01329-2] [Medline: 32241295]
5. Luo T, Wei D, Guo J, Hu M, Chao X, Sun Y, et al. Diagnostic contribution of the DSM-5 criteria for internet gaming disorder. Front Psychiatry 2022;12:777397 [FREE Full text] [doi: 10.3389/fpsyt.2021.777397] [Medline: 35069285]
6. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5th ed. Washington, DC: American Psychiatric Association; 2013.
7. Gentile D. Pathological video-game use among youth ages 8 to 18: a national study. Psychol Sci 2009;20(5):594-602. [doi: 10.1111/j.1467-9280.2009.02340.x] [Medline: 19476590]
8. Cleveland Clinic. Video game addiction: What it is, symptoms & treatment. URL: https://my.clevelandclinic.org/health/diseases/23124-video-game-addiction [accessed 2024-07-03]
9. González-Bueso V, Santamaría JJ, Fernández D, Merino L, Montero E, Ribas J. Association between internet gaming disorder or pathological video-game use and comorbid psychopathology: a comprehensive review. Int J Environ Res Public Health 2018;15(4):668 [FREE Full text] [doi: 10.3390/ijerph15040668] [Medline: 29614059]
10. von der Heiden JM, Braun B, Müller KW, Egloff B. The association between video gaming and psychological functioning. Front Psychol 2019;10:1731 [FREE Full text] [doi: 10.3389/fpsyg.2019.01731] [Medline: 31402891]
11. Rho MJ, Lee H, Lee TH, Cho H, Jung DJ, Kim DJ, et al. Risk factors for internet gaming disorder: psychological factors and internet gaming characteristics. Int J Environ Res Public Health 2017;15(1):40 [FREE Full text] [doi: 10.3390/ijerph15010040] [Medline: 29280953]
12. Choi SW, Kim HS, Kim GY, Jeon Y, Park SM, Lee JY, et al. Similarities and differences among internet gaming disorder, gambling disorder and alcohol use disorder: a focus on impulsivity and compulsivity. J Behav Addict 2014;3(4):246-253 [FREE Full text] [doi: 10.1556/JBA.3.2014.4.6] [Medline: 25592310]
13. Smith KL, Hummer TA, Hulvershorn LA. Pathological video gaming and its relationship to substance use disorders. Curr Addict Rep 2015;2(4):302-309. [doi: 10.1007/s40429-015-0075-6]
14. Marques LM, Uchida PM, Aguiar FO, Kadri G, Santos RIM, Barbosa SP. Escaping through virtual gaming-what is the association with emotional, social, and mental health? A systematic review. Front Psychiatry 2023;14:1257685 [FREE Full text] [doi: 10.3389/fpsyt.2023.1257685] [Medline: 38025467]
15. Škařupová K, Blinka L, Ťápal A. Gaming under the influence: an exploratory study. J Behav Addict 2018;7(2):493-498 [FREE Full text] [doi: 10.1556/2006.7.2018.27] [Medline: 29788755]
16. Pollack CC, Kim J, Emond JA, Brand J, Gilbert-Diamond D, Masterson TD. Prevalence and strategies of energy drink, soda, processed snack, candy and restaurant product marketing on the online streaming platform twitch. Public Health Nutr 2020;23(15):2793-2803 [FREE Full text] [doi: 10.1017/S1368980020002128] [Medline: 32618235]

17. Ip EJ, Urbano EPT, Caballero J, Lau WB, Clauson KA, Torn RA, et al. The video gamer 500: performance-enhancing drug use and internet gaming disorder among adult video gamers. Comput Human Behav 2021;123:106890. [doi: 10.1016/j.chb.2021.106890]
18. Griffiths MD, Lewis AM, Ortiz De Gortari AB, Kuss DJ. Online forums and solicited blogs: innovative methodologies for online gaming data collection. Stud Psychol 2016;15(2):101-122 [FREE Full text]
19. Chi Y, Chen H. Investigating substance use via reddit: systematic scoping review. J Med Internet Res 2023;25:e48905 [FREE Full text] [doi: 10.2196/48905] [Medline: 37878361]
20. Bergstrom K, Poor N. Reddit gaming communities during times of transition. Soc Media Soc 2021;7(2):205630512110101 [FREE Full text] [doi: 10.1177/20563051211010167]
21. Gameopedia. Unique taxonomy in gaming: classifying video game genres. 2023. URL: https://www.gameopedia.com/video-game-genres/ [accessed 2024-07-03]
22. Steam Tags (Steamworks Documentation). URL: https://partner.steamgames.com/doc/store/tags [accessed 2024-07-03]
23. Reuters. China announces rules to reduce spending on video games. URL: https://www.reuters.com/world/china/china-issues-draft-rules-online-game-management-2023-12-22/ [accessed 2024-07-03]
24. Mohammadi M, RezaeiDehaghani A, Mehrabi T, RezaeiDehaghani A. Association between playing computer games and mental and social health among male adolescents in Iran in 2014. Iran J Nurs Midwifery Res 2016;21(2):153-158 [FREE Full text] [doi: 10.4103/1735-9066.178236] [Medline: 27095988]
25. Chen V. Council post: leveling up the gaming gender gap.: Forbes URL: https://www.forbes.com/sites/forbesbusinesscouncil/2023/08/24/leveling-up-the-gaming-gender-gap/ [accessed 2024-07-16]
26. Statista. Video game industry - statistics & facts. URL: https://www.statista.com/topics/868/video-games/ [accessed 2024-07-16]

Abbreviations

APA: American Psychological Association
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)*
IGD: internet gaming disorder
OR: odds ratio
WHO: World Health Organization

Edited by A Mavragani; submitted 08.03.24; peer-reviewed by I Bonilla, D Frohlich; comments to author 16.05.24; revised version received 16.07.24; accepted 12.08.24; published 02.10.24.

Please cite as:
Le N, McMann T, Yang L, Li Z, Cuomo RE, Mackey TK
Detection and Characterization of Online Substance Use Discussions Among Gamers: Qualitative Retrospective Analysis of Reddit r/StopGaming Data
JMIR Infodemiology 2024;4:e58201
URL: https://infodemiology.jmir.org/2024/1/e58201
doi:10.2196/58201
PMID:

Original Paper

Uncovering the Top Nonadvertising Weight Loss Websites on Google: A Data-Mining Approach

Carlos A Almenara[1*], PhD; Hayriye Gulec[2*], PhD

[1]School of Health Sciences, Universidad Peruana de Ciencias Aplicadas, Lima, Peru

[2]Interdisciplinary Research Team on Internet and Society, Faculty of Social Studies, Masaryk University, Brno, Czech Republic

[*]all authors contributed equally

Corresponding Author:
Carlos A Almenara, PhD
School of Health Sciences
Universidad Peruana de Ciencias Aplicadas
Av. Alameda San Marcos 11
Lima, 15067
Peru
Phone: 51 3133333 ext 2803
Fax: 51 3133333
Email: carlos.almenara@upc.pe

Abstract

Background: Online weight loss information is commonly sought by internet users, and it may impact their health decisions and behaviors. Previous studies examined a limited number of Google search queries and relied on manual approaches to retrieve online weight loss websites.

Objective: This study aimed to identify and describe the characteristics of the top weight loss websites on Google.

Methods: This study gathered 432 Google search queries collected from Google autocomplete suggestions, “People Also Ask” featured questions, and Google Trends data. A data-mining software tool was developed to retrieve the search results automatically, setting English and the United States as the default criteria for language and location, respectively. Domain classification and evaluation technologies were used to categorize the websites according to their content and determine their risk of cyberattack. In addition, the top 5 most frequent websites in nonadvertising (ie, nonsponsored) search results were inspected for quality.

Results: The results revealed that the top 5 nonadvertising websites were healthline.com, webmd.com, verywellfit.com, mayoclinic.org, and womenshealthmag.com. All provided accuracy statements and author credentials. The domain categorization taxonomy yielded a total of 101 unique categories. After grouping the websites that appeared less than 5 times, the most frequent categories involved “Health” (104/623, 16.69%), “Personal Pages and Blogs” (91/623, 14.61%), “Nutrition and Diet” (48/623, 7.7%), and “Exercise” (34/623, 5.46%). The risk of being a victim of a cyberattack was low.

Conclusions: The findings suggested that while quality information is accessible, users may still encounter less reliable content among various online resources. Therefore, better tools and methods are needed to guide users toward trustworthy weight loss information.

(JMIR Infodemiology 2024;4:e51701) doi:10.2196/51701

KEYWORDS

consumer health informatics; cyberattack risk; data mining; Google; information seeking; weight loss; online health information; website analysis; digital health; internet search

Introduction

Seeking online health information is prevalent among Americans [1,2]. In 2013, a representative survey of adult internet users in the United States found that 27% had looked online for information about losing or controlling their weight in the last 12 months [3]. General search engines are the first tool for health-related queries [4]. With an average of over 80 billion monthly visits in 2022, Google is the most preferred search engine for such internet-based activity [5].

Although limited, the available evidence points to a lack of quality and comprehensiveness regarding weight loss information in English and Spanish on websites in the United States and the German language web [6-8]. It also indicates the potential risk of accessing information with doubtful credibility, such as deceiving advertisements about commercial and fad diets [9]. These findings are concerning because online health information impacts health decisions and behaviors [3]. Furthermore, low-quality or misleading information might be detrimental for those who have eating disorder-related symptomatology or obesity because previous research identified a higher frequency for the use of weight loss websites among individuals with body image and body weight concerns [10].

Google Trends, an open-source repository, is a popular tool for retrieving the most common search queries about weight loss in previous studies [11-14]. Yet, the Google search engine offers additional "Autocomplete Suggestions" and "People Also Ask" features based on users' queries. To our knowledge, no previous study used search queries that combined the data from Google Trends and the additional features of the Google Search engine to determine the most frequently accessed weight loss websites. This study aimed to identify and describe the characteristics of the top weight loss websites on Google using an automated approach that combined data from Google Trends and the additional features of the Google Search engine.

Methods

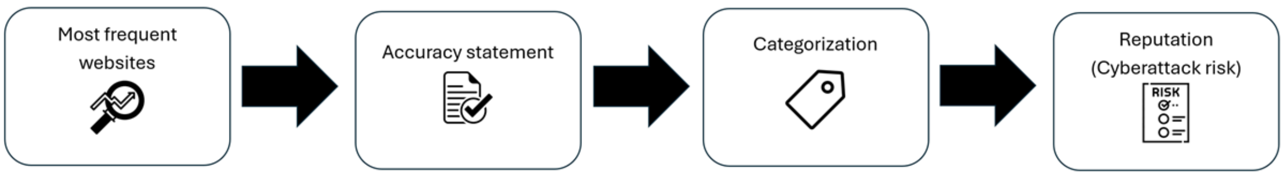
Data Mining

On April 6, 2021, a comprehensive list of 432 unique search queries about weight loss (Multimedia Appendix 1) was retrieved from the Answer Socrates database [15] with the search term "weight loss" (without quotations). Answer Socrates is a database of users' search queries combining Google search suggestions (Autocomplete Suggestions), "People Also Ask" featured questions and Google Trends data. We developed a data-mining software tool (Multimedia Appendix 2) to search for each query in Google and automatically retrieve the results. The tool sets the United States as the default geographical location and English as the default language. In this study, we set the United States as the default location because it is the largest English-speaking country. According to Google Trends, the third largest volume of "weight loss" searches since 2004 has occurred in the United States, after Trinidad and Tobago and South Africa [16]. The tool was developed with Python (Python Software Foundation) programming language and designed to gather Google Search engine results through an application programming interface (API) developed by SerpApi [17]. More precisely, it is a Python-code tool for data mining that automatically searches each query on Google and retrieves the results for each query in a JSON file. Each JSON file contains Google advertising results, "Google Places" results, "People Also Ask" results, and organic search results, including web address, title, and description. Organic results are nonadvertising (ie, nonsponsored) search results that appear because they are relevant to the search terms [18]. By contrast, nonorganic or sponsored or advertising results are paid advertisements and appear above organic results. The Python code tool creates a database (a CSV file) that includes the top 100 organic results for each query (Multimedia Appendix 3). To test for the validity of this software tool, the first author performed a manual verification. The tool was set to Peru as the default geographical location because the first author was located in Peru. English was set as the default language because 10 queries were randomly selected from the total of 432 English queries used in this study. A manual search in Google was done separately for each of those 10 queries to compare the results obtained with the software tool. The manual searches were done with an anonymous web browser (Mozilla Firefox) in Peru, with English as the default language of the search engine. The key data (ie, title, description, URL) of the first 10 nonadvertising results for each query were compared, and no differences were found. The tool was used on April 14, 2021. This study focused solely on organic search results. This decision was based on the fact that advertising or sponsored results have already been studied, although several years ago [6], and they are idiosyncratic because they are regulated by algorithms. In other words, advertising results are tailored to each user and, therefore, do not represent the results retrieved for all users, as is the case for organic results.

Data Analyses

The data analysis procedure is represented in the flowchart below (Figure 1).

Figure 1. Steps of the data analysis procedure.



First, the most frequent websites for the top 5 Google Search results were calculated using Python code (Multimedia Appendix 2). This study focused on the top 5 organic search results because nearly 90% of the clicks occur on Google's first 5 nonadvertising results [7]. This clicking behavior might suggest users' tendency to perceive the information presented at the top as most relevant and helpful. Since we focused on queries looking for answers about weight loss on websites, we expected a similar clicking behavior in this study. Next, the 5 most frequent websites were manually inspected to evaluate their accuracy. Like previous studies [6], we examined whether the websites presented an "accuracy statement" that indicated that the provided information was evidence-based, fact-checked, or, at least, had been reviewed to ensure its trustworthiness. Any statement like this within the website indicated that the website provided an accuracy statement. The websites were also examined to evaluate whether author credentials were disclosed when weight loss information was provided. When the authorship of the weight loss information was given, the website was regarded as providing author credentials.

Then, all the unique websites that appeared in the top 5 results were categorized. There are several methods to categorize websites, including manual and automated processes. A recent study examined several domain classification services and provided directions for future research using website categories [19]. The study found that manual categorization of websites tends to be biased due to disagreements resulting from subjective opinions. In addition, the inherent ambiguity of many categories in taxonomy and the dual nature of many websites contribute to this bias [19]. For example, taxonomies containing thousands of categories and websites with ambiguous content, such as tourism blogs advertising casinos, make it difficult to categorize them manually. On the other hand, the study found that domain classification services vary in coverage, meaning that some websites are categorized because the service indexes them, whereas others do not. Furthermore, the accuracy of the categorization is affected by inconsistent taxonomies. Finally, there is low agreement among these services. Consequently, the authors suggest that researchers can manually examine random subsets of categorizations to determine if the labeling quality is sufficient for the purpose of the research study and rely on classification services with category labels acquired through a well-documented process and incorporated into a thoroughly vetted taxonomy [19]. Based on those results and recommendations, the following domain classification services were tested: BrightCloud, Curlie, Cyren, FortiGuard, McAfee, VirusTotal, and Zvelo. After manually inspecting the results, we found that Zvelo [20] produced the most meaningful and accurate classification labels with a reasonable number of categories (nearly 500).

Zvelo is a domain classification service with more than 13 years of service. The categories are generated by Zvelo's human-supervised artificial intelligence (AI) models, and they currently use fourth-generation AI models (personal communication, January 31, 2023). For each URL (ie, web address), the Zvelo database provides 3 classification groups, each with up to 3 categories [20]. The first group is the primary classification provided by Zvelo; this is the one used in this study. The other two are additional categorizations provided by the Interactive Advertising Bureau (IAB), a marketing-orientated taxonomy [21]. Consistent with the recommendations made by Vallina et al [19], Zvelo was preferred over the IAB taxonomy because it provided more extensive coverage of the websites in the dataset. The manual inspection of the website categories revealed that Zvelo's category labels were also more relevant to the content than the IAB's. This observation is probably because the IAB is a marketing-oriented domain classification service. We also preferred the Zvelo classification due to its state-of-the-art AI technology use. All categories for each URL are available in Multimedia Appendix 4. The categories were retrieved between January 2023 and February 2023.

Finally, each URL was given a reputation score based on Webroot's BrightCloud IP Reputation Services [22]. The reputation score predicts the risk that an IP will deliver a cyberattack (ie, it identifies malicious addresses). Scores are obtained using big data and a machine-learning algorithm. Scores are grouped into 5 categories: High-risk (1-20), Suspicious (21-40), Moderate-risk (41-60), Low-risk (61-80), and Trustworthy (81-100) [19].

Ethical Considerations

Given that no human participants participated in this study and the data were public, the study was deemed exempt from the institutional review board approval process.

Results

After applying the Python-code tool to each JSON file, a CSV file was created containing a database with the 432 search queries and the Top 100 organic search results (Multimedia Appendix 3).

Table 1 shows the count of the most frequent website results, the trustworthiness, the reputation, and the web categorization of the top 5 website results.

Table 1. The 5 most frequent websites within the top 5 Google search results (N=2160).

Website	Total[a], n (%)	Accuracy statement	Author credentials	Reputation (0-100)[b]	Category
healthline.com	197 (9.12)	Yes	Yes	96	Health
webmd.com	91 (4.21)	Yes	Yes	96	Health
verywellfit.com	85 (3.94)	Yes	Yes	88	Nutrition & diet
mayoclinic.org	79 (3.66)	Yes	Yes	100	Health
womenshealthmag.com	61 (2.82)	Yes	Yes	88	Women's health
Other	1647 (76.25)	—[c]	—	—	—
Total	2160 (100)	—	—	—	—

[a]Total number of times the website appeared in the top 5 results.

[b]Cyberattack risk. Scores higher than 80 indicate a trustworthy website.

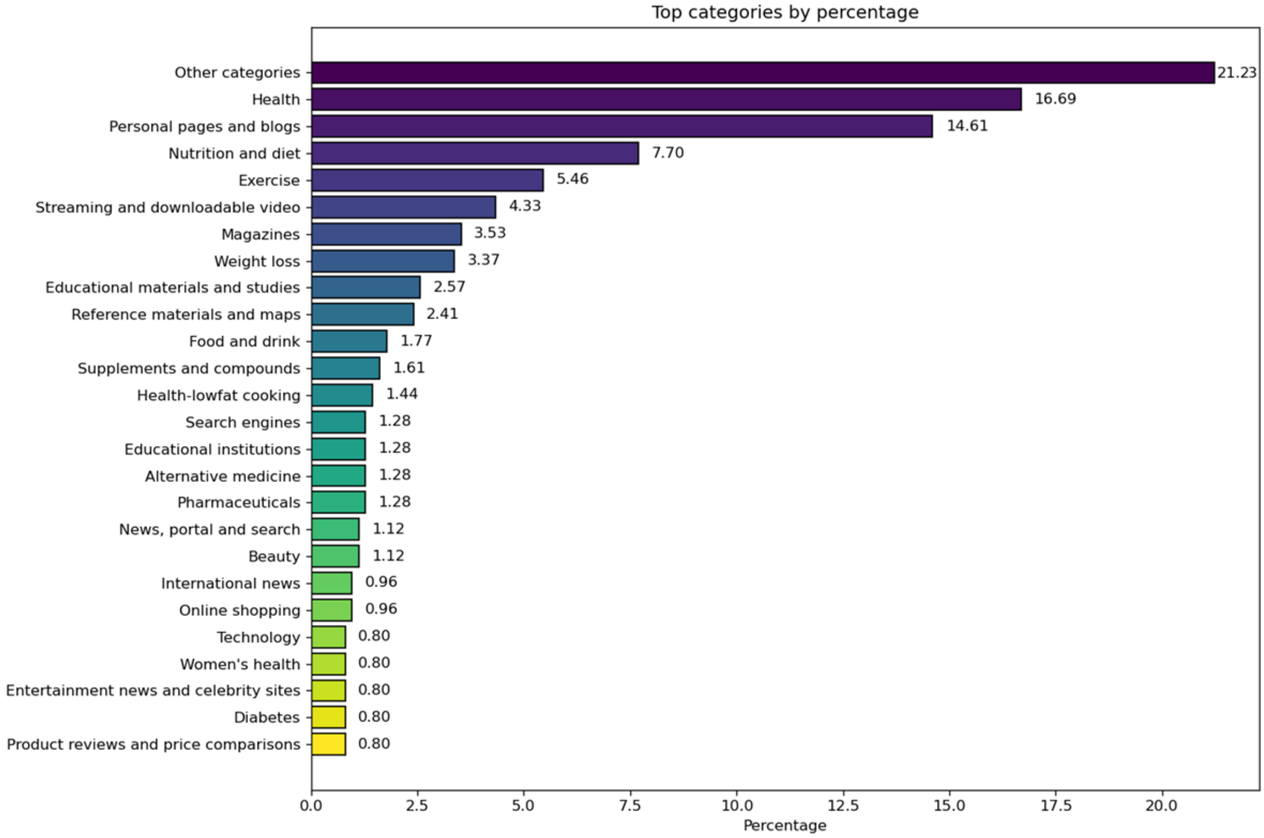
[c]Not applicable.

There were 2160 results (the first 5 organic search results multiplied by 432 queries) and 623 unique websites (Multimedia Appendix 5). The most frequent website was healthline.com, which appeared 197/2160 (9.12%) times in the top 5 organic search results. This was followed by webmed.com, verywellfit.com, mayoclinic.org, and womenshealthmag.com.

These top 5 most frequent websites provided accuracy statements (eg, fact-check statements) and author credentials. The authors of healthline.com, webmd.com, and verywellfit.com were usually health care professionals. In the case of mayoclinic.org, the authorship was usually declared as "Mayo Clinic Staff." womenshealthmag.com had several authors who were described as "freelance writers." In addition, healthline.com, verywellfit.com, and womenshealthmag.com explicitly provided a fact-check statement to assure users that the information was evidence-based and written by health care professionals.

Next, the 623 websites were categorized by the Zvelo categorization taxonomy [23]. In total, 101 unique categories were found (Multimedia Appendix 4). The categories with less than 5 appearances were grouped into "Other Categories." As can be seen in Figure 2, most websites appeared less than 5 times (Other categories: 132/623 times, 21.23%), indicating a large diversity of results.

Figure 2. Frequency and percentage of website categories.



The following most common categories were "Health" (104/623, 16.69%) and "Personal Pages and Blogs" (91/623, 14.61%), followed far behind by "Nutrition and Diet" (48/623, 7.7%) and "Exercise" (34/623, 5.46%). The most common websites in the "Health" category tended to ensure that they provided the most reliable information (eg, healthline.com). By contrast, "Personal Pages and Blogs" had little to no information about the authorship and accuracy of the information provided, although with some exemptions (eg, thehealthy.com).

Table 2 shows the 5 most frequent websites for the most frequent categories (not including the "Other Categories").

Table 2. Examples of the most frequent websites for the top categories.

	Health	Personal pages and blogs	Nutrition and diet	Exercise
1	healthline.com	simple-nourished-living.com	verywellfit.com	shape.com
2	webmd.com	obesitycoverage.com	livestrong.com	muscleandfitness.com
3	mayoclinic.org	thehealthy.com	mediweightloss.com	core-trainingpt.com
4	medicalnewstoday.com	developgoodhabits.com	precisionnutrition.com	sparkpeople.com
5	prevention.com	blog.myfitnesspal.com	trifectanutrition.com	thinwithin.com

With some exemptions (4.01%), most websites ranked 79 or higher (95.99%) in reputation, with a mean reputation score of 89.85 (SD 8.53). This finding indicates that the average reputation score of websites was trustworthy (Multimedia Appendix 4). The reputation scores for the 5 most frequent websites within the top 5 Google search results are displayed in Table 1.

Discussion

Principal Findings

This study identified and described the characteristics of the top weight loss websites on Google using an automated approach that combined data from Google Trends and the additional features of the Google Search engine. The 5 most frequent websites for weight loss were healthline.com, webmd.com, verywellfit.com, mayoclinic.org, and womenshealthmag.com. They provided fact-checked or evidence-based information about weight loss, written by health care professionals with disclosed credentials. We used the Zvelo domain classification service to categorize the search results. The results yielded a wide diversity of website categories. The most common category was "Other Categories," accounting for around one-fifth of all websites. It was followed by "Health," "Personal Pages and Blogs," "Nutrition and Diet," and "Exercise" categories, which altogether accounted for 45% (277/623) of the search results. When we examined the most frequent 5 websites under these categories, only the "Health" category involved websites with fact-checked or evidence-based weight loss information that disclosed the author's credentials. By comparison, the information on authorship and the accuracy of the statements was mainly lacking in the "Personal Pages and Blogs," "Nutrition and Diet," and "Exercise" categories. Finally, we found that the average reputation of all websites was trustworthy, which indicates that there is no security risk and users are very unlikely to be victims of cyberattacks by these weight loss websites.

Comparison With Previous Work

Previous studies found that websites with higher quality and comprehensiveness rank low in the search results [6]. Although we identified that evidence-based and fact-checked information was conveyed by the most frequent websites in the top 5 organic search results, it should be noted that what users retrieve as a result of their search queries may not always lead them to the websites with credible information, given that they represent only a fraction of the total number of websites. Furthermore, websites with quality content may get lost among the myriad results, including advertising results that appear first on websites like Google.

Previous studies adopted the standard website nomenclature and manual categorization of websites based on their source (eg, commercial sites, medical sites, government sites, university sites, news sites, online media sites, and blogs) [6-8]. In this study, we used the Zvelo domain classification service, which uses an AI-based classification engine to apply a topic-based taxonomy for URL classification. Given that we applied this approach for the first time on weight loss websites, there is no previous evidence to compare our findings. Nevertheless, it is noteworthy that earlier studies reported nutrition and exercise as the most frequently addressed topics across the website categories on weight loss. We found that "Personal Pages and Blogs" was a frequent category in the search results, which means that individuals who seek online weight loss information might frequently encounter personal accounts and histories on English-language weight loss websites in the United States. Given that the most frequent websites under "Personal Pages and Blogs," "Nutrition and Diet," and "Exercise" categories have suboptimal quality, users who endorse these websites may be at risk of inadequate and even misleading information.

Future Directions

In this study, we evaluated the 5 most frequent websites under each category and determined whether they disclosed author credentials and an accuracy statement to ensure their quality. As a next step, future studies can examine each topic category based on other quality criteria, such as the frequency and scope of the evidence-based information provided using manual or automated approaches. This information might give a more in-depth understanding of the content and quality of the website categories. For instance, websites in the "Health" category might be more trustworthy because they integrate a more holistic approach by addressing several evidence-based domains (eg, healthy eating, nutrition, diet, exercise, pharmacotherapy, and behavioral change strategies) for successful weight loss.

Cyberattacks have social and psychological consequences, such as the series of emotional reactions that can follow a cyberattack (feeling violated, powerless, angry, rage, grief, shame, etc), and previous research has highlighted the importance of addressing them [24]. Thus, there are diverse opportunities for future studies, such as studying how cyberattacks target weight loss advertisements and the psychological consequences of these attacks for individuals looking to lose weight.

Strengths

Previous studies generated search queries with convenience samples for the terms with which people would search the internet for weight loss information [6-8]. The search terms were then categorized and submitted to the Google Trends repository to obtain the final search queries. The total number of search queries ranged between 26 and 30. In this study, we retrieved 432 unique search queries from Answer Socrates [15], which combined Google search suggestions (Autocomplete suggestions), "People Also Ask" featured questions and Google Trends data. We also used a data-mining software tool to determine the top 5 most frequent websites that appeared among the 5 organic search results. This study provided initial evidence that a data-mining software tool can help identify the most common websites for weight loss. This automated approach might also help to inspect other health-related search queries submitted to the Google search engine.

Limitations

Although this study has strengths, such as the large number of search queries used and the automated approach to retrieve search results and categorize websites, it has some limitations. First, our results are concerned with English-language websites found in the United States; thus, they cannot be generalized to other locations. Further research may focus on different locations for English search queries related to weight loss and compare the Google search results across locations. Second, notwithstanding the fact that actual Google users make the search queries, it is unclear how Google selects them. Therefore, future studies could use different approaches, such as an ecological design, to gather search queries when users search the internet for weight loss information. For example, real-time data capture techniques, including an ecological momentary assessment, can be used in a large sample to study the behavior of individuals while looking for weight loss information [25]. Such a design also has the advantage of tracking various concurrent details, such as the time of day, the season, the geographical location, sociodemographics, anthropometrics, and current mood, to study their association with the search terms. Similarly, case study research can be used to explore this phenomenon because it allows for an in-depth exploration of individual behaviors, contextual influences, and the factors that shape the information-seeking process [26]. This approach would help develop a comprehensive understanding of how individuals interact with weight loss information online, including the challenges they face and the strategies they use to navigate the abundance of available online information.

Third, we used a data-driven and AI-based domain classification approach using Zvelo. Nevertheless, the categories produced by Zvelo were not readily interpretable. There was a broad diversity of website categories and possible overlaps between them (eg, "Nutrition & Diet," "Health-Low Fat Cooking," "Food & Drink"). The interpretability of AI-based domain classification should be determined in future studies. Fourth, we manually inspected websites to find the accuracy statements (eg, fact-check statements). These statements were usually in sections like "About Us." Unfortunately, this method is time-consuming, so future studies can look for a way to automate this task so that all the websites can be examined for accuracy statements. For example, a third-party authenticator can provide a certificate to websites that provide accurate information about weight loss and an API to developers so they can easily access this information.

Fifth, from the data-mining results, it was clear that sponsored results (ie, advertising), which were not included in the analyses, were assigned the top results in Google searches. Weight loss advertising already has many deceiving pieces of information, with hundreds of cases presented to the Federal Trade Commission [27,28]. In addition, Google's online advertising is regulated by algorithms that consider idiosyncratic information about the internet user, such as sociodemographic information, so the ads are better tailored to the user. Future studies could evaluate how this information is displayed to the user, how users interact with it, and how they use it in their everyday lives, particularly in their attempts to lose weight, to evaluate the risks and benefits of interacting with online advertising results. These studies could use an automated approach like the one used in this study.

Therefore, initiatives should be taken to improve web results so that users can get reliable information about weight loss. For example, a web browser plugin that identifies the top health websites can mark them with an icon so that users can quickly identify them. An example of such a web browser plugin is Web of Trust [29], which provides a website reputation rating, and it marks URLs with a green (trusted), yellow (suspicious), red (untrusted), or gray (unknown) icon.

Furthermore, governments can create policies to ensure websites disclose how they handle misinformation and fact-checking. Websites can also be required to label content that has been fact-checked, along with the source of the fact-check. Weight loss information generated by AI, such as ChatGPT, should also be disclosed. Internet users should look up websites for the "accuracy statements" indicating that the weight loss information is "evidence-based," "fact-checked," or at least "reviewed by" to ensure its trustworthiness. Similarly, consumers should inspect whether author credentials are disclosed when weight loss information is provided, such as the name of the person who wrote the information, their profession, academic degrees, and links to further professional information (eg, LinkedIn profile page). Software developers can design algorithms to down-rank or flag potentially misleading content or design web browser plugins to alert users interacting with misinformation.

Conclusions

We developed a data-mining approach to identify and evaluate the top nonadvertising weight loss websites found through Google searches and analyzed 432 search queries. The 5 most frequent websites were healthline.com, webmd.com, verywellfit.com, mayoclinic.org, and womenshealthmag.com. These sites consistently provided fact-checked, evidence-based information with author credentials, indicating high trustworthiness. We also classified websites into various categories using an automated approach and found "Health," "Personal Pages and Blogs," "Nutrition and Diet," and "Exercise" as the most common categories. The findings suggested that while quality information is accessible, users may still encounter less reliable content among various online resources. Therefore, better tools and methods are needed to guide users toward trustworthy weight loss information.

Acknowledgments

This output was supported by the "Systemic Risk Institute" (number LX22NPO5101), funded by the European Union - Next Generation EU (Ministry of Education, Youth and Sports, NPO: EXCELES). The authors declare that generative artificial intelligence was not used in ideation or in the writing of any portion of the manuscript or Python code.

Data Availability

The datasets generated during and/or analyzed during this study are available in the Open Science Framework repository [30].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Queries.
[XLSX File (Microsoft Excel File), 17 KB - infodemiology_v4i1e51701_app1.xlsx]

Multimedia Appendix 2
Python code.
[PDF File (Adobe PDF File), 1731 KB - infodemiology_v4i1e51701_app2.pdf]

Multimedia Appendix 3
Results.
[XLSX File (Microsoft Excel File), 1031 KB - infodemiology_v4i1e51701_app3.xlsx]

Multimedia Appendix 4
Categories.
[XLSX File (Microsoft Excel File), 36 KB - infodemiology_v4i1e51701_app4.xlsx]

Multimedia Appendix 5
Google weight loss total count of top 5.
[XLSX File (Microsoft Excel File), 25 KB - infodemiology_v4i1e51701_app5.xlsx]

References

1. Jacobs W, Amuta AO, Jeon KC. Health information seeking in the digital age: an analysis of health information seeking behavior among US adults. Cogent Soc Sci 2017;3(1):1302785. [doi: 10.1080/23311886.2017.1302785]
2. Rideout V, Fox S, Peebles A, Robb MB. Coping with COVID-19: How young people use digital media to manage their mental health. URL: https://www.commonsensemedia.org/sites/default/files/research/report/2021-coping-with-covid19-full-report.pdf [accessed 2024-10-15]
3. Fox S, Duggan M. Health online 2013. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2021-11-17]
4. Fox S. Online health search 2006. URL: https://www.pewresearch.org/internet/2006/10/29/online-health-search-2006/ [accessed 2021-10-31]
5. Google. Similarweb. 2023. URL: https://www.similarweb.com/website/google.com/ [accessed 2023-02-13]
6. Modave F, Shokar NK, Peñaranda E, Nguyen N. Analysis of the accuracy of weight loss information search engine results on the internet. Am J Public Health 2014;104(10):1971-1978. [doi: 10.2105/AJPH.2014.302070] [Medline: 25122030]
7. Cardel MI, Chavez S, Bian J, Peñaranda E, Miller DR, Huo T, et al. Accuracy of weight loss information in Spanish search engine results on the internet. Obesity (Silver Spring) 2016;24(11):2422-2434 [FREE Full text] [doi: 10.1002/oby.21646] [Medline: 27653438]
8. Meyer S, Elsweiler D, Ludwig B. Assessing the quality of weight loss information on the German language web. Mov Nutr Heal Dis 2020;3:39-52. [doi: 10.22233/9781910443361-3e.9]
9. Passos JA, Vasconcellos-Silva PR, Santos LADS. Cycles of attention to fad diets and internet search trends by Google trends. Cien Saude Colet 2020;25(7):2615-2631 [FREE Full text] [doi: 10.1590/1413-81232020257.23892018] [Medline: 32667545]
10. Almenara CA, Machackova H, Smahel D. Sociodemographic, attitudinal, and behavioral correlates of using nutrition, weight loss, and fitness websites: an online survey. J Med Internet Res 2019;21(4):e10189 [FREE Full text] [doi: 10.2196/10189] [Medline: 30946018]
11. Coogan S, Sui Z, Raubenheimer D. Gluttony and guilt: monthly trends in internet search query data are comparable with national-level energy intake and dieting behavior. Palgrave Commun 2018;4(1):1-9. [doi: 10.1057/s41599-017-0055-7]
12. Madden KM. The seasonal periodicity of healthy contemplations about exercise and weight loss: ecological correlational study. JMIR Public Health Surveill 2017;3(4):e92 [FREE Full text] [doi: 10.2196/publichealth.7794] [Medline: 29237582]
13. Pawar AS, Nagpal S, Pawar N, Lerman LO, Eirin A. General public's information-seeking patterns of topics related to obesity: Google trends analysis. JMIR Public Health Surveill 2020;6(3):e20923 [FREE Full text] [doi: 10.2196/20923] [Medline: 32633725]
14. Teng Y, Huang SW, Li Z, Xie Q, Zhang M, Lou Q, et al. Seasonal variation and trends in the Internet searches for losing weight: an infodemiological study. Obes Res Clin Pract 2020;14(3):225-233. [doi: 10.1016/j.orcp.2020.04.001] [Medline: 32349915]
15. AnswerSocrates. 2023. URL: https://answersocrates.com/ [accessed 2021-04-06]
16. Google Trends. Weight loss. 2023. URL: https://trends.google.com/trends/explore?date=all&q=weight%20loss&hl=en [accessed 2023-01-10]
17. SerpApi. Google search engine results API. 2021. URL: https://serpapi.com/search-api [accessed 2021-04-07]

18. Google. Organic search result. 2023. URL: https://support.google.com/google-ads/answer/6054492 [accessed 2023-05-30]
19. Vallina P, Le PV, Feal Á, Paraschiv M, Gamba J, Burke T, et al. Mis-shapes, mistakes, misfits: an analysis of domain classification services. New York, NY, United States: Association for Computing Machinery; 2020 Presented at: Proceedings of the ACM Internet Measurement Conference; 2020 October 27-29; Virtual Event USA p. 598-618. [doi: 10.1145/3419394.3423660]
20. Zvelo. zveloDB™ URL Database. 2023. URL: https://zvelo.com/zvelodb-url-database [accessed 2023-01-05]
21. Interactive Advertising Bureau. Content taxonomy. 2023. URL: https://www.iab.com/guidelines/content-taxonomy/ [accessed 2023-01-06]
22. BrightCloud. URL or IP Lookup. URL: https://www.brightcloud.com/tools/url-ip-lookup.php [accessed 2023-01-06]
23. Zvelo. zveloLIVE: check a URL. 2023. URL: https://tools.zvelo.com [accessed 2024-01-05]
24. Bada M, Nurse JRC. The social and psychological impact of cyberattacks. In: Emerging Cyber Threats and Cognitive Vulnerabilities. Cambridge, Massachusetts: Academic Press; 2020:73-92.
25. Montag C, Baumeister H. Digital phenotyping and mobile sensing. In: New Developments in Psychoinformatics. Berlin, Heidelberg, Dordrecht, and New York City: Springer; 2022.
26. Yin RK. Case study research and applications: Design and methods. 6th ed. UK: Sage; 2018.
27. Anderson KB. Consumer fraud in the United States, 2011: The third FTC survey. 2013. URL: https://www.ftc.gov/sites/default/files/documents/reports/consumer-fraud-united-states-2011-third-ftc-survey/130419fraudsurvey_0.pdf [accessed 2022-03-13]
28. Federal Trade Commission. Gut check: a reference guide for media on spotting false weight loss claims. 2014. URL: https://www.ftc.gov/business-guidance/resources/gut-check-reference-guide-media-spotting-false-weight-loss-claims [accessed 2023-02-07]
29. Web of Trust. 2023. URL: https://www.mywot.com/ [accessed 2023-07-02]
30. Carlos A Almenara. Uncovering the top non-advertising weight loss websites on Google: A data-mining approach. URL: https://osf.io/xeajh [accessed 2024-11-20]

Abbreviations

AI: artificial intelligence
API: application programming interface
IAB: Interactive Advertising Bureau

Edited by T Mackey; submitted 08.08.23; peer-reviewed by C Collins, R Ciptaningtyas, S Goldstein, C Lopezosa; comments to author 22.07.24; revised version received 21.08.24; accepted 24.10.24; published 11.12.24.

Please cite as:
Almenara CA, Gulec H
Uncovering the Top Nonadvertising Weight Loss Websites on Google: A Data-Mining Approach
JMIR Infodemiology 2024;4:e51701
URL: https://infodemiology.jmir.org/2024/1/e51701
doi:10.2196/51701
PMID:

Original Paper

Application of a Language Model Tool for COVID-19 Vaccine Adverse Event Monitoring Using Web and Social Media Content: Algorithm Development and Validation Study

Chathuri Daluwatte[1], PhD; Alena Khromava[2], MD, MPH; Yuning Chen[1], MS, PhD; Laurence Serradell[3], PhD; Anne-Laure Chabanon[4], PharmD, PhD; Anthony Chan-Ou-Teung[5], MS; Cliona Molony[1], PhD; Juhaeri Juhaeri[6], PhD

[1]Digital Data, Sanofi, Cambridge, MA, United States
[2]Epidemiology and Benefit-Risk Department, Sanofi, Toronto, ON, Canada
[3]Epidemiology and Benefit-Risk Department, Sanofi, Lyon, France
[4]Global Pharmacovigilance Department, Sanofi, Lyon, France
[5]Digital Data, Sanofi, Lyon, France
[6]Epidemiology and Benefit-Risk Department, Sanofi, Bridgewater, NJ, United States

Corresponding Author:
Alena Khromava, MD, MPH
Epidemiology and Benefit-Risk Department
Sanofi
1755 Steeles Ave West
Toronto, ON, M2R 3T4
Canada
Phone: 1 4166672753
Email: alena.khromava@sanofi.com

Abstract

Background: Spontaneous pharmacovigilance reporting systems are the main data source for signal detection for vaccines. However, there is a large time lag between the occurrence of an adverse event (AE) and the availability for analysis. With global mass COVID-19 vaccination campaigns, social media, and web content, there is an opportunity for real-time, faster monitoring of AEs potentially related to COVID-19 vaccine use. Our work aims to detect AEs from social media to augment those from spontaneous reporting systems.

Objective: This study aims to monitor AEs shared in social media and online support groups using medical context-aware natural language processing language models.

Methods: We developed a language model–based web app to analyze social media, patient blogs, and forums (from 190 countries in 61 languages) around COVID-19 vaccine–related keywords. Following machine translation to English, lay language safety terms (ie, AEs) were observed using the PubmedBERT-based named-entity recognition model (precision=0.76 and recall=0.82) and mapped to Medical Dictionary for Regulatory Activities (MedDRA) terms using knowledge graphs (MedDRA terminology is an internationally used set of terms relating to medical conditions, medicines, and medical devices that are developed and registered under the auspices of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use). Weekly and cumulative aggregated AE counts, proportions, and ratios were displayed via visual analytics, such as word clouds.

Results: Most AEs were identified in 2021, with fewer in 2022. AEs observed using the web app were consistent with AEs communicated by health authorities shortly before or within the same period.

Conclusions: Monitoring the web and social media provides opportunities to observe AEs that may be related to the use of COVID-19 vaccines. The presented analysis demonstrates the ability to use web content and social media as a data source that could contribute to the early observation of AEs and enhance postmarketing surveillance. It could help to adjust signal detection strategies and communication with external stakeholders, contributing to increased confidence in vaccine safety monitoring.

(JMIR Infodemiol 2024;4:e53424) doi:10.2196/53424

KEYWORDS

adverse event; COVID-19; detection; large language model; mass vaccination; natural language processing; pharmacovigilance; safety; social media; vaccine

Introduction

An adverse event (AE) is defined by the US Food and Drug Administration (FDA) as any undesirable experience associated with the use of a medical product (including vaccines) in a patient [1]. It can be challenging to assess uncommon or rare AEs in clinical trials due to the low number of patients enrolled in the trials and strict inclusion criteria. Therefore, postmarketing surveillance in a real-world setting is important to gain knowledge of any AE for manufacturers and regulatory bodies, including the FDA and European Medicines Agency (EMA), as well as international organizations such as the World Health Organization (WHO). New safety signals are defined by the EMA as "information on a new or known AE that may be caused by a medicine and requires further investigation" [2]. Organizations (pharmaceutical companies, drug regulators, distributors, wholesalers, and retailers) conduct postmarket pharmacovigilance surveillance both passively and actively, using systems including postauthorization safety studies, as well as voluntary and mandatory surveillance such as the US Centers for Disease Control and Prevention (CDC) or FDA Vaccine Adverse Event Reporting System (VAERS), MedWatch, Eudravigilance, and the WHO's Vigibase. The reporting of suspected or observed AEs is mandatory for manufacturers and in many countries for health care professionals, however, the public may be unaware of AE reporting systems or feel a lack of obligation to report AEs, which in some cases may lead to delayed and incomplete records [3-6]. In addition, there can be a lag time between the reporting of an AE to the regulators and the information being available to the vaccine manufacturer from public sources. With the advent of the COVID-19 pandemic, with new vaccines being developed at speed and vaccines introduced in mass campaigns, assessing very rare AEs at the time of emergency use authorization has been difficult. Postmarketing AE reporting systems have become central to determining AEs and there have been some endeavors to decrease the lag time between reporting of an AE and information being available to the public.

Social media [7] has seen recent unprecedented growth in the numbers of users worldwide and in large populations of patients actively involved in sharing and posting health-related information [7]. This wealth of information has consequently led to data from such discussions being increasingly used for monitoring AEs [7-10], with the advantage of these occurring closer to real time than traditional postmarketing AE reporting systems. Many study groups have made use of natural language processing (NLP) and artificial intelligence methods [7,11,12], including the use of quantum computing [10,13], to observe AEs from social media data [7,10-12] and have had promising outcomes [10]. The COVID-19 pandemic has been at the heart of discussions on social media and the strong patient voice, discussing every concern surrounding COVID-19 vaccines in real time, has been documented in many studies [14-21]. Of note, Portelli et al [22] developed a tool that collected and analyzed public reaction to specific COVID-19 vaccines on 650,000 English language X (formerly Twitter; Twitter, Inc) posts (formerly tweets) since December 2020, including sentiment and AEs. Using a symptom extraction module, they showed news coverage had a high impact on topics discussed.

Safety surveillance will continue to evolve as there is currently a delay in reporting of AEs. A context-aware language model (LM), unlike a dictionary method, increases the sophistication of methods, for example, the interpretation of "corona" as a medical term and a nonmedical term [23]. Our work aims to detect AEs from social media to augment those from spontaneous reporting systems. The ability to monitor AEs on social media in real time has the potential to enhance postmarketing surveillance. Therefore, our objective was to monitor AEs and the related trends shared in social media and online support groups associated with COVID-19 vaccines using medical context-aware LMs.

The safety signals reported here have been denoted as AEs as they are mapped to Medical Dictionary for Regulatory Activities (MedDRA) preferred terms (PTs). However, they are not fully aligned with the definition of AEs that are subject to reporting to health authorities according to applicable regulations.

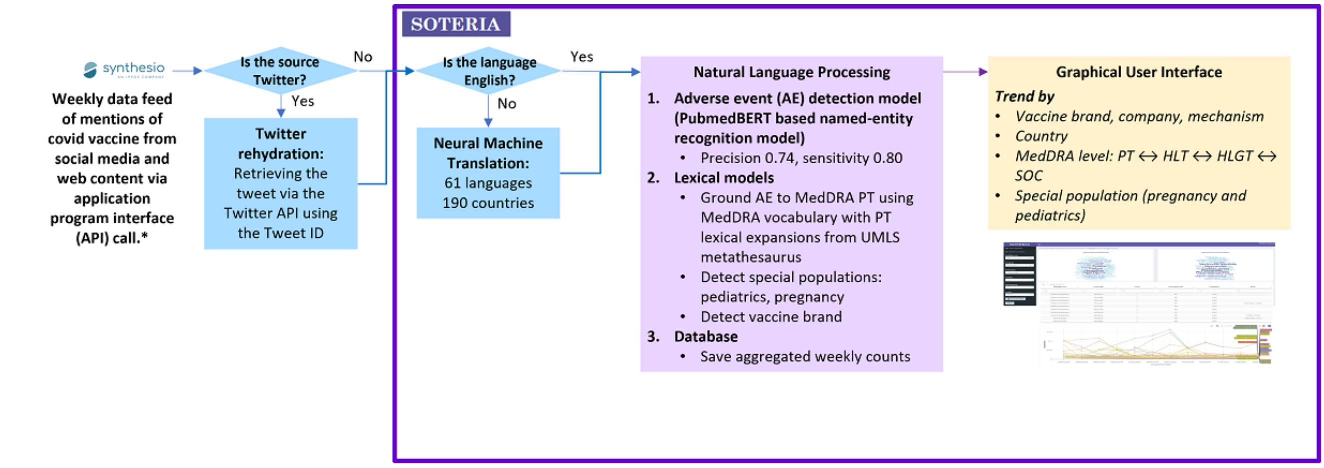
Methods

Overview

Our LM-powered web app is referred to as the Soteria web app from here onwards (Figure 1). Following machine translation to English, lay language AEs related to the safety of COVID-19 vaccines were detected using a named-entity recognition model and mapped to the International Council for Harmonization (ICH) MedDRA standards. Visual analytics as word clouds and line graphs were available on the graphical user interface to analyze across periods any trending of AEs (counts, proportions, and ratios) by COVID-19 vaccine brand, mechanism (messenger ribonucleic acid [mRNA], adenovirus vector, and protein), country, or special population (pediatrics or pregnant women). These can also be grouped by MedDRA hierarchy levels—system organ class (SOC), high-level group term (HLGT), high-level term (HLT), and PT, according to the latest version of MedDRA (23.0-24.0 in this analysis) as per Maintenance and Support Services Organization recommendations.

Contextual lexicons were generated to describe the pediatric population, pregnant population, and vaccine brand. Fuzzy matching detected if these topics were co-occurring with a medDRA PT mention.

Figure 1. Soteria: language model powered web app to analyze web content related to COVID-19 vaccine AEs. AE: adverse event; API: application programming interface; HLGT: high-level group term; HLT: high-level term; MedDRA: Medical Dictionary for Regulatory Activities; mRNA: messenger ribonucleic acid; PT: preferred term; SOC: system organ class; UMLS: Unified Medical Language System. *Synthesio's global partnerships guarantee that customers can complete their datasets with mentions from dozens of geo-specific social media and niche websites. Adhering to Twitter platform agreements and policies, Synthesio shares the ID, but not the tweets via the API. Using tweet IDs, tweets were recovered using the Twitter API, a process called "Twitter rehydration.".



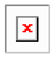
Soteria Web App

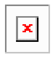
Neural Machine Translation

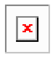
Non-English content in the data stream was translated to English before sending it to the LM for AE observation or detection. For each non-English sentence, the Amazon Translate neural machine translation service or Helsinki-NLP [24] translation models were used according to the source language (Multimedia Appendix 1) and the data source (data from X were translated using Helsinki-NLP, respecting X application programming interface [API] user agreement policies on not sending X posts to third parties). The translations from these machine translation models are continuously evaluated and validated by open-source communities [24] and Amazon Web Services. The Amazon Translate neural machine translation service is an off-the-shelf, usage-based service and Helsinki-NLP is noncommercial and open-source. Non-English sentences written using the standard English alphabet were removed due to known poor performance with machine translation models. COVID-19 vaccines monitored via the Soteria web app are listed in Multimedia Appendix 2.

AE Observation or Detection Model

We fine-tuned the pretrained LM, PubmedBERT [25], to perform a token-level classification task to obtain a named-entity recognition model using 2 publicly available datasets, adverse drug events (ADE)-Corpus-V2 [26] and psychiatric treatment adverse reactions (PsyTAR) [27]. The ADE-Corpus-V2 data contained 4271 sentences with AEs and 16,625 without, the PsyTAR data contained 4813 ADE mentions. The 2 datasets were combined and "machine labeled" for all ADE using the inside, outside, and beginning format and split into 70% training, 20% validation, and 10% test sets. The training was in 3 epochs with a batch size of 32 with a loss function of categorical cross-entropy. The final model's performance was evaluated on the test set using precision (the proportion of sentences that had an AE among all sentences the algorithm had identified an AE), recall (the proportion of sentences the algorithm identified an AE among all sentences with an AE), and F_1-score (harmonic mean of the precision and recall, which is a measure of accuracy).







All analyses were performed in Python (version 3; Python Software Foundation) and with the module torch [28]. This named entity recognition model detected AEs in the Soteria web app (Figure 1).

The validation set was used to search for the hyperparameters that yielded the best performance for this set. The final model was trained on the entire dataset (training set plus evaluation set) and evaluated on a separate test set. Fivefold cross-validation was used for this process. After the final model had been trained, in the test phase, a new separate test set (not used during the training or validation process) provided an unbiased estimate of the model's performance on unseen data. In order to account for the variability introduced by the random split, the model was trained and evaluated on each fold separately, with the results averaged across all folds to obtain a final estimate of the model's performance [29].

MedDRA PT Lexical Expansion

We first performed a lexical expansion of MedDRA PTs using the Unified Medical Language System (UMLS) metathesaurus [30]. Each MedDRA PT was mapped with UMLS synonyms that have the same concept unique identifier but were from a different vocabulary other than MedDRA (Multimedia Appendix 3).

Lay language safety terms detected by the named-entity recognition model in the Soteria web app were mapped to MedDRA PT using these lexical expansions derived from the UMLS metathesaurus (Figure 1) using fuzzy matching. This mapping of AEs into MedDRA terminology allowed the harmonization of terms for a better understanding of patients' chatting.

AE Trend Generation

Using mapped MedDRA terms, reports were generated in the Soteria web app as weekly or monthly AE count (for period and cumulative), proportion, ratio, and 95% CI around the ratio.

These metrics can be calculated for combinations of groupings by country, mechanism (mRNA, adenovirus vector, and protein), COVID-19 vaccine brand names, and MedDRA levels (PT, HLT, HLGT, and SOC; Figure 2).

These trends are presented as word clouds, tables with monthly top 50 AEs, and time trend line plots (Figure 3).

Figure 2. Detecting AEs for combinations of grouping by country, mechanism, brand names, and MedDRA levels (preferred term, high-level term, high-level group term, and system organ class). AE: adverse event; HLGT: high-level group term; HLT: high-level term; MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term; SOC: system organ class.

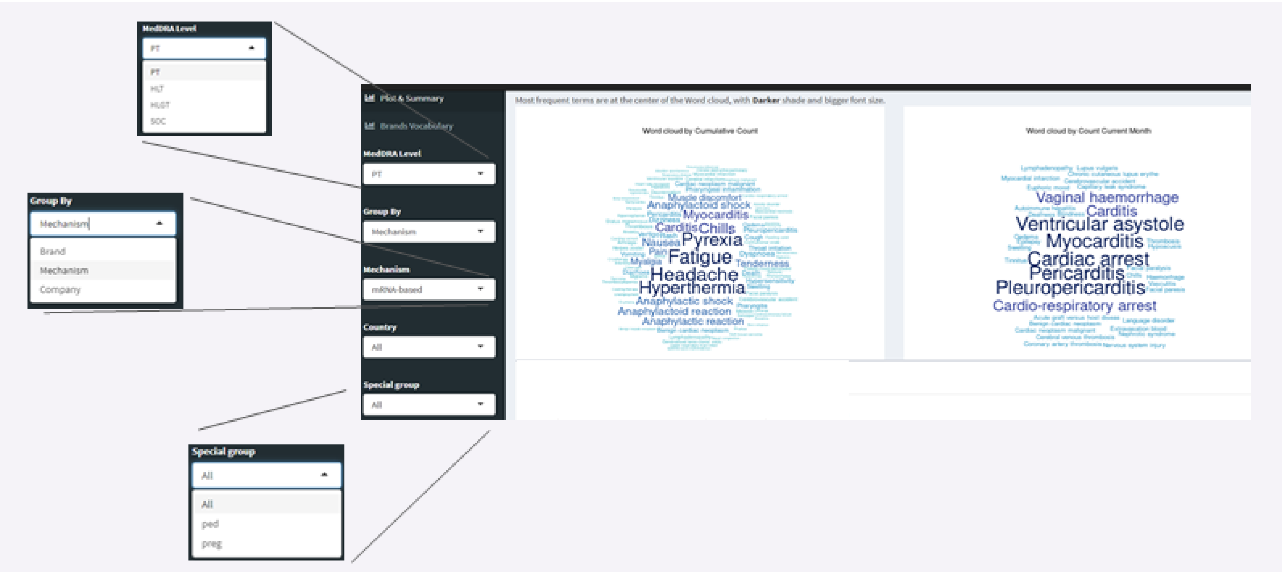
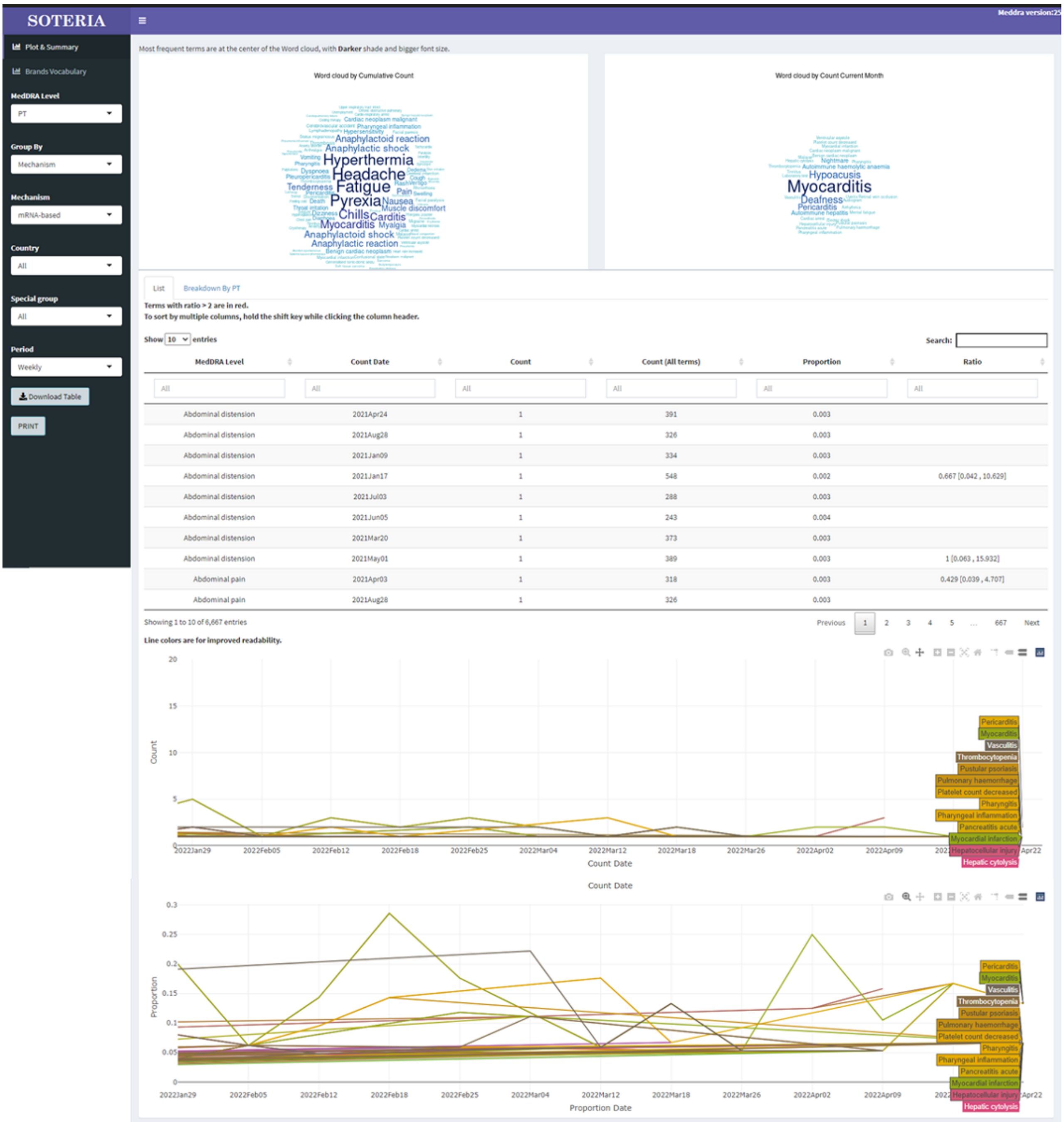


Figure 3. Soteria visual analytics. MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term.



Data Source

We obtained mentions of COVID-19 vaccine–related keywords from web content and social media via the API of the social listening tool from Synthesio Ltd using license-based data access via the API. Details of this tool can be obtained from Synthesio. Overall, mentions were collected from 190 countries in 61 different languages using a data query to identify mentions of a COVID-19 vaccine using COVID-19 vaccine–related keywords (Multimedia Appendix 4).

The social media types (as defined in the Soteria API) included are forums (excluding press releases), X, social networks, and comments and consumer opinions in the Soteria web app data stream via the Synthesio API. Adhering to X platform agreements and policies, Synthesio does not share the X posts via the API but only the X post ID. Using these X post IDs, we recovered the X post using the X API, a process at that time called "Twitter rehydration." Synthesio Ltd does not share a list of social media platforms and websites. Considering the data volume, an analysis of the social media platform and website was not performed. While Synthesio Ltd API was used to collect social media and web content data for the Soteria web app, this data stream can be easily replaced by a similar provider's API, X API, or a custom-built program to scrape web data (the authors do not recommend custom-built programs due to the nontrivial nature of the task in terms of technology, privacy, and compliance).

Data collection started on November 12, 2020, with automatic periodic weekly analysis of posts from the prior week and concatenation with all historic aggregated counts. In this analysis, results until April 2022 are presented (except for cumulative word counts, which are from October 2022).

Ethical Considerations

The data analyzed does not contain any personal information but only the text data that mention a vaccine based on our query. The processing is an in-memory process that is completed within 24 hours and text data are not retained. Only aggregated counts of AEs were retained. This type of analysis does not require an institutional review board or ethics committee review. Sanofi follows the General Data Protection Regulation (GDPR) and Personal Information Protection Law (PIPL) and similar country-related policies for data protection. Synthesio also follows GDPR and other country-related policies for data protection and adheres to platform agreements or policies of all platforms they query from. For data and API acquired via X, Sanofi adhered to the X developer agreement and policy.

Results

AE Detection Model

The AE detection model had precision=0.76, recall=0.82, and F_1-score=0.79 when evaluated on the test dataset (the AE Observation or Detection Model section provides details on the test dataset). This meant that 76% of the results were relevant and out of all positive predictions that could have been made, 82% were correct, resulting in 79% accuracy. Using the AE detection model, between November 2020 and December 2021, around 1 AE was observed at the MedDRA PT level for every 500 COVID-19 vaccine mentions and in 2022, this rate decreased to about 1 AE for every 2000 mentions. A considerably large portion of these AEs came from X data. Not every COVID-19 mention was associated with an AE.

AE Trends Analyses

Number of AEs by Country

The countries with most AEs observed were the United States (>15,000), United Kingdom (~5000), Italy (~2000), France (~2000), Australia (~2000), Japan (~1000), Singapore (~800), Philippines (~650), and Canada (~600; Figure 4). The number of AEs observed varied each week and were less frequent in 2022 compared with 2021 (the distribution as of April 2022 is shown in Figure 5).

Figure 4. Distribution of number of mentions with an AE (detected as MedDRA PT) between November 2020 and April 2022 by country. AE: adverse event; MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term.

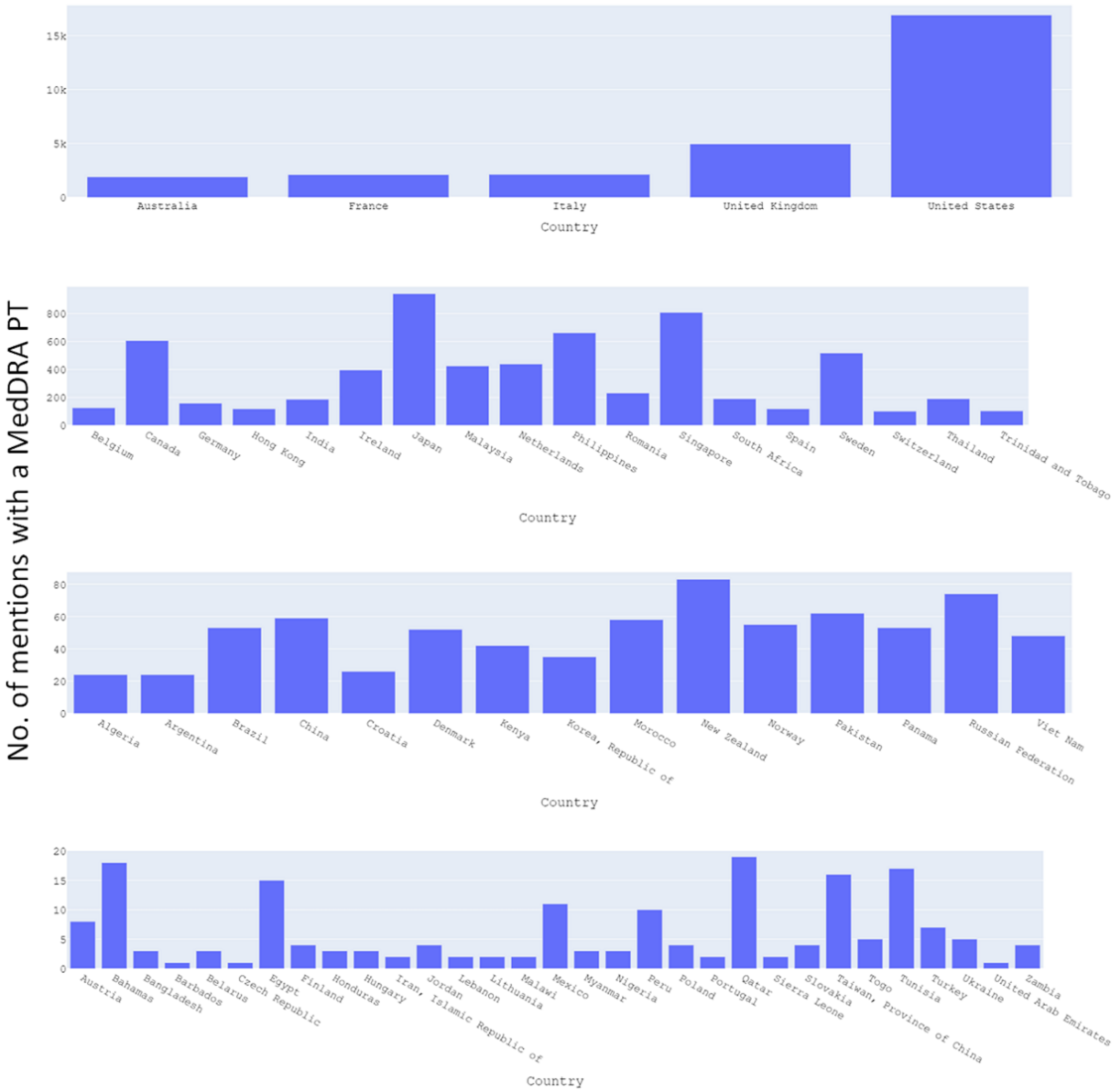
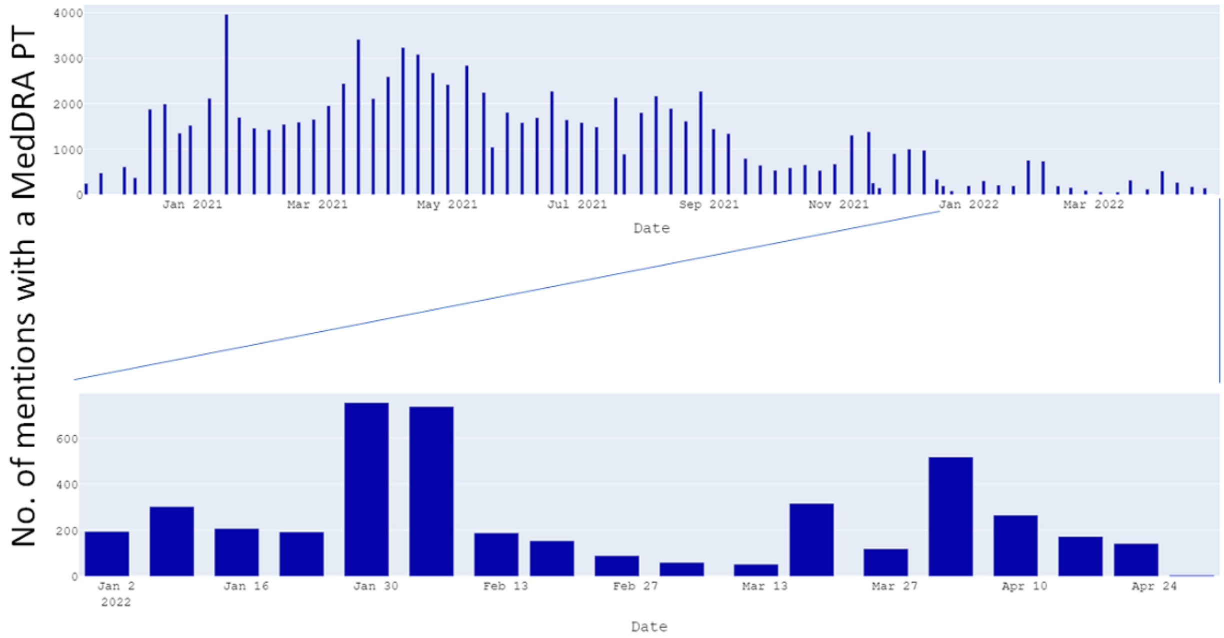


Figure 5. Distribution of number of mentions with an AE (detected as an MedDRA PT) between November 2020 and April 2022. AE: adverse event; MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term.



Number of AEs by Type of COVID-19 Vaccine (Platform and Brand)

The AEs observed were mostly related to mRNA vaccines and adenovirus vector vaccines while some AEs were related to inactivated virus vaccines and protein vaccines (Figure 6). AEs related to mRNA vaccines were first observed in December 2020 with a peak of 1400 in January 2021. AEs related to adenovirus vector vaccines increased from March 2020 with peaks of approximately 1200 in April and May 2021. AEs observed for mRNA vaccines stayed relatively high and stable during 2021, while in adenovirus vector vaccines' decreased in the second quarter and over the year (Figure 6).

Most AEs between November 2020 and April 2022 were associated with vaccines approved by the health authorities, that is, the mRNA vaccines from Pfizer-BioNTech and Moderna, and the adenovirus vector vaccines from AstraZeneca and Janssen (Johnson & Johnson). There were also many AEs where the administered vaccine names or pharmaceutical companies were not mentioned (Figure 7).

Figure 6. Distribution of number of mentions with an AE (detected as MedDRA PT) between November 2020 and April 2022 by the vaccine platform. AE: adverse event; MedDRA: Medical Dictionary for Regulatory Activities; mRNA: messenger ribonucleic acid; PT: preferred term.

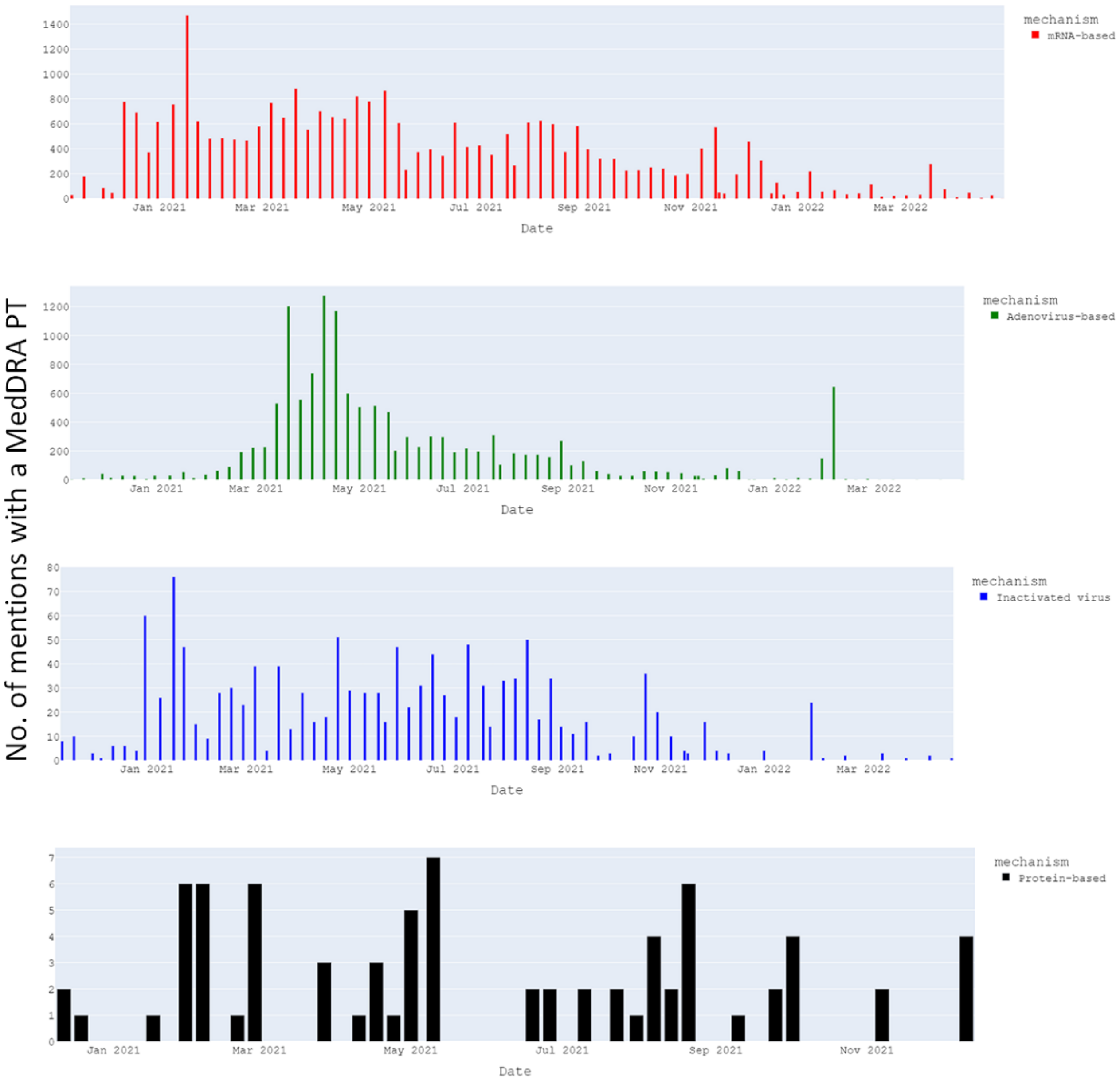
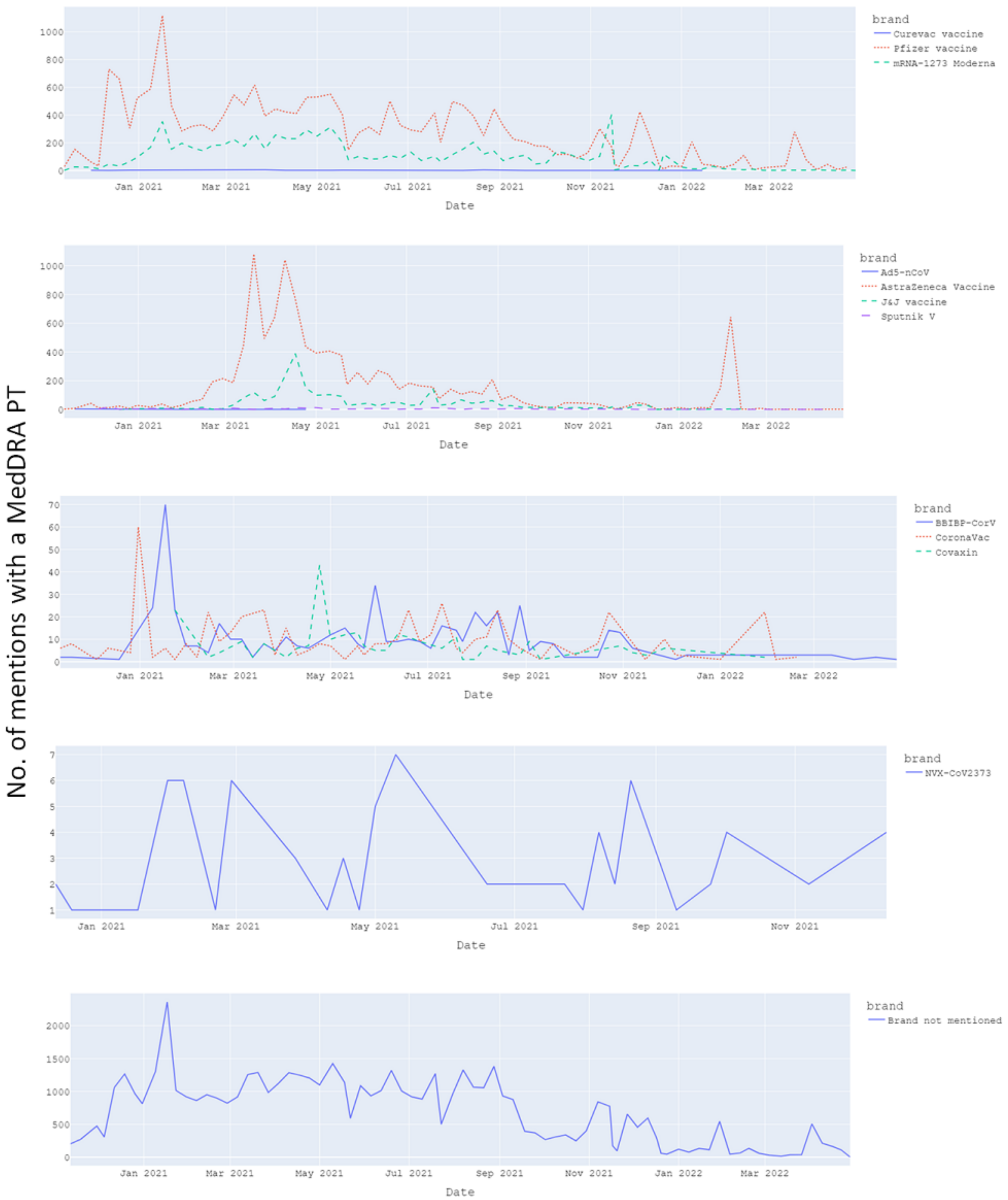


Figure 7. Distribution of number of mentions with an AE (detected as MedDRA PT between November 2020 and April 2022 by vaccine brand. AE: adverse event; J&J: Janssen (Johnson & Johnson); MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term; mRNA: messenger ribonucleic acid; NVX: Novovax.



Types of AEs

COVID-19 Vaccine Platform

The most frequently observed AEs for mRNA vaccines were headache, fatigue, pyrexia or hyperthermia, chills, nausea, pain or tenderness, and myalgia or muscle discomfort. The next most frequently observed AEs were those identified in the postmarketing setting—myocarditis and anaphylactic reactions (Multimedia Appendix 5).

The most commonly observed AE for adenovirus vector vaccines was thrombosis, which was also the AE identified in the postmarketing setting. The number of headaches and pyrexia, hyperthermia, or body temperature AEs all increased over time. Thrombocytopenia or platelet count AEs were associated with a specific syndrome “thrombosis with thrombocytopenia” and decreased over time (Multimedia Appendix 5).

COVID-19 Vaccine Brand

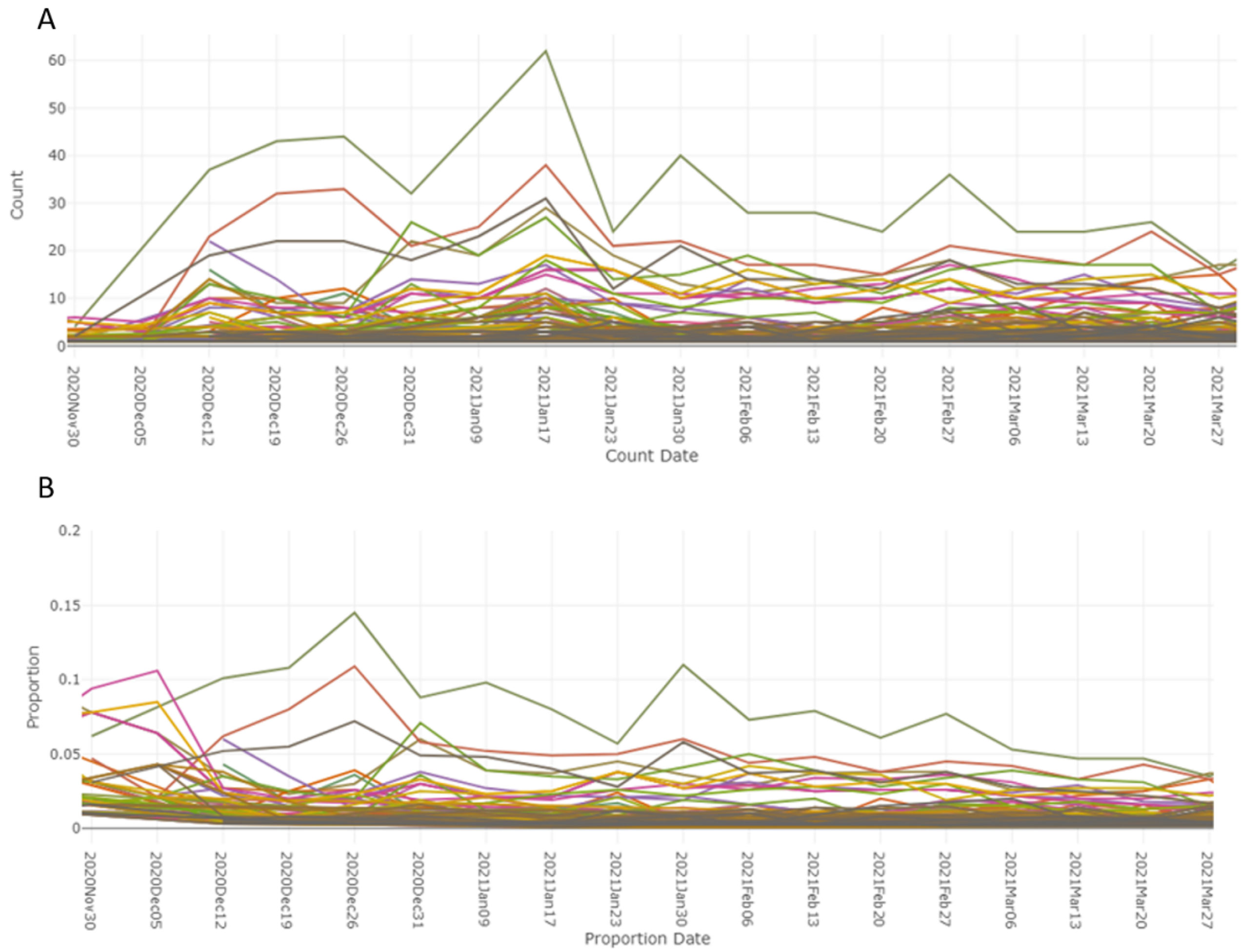
The AE “thrombosis” associated with the AstraZeneca vaccine (adenovirus vector) started to emerge in March 2021, and the word “COVID-19” related to the Sinopharm vaccine (inactivated virus) started to emerge in June 2021 (Multimedia Appendix 6).

The AEs observed associated with the Pfizer-BioNTech and Moderna vaccines (both mRNA) were similar, but a higher number of patients mentioned anaphylactic reactions regarding the Pfizer-BioNTech vaccine than the Moderna vaccine (Multimedia Appendix 7).

The AEs observed associated with AstraZeneca and Janssen vaccines (both adenovirus vectors) were different, with thrombosis being the most reported AE for the AstraZeneca vaccine and hyperthermia or pyrexia, fatigue, and headache followed by thrombosis and chills for the Janssen vaccine (Multimedia Appendix 7).

Anaphylactic reaction AEs started to trend in mid-December 2020 (under HLT “Anaphylactic and anaphylactoid responses”) shortly after the introduction of the mRNA vaccines (Figure 8).

Figure 8. AEs from social media of mRNA COVID-19 vaccines from the time of launch in December 2020. (A) Count and (B) proportion. AE: adverse event; mRNA: messenger ribonucleic acid.



Thrombo-embolic AEs started to trend on March 13, 2021, for adenovirus vector COVID-19 vaccines with HLTs including pulmonary thrombotic and embolic conditions, nonsite-specific embolism and thrombosis.

Myocarditis or pericarditis AEs started to trend for mRNA COVID-19 vaccines on May 1, 2021, with an increased ratio (proportion of noninfectious myocarditis HLT of the current period as compared to the previous one) on May 29, 2021. The proportion of myocarditis (count of myocarditis HLT divided by the count of all HLTs within the period) was the highest from September 2021 onwards (Multimedia Appendix 5).

Subgroup Analyses

Pediatric Population

The most frequently observed AEs as of April 2022 for the pediatric population were carditis, myocarditis, hyperthermia, fatigue, and headache, followed by chills, anaphylaxis, and tenderness (Multimedia Appendix 8).

Pregnant Women

The most frequently observed AEs as of April 2022 for the pregnant women population were hyperthermia, fatigue, pyrexia, and headache, followed by chills, nausea, tenderness, and spontaneous abortion (Multimedia Appendix 8).

Discussion

Principal Findings

Monitoring the web and social media provides opportunities to observe both AEs and patient concerns around the safety of COVID-19 vaccines. We developed an LM powered web app (Soteria) to analyze web content related to COVID-19 vaccine AEs.

Using the Soteria web app, we were able to observe AEs associated with COVID-19 vaccines by country, vaccine brand, and by MedDRA level (PT, HLT, HLGT, and SOC) using data from social media and web content. Because social media and web content data are readily available and can be accessed and analyzed quickly, the Soteria web app could observe AEs in real time, much faster than detection using traditional spontaneous reporting systems, which have a larger time gap between the occurrence of an AE and the availability of data for analysis.

Results from our analyses were, in general, consistent with those from other sources. For example, our study showed that the number of AEs observed for COVID-19 mRNA vaccines remained relatively high and stable during 2021, aligning with the first COVID-19 vaccination campaigns, while for COVID-19 adenovirus vector vaccines the number of AEs decreased over the year, again consistent with the decreasing use of these types of vaccines. Widely known AEs for specific COVID-19 vaccines were also consistent with our findings using the Soteria web app, such as the thrombosis related to the AstraZeneca vaccine that started to emerge from March 2021 [31]. Other AEs observed using the Soteria app (anaphylactic reaction, myocarditis, and thrombosis) were also consistent with those communicated by different Health Authorities shortly before or concurrently [32-34]. Over time, there were fewer mentions of AEs observed in social media, reflecting the reduction of COVID-19 mass vaccination campaigns.

When a new drug or a vaccine is released onto the market, the only safety concerns reported will be those arising during a clinical trial, which typically has limits due to the small patient numbers, is conducted over a limited time period and has a long list of patient exclusion criteria. It is, therefore, essential to capture AEs that occur after the trial in a real-world setting. Pharmacovigilance systems have been set up to capture this information, including the Eudravigilance reporting system in Europe and the CDC or FDA VAERS in the United States. It is mandatory for manufacturers to report any AEs, but voluntary for some health care professionals and the public. Some patients may not know how to report AEs or know about the systems in place for reporting AEs; therefore, social media is an important data source that can be used to harness social reaction and, with the use of LM to observe AEs, as in this analysis. In addition, social media data can provide information on AEs in real time without any filter or having to wait for reporting through standard AE systems, which can take at least 2 months. This is particularly important in situations such as the COVID-19 pandemic when new vaccines released onto the market may not have completed lengthy trials, which is when many AEs are discovered. In this analysis, the data were collected soon after publication, that is, collated every week automatically and the use of the MedDRA terms allowed the identification of AEs using the same terms as traditional VAERS and Eudravigilance reporting systems.

There can be bias in pharmacovigilance systems, as not all patients are aware of spontaneous pharmacovigilance reporting. However, social media is widely accessible, and patients discuss AEs, especially during a pandemic. This provides the opportunity to detect AEs from social media to augment the bias in spontaneous reporting. However, social media is not structured to capture AEs the way spontaneous reporting systems are and even with context-aware LMs, false positives and false negatives can occur. Previous analyses of social media have been undertaken for AEs relating to Zika, Ebola, and dengue viruses. There is a suggestion, however, that illnesses less prevalent in the news may be better for prediction as there is less influence and bias through media outlets [35]. Although, some rare illnesses with a smaller population would need sufficient social media comments for meaningful analysis. Studies using social media listening have tended to focus on vaccine hesitancy and sentiment [14-22]. The work presented here using the Soteria web app has some similarities to previous studies, especially the work from Portelli et al [22], which, as our analysis, used continuous data collection and processing, global data collection, and used a transformer-pretrained LM for AE detection. However, our Soteria web app differs in that it includes a multitude of social media listening, including patient blogs and forums beyond X, it also uses AE coding using ICH standards (MedDRA) rather than focusing on sentiment. AAs well, it does not focus only on English but uses translation models to be able to include mentions from 190 countries in 61 different languages. Finally, the generation of trends using multiple dimensions (vaccine brand, mechanism, country, different hierarchy levels MedDRA, and special populations: pediatrics or pregnant women) and combinations of these dimensions (eg, AEs into a pediatric group who received Pfizer-BioNTech vaccine) can be undertaken via the graphical user interface using visual analytics. For example, the top serious AEs reported via VAERS for children aged 5-11 years who received the Pfizer-BioNTech COVID-19 vaccine, according to a published study [32], were incorrect dose, vomiting, fever, and headache, and in a parallel study for children aged 12-17 years were dizziness, syncope, nausea, headache, and fever [33]. These AEs align with the AEs observed using the Soteria web app. In addition, preliminary findings of mRNA COVID-19 vaccine safety in pregnant women have shown that the most frequently reported pregnancy-related AE was spontaneous abortion, again aligning with the AE observed in this analysis [34].

It is important to note that these AEs are not necessarily safety signals, as safety signals have a very specific definition: "information on a new or known AE that may be caused by a medicine and requires further investigation" (EMA) [36]. In fact, 1 study showed that the current methods of signal detection using social media did not perform well and could not be used to replace or integrate with the current pharmacovigilance activities [5]. However, these AEs observed from social media can potentially be of importance to adjust signal detection, assessment strategies, and communication with external stakeholders. For example, AEs observed from social media could augment and optimize existing signal detection processes in place, or even become the focus of signal assessment that traditionally uses data sources such as electronic health records. AEs observed via the Soteria app have the potential to inform the companies earlier than those of traditional postmarketing AE reporting systems, therefore, allowing early and timely alerts of rare AEs.

Limitations

This study has limitations. The chatting habits are different between countries, for example, there was a higher availability of chats from the United States possibly leading to country bias. In addition, the frequency of chats may be affected by the media coverage within that country. As well, although social media and X reposts and reshares are only counted once, the same person may post the same AE several times, leading to over-reporting (although this can also occur within the VAERS system). Also, there may be false negatives due to incorrect translations (non-English sentences written using the standard English alphabet were removed as we were aware of poor performance with machine translation models at the time). Finally, these posts are not true diagnoses, and the people providing the chat may not experience the events or be aware of their medical diagnosis.

While BERT was the only available LM at the time of this work, this entire pipeline can be redone with novel LMs available today. Similarly, benchmark datasets for AE detection in social media are now available that can be used to measure the performance of the model. Further external validation of the model using these benchmark datasets could enhance the reliability of the model's performance claims.

Future studies could include an artificial intelligence–based signal detection (instead of AE detection) with validation using more traditional methods and more commonly used data sources, such as VAERS, claims, and electronic medical records databases. Similar tools could be developed to monitor the safe use of vaccines other than COVID-19 vaccines or drugs.

Of note, as of November 23, 2022, X has not enforced their COVID-19 misinformation policy. A comparison of data before and after this date could be of interest, for example, to determine if the removal of this policy influenced vaccine AEs discussed within X.

Conclusions

The application of LM to monitor web and social media data provides opportunities to observe AEs associated with COVID-19 vaccines faster compared to the traditional spontaneous reporting systems, which have a longer lag time between the occurrence of AEs and the availability of data. This gives the potential to enhance postmarketing surveillance. While AEs are not necessarily signals that require further analyses to confirm, they could help to adjust signal detection strategies by refocusing signal assessment on observed AEs and help to improve communication with external stakeholders, contributing to increased confidence in vaccines' monitoring and safety. While "chatting" regarding AEs following COVID-19 vaccination is decreasing in social media, our LM-based AE detection model can be applied to other vaccines and medicines.

Acknowledgments

The authors would like to thank Synthesio (New York), for use of their application programming interface (API)–based social listening tool, Corinne Jouquelet-Royer (Sanofi), Eng-Soon Chan (Sanofi), Prithvi Kamath (Sanofi), and Kiran Mahadeshwar (Sanofi), for their help with the paper, as well as the team at Medical Dictionary for Regulatory Activities (MedDRA). Medical writing assistance was provided by Ella Palmer, PhD, of inScience Communications, Springer Healthcare Ltd (London). This study and medical writing support for the preparation of this study were funded by Sanofi.

Data Availability

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

Authors' Contributions

CD, AK, AL-C, and JJ contributed to the concept or design of the study. CD, AK, YC, AL-C, AC-O-T, CM, and JJ contributed to acquisition of data. CD, AK, YC, LS, AL-C, and JJ contributed to analysis of data. CD, AK, LS, AL-C, CM, and JJ contributed to data interpretation. All authors read and approved the final version of the paper.

Conflicts of Interest

AK, LS, AL-C, AC-O-T, and JJ are employees of Sanofi and may hold shares and stock options in the company. CD, YC, and CM were employees of Sanofi at the time of the study and may have held shares and stock options in the company. CD was an employee of Alexion, AstraZeneca Rare Disease at the time of publication and may hold shares and stock options in the company. The opinions expressed are CD's own and not those of Alexion, AstraZeneca Rare Disease.

Multimedia Appendix 1

Languages available via Synthesio Ltd API and translated by Amazon Web Services translation or Helsinki-NLP model.
[DOCX File , 16 KB - infodemiology_v4i1e53424_app1.docx]

Multimedia Appendix 2
Vaccines monitored in Soteria app as of April 2022.
[DOCX File , 22 KB - infodemiology_v4i1e53424_app2.docx]

Multimedia Appendix 3
MedDRA lexical expansion.
[DOCX File , 197 KB - infodemiology_v4i1e53424_app3.docx]

Multimedia Appendix 4
COVID-19 vaccine related keywords used to query web content and social media (as of April 2022).
[DOCX File , 111 KB - infodemiology_v4i1e53424_app4.docx]

Multimedia Appendix 5
Word cloud by distribution of number of mentions with an AE (detected as MedDRA PT) between November 2020 and April 2022 by the vaccine platform. AE: adverse event; MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term.
[PNG File , 1072 KB - infodemiology_v4i1e53424_app5.png]

Multimedia Appendix 6
Word cloud by distribution of number of mentions with an AE (detected as MedDRA PT) between November 2020 and April 2022 by vaccine brand. AE: adverse event; MedDRA: Medical Dictionary for Regulatory Activities; PT: preferred term.
[PNG File , 1674 KB - infodemiology_v4i1e53424_app6.png]

Multimedia Appendix 7
Word cloud by cumulative count; selection criteria: PT, company. Extraction October 22, 2022. (A) Pfizer-BioNTech vaccine, (B) Moderna vaccine, (C) AstraZeneca vaccine, and (D) Janssen vaccine. PT: preferred term.
[PNG File , 748 KB - infodemiology_v4i1e53424_app7.png]

Multimedia Appendix 8
Word cloud by cumulative count. Selection criteria: PT, mRNA vaccines. Extraction April 2022. (A) Pediatric population and (B) pregnant women. mRNA: messenger ribonucleic acid; PT: preferred term.
[PNG File , 292 KB - infodemiology_v4i1e53424_app8.png]

References

1. IND application reporting: safety reports. US Food and Drug Administration (FDA). 2023. URL: https://www.fda.gov/drugs/investigational-new-drug-ind-application/ind-application-reporting-safety-reports [accessed 2024-11-20]
2. Signal management. European Medicines Agency. 2023. URL: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management [accessed 2024-11-20]
3. Gavrielov-Yusim N, Kürzinger ML, Nishikawa C, Pan C, Pouget J, Epstein LB, et al. Comparison of text processing methods in social media-based signal detection. Pharmacoepidemiol Drug Saf 2019;28(10):1309-1317. [doi: 10.1002/pds.4857] [Medline: 31392844]
4. Colilla S, Tov EY, Zhang L, Kurzinger ML, Tcherny-Lessenot S, Penfornis C, et al. Validation of new signal detection methods for web query log data compared to signal detection algorithms used with FAERS. Drug Saf 2017;40(5):399-408. [doi: 10.1007/s40264-017-0507-4] [Medline: 28155198]
5. Caster O, Dietrich J, Kürzinger ML, Lerch M, Maskell S, Norén GN, et al. Assessment of the utility of social media for broad-ranging statistical signal detection in pharmacovigilance: results from the WEB-RADR project. Drug Saf 2018;41(12):1355-1369 [FREE Full text] [doi: 10.1007/s40264-018-0699-2] [Medline: 30043385]
6. Kürzinger ML, Schück S, Texier N, Abdellaoui R, Faviez C, Pouget J, et al. Web-based signal detection using medical forums data in france: comparative analysis. J Med Internet Res 2018;20(11):e10466 [FREE Full text] [doi: 10.2196/10466] [Medline: 30459145]
7. Roosan D, Law AV, Roosan MR, Li Y. Artificial intelligent context-aware machine-learning tool to detect adverse drug events from social media platforms. J Med Toxicol 2022;18(4):311-320 [FREE Full text] [doi: 10.1007/s13181-022-00906-2] [Medline: 36097239]
8. Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, et al. Utilizing social media data for pharmacovigilance: a review. J Biomed Inform 2015;54:202-212 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.004] [Medline: 25720841]

9. Roche V, Robert JP, Salam H. AI-based approach for safety signals detection from social networks: application to the levothyrox scandal in 2017 on doctissimo forum. SSRN Electron J 2017;2022:36. [doi: 10.2139/ssrn.3993037]
10. Wang X, Wang X, Zhang S. Adverse reaction detection from social media based on Quantum Bi-LSTM with attention. IEEE Access 2023;11(99):1. [doi: 10.1109/access.2022.3151900]
11. Aronson JK. Artificial intelligence in pharmacovigilance: an introduction to terms, concepts, applications, and limitations. Drug Saf 2022;45(5):407-418. [doi: 10.1007/s40264-022-01156-5] [Medline: 35579806]
12. Huang JY, Lee WP, Lee KD. Predicting adverse drug reactions from social media posts: data balance, feature selection and deep learning. Healthcare (Basel) 2022;10(4):618 [FREE Full text] [doi: 10.3390/healthcare10040618] [Medline: 35455795]
13. Wang X, Huang W, Zhang S. Social media adverse drug reaction detection based on Bi-LSTM with multi-head attention mechanism. 2021 Presented at: Intelligent Computing Theories and Application: 17th International Conference, ICIC 2021, Shenzhen, China, August 12–15, 2021, Proceedings, Part III; August 12, 2021; Shenzhen, China p. 57-65. [doi: 10.1007/978-3-030-84532-2_6]
14. Hussain Z, Sheikh Z, Tahir A, Dashtipour K, Gogate M, Sheikh A, et al. Artificial intelligence-enabled social media analysis for pharmacovigilance of COVID-19 vaccinations in the United Kingdom: observational study. JMIR Public Health Surveill 2022;8(5):e32543 [FREE Full text] [doi: 10.2196/32543] [Medline: 35144240]
15. Yan C, Law M, Nguyen S, Cheung J, Kong J. Comparing public sentiment toward COVID-19 vaccines across Canadian cities: analysis of comments on Reddit. J Med Internet Res 2021;23(9):e32685 [FREE Full text] [doi: 10.2196/32685] [Medline: 34519654]
16. Kwok SWH, 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;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: 33886492]
17. Benis A, Chatsubi A, Levner E, Ashkenazi S. Change in threads on Twitter regarding influenza, vaccines, and vaccination during the COVID-19 pandemic: artificial intelligence-based infodemiology study. JMIR Infodemiology 2021;1(1):e31983 [FREE Full text] [doi: 10.2196/31983] [Medline: 34693212]
18. 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;7(12):e32814 [FREE Full text] [doi: 10.2196/32814] [Medline: 34665761]
19. Liew TM, Lee CS. Examining the utility of social media in COVID-19 vaccination: unsupervised learning of 672,133 Twitter posts. JMIR Public Health Surveill 2021;7(11):e29789 [FREE Full text] [doi: 10.2196/29789] [Medline: 34583316]
20. Muric G, Wu Y, Ferrara E. COVID-19 vaccine hesitancy on social media: building a public Twitter data set of antivaccine content, vaccine misinformation, and conspiracies. JMIR Public Health Surveill 2021;7(11):e30642 [FREE Full text] [doi: 10.2196/30642] [Medline: 34653016]
21. DeVerna MR, Pierri F, Truong BT, Bollenbacher J, Axelrod D, Loynes N, et al. CoVaxxy: a collection of English-language Twitter posts about COVID-19 vaccines. 2021 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; 2021 June 04; California p. 992-999. [doi: 10.1609/icwsm.v15i1.18122]
22. Portelli B, Scaboro S, Tonino R, Chersoni E, Santus E, Serra G. Monitoring user opinions and side effects on COVID-19 vaccines in the twittersphere: infodemiology study of tweets. J Med Internet Res 2022;24(5):e35115 [FREE Full text] [doi: 10.2196/35115] [Medline: 35446781]
23. Zhou W, Zhang S, Poon H, Chen M. Context-Faithful Prompting for Large Language Models. Singapore: Association for Computational Linguistic; 2023:14544-14556.
24. Tiedemann J, Thottingal S. OPUS-MT – Building Open Translation Services for the World. Sheffield, United Kingdom: European Association for Machine Translation; 2020:2020.
25. Gu Y, Tinn R, Cheng H, Lucas M, Usuyama N, Liu X, et al. Domain-specific language model pretraining for biomedical natural language processing. ACM Trans Comput Healthcare 2021;3(1):1-23. [doi: 10.1145/3458754]
26. Gurulingappa H, Rajput AM, Roberts A, Fluck J, Hofmann-Apitius M, Toldo L. Development of a benchmark corpus to support the automatic extraction of drug-related adverse effects from medical case reports. J Biomed Inform 2012;45(5):885-892 [FREE Full text] [doi: 10.1016/j.jbi.2012.04.008] [Medline: 22554702]
27. Zolnoori M, Fung KW, Patrick TB, Fontelo P, Kharrazi H, Faiola A, et al. The PsyTAR dataset: from patients generated narratives to a corpus of adverse drug events and effectiveness of psychiatric medications. Data Brief 2019;24:103838 [FREE Full text] [doi: 10.1016/j.dib.2019.103838] [Medline: 31065579]
28. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: an imperative style, high-performance deep learning library. arXiv 2019 [FREE Full text]
29. Beretta G, Marelli L. Fast-tracking development and regulatory approval of COVID-19 vaccines in the EU: a review of ethical implications. Bioethics 2023;37(5):498-507 [FREE Full text] [doi: 10.1111/bioe.13151] [Medline: 36905651]
30. Bodenreider O. The unified medical language system (UMLS): integrating biomedical terminology. Nucleic Acids Res 2004;32(Database issue):D267-D270 [FREE Full text] [doi: 10.1093/nar/gkh061] [Medline: 14681409]
31. Signal assessment report on embolic and thrombotic events (SMQ) with COVID-19 vaccine (ChAdOx1-S [recombinant]) – vaxzevria (previously COVID-19 vaccine astraZeneca) (other viral vaccines). European Medicines Agency. URL: https:/

/www.ema.europa.eu/en/documents/prac-recommendation/signal-assessment-report-embolic-thrombotic-events-smq-covid-19-vaccine-chadox1-s-recombinant_en.pdf [accessed 2024-11-20]

32. Hause AM, Baggs J, Marquez P, Myers TR, Gee J, Su JR, et al. COVID-19 vaccine safety in children aged 5-11 years - United States, November 3-December 19, 2021. MMWR Morb Mortal Wkly Rep 2021;70(5152):1755-1760 [FREE Full text] [doi: 10.15585/mmwr.mm705152a1] [Medline: 34968370]
33. Hause AM, Gee J, Baggs J, Abara WE, Marquez P, Thompson D, et al. COVID-19 vaccine safety in adolescents aged 12-17 years - United States, December 14, 2020-July 16, 2021. MMWR Morb Mortal Wkly Rep 2021;70(31):1053-1058 [FREE Full text] [doi: 10.15585/mmwr.mm7031e1] [Medline: 34351881]
34. Shimabukuro TT, Kim SY, Myers TR, Moro PL, Oduyebo T, Panagiotakopoulos L, CDC v-safe COVID-19 Pregnancy Registry Team. Preliminary findings of mRNA covid-19 vaccine safety in pregnant persons. N Engl J Med 2021;384(24):2273-2282 [FREE Full text] [doi: 10.1056/NEJMoa2104983] [Medline: 33882218]
35. Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. Annu Rev Public Health 2020;41:101-118 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094402] [Medline: 31905322]
36. Signal management. European Medicines Agency. URL: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management#:~:text=A%20safety%20signal%20is%20information,medicine%20and%20requires%20further%20investigation [accessed 2024-11-20]

Abbreviations

ADE: adverse drug events
AE: adverse event
API: application programming interface
CDC: Centers for Disease Control and Prevention
EMA: European Medicines Agency
FDA: US Food and Drug Administration
GDPR: General Data Protection Regulation
HLGT: high-level group term
HLT: high-level term
ICH: International Council for Harmonization
LM: language model
MedDRA: Medical Dictionary for Regulatory Activities
mRNA: messenger ribonucleic acid
NLP: natural language processing
PIPL: Personal Information Protection Law
PsyTAR: psychiatric treatment adverse reactions
PT: preferred term
SOC: system organ class
UMLS: Unified Medical Language System
VAERS: Vaccine Adverse Event Reporting System
WHO: World Health Organization

Edited by T Mackey; submitted 06.10.23; peer-reviewed by X Liu, K Liew, S Scaboro; comments to author 30.04.24; revised version received 03.06.24; accepted 08.10.24; published 20.12.24.

Please cite as:
Daluwatte C, Khromava A, Chen Y, Serradell L, Chabanon AL, Chan-Ou-Teung A, Molony C, Juhaeri J
Application of a Language Model Tool for COVID-19 Vaccine Adverse Event Monitoring Using Web and Social Media Content: Algorithm Development and Validation Study
JMIR Infodemiology 2024;4:e53424
URL: https://infodemiology.jmir.org/2024/1/e53424
doi:10.2196/53424
PMID:39705077

Original Paper

Navigating Awareness and Strategies to Support Dementia Advocacy on Social Media During World Alzheimer's Month: Infodemiology Study

Juanita-Dawne Bacsu[1], PhD; Sarah Anne Fraser[2], PhD; Ali Akbar Jamali[3], PhD; Christine Conanan[1], BSN; Alison L Chasteen[4], PhD; Shirin Vellani[5], PhD; Rory Gowda-Sookochoff[3], BCom; Corinne Berger[3], PhD; Jasmine C Mah[6], MD, MSc; Florriann Fehr[1], PhD; Anila Virani[1], BSN; Zahra Rahemi[7], PhD; Kate Nanson[1], BSN; Allison Cammer[3], PhD; Melissa K Andrew[6], PhD; Karl S Grewal[3], BSc (Hons); Katherine S McGilton[5], PhD; Samantha Lautrup[1]; Raymond J Spiteri[3], PhD

[1]Population Health and Aging Rural Research Centre, School of Nursing, Thompson Rivers University, Kamloops, BC, Canada
[2]Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada
[3]Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada
[4]Department of Psychology, University of Toronto, Toronto, ON, Canada
[5]Toronto Rehabilitation Institute, Knowledge, Innovation, Talent, Everywhere Research Institute, University Health Network, Toronto, ON, Canada
[6]Division of Geriatrics, Department of Medicine, Dalhousie University, Halifax, NS, Canada
[7]School of Nursing, Clemson University, Greenville, SC, United States

Corresponding Author:
Juanita-Dawne Bacsu, PhD
Population Health and Aging Rural Research Centre
School of Nursing
Thompson Rivers University
805 TRU Way
Kamloops, BC, V2C 0C8
Canada
Phone: 1 1 250 371 5538
Email: jbacsu@tru.ca

Abstract

Background: Understanding advocacy strategies is essential to improving dementia awareness, reducing stigma, supporting cognitive health promotion, and influencing policy to support people living with dementia. However, there is a dearth of evidence-based research on advocacy strategies used to support dementia awareness.

Objective: This study aimed to use posts from X (formerly known as Twitter) to understand dementia advocacy strategies during World Alzheimer's Awareness Month in September 2022.

Methods: Posts were scraped from X during World Alzheimer's Awareness Month from September 1, 2022, to September 30, 2022. After applying filters, 1981 relevant posts were analyzed using thematic analysis, and measures were taken to support trustworthiness and rigor.

Results: Our study revealed a variety of advocacy strategies, including sharing the voices of lived experience, targeting ethnic and cultural groups, myth-busting strategies, and political lobbying. Although a range of strategies were identified, further research is needed to examine advocacy strategies within different countries and political contexts. Furthermore, the impact of specific strategies on stigma reduction, cognitive health promotion, and policy change needs to be scientifically evaluated.

Conclusions: Our study offers valuable insight into strategies to bolster dementia advocacy and awareness campaigns to support people living with dementia. Findings from our research may provide critical insight for policymakers, organizations, and health professionals working to reduce dementia-related stigma and increase the uptake of risk-reduction activities to support the promotion of cognitive health.

(JMIR Infodemiology 2024;4:e63464) doi:10.2196/63464

KEYWORDS

dementia; Alzheimer disease; advocacy; stigma; myths; awareness; social media; political lobbying; lobbying; X; Twitter; tweet; thematic; promotion; campaign; geriatric; aging

Introduction

Dementia is an escalating global health concern. Over 55 million people worldwide live with dementia, with nearly 10 million new cases each year [1]. Dementia is the seventh leading cause of death and an important cause of disability [2]. Age is the greatest risk factor for dementia, and as the population ages, the number of people with dementia is expected to increase significantly. By 2050, it is estimated that approximately 153 million people will have dementia worldwide [3]. Although dementia is a common global health issue, stigma and myths about dementia are widespread.

Dementia-related stigma is pervasive and has been documented by numerous literature reviews [4-6]. Stigma refers to a characteristic or attribute that is socially discrediting (in this case, the diagnosis of dementia) and may consist of stereotyping, labeling, loss of status, and discrimination [7]. Studies show that dementia-related stigma often leads to social distancing [8], internalized shame and embarrassment [5], devaluing the lives of people with dementia [9], health care discrimination [10-12], and assumptions that people with dementia should be institutionalized [13]. Consequently, dementia-related stigma creates critical barriers to accessing support and receiving a diagnosis [14].

There are significant misconceptions and negative connotations surrounding dementia. For example, a study of 291 news media articles found derogative wording was often used to describe dementia [15]. Recently, an international survey reported that two-thirds of people erroneously believe that dementia is a normal part of aging [16]. The survey further revealed that 1 in 4 people believe there is nothing one can do about dementia [16]. However, research suggests that approximately 40% of all dementia cases may be prevented through modifiable risk factors [17,18]. Specifically, the World Health Organization published guidelines on 12 modifiable risk factors of dementia ranging from psychosocial factors (eg, social engagement and depression) to lifestyle factors (eg, diet and physical activity) [19]. This research sheds light on the need for more dementia advocacy and awareness to support the promotion of cognitive health.

Globally, organizations are striving to reduce stigma through dementia advocacy and awareness campaigns. For example, the World Health Organization has published numerous global reports advocating for the need to reduce stigma and increase dementia awareness [19,20]. In addition, Alzheimer's Disease International annually hosts the World Alzheimer's Day and Month awareness campaigns, as well as publishing annual reports. Despite these initiatives, there remains a paucity of evidence-informed research on specific strategies to support dementia advocacy and reduce dementia-related stigma [4]. This is concerning given that in two separate Dementia Priority Setting Partnerships, older adults living with dementia, care partners, and clinicians emphasized that the top research priorities were to assess effective strategies to target and reduce stigma and discrimination towards persons living with dementia [21].

There is a growing body of literature examining dementia-related stigma on social media [22-26]. For example, a study by Oscar et al [22] used machine learning to identify keywords and phrases commonly used on X (formerly known as Twitter) to perpetuate dementia-related stigma. Another study conducted a content analysis of dementia-related tweets and found stigmatization to be a main theme [26]. In a study of Dutch language tweets, Creten et al [23] identified that dementia-related ridicule contributed to the stigma of dementia on X. Although existing studies focus on documenting dementia-related stigma, there is a paucity of research focusing on specific actions to reduce dementia-related stigma. Consequently, Bartmess et al [25] described the need for further research to examine actions to reduce dementia-related stigma on social media.

Recently, we conducted a study on dementia discourse during Alzheimer's Awareness Month in Canada that identified educational, promotional, and fundraising tweets being used as actions to enhance dementia awareness [27]. However, this project also shed light on the need for research to examine awareness campaigns outside of Canada to identify advocacy strategies being used to reduce dementia-related stigma globally. Accordingly, the goal of this study is to identify dementia advocacy and awareness strategies used on X during World Alzheimer's Month in September 2022. Findings from our research may provide critical insight for policymakers, community leaders, and health professionals working to reduce dementia-related stigma among diverse population groups.

Methods

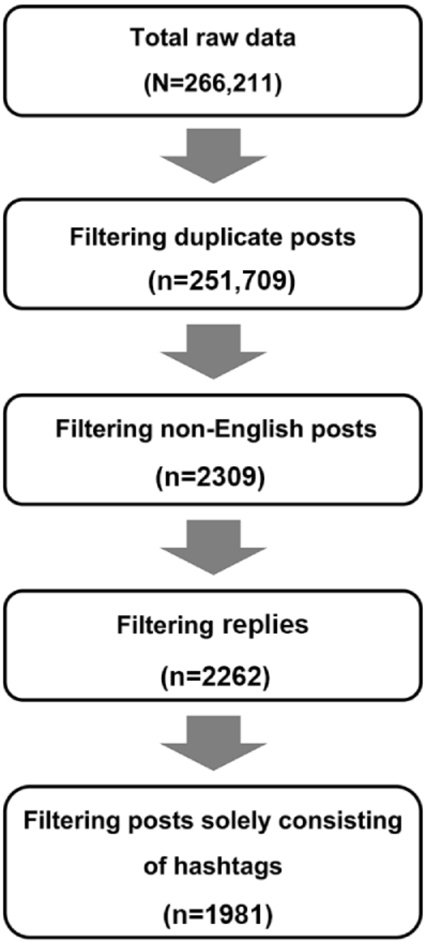
Data Collection

From the platform X, we purchased a developer account to acquire access to the official application programming interface (API). The API allows developers to access data from X and create content from the platform. We used the Python-based Tweepy library [28] and X's API to collect tweets posted from September 1, 2022, to September 30, 2022. The dates selected were based on World Alzheimer's Month, an international campaign organized by Alzheimer's Disease International that takes place every September.

Our search terms focused on examining social media discourse during World Alzheimer's Month. More specifically, our full set of search terms consisted of "World Alzheimer's Month" OR "World Alzheimers Month" OR "WorldAlzMonth" OR "World Alzheimers Awareness Month" OR "World Alzheimer's Awareness Month" OR "World Alzheimer's Day" OR "Dementia Awareness" OR "Alzheimer's Awareness."

Our initial search identified 266,211 posts. We removed the posts not written in English (14,502 non-English tweets), which left us with a total of 251,709 remaining tweets. In addition, we removed posts with duplicated content, posts solely consisting of hashtags, and reply posts. We also removed replies because they are often missing important information and contextual content. The data filtering process is documented in Figure 1. The remaining 1981 posts were used for thematic analysis.

Figure 1. X (formerly known as Twitter) data filtering process.



Data Analysis

We analyzed the posts by conducting a 5-stage thematic analysis process outlined by Braun and Clarke [29,30]. This process encompassed (1) reviewing the data, (2) developing initial codes through comprehensive data examination, (3) identifying preliminary themes, (4) refining these themes, and (5) assigning labels to the themes [31]. To analyze the data, we used both inductive (originating from the data, bottom-up approach) and deductive (theory-driven, top-down approach) techniques. First, an initial coding team was formed with 5 of our coauthors who had diverse interdisciplinary expertise and backgrounds, including psychology, computer science, nursing, and population health. This team was formed to develop and pilot test a codebook that was created by reading and reviewing posts (ie, data review and familiarization) to develop an initial list of the codes. To begin the coding process, an inductive bottom-up (data-driven) approach was used, where the researchers created the codes as they analyzed the data. For the pilot test, the coding team independently coded the posts and then compared their codes as a team. A final codebook was created which consisted of 9 codes including (1) stopping stigma, (2) sponsorship (fundraising), (3) simple or basic posts (no educational content), (4) promotional (marketing), (5) lived experiences (family, friend, or person with dementia), (6) fostering stigma, (7) resources (supports), (8) informative (educational messages), and (9) exploitation (capitalistic and political posts). In addition, 2 of our codes (promotional and exploitation) may appear to be similar, so we have provided a description to help differentiate the codes. For example, the promotional code refers to tweets promoting activities, events, conferences, webinars, websites, or informational materials. In contrast, the exploitation code refers to tweets aiming to make a profit from Alzheimer's Awareness Month for capitalistic gains, such as moving companies, assisted-living homes, or drug companies.

The full team of coauthors participated in practice coding activities using the finalized codebook. The first activity consisted of the trainee practice coding exercise, designed as an introductory task for participants with minimal coding experience. The primary objective was to familiarize trainees with the basic concept of coding and the use of the codebook and its codes. The trainees collectively coded 50 posts. The second practice involved independently coding 75 posts and then comparing individual codes to an answer sheet developed by the initial coding team. The third practice activity involved a team coding exercise in which the full group collectively coded and discussed any coding uncertainties, challenges, questions, or disagreements for 75 posts. After the practice coding was completed, 9 pairs of coauthors (ie, 18 coders) independently coded the same posts (approximately 208 posts per coder) and then met to compare codes and determine the final code for each post through discussion and pair consensus. Any coding discrepancies were resolved through group consensus.

After the coding, the research team used an inductive strategy to sort through the data, aiming to identify prevalent themes by developing a thematic map by grouping the codes into themes. Subsequently, the entire team engaged in a thorough review of the data to ascertain that all relevant themes were captured, clearly articulated, and unique from the other themes [29]. The final stage involved collective discussions within the team to review the clarity and conciseness of each theme name.

Trustworthiness and Rigor

The criteria created by Lincoln and Guba [32] for trustworthiness were applied in this study to ensure credibility, dependability, and confirmability from data collection to analysis. Credibility was conducted using peer debriefing, member checking, and triangulation. Specifically, credibility was applied through peer debriefing by facilitating practice coding exercises to ensure comprehension of the codes [32]. Furthermore, a codebook was developed to outline the code definitions, keywords, and Twitter examples. This codebook helped to ensure a systematic approach to the coding of the posts. In addition, member checking was implemented to evaluate intercoder reliability by independently evaluating each post by a pair of coders to cross-check the interpretation of the posts [33]. We determined the intercoder reliability by calculating the percentage of agreement between each of the 9 different pairs of coders (ie, 2 coders for each post) and then taking each pair's percentage of agreement numbers to calculate the overall group average, which was 80% agreement across the coded tweets (Table 1). Less experienced coders were paired with senior coders to demonstrate varying levels of experience with qualitative coding [34]. Researcher triangulation was achieved through the multidisciplinary expertise of the research team [35], composed of experts from various fields such as nursing, nutrition, psychology, public health, gerontology, computer science, and geriatrics. To establish dependability and confirmability, the research process was clearly documented with a comprehensive audit trail [36]. The audit trail was maintained by keeping a record of the research methods and tools, such as the search terms (more details in the Data Collection section), the filtering process, and the thematic map. Regular team meetings with peer debriefing were instrumental in all steps of the research processes, ranging from practice coding activities to theme development.

Table 1. Intercoder reliability (weighted percentages).

Coding partners for thematic analysis	Percentage of agreement (%)
KSG and CC	88
RJS and CB	94
SAF and AAJ	52
RG-S and JDB	93
AC and KN	67
FF and KSM	92
MKA and ZR	78
ALC and SV	96
JCM and SL	59
Intercoder reliability average	80

Ethical Considerations

This study did not require ethics approval because it was based solely on anonymized, secondary data collected from publicly available tweets, which are not classified as human participant research [31,37]. We omitted all usernames and handles before any data analysis to help protect the anonymity of the users. However, we did not anonymize tweet content regarding celebrities, politicians, and national dementia advocates. For example, a new report on ethics and social media research published at the University of Aberdeen notes that celebrities, politicians, and public figures seeking to share their messages as widely as possible do not require anonymity [38]. Accordingly, we did not anonymize posts about celebrity champions (celebrities or athletes), politicians, or nationally known dementia advocates.

Results

Based on our thematic analysis, 4 main social media advocacy strategies were identified, including (1) sharing the voices of lived experience, (2) targeting ethnic and cultural groups, (3) myth-busting strategies, and (4) political lobbying.

Strategy 1: Sharing the Voices of Lived Experience

A predominant strategy used for dementia advocacy focused on sharing the voices of lived experience, including people living with dementia and their family care partners. For example, this strategy called on people living with dementia to spread awareness and often shared insights to support dementia advocacy and leadership actions. This strategy is highlighted in the following posts:

> *Share your story with us this #worldalzheimersmonth*

> *VOICES OF LIVED EXPERIENCE: Author & advocate Myrna Norman shares her insights about how more people living with dementia might be engaged in leadership & research opportunities. Read Advocacy in Action.*

World Alzheimer’s Month also enabled people living with dementia to share their lived experiences to improve the general public’s understanding of dementia. These experiences challenged negative stereotypes in describing that it is possible to live well with dementia. Lived experiences often included sharing personal challenges as well as support to enhance the quality of life, ranging from technology to an optimistic mindset.

A #WorldAlzheimersMonth message of support from Patient & Family Advisory Council member:

> *I am a person living with vascular dementia for 10+ years. I live independently & I live well. Some days can be challenging, but I use technologies to help.*

> *After living with dementia for almost 5 years, I came to realize my life can be good whenever people that I work with practice what I called the Four As - Acknowledge, Accept, Accommodate, and Adapt.*

Many celebrity champions (including famous Olympians and actors) used their celebrity status and fame to help build a platform to raise dementia awareness. Examples of celebrity champions included Mariah Bell, Natalie Morales, Selenis Leyva, Izzy Diaz, Oscar Nunez, Seth Rogen, Lauren Miller, Vicky McClure, and Peter Gallagher, who shared their family members' stories and their own experiences as care partners in raising awareness and advocate for people living with dementia.

Actor Peter Gallagher shares his personal experience with Alzheimer disease. As the awareness month nears its end, we should always remember to fight the stigma surrounding Alzheimer’s and raise awareness.

> *#ENDALZ Celebrity Champion and #TheOffice star Oscar Nunez raises Alzheimer's awareness for his dad and grandmother. Thank you, Oscar, for making a difference in the fight to end Alzheimer's! #HispanicHeritageMonth.*

Researchers also shared stories and encouraged people to support research and awareness of people living with dementia.

> *As #World Alzheimer's Month draws to a close, we are sharing this video of inspirational Research Champion who explains why encouraging others to take part in dementia research is so important to her.*

Strategy 2: Targeting Ethnic and Cultural Group Awareness

Another strategy to enhance dementia advocacy was to specifically target different cultural, ethnic, and linguistic groups. This took the form of raising awareness, highlighting disparities, and using creative hashtags to increase the visibility of the posts. For example, awareness materials and activities in various languages were showcased, reaching and engaging individuals from various ethnic and cultural backgrounds. A wide range of materials also included culturally tailored content, tips, and ideas to raise awareness in diverse communities. The following posts demonstrated ethnically and culturally targeted awareness strategies:

> *Thank you to all friends and colleagues in Greece for translating the 12 risk factors for #dementia #alzheimers in Greek efkaristo' poli' #worldalzmonth if you are Greek please share!*

> *It’s September It's World Alzheimer's Month These are our list of activities in mix culture – languages [including Dutch, English, and Bahasa Indonesia]… @alzi_indonesia @AlzDisInt*

> *In honor of #WorldAlzheimersMonth we would like to highlight Dr. Tiffany Chow, who shares her insights on the prevalence of Alzheimer's disease among the AANHPI (Asian American, Native Hawaiian & Pacific Islander) community.*

Efforts to enhance dementia awareness also included highlighting existing disparities between different cultural and ethnic groups in terms of dementia incidence and access to care.

> *Older African Americans are twice as likely to have Alzheimer’s or other dementia as older white Americans, while older Hispanic adults are 1.5 times as likely to have Alzheimer's. Awareness and education are vital.*

Posts also incorporated the usage of hashtags to target specific cultural and ethnic groups, such as “#HispanicHeritageMonth.” In addition, partnerships with cultural organizations were also used to promote awareness. These hashtags and partnership strategies are highlighted in the following posts:

> *We are especially grateful for the wonderful partnership of @ProgresoLatino in helping to raise Alzheimer's awareness among the Latino community in our area. Together, we are making a difference. #HispanicHeritageMonth #AlzRI 5/5*

> *Happening today 9/24 at Noon #WorldAlzheimersMonth #HispanicHeritageMonth #ENDAlz*

As such, culturally and ethnically targeted awareness materials and tips increased the accessibility of knowledge and available support and may resonate better within different groups.

Strategy 3: Myth Busting

Another dementia advocacy strategy was to address misinformation about dementia. Most examples included educational-style lessons or posts alluding to evidence or quotes from research related to treatment options or symptoms of Alzheimer disease. This myth-busting related to education is demonstrated in the following posts:

Many people still wrongly believe that dementia is a normal aging. This alone highlights how important public awareness campaigns, like World Alzheimer’s Month, are for changing perceptions and increasing public knowledge around dementia and Alzheimer disease.

> *This year the theme for World Alzheimer’s Month is “Know Dementia, Know Alzheimer’s”. Dementia is an umbrella term for a collection of symptoms caused*

by disorders affecting the brain & impacting memory, thinking, behaviour & emotion. The most common is Alzheimer's disease.

Many tweets focused on combating myths and false assumptions about dementia. For example, these tweets would often challenge dementia myths by providing facts, information, and statistics about dementia. The following are related examples:

Myth - dementia does not just affect older adults! Age is a risk factor - but in rare cases, it affects younger adults too... Scientists estimate that 260 in 100,000 people (aged 30-64) develop early onset dementia!

Myth - there are no treatments available for people with Alzheimer's disease! There has been lots of progress on treatments for people with Alzheimer's, including medications and coping strategies to help manage behavioral symptoms.

Strategy 4: Political Lobbying

Political lobbying of governments was also used to increase dementia advocacy. This strategy often included lobbying members of the government to make dementia a national priority. For example, this strategy included targeted political advocacy to support the development of national dementia plans for countries without any national plan. This political lobbying for national dementia plans is highlighted in the following posts:

When countries have a national #dementia plan, they are more likely to have robust strategies for post-diagnostic support in place. During #WorldAlzMonth, #KnowDementia & #KnowAlzheimers better by reading about the importance of national dementia plans.

…EmmanuelMacron France now does not have a dedicated national dementia plan.

Political lobbying also included advocating for increased government support and funding to enhance dementia research. For example, posts highlighted the need for larger clinical trials with diverse populations to support dementia research. This advocacy for government funding to support dementia research is demonstrated in the following posts:

A call for the government to reduce #health #inequalities in #dementia - from prioritizing funding for scientific innovation, through to facilitating larger and more diverse dementia trials

Will you lend your voice for research? This #WorldAlzMonth, we're asking MPs across the UK to write to @trussliz [the Prime Minister of the United Kingdom in 2022], calling for people living with dementia & research… to remain a political priority. Write to your MP now…

Discussion

Principal Findings

Our study examined advocacy strategies being used during World Alzheimer's Awareness Month. This study is among the first to examine dementia advocacy used in a global awareness campaign on social media. Understanding advocacy strategies is essential to improving dementia awareness, reducing stigma, supporting cognitive health promotion, and improving policymaking for people living with this condition and their care partners. Our study revealed a variety of advocacy strategies, including sharing the voices of lived experience, targeting diverse ethnic and cultural groups, myth-busting strategies, and political lobbying.

Our study identified that political lobbying was a key strategy used to conduct dementia advocacy with governments. However, it is important to note that the political context differs by country and often requires different advocacy efforts. For example, our study found that some countries without national dementia strategies (eg, France) tended to focus on the issue of advocacy by endorsing their governments to develop a national dementia plan, whereas other lobbying initiatives focused on financial advocacy and lobbying governments for increased funding and resources to support dementia research. Consequently, further research is needed to examine the political climate and policy-related issues that are relevant to supporting dementia advocacy within different countries, jurisdictions, and political contexts.

Other strategies identified in our study involved myth-busting, targeting different ethnic and cultural groups, and sharing lived experiences. These strategies focused on increasing dementia awareness and education among the public, including diverse groups. Seetharaman and Chaudhury [39] note that sharing lived experiences of dementia helps to challenge common stereotypes and misconceptions about dementia. In particular, this study identified celebrities (including famous athletes, Olympians, and actors) and national dementia advocates sharing their voices of lived experience to challenge the stigma of dementia. This finding contrasts with our previous research [27] that reported a lack of lived experience and celebrity advocates during Alzheimer's Awareness Month in Canada. Through celebrity endorsement and the voices of lived experience, people living with dementia and care partners can formulate a collective identity to bring about social change and networking opportunities to challenge dementia-related stigma [40]. Consequently, Weetch et al [41] note that dementia advocacy may offer well-being benefits for those who engage through the activity itself and the development of new social networks.

Similar to existing research in mental health [31] and dementia [27], our study found several cursory posts that made basic reference to the awareness campaign. Although these posts supported basic awareness, they also shed light on the opportunity for more targeted advocacy strategies to enhance dementia education, counter misconceptions, and reduce dementia-related stigma [27]. Consequently, having more targeted posts may decrease the risk of having posts being disregarded and perceived as pointless.

Limitations

Although our study was conducted in a comprehensive manner, it is not without limitations. More specifically, our study focused solely on textual data and excluded images and photos on X. However, this exclusion limits our study's ability to draw conclusions based on image content to inform dementia advocacy and awareness-raising efforts. For example, a recent study suggests that a post with an image contributes to higher user engagement on social media [42]. This research suggests that posts with high-quality, professional images and human faces are more likely to lead to higher user engagement on the X platform [42]. Further research is needed to examine the impact of images on user engagement during World Alzheimer's Awareness Month. This research on image content may provide critical insight to inform novel interventions to raise awareness.

In addition, our study did not include any evaluation to assess the impact of the different advocacy strategies. Although a range of strategies was identified, further research is needed to evaluate the impact of the specific strategies on supporting dementia advocacy. For example, research is needed to examine whether these social media strategies effectively work to reduce dementia-related stigma. Moreover, studies are required to assess whether dementia advocacy on social media advocacy improves dementia awareness and advances cognitive health promotion. In addition, further research is needed to examine whether dementia advocacy on social media contributes to and/or results in policy and program change to support people living with dementia.

Another limitation is that our data were restricted to posts in the English language. This limitation is particularly relevant given that approximately 60% of people with dementia live in low- and middle-income countries [43], where English may not be the predominant language. Consequently, this language restriction may have overlooked novel strategies employed to target different cultural, linguistic, and ethnic groups. Accordingly, further research is needed to examine dementia advocacy strategies used in non-English-speaking countries.

Conclusions

Our study examined dementia advocacy strategies used during World Alzheimer's Awareness Month in September 2022. Understanding advocacy strategies is vital to reducing dementia-related stigma, countering myths, and supporting the promotion of cognitive health. Furthermore, advocacy is important to ensuring that dementia care is prioritized and supported by governments. By examining different advocacy strategies, our study offers valuable insight into ways to bolster dementia advocacy and awareness campaigns to support people living with dementia.

Our study identified a variety of advocacy strategies ranging from myth-busting to political lobbying. Although a range of strategies was identified, further research is needed to examine advocacy strategies within different countries and political contexts. Furthermore, evaluation research is needed to assess the impact of specific strategies on stigma reduction, cognitive health promotion, and policy change. Findings from our study may provide critical insight for policymakers, researchers, and health professionals working to reduce dementia-related stigma and increase the uptake of risk-reduction activities to support cognitive health promotion.

Acknowledgments

This work was supported by the Canada Research Chairs Program (grant 451097). This work was also supported by Team 15 in the Canadian Consortium on Neurodegeneration in Aging (CCNA). CCNA is supported by a grant from the Canadian Institutes of Health Research with funding from several partners, including the Saskatchewan Health Research Foundation, the Centre for Aging and Brain Health, and the Alzheimer Society of Canada. Funding was provided by the Open Access Publishing Grant and the Undergraduate Research Apprenticeship Fund at Thompson Rivers University.

Data Availability

The data sets generated during and/or analyzed during this study are not publicly available but are available from the corresponding author on reasonable request.

Authors' Contributions

JDB and RJS conceived the idea for the study. JDB, AAJ, and RJS devised the X scraping approach, and AAJ scraped for tweets. JDB, RG-S, AAJ, CB, and SAF devised the codebook. All of the coauthors coded the posts, contributed example tweets, and participated in the thematic analysis. JDB wrote the first draft of the manuscript with theme-writing contributions from CC, SV, KSM, JCM, AV, SAF, ALC, AC, ZR, FF, KN, and CB. All authors reviewed the final manuscript.

Conflicts of Interest

MKA received grants from Sanofi, GSK, Prizer, Merck, Icosavac; all unrelated to the present study.

References

1. Fact sheets: Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2024-05-22]
2. The top 10 causes of death. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [accessed 2024-06-05]
3. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. Lancet Public Health 2022;7(2):e105-e125 [FREE Full text] [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]

4. Bacsu JD, Johnson S, O'Connell ME, Viger M, Muhajarine N, Hackett P, et al. Stigma reduction interventions of dementia: a scoping review. Can J Aging 2022;41(2):203-213. [doi: 10.1017/S0714980821000192] [Medline: 34253273]
5. Philip G, Savundranayagam MY, Kothari A, Orange JB. Exploring stigmatizing perceptions of dementia among racialized groups living in the anglosphere: a scoping review. Aging and Health Research 2024;4(1):100170 [FREE Full text] [doi: 10.1016/j.ahr.2023.100170]
6. Werner P, Vermeulen P, Van Gorp B, Simonsen P. Dementia and Society: From History to Intervention. Cambridge, UK: Cambridge University Press eBooks; 2022.
7. World Alzheimer report 2012. Alzheimer Disease International. URL: https://www.alz.co.uk/research/WorldAlzheimerReport2012.pdf [accessed 2024-05-25]
8. Stites SD, Rubright JD, Karlawish J. What features of stigma do the public most commonly attribute to Alzheimer's disease dementia? Results of a survey of the U.S. general public. Alzheimers Dement 2018 Jul;14(7):925-932 [FREE Full text] [doi: 10.1016/j.jalz.2018.01.006] [Medline: 29602733]
9. Bacsu JD, Fraser S, Chasteen AL, Cammer A, Grewal KS, Bechard LE, et al. Using Twitter to examine stigma against people with dementia during COVID-19: infodemiology study. JMIR Aging 2022 Mar 31;5(1):e35677 [FREE Full text] [doi: 10.2196/35677] [Medline: 35290197]
10. Bacsu JR, O'Connell ME, Wighton MB. Improving the health equity and the human rights of Canadians with dementia through a social determinants approach: a call to action in the COVID-19 pandemic. Can J Public Health 2022 Apr;113(2):204-208 [FREE Full text] [doi: 10.17269/s41997-022-00618-8] [Medline: 35239172]
11. Ménard A, O'Sullivan T, Mulvey M, Belanger C, Fraser S. Perceptions of hospital care for persons with dementia during the COVID-19 Pandemic: a social media sentiment analysis. Gerontologist 2024 Jul 01;64(7):gnad155. [doi: 10.1093/geront/gnad155] [Medline: 37943714]
12. Werner P, Giveon SM. Discriminatory behavior of family physicians toward a person with Alzheimer's disease. Int Psychogeriatr 2008 Aug;20(4):824-839. [doi: 10.1017/S1041610208007060] [Medline: 18341751]
13. Werner P. Discriminatory behavior towards a person with Alzheimer's disease: examining the effects of being in a nursing home. Aging Ment Health 2008 Nov;12(6):786-794. [doi: 10.1080/13607860802380649] [Medline: 19023730]
14. Bacsu J, Mateen FJ, Johnson S, Viger MD, Hackett P. Improving dementia care among family physicians: from stigma to evidence-informed knowledge. Can Geriatr J 2020 Dec;23(4):340-343 [FREE Full text] [doi: 10.5770/cgj.23.426] [Medline: 33282053]
15. Sm-Rahman A, Lo CH, Jahan Y. Dementia in media coverage: a comparative analysis of two online newspapers across time. Int J Environ Res Public Health 2021 Oct 08;18(19):10539 [FREE Full text] [doi: 10.3390/ijerph181910539] [Medline: 34639840]
16. World Alzheimer report 2019. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2019/ [accessed 2019-06-24]
17. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec 16;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
18. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
19. Global status report on the public health response to dementia. World Health Organization. URL: https://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=1962&context=srhreports [accessed 2024-06-05]
20. Global action plan on the public health response to dementia: 2017-2025. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/259615/?sequence=1 [accessed 2024-06-07]
21. Bethell J, Commisso E, Rostad HM, Puts M, Babineau J, Grinbergs-Saull A, et al. Patient engagement in research related to dementia: a scoping review. Dementia (London) 2018 Nov;17(8):944-975. [doi: 10.1177/1471301218789292] [Medline: 30373460]
22. Oscar N, Fox PA, Croucher R, Wernick R, Keune J, Hooker K. Machine learning, sentiment analysis, and tweets: an examination of Alzheimer's disease stigma on Twitter. J Gerontol B Psychol Sci Soc Sci 2017 Sep 01;72(5):742-751. [doi: 10.1093/geronb/gbx014] [Medline: 28329835]
23. Creten S, Heynderickx P, Dieltjens S. The stigma toward dementia on Twitter: a sentiment analysis of Dutch language tweets. J Health Commun 2022 Oct 03;27(10):697-705. [doi: 10.1080/10810730.2022.2149904] [Medline: 36519829]
24. Alvarez-Mon MA, Llavero-Valero M, Sánchez-Bayona R, Pereira-Sanchez V, Vallejo-Valdivielso M, Monserrat J, et al. Areas of interest and stigmatic attitudes of the general public in five relevant medical conditions: thematic and quantitative analysis using Twitter. J Med Internet Res 2019 May 28;21(5):e14110 [FREE Full text] [doi: 10.2196/14110] [Medline: 31140438]
25. Bartmess M, Talbot C, O'Dwyer ST, Lopez RP, Rose KM, Anderson JG. Using Twitter to understand perspectives and experiences of dementia and caregiving at the beginning of the COVID-19 pandemic. Dementia (London) 2022 Jul;21(5):1734-1752 [FREE Full text] [doi: 10.1177/14713012221096982] [Medline: 35549466]
26. Cheng TY, Liu L, Woo BK. Analyzing Twitter as a platform for Alzheimer-related dementia awareness: thematic analyses of tweets. JMIR Aging 2018 Dec 10;1(2):e11542 [FREE Full text] [doi: 10.2196/11542] [Medline: 31518232]

27. Bacsu JD, Cammer A, Ahmadi S, Azizi M, Grewal KS, Green S, et al. Examining the Twitter discourse on dementia during Alzheimer's Awareness Month in Canada: infodemiology study. JMIR Form Res 2022 Oct 26;6(10):e40049 [FREE Full text] [doi: 10.2196/40049] [Medline: 36287605]
28. Tweepy. URL: https://www.tweepy.org/ [accessed 2024-11-04]
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Braun V, Clarke V. Thematic Analysis: A Practical Guide. New York, NY: Sage Publications; 2022.
31. Makita M, Mas-Bleda A, Morris S, Thelwall M. Mental health discourses on Twitter during Mental Health Awareness Week. Issues Ment Health Nurs 2021 May;42(5):437-450. [doi: 10.1080/01612840.2020.1814914] [Medline: 32926796]
32. Lincoln Y, Guba EG. Naturalistic Inquiry. New York, NY: SAGE Publications; 1985.
33. O'Connor C, Joffe H. Intercoder reliability in qualitative research: debates and practical guidelines. Int J Qual Methods 2020 Jan 22;19:160940691989922 [FREE Full text] [doi: 10.1177/1609406919899220]
34. Giesen L, Roeser A. Structuring a team-based approach to coding qualitative data. Int J Qual Methods 2020;19:160940692096870 [FREE Full text] [doi: 10.1177/1609406920968700]
35. Raskind IG, Shelton RC, Comeau DL, Cooper HLF, Griffith DM, Kegler MC. A review of qualitative data analysis practices in health education and health behavior research. Health Educ Behav 2019;46(1):32-39 [FREE Full text] [doi: 10.1177/1090198118795019] [Medline: 30227078]
36. Tobin GA, Begley CM. Methodological rigour within a qualitative framework. J Adv Nurs 2004;48(4):388-396. [doi: 10.1111/j.1365-2648.2004.03207.x] [Medline: 15500533]
37. Do I need ethics approval for social media research? JMIR Publications. URL: https://support.jmir.org/hc/en-us/articles/115001620728-Do-Ineed-ethics-approval-for-social-media-research [accessed 2024-11-21]
38. Townsend L, Wallace C. Social media research: a guide to ethics. University of Aberdeen. URL: https://www.gla.ac.uk/media/Media_487729_smxx.pdf [accessed 2024-11-01]
39. Seetharaman K, Chaudhury H. 'I am making a difference': understanding advocacy as a citizenship practice among persons living with dementia. J Aging Stud 2020;52:100831. [doi: 10.1016/j.jaging.2020.100831] [Medline: 32178801]
40. Talbot CV, O'Dwyer ST, Clare L, Heaton J, Anderson J. How people with dementia use Twitter: a qualitative analysis. Comp Hum Behav 2020;102:112-119 [FREE Full text] [doi: 10.1016/j.chb.2019.08.005]
41. Weetch J, O'Dwyer S, Clare L. The involvement of people with dementia in advocacy: a systematic narrative review. Aging Ment Health 2021;25(9):1595-1604. [doi: 10.1080/13607863.2020.1783512] [Medline: 32578451]
42. Li Y, Xie Y. Is a picture worth a thousand words? An empirical study of image content and social media engagement. J Mark Res 2019;57(1):1-19 [FREE Full text] [doi: 10.1177/0022243719881113]
43. World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International. URL: https://www.alzint.org/u/WorldAlzheimerReport2015.pdf [accessed 2024-05-27]

Abbreviations

API: application programming interface

Edited by T Mackey; submitted 21.06.24; peer-reviewed by C Hung Lo, S Creten; comments to author 27.09.24; revised version received 07.11.24; accepted 10.11.24; published 27.12.24.

Please cite as:
Bacsu JD, Fraser SA, Jamali AA, Conanan C, Chasteen AL, Vellani S, Gowda-Sookochoff R, Berger C, Mah JC, Fehr F, Virani A, Rahemi Z, Nanson K, Cammer A, Andrew MK, Grewal KS, McGilton KS, Lautrup S, Spiteri RJ
Navigating Awareness and Strategies to Support Dementia Advocacy on Social Media During World Alzheimer's Month: Infodemiology Study
JMIR Infodemiology 2024;4:e63464
URL: https://infodemiology.jmir.org/2024/1/e63464
doi:10.2196/63464
PMID:39729354

Original Paper

The Journey of Engaging With Web-Based Self-Harm and Suicide Content: Longitudinal Qualitative Study

Zoë Haime[1], BSc, MSc, PhD; Laura Kennedy[2,3,4], BA, MPhil, PhD; Lydia Grace[3], BSc, MSc, PhD; Rachel Cohen[1,5], BA, MA, PhD; Jane Derges[1], MSc, PhD; Lucy Biddle[1], BA, MA, PhD

[1]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom
[2]Centre for Society and Mental Health, Kings College, London, United Kingdom
[3]Samaritans, Surrey, United Kingdom
[4]Department of Health Service and Population Research, King's College London, London, United Kingdom
[5]Health and Social Services Group, Welsh Government, Cardiff, United Kingdom

Corresponding Author:
Zoë Haime, BSc, MSc, PhD
Population Health Sciences
Bristol Medical School
University of Bristol
Canynge Hall
39 Whatley Road
Bristol, BS8 2PS
United Kingdom
Phone: 44 01179289000
Email: zoe.haime@bristol.ac.uk

Abstract

Background: Self-harm and suicide are major public health concerns worldwide, with attention focused on the web environment as a helpful or harmful influence. Longitudinal research on self-harm and suicide–related internet use is limited, highlighting a paucity of evidence on long-term patterns and effects of engaging with such content.

Objective: This study explores the experiences of people engaging with self-harm or suicide content over a 6-month period.

Methods: This study used qualitative and digital ethnographic methods longitudinally, including one-to-one interviews at 3 time points to explore individual narratives. A trajectory analysis approach involving 4 steps was used to interpret the data.

Results: The findings from 14 participants established the web-based journey of people who engage with self-harm or suicide content. In total, 5 themes were identified: initial interactions with self-harm or suicide content, changes in what self-harm or suicide content people engage with and where, changes in experiences of self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement, the disengagement-reengagement cycle, and future perspectives on web-based self-harm or suicide content engagement. Initial engagements were driven by participants seeking help, often when offline support had been unavailable. Some participants' exposure to self-harm and suicide content led to their own self-harm and suicide behaviors, with varying patterns of change over time. Notably, disengagement from web-based self-harm and suicide spaces served as a protective measure for all participants, but the pull of familiar content resulted in only brief periods of disconnection. Participants also expressed future intentions to continue returning to these self-harm and suicide web-based spaces, acknowledging the nonlinear nature of their own recovery journey and aiming to support others in the community. Within the themes identified in this study, narratives revealed that participants' behavior was shaped by cognitive flexibility and rigidity, metacognitive abilities, and digital expertise. Opportunities for behavior change arose during periods of cognitive flexibility prompted by life events, stressors, and shifts in mental health. Participants sought diverse and potentially harmful content during challenging times but moved toward recovery-oriented engagements in positive circumstances. Metacognitive and digital efficacy skills also played a pivotal role in participants' control of web-based interactions, enabling more effective management of content or platforms or sites that posed potential harms.

Conclusions: This study demonstrated the complexity of web-based interactions, with beneficial and harmful content intertwined. Participants who demonstrated metacognition and digital efficacy had better control over web-based engagements. Some attributed these skills to study processes, including taking part in reflective diaries, showing the potential of upskilling users. This study also highlighted how participants remained vulnerable by engaging with familiar web-based spaces, emphasizing the responsibility of web-based industry leaders to develop tools that empower users to enhance their web-based safety.

(JMIR Infodemiology 2024;4:e47699) doi:10.2196/47699

KEYWORDS

suicide; self-harm; online; longitudinal; qualitative

Introduction

Background

Self-harm and suicide are major global public health concerns, with >700,000 people worldwide dying by suicide each year [1]. Attention has increasingly focused on the role of the web environment in triggering, exacerbating, or normalizing self-harm and suicide [2-4]. The amount of suicide-related information accessible on the web has grown [5], and graphic content depicting self-harm is increasingly available on social networking platforms [6]. Research shows that self-harm and suicide–related internet use is common among young people [7], particularly those who are under psychiatric care [8] and who go on to die by suicide [9].

There is a range of self-focused and social motivations for engaging with web-based self-harm and suicide content. These include accessing ongoing peer support or immediate help during a crisis [2,10,11], documenting recovery from self-harm [1,8], and researching suicide methods [12]. Moreover, research has shown that the ways in which people interact with web-based self-harm and suicide content vary depending on their level of distress [11,13].

The diversity in self-harm and suicide material complicates the experiences of content engagement. Research has identified these content interactions as being both a public health concern and a possible preventative measure [3,14], and studies have recognized the potential for engagement to have both benefits and costs [15]. Content with the potential to harm includes information on high-lethality suicide methods [16], prosuicide websites that may encourage suicide [13], and content describing novel methods of self-harm [17]. Benefits associated with accessing content include the role of the online community in peer support, validation and acceptance of one’s own self-harm or suicide feelings, and the opportunity for altruism when helping others [2,10,18-20]. These benefits may be particularly valuable given existing gaps in mental health care services and the widespread stigma that people who self-harm or experience suicidal thoughts encounter offline [2]. However, a recent review suggested that the impact of engaging with particular types of web-based self-harm or suicide–related content varies both between and within individuals, with content that benefits some having negative consequences for others and vice versa [15]. The review also identified only 4 longitudinal studies on the impact of self-harm and suicide–related internet use. Of these studies, 2 identified preventative effects of suicide prevention websites and web-based health forums on suicidal ideation [21,22]. One study showed minimal effects of search engine helpline notices on future suicide queries [23], and another study found that exposure to self-harm on Instagram predicted suicidal ideation and self-harm–related outcomes [17]. However, none of these studies used qualitative methods with their participants, emphasizing the current paucity of evidence on how self-harm and suicide–related web-based behavior evolves and the long-term effects and experiences of engaging with such content from the user’s perspective, including whether these are brief or permanent.

Objectives

The aim of this study was to explore the motivations for and consequences of viewing, searching for, and posting web-based self-harm or suicide content over a longitudinal period. Specifically, this study builds on existing knowledge by using qualitative and digital ethnographic methods to explore individual narratives of web-based engagement. Exploration of “significant moments” and points of transition within the web journey could also have substantial implications for the prevention of suicide and reduction of self-harm [24].

Methods

Design

This was a 6-month qualitative ethnographic study that investigated the stability and change in engagement with web-based self-harm and suicide content. This involved 3 one-to-one interviews and daily diary completion by participants over the study duration. We selected a 6-month time frame to ensure that we could observe changes over time in web-based engagement and associated behaviors [25,26] while also remaining mindful of the considerable commitment required for this ethnographic approach to maintain retention of participants.

Ethical Considerations

Ethics approval was obtained from the University of Bristol Faculty of Health Sciences ethics committee (reference: 117491). All participants provided written informed consent before participation, and were informed that they could withdraw from the study (including data withdrawal up to the time of analysis), without giving a reason. During consent, participants were assigned a participant ID used to identify their data and ensure anonymity. They were also informed that their data would be held confidentially and securely by the University of Bristol according to its duties and obligations under GDPR and the Data Protection Act. All participants were also compensated for their time, receiving a total of up to £75 (US $94.79) for full study completion.

Sampling and Recruitment

UK residents aged ≥16 years who were able to communicate in English and had experience engaging with web-based self-harm or suicide content were eligible. This included posting images, videos, memes, forum posts, blog posts, recovery posts, or comments related to self-harm or suicide or engaging with others' self-harm or suicide–related content through reposting and reblogging, quoting, liking, sharing, saving, subscribing to, or commenting. They did not need to have previous experience with self-harm or suicidal thoughts or behaviors.

Potential participants responded to advertisements posted between November 2021 and April 2022 on social media platforms (Facebook, Twitter [subsequently rebranded as X], and Reddit subreddits ["r/AdultSelfHarm," "r/StopSelfHarm," "r/BPD," "r/MentalHealthUK," and "r/malementalhealth"]), via Tellmi—a UK-based young person mental health app), and through charity websites and newsletters (Samaritans, SMaRteN, The McPin Foundation, and MQ Mental Health Research). Advertisements were posted once to platforms or sites until the end of recruitment in April 2022; however, due to web-based posting and reposting, it is possible that they were also shared elsewhere by others. Permission was sought from moderators or administrators before posting. Advertisements included a link to an expression of interest form in which participants consented via completion to the collection of brief demographic information, if and when the person last self-harmed, the way they were engaging with web-based self-harm or suicide content, and what platforms they used. All respondents had engaged with web-based self-harm or suicide content in some way.

This information was used to sample a diverse range of participants from those who expressed interest and target recruitment advertisements. Potential participants were sent the study information sheet via email, and those who were still interested in participating completed a consent form. Interviews were then arranged via email. The demographic data of those who did not participate were deleted. Once 14 baseline interviews had been conducted, the study team considered that there was good participant diversity in ethnicity and sufficient gender diversity. In addition, we had a broad range of platforms and apps represented in participant use. The authors also identified high-quality dialogue data sufficient for analysis and consistent themes to address the research aims. This resulted in the data achieving good information power [27], and therefore, recruitment was terminated. Information power was used as an alternative to data saturation in this study as the diverse nature of participant narratives meant that we were unlikely to reach a point of saturation.

Data Collection

Written consent to participate was provided by participants before entering the study. Participants were also required to complete a mandatory safety plan, including contact details for someone who could support them, their general practitioner's details (in case serious safety concerns arose), and a self-care plan that was individually designed by each participant to suit their needs (Multimedia Appendix 1). Study information was sent to the parents or guardians of those aged 16 to 18 years as a transparency measure. However, formal parental or guardian consent was not deemed a requirement by the ethics committee given the ages of the participants involved. As part of the study, a distress protocol was developed with a clinician to manage the risk of worsening mental health or increased self-harm or suicidal thoughts as a result of participation in the study. According to the protocol, participants would first be referred to their own safety plan if their mental health declined as a result of the study. A hierarchy of responses was specified in cases of more serious distress, including the options of offering follow-up support from UK suicide charity "Samaritans" or calling upon the advice of a named senior clinician. However, study-induced distress was not reported by participants during the study, and therefore, such responses were not actioned by researchers.

One-to-one interviews were conducted at baseline and the 3- and 6-month time points via Zoom (Zoom Video Communications) with just the researcher and participant present. The interviews were open-ended and flexible, using probing techniques where appropriate, and structured loosely using a topic guide. The main topics explored were "history of self-harm and suicide feelings"; "current and historic web-based activity related to self-harm and suicide content"; "patterns, motivations, and impact of web-based content engagement"; "critical moments in the web-based content engagement journey"; "keeping safe on the web"; and "experiences of web-based moderation and blocking." The topic guides were originally refined using feedback from 2 lived-experience experts. Throughout the study, the topic guides continued to be iteratively adapted between interviews, grounding question modifications in the study data. The interviews were conducted by ZH, LK, or LB and lasted between 35 and 80 minutes (with baseline interviews averaging 65 [SD 8.55] min and follow-up interviews averaging 45 [SD 2.87] min). They were audio recorded using an encrypted device and then transcribed.

Diaries

Participants completed daily diaries independently between interviews. These diaries served as an ethnographic tool and were introduced at the end of the baseline interview. Blank digital templates were then provided periodically via email. Each covered a 4-week period and had 3 main components (daily recording of content engagement, mood ratings, and a weekly reflection of content impact). Each participant was asked to complete 5 diaries in total. Entries were used to formulate personalized follow-up interview schedules in which further information or clarifications could be sought from participants.

Measures

Self-reported mental well-being data were collected from participants at baseline and monthly intervals to coincide with diary data collection. This was done via surveys on SurveyMonkey and included validated measures for assessing anxiety, depression, and psychological well-being (Multimedia Appendix 2 [28-31]). These data were used to characterize the sample and identify whether changes in mental health and mood reported by participants during the study interviews and in the diary data were reflected in outcome measure scores.

Data Analysis

Descriptive Analysis

Participant baseline demographic characteristics were reported as proportions or frequencies, as appropriate. Individual trajectories for well-being measures were represented visually using line graphs.

Qualitative Analysis

A trajectory analysis approach [32] was undertaken to interpret interview data temporally using the following steps:

1. Baseline interviews were transcribed, and then, through coding, themes were derived deductively from topic guide questions and inductively from the data themselves. ZH, LK, and LB separately listed preliminary themes and then refined and revised them collaboratively (Table S1 in Multimedia Appendix 2).
2. Initial matrices were produced for each participant, which included data from the baseline and the 3- and 6-month interviews. These were ordered so that each row was dedicated to a theme established in the previous step. Time points were then assigned to each column. Web-based engagement time points included "initiation," "historic," "current," "never," and periods of "disengagement and reengagement." These time points were adapted from the original trajectory approach [31] to preserve the "chronological flow" of the data collected during this study. This allowed us to acknowledge historical content engagement and the nonlinear flow of participant journeys as the levels of engagement fluctuated, ceased, and restarted. This also enabled the inclusion of participants who were only interviewed at baseline (due to dropout) as their data included information about past experiences. Data were formatted according to a "key" using text color to denote the site or platform used and highlighting whether it was related to a significant web event. An event was deemed to be "significant" if the participant recalled it as such or if the researchers found evidence within the narrative that it had a significant impact on the participant's thinking or behavior. Matrices were developed by extracting relevant quotes or context summaries for 2 participants by ZH, LK, and LB, and once consistency in interpretation was achieved, ZH and LK separately constructed the remaining initial matrices, with ongoing discussion between the researchers to ensure that all the data were captured.
3. Second matrices were then constructed for each participant. These were ordered with the initial themes as column headings. Each row represented an web-based platform or site used by the individual and included condensed versions of the "journey" that participants had experienced for each theme. The comparison allowed us to explore possible patterns in theme content by platforms or sites used. Second matrices were created by ZH for each participant and reviewed by LK and LB.
4. With all matrices complete, ZH, LK, and LB met to discuss similarities and differences across participant trajectories, noting trends, patterns, and outliers. Member checking of transcripts did not occur in this study due to funding and time constraints. During qualitative meeting discussions, overarching longitudinal themes were finalized.

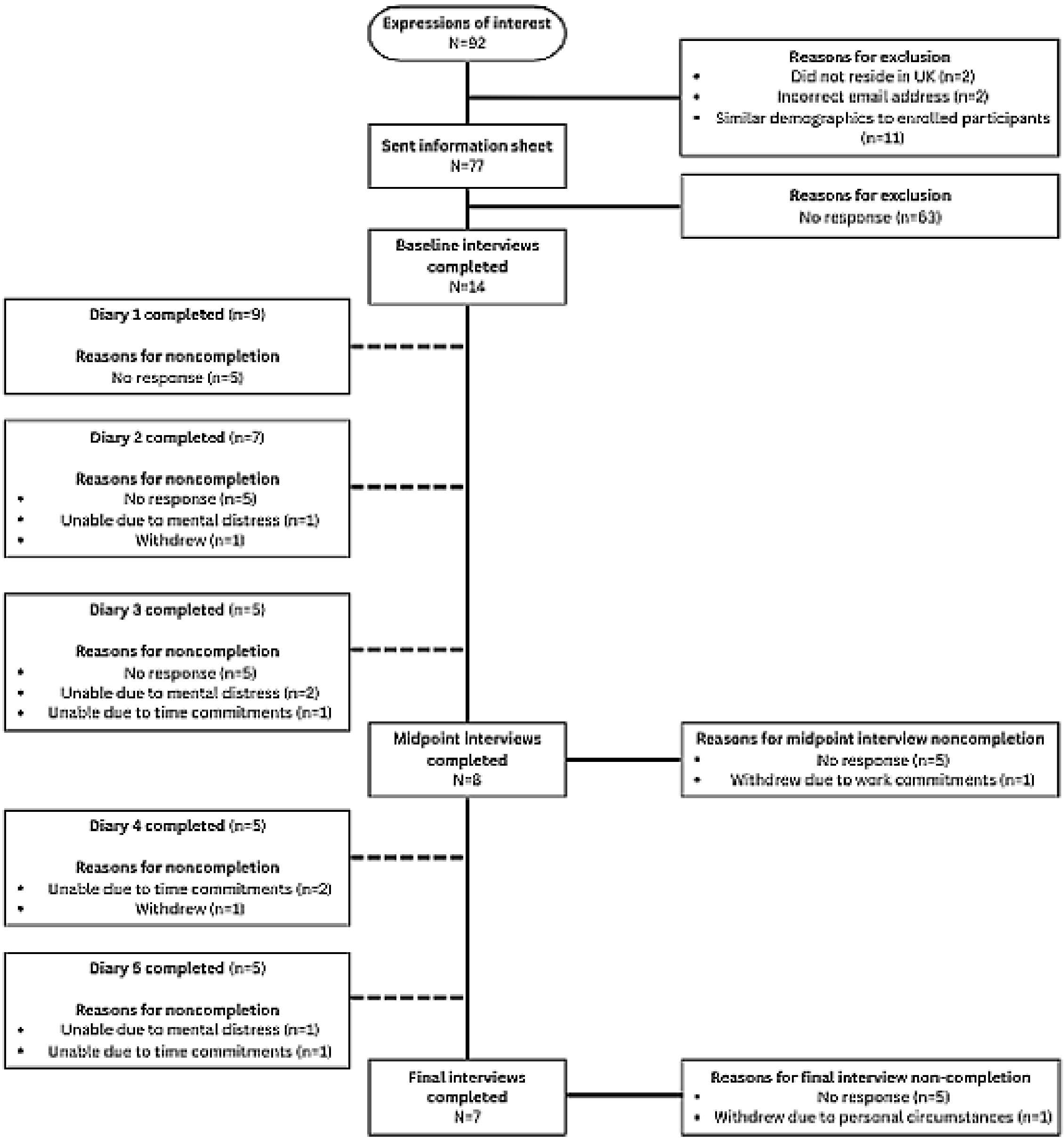
Results

Participant Flow

The participant flow through the study is shown in Figure 1. There were 92 expressions of interest. Of the 77 individuals who were sampled and sent study information, 63 (82%) did not respond and 14 (18%) took part. Data from the expression of interest forms showed that participants were less likely to respond to the research invite if they were younger (aged 16-24 years), had never hurt themselves on purpose, or had self-harmed in the last week.

Of the 14 participants who completed a baseline interview, 8 (57%) completed a midpoint interview, and 7 (50%) also completed the end-point interview. On the basis of preliminary observations of demographic characteristic data from the final sample, it appears that participants of non-British ethnicity may have had a lower likelihood of completing the study compared with those of British ethnicity. However, it should be noted that this observation was not tested statistically (Table S2 in Multimedia Appendix 2). Throughout the study, participants regularly completed their diaries, with study completers returning 77% (27/35) of the distributed diaries.

Figure 1. Participant flow through the study. Note: diaries were considered completed if one or more responses were provided.



Participant Characteristics

The characteristics of those who completed the baseline interviews (N=14) are displayed in Table 1. Of the 14 participants, 4 (29%) self-identified as male and 10 (71%) self-identified as female. Their ages ranged from 16 to 52 years, with 18 to 24 years being the most prevalent age group represented. There was a range of ethnicities, with almost half (6/14, 43%) of the participants being from global majority groups. Participants had engaged with self-harm or suicide–related content on a wide variety of sites and platforms.

Table 1. Participant characteristics at baseline (N=14).

Demographic variables	Participants, n (%)
Gender	
Man	4 (29)
Woman	10 (71)
Nonbinary	0 (0)
Ethnicity	
Asian British	2 (14)
Asian other	2 (14)
Black British	1 (7)
Black other	0 (0)
White British	7 (50)
White other	1 (7)
Mixed	1 (7)
Age (y)	
16-17	1 (7)
18-24	7 (50)
25-35	0 (0)
36-45	4 (29)
46-54	2 (14)
≥55	0 (0)
Have you ever hurt yourself on purpose?	
Yes	14 (100)
No	0 (0)
Website or platform used to access content[a]	
Instagram	3 (21)
Facebook	5 (36)
TikTok	1 (7)
Twitter	6 (43)
Tumblr	3 (21)
Weibo	1 (7)
Discord	1 (7)
WhatsApp	1 (7)
YouTube	3 (21)
Suicide forums	3 (21)

[a]Participants were able to select more than one option.

Participant IDs

Participant IDs were assigned during the consent process to ensure anonymity. As participants were aware of their assigned IDs, these were changed in the manuscript (see further details in Multimedia Appendix 2).

Descriptive Analysis Results

Individual line graphs for each well-being measure demonstrated fluctuations in mental health throughout the 6-month study period that reflected participant journeys recalled through interviews (Multimedia Appendix 2). One participant, IDB, scored poorer at 6 months on the Entrapment Scale–Short Form (which measures feelings of entrapment in a concise manner) than at baseline; however, the decline was minimal. All other study participants (13/14, 93%) improved from baseline or had no change in total score at the study end point in all quantitative measures, although no statistical analysis of change was undertaken.

Longitudinal Qualitative Analysis

Overview

The themes developed following trajectory analysis included (1) initial engagements with web-based self-harm or suicide content, (2) changes in what self-harm or suicide content people engage with and where, (3) changes in self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement, (4) the disengagement-reengagement cycle, and (5) future perspectives on self-harm and suicide content engagement. The themes and their constituent subthemes are summarized in Textbox 1.

Within these themes, fluctuations in mental health and control were identified as significant factors impacting behavioral and emotional responses to web-based content and, therefore, will be further explored in the following sections.

Textbox 1. Themes and subthemes.

Initial engagement with web-based self-harm or suicide content

- Motivations for initial web-based self-harm or suicide content engagement
- Experience of engaging with self-harm or suicide content for the first time

Changes in what self-harm or suicide content people engage with and where

- Changes in types of web-based self-harm or suicide content engagement over time
- Balancing curiosity and control
- Changes in posting web-based self-harm or suicide content over time

Changes in self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement

- Personal risk associated with web-based self-harm or suicide content engagement
- The precipitative and protective effects of engagement with self-harm or suicide content on self-harm or suicide behavior

The disengagement-reengagement cycle

- Disengagement from web-based self-harm and suicide content
- Reengagement with web-based self-harm and suicide content
- Longer periods of disengagement
- Limiting content engagement: strategies

Future perspectives on self-harm and suicide content engagement

Initial Engagement With Web-Based Self-Harm or Suicide Content

Motivations for Initial Web-Based Self-Harm or Suicide Content Engagement

Our first theme captured historical accounts of engaging with web-based self-harm and suicide content. Participants in this study, most of whom (12/14, 86%) had already self-harmed, initially engaged with web-based self-harm or suicide content following attempts to seek help offline during mental health declines. Those who attended mental health services and received new or changes in diagnoses generally reported leaving unsatisfied, citing reasons that included lack of support, inadequate availability, or feelings of being "dismissed" (IDH; baseline interview) due to a perception of low risk. Some were unable to access services at all or felt that attending was not worthwhile. These mental health declines alongside gaps in service provision were the common catalysts for initial web-based searches for self-harm and suicide content. While some of these searches were motivated by a desire to seek help, they varied among participants, with some also seeking information on self-harm and suicide methods:

> *So, I'd been to the doctors...I'd already tried looking for help, I was waiting for a referral to the CMHT [Community Mental Health Team]...And then within a couple of days I'd started lightly [cutting] on my hand then I moved up to my arm, and then I was looking for support groups online, just general support groups.* [IDB; baseline interview]

Experience of Engaging With Self-Harm or Suicide Content for the First Time

The experience of initially encountering self-harm or suicide content on the web is captured through the participant responses in Table 2. Only 14% (2/14) of the participants recalled first coming across content unintentionally, with most (12/14, 86%) describing purposeful searches to access material. While most of these searches were for help and support, 14% (2/14) of the participants reported seeking information about methods for self-harm or suicide, and 7% (1/14) of the participants were uncertain about what they were hoping to find but acknowledged that support-focused sites were unhelpful to them at that point.

Table 2. Quotations related to the experience of first encountering web-based self-harm or suicide content.

Reaction	Description	Quotations
Negative	First engagement with web-based self-harm or suicide content produced a negative response.	• "That's not what I was looking for [support sites], I didn't want help, at that point I was beyond help." [IDH; baseline interview] • "...I was researching [a suicide method]...what's required and the best way to manage [that]...It was scary. It'd have been really easy just to have thought, well, actually, I know more about it now and I can do that." [IDC; baseline interview] • "I received a picture on WhatsApp of someone, of a friend at the time who was self-harming and she basically just sent me a picture of her scars. I think that that image has stayed with me until today, and I think it's one of the reasons why I'm so careful because it's not something that I want to see again." [IDI; baseline interview]
Mixed	Participants experienced both positive and negative responses to the first engagement with web-based self-harm or suicide content.	• "...because people were experiencing very similar things to what you were experiencing you wanted to have more of that. It was a good environment in one respect, but it was a very toxic environment in the next because you were listening and you were going, 'Oh, I've been through that.' But it wasn't helping. It was actually pushing you down a bit because you were getting ideas [about how to self-harm]." [IDG; baseline interview] • "I think I was just surprised that there was so much content out there. And yeah, that they haven't been removed, and I think...I guess a sort of comfort knowing that there were others out there who were also going through tough times...And I think, I guess also shocked at how severe some [images of self-harm] are yeah." [IDL; baseline interview]
Positive	First engagement with web-based self-harm or suicide content produced a positive response.	• "I applied to go onto that [Facebook] group just so that I could reach out to people and find out more from survivor-led experiences. And people offered support to each other, and I felt that was quite a good thing to do." [IDA; midpoint interview] • "It made me feel a lot less alone just knowing, even if they were anonymous people out on the internet that could be wherever in the world, that there were other people, and I wasn't the only person feeling like this. It was so beneficial, especially as a young teen." [IDF; baseline interview]

Some participants sought support-related content, and others not intending to access self-harm or suicide content at all unintentionally came across graphic content (eg, images of fresh self-harm) or suicide method descriptions during their first engagements. Those whose initial interactions were with this type of self-harm or suicide content described feelings of distress even when this was the content they were seeking out. Some of these participants (2/14, 14%) recognized that this content could inadvertently validate and trigger their own self-harm and suicide feelings and behaviors, making them feel more at risk. In cases in which participants first engaged with web-based self-harm or suicide content in a discussion forum or peer support group, they were more likely to respond positively, describing how they felt less alone and were able to share experiences with others. However, some participants (2/14, 14%) had mixed emotions—it was comforting to know similar others existed, but processing extreme content was challenging and subjected them to information about novel self-harm and suicide methods, revealing their lack of control over what they were exposed to.

Changes in What Self-Harm or Suicide Content People Engage With and Where

Changes in Types of Web-Based Self-Harm or Suicide Content Engagement Over Time

All participants continued interacting with online self-harm or suicide content after their initial encounter even if it had been a negative experience. In cases in which they had positive initial engagements, participants continued to use the same platforms to access self-harm or suicide content in the long term. When those platforms or sites became obsolete, they sought out equivalent content in other web locations. Participants who had negative initial interactions accessed different platforms or sites searching for self-harm or suicide material that resonated with them.

Although participants had self-harm and suicide content that they accessed in a stable and routine manner, many also described occasions when they would change what they were accessing. Most participants (12/14, 86%) explained that different content satisfied different needs depending on their current mental state or mood. Examples of this can be found in Textbox 2.

Textbox 2. Web-based self-harm and suicide content accessed during mental health changes.

Change in content accessed due to mental health declines

- "On a good day I can be in there and I can be supporting others and helping them and building them up. And then on a bad day I'll be the one looking for support and asking for somebody to you know pick me up a little bit. So, it very much depends on what mood I'm in that day to be honest." [IDB; baseline interview]
- "If you are depressed and you start like looking at videos that are to do with that sort of thing it's so easy for you to be in the spiral of just like looking at more and more content about suicide and stuff like that..." [IDK; baseline interview]

Change in content accessed due to suicidal feelings or intentions

- "When I am thinking about self-harm, I will just look it up online. I go to the text service when I have suicidal feelings." [IDF; baseline interview]
- "That's when ['if I'm in a really bad crisis'] I'm more seeking it out, so Tik Tok I'm not actively seeking out that content [ok] but that's when I'm actually seeking it out, thinking I want to die, that's when I start accessing suicide forums and stuff." [IDK; baseline interview]

Change in content accessed due to mental health improvements

- "I've reached a place where I feel like I want to kind of, hear more about recovery and things like that. I think that's why I found this sort of ['recovery-based images'] content useful to look at. And I think that, I don't think I'm triggered by it, but I also don't want to interact with that kind of content where people are talking about their own [recovery] journeys because I'm not in that kind of place or not in a place where I want to hear about that kind of stuff at the moment. So, yeah, kind of like more interactions with the positive stuff, I think." [IDI; final interview]

Dips in mental health often resulted in changes in the way participants engaged on the web, such as posting their own self-harm or suicide material rather than just interacting with others' content. In cases in which participants experienced sustained episodes of poorer mental health, self-harm and suicide content was also seemingly accessed more frequently and sometimes uncontrollably through "habit" (IDG) or "addiction" (IDC and IDH), with 21% (3/14) of the participants describing it as falling down a "rabbit hole" (IDJ, IDL, and IDF). In total, 7% (1/14) of the participants reported how this compulsive engagement with self-harm or suicide content interrupted elements of their usual social and occupational functioning:

> *Even through work time I would take ten minutes and just read some of it.* [IDH; baseline interview]

Directly questioning participants about web-based engagement when feeling "actively suicidal" elicited similar reported changes in behavior. A couple of participants described engaging with different content—notably turning to web-based suicide organizations and charities or friends and family members offline when they needed support for suicidal thoughts rather than their usual web-based resources for self-harm or suicide content. However, another 21% (3/14) of the participants described how prominent suicidal thoughts were more likely to result in them returning to prosuicide forums, where they would seek or check resources for their own suicide plans.

Improvements in mental health saw participants more likely to transition to web content of a recovery-based nature while often sticking to the same web-based locations. Some participants (3/14, 21%) also attempted to limit web engagement with greater use of offline resources such as community help centers or taking part in meaningful activities.

Balancing Curiosity and Control

Other participants came across content unexpectedly in their web journey or seemed to spontaneously seek out different self-harm or suicide content due to "curiosity." Some described the ability to negotiate novel self-harm and suicide content with a developed sense of control over time, skipping over or avoiding engaging with content that was undesirable to them:

> *Being able to scroll past content with trigger warnings of self-harm pics has been quite a new thing. Like in the last year-ish, before then I wouldn't have been able to have done that. I'd have looked.* [IDJ; baseline interview]

However, others described tensions between curiosity and control and how that curiosity led them to seek out different self-harm or suicide content. For example, 14% (2/14) of the participants, who read a news article on a person's death by suicide that referenced web-based prosuicide forums, went on to search for them:

> *...I saw it [article on death-by-suicide of person who used pro-suicide forums] in the news. When you see something in the news, especially on the BBC website you know...it's quite serious stuff. So, then you end up looking further. Now sometimes you have to be careful because you get drawn into it and I think you have to sort of say to yourself, "I'm only going to spend a few minutes doing this..."* [IDE; midpoint interview]

The functions of social media sites (eg, hashtags or algorithms) could also enable unintentional content encounters, making control over engagements less feasible:

> *I guess sometimes that like tags on social media and...it's usually by chance, I don't actively go and seek them, but sometimes it appears and then I kind of just go down a rabbit hole of looking at more of such content. Even though I didn't do it intentionally.* [IDL; baseline interview]

Another participant explained that, in transitioning from self-harm and suicide content that no longer resonated, they had less control over what they engaged with:

> *I think recently, it's like I don't know what I'm looking for, but it's like I know that I haven't been able to*

> *find it...So, I think it's normally looking through my explore page instead of searching for anything in particular...* [IDI; midpoint interview]

Changes in Posting Web-Based Self-Harm or Suicide Content Over Time

For some participants (5/14, 36%), posting content seeking help and support regarding self-harm or suicide feelings or looking for ways to stay safe while self-harming was an early action in their web journey (IDA, IDB, IDC, IDD, and IDG). Others also posted detailed descriptions of suicide methods they were considering on discussion forums (IDH, IDK, and IDD), blog posts detailing their own self-harm and suicide feelings (IDN), and images of quotes on Instagram with captions about their mental health (IDI). One participant sent images of their own self-harm via direct messaging after other users requested them (IDK).

A total of 21% (3/14) of the participants in the study refrained from publishing their own content publicly (IDF, IDL, and IDJ). Of these 3 participants, 2 (67%) posted content privately (meaning that it was posted on the web but was only visible to them). Both participants described this as their way to "vent" (IDF) or "rant" (IDL) when upset and an opportunity to document their journey.

Notably, all 3 "observation-only" participants mentioned valuing their anonymity in the web space and refraining from online community interactions. They also emphasized that the potential for posts to negatively affect others deterred them from posting self-harm or suicide content publicly:

> *I always felt quite conflicted about reposting other people's content related to it [photos or videos of fresh self-harm]. I feel like it's one thing for me to look at it because they've posted it...versus me reblogging it to my own blog. I don't know. It's odd to explain it but it just felt weird.* [IDF; baseline interview]

Another participant reported posting content in one context (asking for support on a Facebook group) but not posting "graphic images" (IDC) due to fear that it may cause harm to children. This particular concern for young people viewing content was echoed by IDF, IDH, and IDK.

IDN, who initially described making public blog posts about their own self-harm, later made these private due to a realization that the material may negatively affect others as well as an attempt to maintain anonymity. IDI also reported a change in posting behavior during and as a result of taking part in the study. After initially posting about their experiences in an attempt to raise mental health awareness, they reflected on their tendency to put a "positive spin" (IDI) on content, and by the 6-month follow-up interview, they had reduced the frequency of their posts as they began to question their own authenticity. They considered that, if they posted about their negative experiences, it would likely have a harmful effect on others, and so they refrained from posting.

Finally, one participant also noted that access to psychological therapy reduced their need to post on the web for support:

> *[I haven't posted] for quite some time actually. I can't remember the last time I did that. It would be over a month ago easily. Yeah, I haven't needed to really.* [IDA; final interview]
>
> *Why do you think that is?* [Interviewer; final interview]
>
> *Because I could handle whatever I was thinking probably on my own or bring it to the next...because I'm having weekly sessions with my psychologist...* [IDA; final interview]

Changes in Self-Harm or Suicide Behaviors Associated With Web-Based Self-Harm or Suicide Content Engagement

Personal Risk Associated With Web-Based Self-Harm or Suicide Content Engagement

As described previously, some participants identified risks after their initial exposure to web-based self-harm or suicide content (Table 2). Others recognized potentially harmful consequences after longer periods of engagement. Some thought that the content they engaged with gave them implicit "permission" to carry out similar self-harm or suicide behaviors (IDG, IDH, IDJ, IDK, and IDN):

> *It makes it [completing suicide] feel less scary and like being able to hear people talk about what happened to them, them saying it's not that bad, like it wasn't...it just felt like nothing, it makes it feel a lot easier to do it if you know what I mean?* [IDK; baseline interview]

Some found that their own self-harm or suicidal behaviors were influenced by self-harm and suicide information they had gathered on the web (IDJ, IDK, IDL, and IDB):

> *...there were some posts which would link to other websites where you could get resources [information on overdose statistics]. I'd say definitely at the start of my mental health journey that was quite a turning point for me. Because it was just an idea and then it became a possible thing to do.* [IDJ; baseline interview]

Another participant experienced feelings of jealousy over the self-harm people had engaged in, which resulted in them feeling the need to escalate their self-harm behaviors:

> *I think that was that self-comparison to myself...maybe I'm being too scared or I'm not trying hard enough...* [IDL; baseline interview]

The Precipitative and Protective Effects of Engagement With Self-Harm or Suicide Content on Self-Harm and Suicide Behavior

The feelings and behaviors that participants experienced following engagement with web-based self-harm or suicide content are shown in Table 3. Content could be precipitative or protective for participants depending on when they encountered it in their journey. Several participants (5/14, 36%) recalled engaging in self-harm and suicide behavior as a result of engaging with web-based content. A few of these participants

(4/14, 29%) went on to describe changes in their self-harm and suicide behaviors related to content engagement over time, implying that the effect could change from precipitative to protective. This included living vicariously through others' self-harm images (IDF and IDL), finding content engagement relaxing (IDC), and using content searches as a way to delay or stop their own suicidal behavior (IDD). One participant suggested that such changes in behavior were due to building their own mental resilience over time:

> *I wouldn't say the internet content changed, it would be more like I've changed to deal with what the internet's providing me.* [IDJ; baseline interview]

Another participant recognized the need to consistently "check-in" with their own mental health before engaging on the web:

> *It's just about how I'm feeling, like do I feel like I have the capacity to deal with the internet really, do I actually want to look at what people are saying and what people are posting.* [IDF; midpoint interview]

For some, there was less consistency regarding whether engagements with self-harm and suicide content would result in helpful or harmful circumstances. This was exemplified by one participant who stated that their searches were usually protective and kept them occupied when their suicidal thoughts were most intense:

> *I think there is a part of me that does it [conducts searches for self-harm and suicide content] to buy time.* [IDB; final interview]

However, this participant also reported attempting a new form of self-harm at the midpoint interview after learning about it through a peer support group on Facebook.

Table 3. Precipitative and protective effects of web-based self-harm or suicide content engagement identified by participants.

Factor and description	Quotations
Precipitative factors	
Self-harm or suicide behavior as a consequence of engaging with web-based self-harm and suicide content	• "It could also be really detrimental because many times, I would just come away feeling much more triggered than previously and then would engage in the behaviour [self-harm]." [IDF; baseline interview] • "One of the [posts] got taken out of a group [by me] because it was talking about bloodletting and since then, I've bought syringes and needles to try and do it myself." [IDB; midpoint interview] • "How did you then cope with the fallout of what you'd seen [distressing self-harm and suicide content]?" [Interviewer; baseline interview] • "I coped by self-harming. Yeah, and I write lots as well. So yeah, writing about how I feel and what I saw." [IDC; baseline interview]
Protective factors	
Vicarious experiences through self-harm or suicide content	• "It would mainly be trying to vicariously live out things through other people. So, I had a particular urge but wasn't in a position where I felt like I could self-harm or necessarily wanted to and almost living those experiences through somebody else's experience which was one of the ways that it [viewing self-harm material] could be really beneficial for me because it could almost meet that urge without me having to engage in the behaviour." [IDF; baseline interview]
Delaying or stopping own self-harm or suicide behavior	• "I don't really need to research it [suicide method] anymore. Sometimes, I do it anyway and I just re-re-search, re-read it and re-check my facts but it can be a way of preventing me from doing anything." [IDB; final interview] • "How do you mean?" [Interviewer; final interview] • "It's like there are levels to it, aren't there? That's what I find anyway. It starts with thoughts, then it turns to urges and once you get to that urge stage, you need to feel like you're doing something, whereas, re-searching it [suicide method] is better than actually putting the tablets in your mouth. It gives you that extra step before you get to that point, if you see what I mean." [IDB; final interview]
Calming effect	• "How did you feel [coming across images of self-harm]?"[Interviewer; midpoint interview] • "Quite relaxed because that's what I do [self-harm], so I could identify with them, those people who'd done things like that." [IDC; midpoint interview]

Disengagement-Reengagement Cycle

Disengagement From Web-Based Self-Harm and Suicide Content

Most participants (8/14, 57%) reported entering a cycle of disengagement and reengagement during their web-based self-harm and suicide content journey. Disengagement was usually temporary, with participants choosing to have "no phone days," deleting their accounts, finding offline activities to take part in, or being forced to disengage due to lack of internet access.

Most often, disengagement was purposeful but impulsive. It would usually occur during periods of compulsive engagement when participants recognized a lapse in their control or as a reaction to a significant life event that resulted in mental health decline. Life events that occurred during this study included suicide bereavement, hospitalization, bullying or victimization, and experiences of exam- or work-related stress. The act of intentionally disengaging from self-harm or suicide content was usually a conscious decision to reclaim control over their web-based actions.

A total of 14% (2/14) of the participants reflected on changes in their disengagement behavior while in the study (IDC and IDI). Previously, similar to other participants, they reported a tendency to compulsively access content during periods of poorer mental health followed by impulsive disengagement. However, at the 6-month interview, both participants described an improved ability to recognize their patterns of web-based behavior (Table 4). This understanding and insight empowered the participants to purposefully disengage during declining mental health episodes as a strategic means of regaining control over their behavior.

However, when other participants were forced to disengage from content because of intermittent internet access or physical health problems, they were often left with feelings of loss. Although one participant described this unintentional disengagement as an opportunity for brief respite from self-harm and suicide content engagement, its existence remained a reassuring presence:

> *I knew I could access them if I needed to, but I thought, "No, I'm having a week off and I'm going to try and distance myself from this as much as I can."* [IDA; final interview]

Table 4. Reasons for disengagement from web-based self-harm or suicide content—from final interviews.

Participant ID and reason for disengagement	Quote
IDC	
Mental well-being	"It felt like I needed to look after myself and that I needed that break to try and keep myself safe. One of the things that this research has taught me and helped me understand, it's helped me understand more about how social media impacts upon me. So, I think social media can be a source for good. I think you also need to recognise that sometimes you need a break."
Regaining control	"I really crashed down, and it scared me because I'd had a lovely weekend. Things are generally a lot better, and it scared me in that I can still crash down and fluster myself. I didn't trust myself to be researching suicide, self-harm...And there was a part of me that knew that I wanted to live, there was a part of me that knew I could spiral out of control, and I didn't want to spiral out of control. And I've alluded to the fact that I've learnt personally a lot about myself during the six month's research and how I use social media. And for me that Monday when I made that decision [to disengage from self-harm content] it was really positive."
IDF	
Regaining control	"I think just to prevent myself from falling down a rabbit hole and looking at content that I know wasn't good for me. And just feeling like so I've always been one of those people that I sort of like to sort of physically remove myself from things and remove things from me. So that's one of the reasons why I do that."
Mental well-being	"So, I think it was about a month ago now and someone who was quite active in Twitter (X) and the mental health recovery community passed away from what I feel was suicide. That's not been confirmed but when all of that happened, I did take a couple of days off the internet just to, I guess, process things there."

Reengagement With Web-Based Self-Harm and Suicide Content

Participants described various reasons for reengagement with web-based self-harm and suicide content, including a "fear of missing out" with the community (IDA, IDC, IDM, and IDK), wanting to use the site or platform to access other types of content (IDE, IDF, and IDI), procrastination or boredom (IDI and IDK), and the need to perform web-based responsibilities (eg, work or moderating roles within self-harm and suicide communities; IDI, IDA, and IDB). Some participants (5/14, 36%) claimed that they weighed the advantages and disadvantages of web-based content engagement before reengaging. Several participants (3/14, 21%) felt that the benefits of reengaging with self-harm and suicide content, such as feeling part of a community, were enough to justify the potential risks. As this participant noted, while the experience could be upsetting at times, it was still considered useful in light of the rewards of engagement:

> *With Twitter [X], I deleted that as well, but I felt like actually I missed the community and felt out of touch with people, so I actually found that useful [reengaging], as much as sometimes it's upsetting, it was useful.* [IDK; baseline interview]

There were also differing accounts of reengagement due to mental health improvements and declines. One participant described feeling more in control once their mental health was stable:

> *I think I was in a better place emotionally and with my mental health...And I just felt stronger, I genuinely felt stronger and more positive. It's a better time of year for me...I've started some new medication...So, I think that's a factor as well and me feeling stronger to go back online. I just felt ready.* [IDC; final interview]

Similarly, another participant felt that they were more able to view and contribute to self-harm and suicide content in a positive way when their mental health improved:

> *When my mental state is better, and I can go back on. I feel like I can share, and I can help someone.* [IDM]

Alternatively, some participants (2/14, 14%) described past reengagement with self-harm and suicide content to "punish"

(IDF; baseline interview) themselves for thinking about or carrying out self-harm behaviors:

> *I think at that time I was kind of trying to make myself feel worse, because it was like, "You need to feel more guilty for what you're doing."* [IDI; final interview]

However, these participants described differences in their reengagements over time. IDI reported how their reengagement behavior changed during the study. When feeling low, they now went on the web and sought out non–self-harm or suicide content.

Other participants also described attempts to engage with self-harm or suicide content differently during the reengagement period with the aim of regaining control. This included observing interactions rather than actively participating or limiting engagements with specific content on platforms or sites:

> *Recently I've just been viewing [prosuicide threads] and I've got to fight the urges [to interact].* [IDH; baseline interview]

However, most participants who disengaged briefly would return to their usual use of web-based content. This reengagement process highlighted weaknesses in participants' ability to exercise control over web-based actions, leaving users vulnerable to reencountering triggering content on the web and beginning the disengagement-reengagement cycle again:

> *I basically quit Tik Tok for three weeks because I was like I just can't deal with it anymore because it's just so hard to block everything and I was also thinking is it actually good for my mental health and it's not...* [IDK; baseline interview]
>
> *...but you are back on TikTok now, is that right?* [Interviewer; baseline interview]
>
> *I think I was just bored really, and I thought do you know what I'll just download it for the afternoon, and...* [IDK; baseline interview]

Longer Periods of Disengagement

In total, 14% (2/14) of the participants in this study disengaged for up to a month before reengaging with specific platforms. One of these participants disengaged after a second death by suicide in their Twitter community. Notably, an earlier death by suicide of another member of the same community had increased their frequency of accessing the platform.

During their Twitter disengagement, the participant continued their engagement with a self-harm support group on Facebook, where they felt less connected:

> *I think because I haven't known them [Facebook users] so long and there's certain people [on Twitter, subsequently rebranded as X]...who post frequently, several times possibly in a day...I think the more you get to know people and recognise the handles, I know it sounds bizarre, but you feel yourself becoming closer to them.* [IDC; final interview]

Despite this, they also reported beginning to reengage with Twitter toward the end of the study:

> *I think just because I feel a bit better, I wanted to check-in on other things on there on my newsfeed, wall thing.* [IDC; final interview]

One other participant disengaged twice from a prosuicide forum. First, they described disengaging following an article on the parents of forum members who had died by suicide. The participant reached out to the parents, and the resulting relationship led to their disengagement:

> *...they've told me I need to get off the site.* [IDH; baseline interview]

However, they reengaged shortly after this event after wanting to check whether "they [the site] put the resources [suicide methods literature] back" (IDH; baseline interview) following their removal after the media article publication.

At the midpoint interview, IDH had again disengaged from and reengaged with the forum following the death by suicide of a relative. On describing their reengagement, they reported that "it was to check [that] the sources of getting stuff [suicide materials]...are still available" (IDH; midpoint interview) as they were aware of scams related to sourcing materials and wanted to verify that their plans would still be viable.

Limiting Content Engagement: Strategies

After spending time engaging with web-based self-harm and suicide content, half (7/14, 50%) of the participants began to develop strategies to limit their content engagement. These included less "arbitrary 'liking'" to curate their feeds (IDI), clearing search histories to "remove temptation" (IDJ), "blocking" or "muting" certain terms or phrases—such as "suicide" and "self-harm" (IDC, IDF, and IDK)—closing their direct messages so that other users were unable to message them (IDI and IDH), "self-banning" so that they were unable to post (IDH), distracting themselves with positive web-based content (IDE, IDI, IDC, and IDF), "starting new accounts" to avoid tailored algorithms (IDK), and distancing themselves from a self-assigned "role" such as being a mental health advocate (IDJ and IDI). In this study, we observed that younger participants predominantly used these strategies, possibly because of the enhanced accessibility to safety features on the platforms or sites they frequented or their proficiency in digital skills. However, it is noteworthy that most participants regardless of age reported learning digital safety methods of limiting engagement over time through their experiences on the web:

> *As I've got older I've realised that actually you know what you see online can really impact on you, and that you know no-one's going to police it for me so I have to be sensible about the types of people that I follow and the types of things that I do online. I think that's something that came with sort of getting a bit older.* [IDF; midpoint interview]

Future Perspectives on Web-Based Self-Harm and Suicide Content Engagement

When asked, none of the participants reported a desire to disengage from web-based self-harm or suicide content entirely in the future. Many alluded to the nonlinear nature of their engagement, recognizing difficulties during previous attempts to disengage. Some also described a sense of comfort and reassurance knowing that content continued to exist on the web:

> *It's a cushion for people who need that.* [IDA; midpoint interview]

In addition, others reported a desire to "give back" and described having a peer support role themselves as a future goal following their recovery (IDI, IDM, and IDB):

> *I'm looking forward to where I improve myself, and maybe be able to talk to more people and if possible, reach out to them, and offer that help.* [IDM; final interview]

> *I'm also very passionate about sharing stuff I've learnt. When people are in that place that I remember being in and you can see it from their posts, I think, "I've just learnt about something that will help them. I'll pass that on to them." It's helping to build that confidence back up to do those posts and say those things on there.* [IDB; final interview]

Some participants in this study (3/14, 21%) also highlighted that they were unable to find alternative web-based or offline spaces that satisfied their current needs. One participant mentioned that disengaging from their current preferred site or platform could be detrimental and so expressed no wish to "move on":

> *What I'm trying to say is that there is nowhere for people when they come off that website. There's no safe space. There's nowhere. If you've been on that particular site [prosuicide forum] for the reason of wanting to die and you didn't, there's nowhere. You'll go on something and just get these silly comments or things where there's lack of understanding that just escalates a situation.* [IDH; midpoint interview]

A few participants in this study (3/14, 21%) did recognize the potential costs associated with continuing to engage in web-based spaces with self-harm and suicide content but compromised, stating that "I do feel that the benefits outweigh the risks" (IDC; baseline interview). For these individuals, the draw of the positive aspects of such content was strong enough to justify the potential negative consequences. Other participants (2/14, 14%) struggled to weigh the risks and benefits of engaging with self-harm and suicide content as they felt that the positive and negative aspects of engaging with content were more intertwined, making it difficult to control what they were exposed to:

> *I'd say that online is very complicated, depending on what you feed your mind, because it has both positive and negative information, so sometimes it's good to your mind, and sometimes not. Also, if you are coming across lots of negative things in a group, that can be harmful, like self-harm pictures. But it's also good to look in those groups for people who are offering help for those things, so that you are learning how to help yourself.* [IDM; final interview]

Ultimately, this resulted in both sets of participants remaining vulnerable to the negative effects of harmful content as they continued to engage with web-based self-harm and suicide material.

Discussion

This study showed that those engaging with web-based self-harm and suicide content experienced nonlinear journeys that were characterized by 5 key themes: "initial engagements with web-based self-harm or suicide content," "changes in what self-harm or suicide content people engage with and where," "changes in self-harm or suicide behaviors associated with web-based self-harm or suicide content engagement," "the disengagement-reengagement cycle," and "future perspectives on web-based self-harm and suicide content engagement."

Cognitive Flexibility Versus Cognitive Rigidity

Constructs that may explain behavior change and maintenance within these themes are cognitive flexibility and its counterpart, cognitive rigidity [33]. Cognitive flexibility refers to an openness in thinking and behavior, which allows an individual to consider alternative perspectives and approaches. In contrast, cognitive rigidity is the tendency to adhere to specific thought and behavior patterns, making it challenging to change one's mindset or actions [33]. Previous research has identified a relationship between cognitive rigidity and suicidal ideation [34] and between cognitive rigidity and self-harm [35]. Another study showed that cognitive flexibility can result in engagement in multiple methods of self-harm [36]. This indicates that the construct of cognitive flexibility may provide important insights into the behavior changes over time associated with web-based self-harm and suicide content engagement. This discussion will explore the ways in which cognitive flexibility was impacted by participants' mental health and control over decision-making and how this influenced their web journeys.

Previous research has identified gaps in clinical support as a key motivator for web-based self-harm and suicide content engagement [2]. The causes for initial engagement in this study were consistent with this, with participants reporting a lack of support but also a reluctance to engage with clinical services due to previous experiences. This suggests a high level of cognitive flexibility among participants during their first engagement with web-based content, with mental distress and a lack of alternative resources potentially triggering participants to be more open to web-based options. This emphasizes the critical need for accessible offline options during the early stages of mental health decline, preventing vulnerable people from resorting to web-based avenues where they may lack the control or knowledge to engage safely.

When participants were unable to find content that was immediately desirable to them, they explored different self-harm or suicide–related material on the web. Often, this led to spontaneous browsing of self-harm and suicide–related links or hashtags, a behavior characterized as "pessimistic browsing" [13]. While this reflects a high level of cognitive flexibility among participants, it also indicates what might be a lack of behavioral control, making participants vulnerable to potentially harmful encounters. Later on in web journeys, when browsing routines had been established, some reported similar bouts of “pessimistic browsing” and harmful behaviors that they considered spontaneous. These episodes of cognitive flexibility were usually triggered by unexpected exposure to web-based self-harm or suicide content, where impulsive tendencies resulted in exploring this novel content further or, in one case, in trying a new self-harm method. This indicates that unexpected engagements with self-harm or suicide content may act as a stimulus for activating cognitive flexibility, resulting in changes in behavior [37]. When experiencing poor control, this cognitive flexibility may lead to a willingness to engage in potentially unhelpful or harmful behaviors when engaging with self-harm or suicide content [38].

Outside of episodes of cognitive flexibility, participants largely accessed web-based self-harm or suicide content in a routine pattern while also reporting a greater feeling of behavioral control. This cognitive rigidity often worked as a coping mechanism allowing for regular engagement with resources of help and support. However, in instances in which content included images or videos of “fresh self-harm,” suicide, or self-harm and suicide method information, repeated engagements were more likely to have negative effects on participant well-being and sometimes led to increased severity of harm to themselves. This shows that, while some perceived their cognitive rigidity as a form of control, it may ultimately have diminished their ability to make decisions to protect themselves and seek alternative positive coping mechanisms [39].

Similarly, participants reported increased engagement with self-harm or suicide content during dips in their mental health, which were prevalent in this study, as indicated by fluctuations in their well-being measure outcomes over time. These engagements, recalled as “habitual” or “addictive,” highlighted a loss of control during these mental health dips. Previous research has shown a relationship between cognitive inflexibility and addictionlike behaviors [40,41], and a more recent study [42] has indicated that distress-driven impulsivity, in which a person is likely to make rash decisions due to a negative mental state, alongside cognitive rigidity, can lead to addictionlike eating behavior. This emphasizes the potential risk of overreliance on web-based self-harm and suicide content as a coping strategy, particularly during periods of mental health decline, when participants may become more vulnerable to the content they are engaging with. The addictive nature of this behavior also has the potential to negatively impact other important aspects of people’s lives, such as social or occupational functioning [43].

Disengaging and Reengaging

Key to self-preservation during the web journey was participants’ ability to disengage from web-based spaces. Most participants recorded disengagements in their web journeys in response to life experiences or stressors, such as work stress, bereavement, or a rapid deterioration in mental health. This indicates a resurgence of cognitive flexibility, which reflects previous research showing that individuals become more open to alternative solutions when their perspective is challenged by a significant life event [44]. Although participants demonstrated disengagement attempts from the content during these times, they were usually temporary. This represents a brief state of cognitive flexibility, with reengagement often occurring within days. When disengagement was longer, it tended to coincide with more significant life events such as bereavements, which may indicate more prolonged changes to behavior following extreme circumstances and mental health declines.

Participants also reported that their mental state dictated whether they returned to more helpful or harmful content during the reengagement period. Participants experiencing poorer mental health were more likely to reengage with content they described as “negative” as a type of self-punishment or as a preventative measure against potentially worsening self-harm or suicide behavior. They were also more likely to post their own content, which included help-seeking comments, suicide method inquiries, and “depressive” blog posts. This showed that, although some participants attempted to use their online communities for help during mental health dips, others could find themselves returning to potentially unhelpful or harmful situations. This reflects previous research showing that “active” suicidal ideation is associated with greater cognitive rigidity compared to “passive” suicidal ideation [45]. Often, when reengaging during mental health declines, use would also regress to “addictive” or “habitual” engagements. However, when experiencing mental health improvements during reengagement periods, those who had previously engaged with more “positive” or “recovery-based” content would be more likely to return to this material. This indicates that cognitive rigidity is influenced by mental health state and that, when experiencing mental health changes, participants’ well-being is reliant on earlier web-based encounters with self-harm and suicide content.

Upskilling Users

Despite this, some participants did experience lasting adaptations to the ways in which they interacted with the content. These more enduring changes were attributable to the skills that participants reported developing in digital efficacy and metacognition. Digital efficacy skills include the ability to use web-based safety mechanisms such as muting, blocking, and self-banning. Participants with digital efficacy skills in the study felt safer and more protected, which acted as a preventative measure against cognitive rigidity. In this study, these participants were likely to be younger, which reflects research showing that digital literacy skills are significantly better in younger cohorts [46]. Despite this, evidence also shows that digital literacy skills can be built over time [47]. This is consistent with the experiences of some participants in this study who reported that their web experiences prompted them to organically develop digital skills and strategies to stay safe over time. This finding has important implications for industry leaders, who should be encouraged to consider ways in which they can empower users by improving accessibility to safety mechanisms on their platforms and sites.

Metacognition skills, or the ability to reflect on one’s own thoughts and behaviors to change one’s responses, were evident in some of the participants [48]. Specific metacognitive abilities such as self-awareness and self-regulation resulted in greater control over their cognitive flexibility. Some described gaining metacognition skills such as self-awareness before their participation in the study, which allowed them to recall changing their responses to content from self-harm behavior to vicarious viewing of material. Others identified metacognitive skills gained through therapeutic input as well as through monitoring web-based behavior and reflecting on it during the study. This may reflect a Hawthorne effect [49] in which participant behavior shifts due to their awareness of being observed in a study. Several diary and ecological momentary assessment interventions have resulted in improved metacognitive skills [50] (Haime, Z, unpublished data, January 2024), and it is possible that metacognition was acquired in this study as a result of completing the research diary. In cases in which participants' metacognition developed during the study, they noted improvements in their mental health, also indicated by improvements in their well-being outcomes over time. This resulted in the type of self-harm and suicide material they engaged or reengaged with changing from "negative" or "depressive" to "recovery-based" or "positive" in nature. This shows that self-awareness and control while experiencing mental health improvements lead to positive content engagement during periods of cognitive flexibility in this population and has important implications for the development of future target metacognitive interventions.

Remaining Vulnerable on the Web

However, shifts toward recovery-based content did not necessarily mean that participants were able to fully disengage from their previous self-harm and suicide material. Sometimes, as recovery-based content coexisted alongside more harmful content in web spaces, there was no alternative place in which to access it. On the other hand, some participants expressed a strong connection with the communities they had previously engaged with and reported intentions to remain active in these spaces with a desire to provide support to others. While this altruistic act had benefits, including the ability to continue drawing on support when needed, it left them vulnerable to potentially triggering content. These findings emphasize the strength of web-based self-harm and suicide spaces as a source of comfort and security, which is consistent with previous research on engagement motivators [2,7,10]. Thus, although participants became more aware of the negative outcomes of engaging with web-based self-harm and suicide content and were better able to manage them, the perceived benefits of being involved in a community of like-minded individuals with similar experiences often outweighed the potential costs.

Limitations

Participants in this study used a diverse range of web-based platforms to access self-harm and suicide content, meaning that attempts to identify patterns in behavior related to the sites used were challenging. However, as common behaviors were observed across participants, it was possible to draw conclusions more broadly about how people engage with web-based self-harm and suicide content over time. Diaries in this study were completed daily by participants, but many had missing entries or were filled out retrospectively. This diluted the advantages of "in-the-moment" diary data capture and resulted in some interview topic guides being less informed by participant data. Despite this, participants reported finding the diaries largely acceptable, and some reported additional benefits to their metacognitive ability related to their completion [51].

While visually observing quantitative data allowed us to identify patterns consistent with participant-reported mental health fluctuations and slight improvements toward the end of the study, our inability to conduct statistical analyses prevented us from identifying any significant differences in participant well-being changes. However, the rich qualitative data and trajectory analysis provided valuable insights into the individual pathways and factors influencing web-based engagement.

In terms of participant characteristics, this study had an underrepresentation of male individuals. Although steps were taken to target male-orientated web spaces for recruitment, uptake remained poor. Furthermore, responses to recruitment were limited, which resulted in possible selection bias and may have affected the representativeness of the sample. In addition, we did not collect data on the educational level or socioeconomic status of the participants involved, limiting our understanding of how demographic characteristics may affect web-based experiences. Half of those recruited at baseline were also lost to follow-up. Strategies were undertaken to limit attrition, including at least 3 attempts to communicate with participants before they were considered lost to follow-up. High attrition rates are consistent with longitudinal studies of self-harm and may represent a selection bias among study completers [52]. Finally, although cognitive flexibility provides a useful framework with which to interpret our findings, it is important to acknowledge that there may be alternative explanations.

Future Implications

The findings of this study have shown that there are ongoing challenges in navigating the web environment for those engaging with self-harm and suicide content. A key priority for future research should be to establish how engaging with web-based content can be better managed in this population. Consequently, the following should be considered:

1. Inaccessibility to offline support was a significant motivator for participants' willingness to explore web-based self-harm and suicide–related resources. Therefore, the availability of offline help and support is necessary to limit or moderate initial web-based engagements.
2. This study offers evidence that greater metacognition and digital efficacy can positively influence web-based behavioral control. As individuals are unlikely to completely disengage from web-based content, it is important to prioritize upskilling users. Therefore, interventions should be developed focusing on improving digital literacy and metacognitive skills, such as the diary-based reflections used in this study.
3. A deeper examination of the perceived benefits of web-based engagement is necessary to ensure that these needs can be met in a safer manner both on the web and offline. In addition, it is crucial to critically evaluate the helpfulness of these perceived benefits, such as the impact of "vicarious living" through observing others self-harm.
4. Web-based industry leaders need to produce more tools that empower individuals to regain control of their

web-based engagement and improve the safety of web-based spaces where self-harm and suicide content is available. This may include changes to the functions of social media, such as providing further control and management options to users over algorithms and hashtags.

Conclusions

A balance between cognitive flexibility and rigidity seems necessary to protect individuals when engaging with self-harm and suicide content on the web. While cognitive flexibility may be helpful in certain situations such as exploring new coping strategies, it can also leave individuals vulnerable to harmful content. On the other hand, cognitive rigidity, or the tendency to repeatedly engage with the same type of content, can lead to desensitization, potential impairments in functioning, and an increased severity of harm to oneself. Cognitive rigidity can also prevent people from engaging in harmful behaviors and allow them to consistently engage with content that is helpful and positive. Although life events and changes in mental health state could trigger cognitive flexibility resulting in behavior changes, these were unlikely to become permanent unless participants developed skills such as digital efficacy and metacognition that gave them greater control over their behavior. Despite this, even with improved skills for recognizing and managing web-based risks, individuals still perceived that the benefits of web spaces outweighed the costs, making it difficult to fully disengage.

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material including data confidentiality statement, additional tables, and descriptive analysis line graphs.
[DOCX File , 43 KB - infodemiology_v4i1e47699_app1.docx]

Multimedia Appendix 2
Example safety plan.
[DOCX File , 64 KB - infodemiology_v4i1e47699_app2.docx]

References

1. Suicide worldwide in 2019: global health estimates. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240026643 [accessed 2024-02-25]
2. Brown RC, Fischer T, Goldwich DA, Plener PL. "I just finally wanted to belong somewhere"—qualitative analysis of experiences with posting pictures of self-injury on Instagram. Front. Psychiatry 2020 Apr 21;11:274 [FREE Full text] [doi: 10.3389/fpsyt.2020.00274] [Medline: 32372983]
3. Lavis A, Winter R. #Online harms or benefits? An ethnographic analysis of the positives and negatives of peer-support around self-harm on social media. J Child Psychol Psychiatry 2020 Aug 27;61(8):842-854. [doi: 10.1111/jcpp.13245] [Medline: 32459004]
4. Marchant A, Hawton K, Stewart A, Montgomery P, Singaravelu V, Lloyd K, et al. Correction: a systematic review of the relationship between internet use, self-harm and suicidal behaviour in young people: the good, the bad and the unknown. PLoS One 2018 Mar 1;13(3):e0193937 [FREE Full text] [doi: 10.1371/journal.pone.0193937] [Medline: 29494676]
5. Biddle L, Derges J, Mars B, Heron J, Donovan JL, Potokar J, et al. Suicide and the internet: changes in the accessibility of suicide-related information between 2007 and 2014. J Affect Disord 2016 Jan 15;190:370-375. [doi: 10.1016/j.jad.2015.10.028] [Medline: 26546772]
6. Miguel EM, Chou T, Golik A, Cornacchio D, Sanchez AL, DeSerisy M, et al. Examining the scope and patterns of deliberate self-injurious cutting content in popular social media. Depress Anxiety 2017 Sep 29;34(9):786-793. [doi: 10.1002/da.22668] [Medline: 28661053]
7. Mars B, Heron J, Biddle L, Donovan JL, Holley R, Piper M, et al. Exposure to, and searching for, information about suicide and self-harm on the internet: prevalence and predictors in a population based cohort of young adults. J Affect Disord 2015 Oct 01;185:239-245 [FREE Full text] [doi: 10.1016/j.jad.2015.06.001] [Medline: 26150198]
8. Nesi J, Burke TA, Lawrence HR, MacPherson HA, Spirito A, Wolff JC. Online self-injury activities among psychiatrically hospitalized adolescents: prevalence, functions, and perceived consequences. Res Child Adolesc Psychopathol 2021 Apr 06;49(4):519-531 [FREE Full text] [doi: 10.1007/s10802-020-00734-4] [Medline: 33404950]
9. Rodway C, Tham SG, Richards N, Ibrahim S, Turnbull P, Kapur N, et al. Online harms? Suicide-related online experience: a UK-wide case series study of young people who die by suicide. Psychol Med 2022 May 19;53(10):4434-4445. [doi: 10.1017/s0033291722001258]

10. Biddle L, Derges J, Goldsmith C, Donovan JL, Gunnell D. Online help for people with suicidal thoughts provided by charities and healthcare organisations: a qualitative study of users' perceptions. Soc Psychiatry Psychiatr Epidemiol 2020 Sep 10;55(9):1157-1166 [FREE Full text] [doi: 10.1007/s00127-020-01852-6] [Medline: 32157324]
11. Tucker IM, Lavis A. Temporalities of mental distress: digital immediacy and the meaning of 'crisis' in online support. Sociol Health Illn 2019 Oct 10;41 Suppl 1(S1):132-146. [doi: 10.1111/1467-9566.12943] [Medline: 31599990]
12. Mok K, Jorm AF, Pirkis J. Who goes online for suicide-related reasons? Crisis 2016 Mar;37(2):112-120. [doi: 10.1027/0227-5910/a000366] [Medline: 27232427]
13. Biddle L, Derges J, Goldsmith C, Donovan JL, Gunnell D. Using the internet for suicide-related purposes: contrasting findings from young people in the community and self-harm patients admitted to hospital. PLoS One 2018 May 24;13(5):e0197712 [FREE Full text] [doi: 10.1371/journal.pone.0197712] [Medline: 29795637]
14. Boyce N. Pilots of the future: suicide prevention and the internet. Lancet 2010 Dec 04;376(9756):1889-1890. [doi: 10.1016/s0140-6736(10)62199-x] [Medline: 21155079]
15. Brennan CA, Saraiva S, Mitchell E, Melia R, Campbell L, King N, et al. Self-harm and suicidal content online, harmful or helpful? A systematic review of the recent evidence. J Public Ment Health 2022 Jan 05;21(1):57-69. [doi: 10.1108/jpmh-09-2021-0118]
16. Gunnell D, Derges J, Chang SS, Biddle L. Searching for suicide methods: accessibility of information about helium as a method of suicide on the internet. Crisis 2015 Sep;36(5):325-331. [doi: 10.1027/0227-5910/a000326] [Medline: 26502782]
17. Arendt F, Scherr S, Romer D. Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. New Media Soc 2019 May 27;21(11-12):2422-2442. [doi: 10.1177/1461444819850106]
18. Picardo J, McKenzie SK, Collings S, Jenkin G. Suicide and self-harm content on instagram: a systematic scoping review. PLoS One 2020;15(9):e0238603 [FREE Full text] [doi: 10.1371/journal.pone.0238603] [Medline: 32877433]
19. Eichenberg C, Schott M. An empirical analysis of internet message boards for self-harming behavior. Arch Suicide Res 2017 Nov 15;21(4):672-686. [doi: 10.1080/13811118.2016.1259597] [Medline: 27845618]
20. Harris KM, McLean JP, Sheffield J. Suicidal and online: how do online behaviors inform us of this high-risk population? Death Stud 2014 Oct 18;38(6-10):387-394. [doi: 10.1080/07481187.2013.768313] [Medline: 24666145]
21. Scherr S, Reinemann C. First do no harm: cross-sectional and longitudinal evidence for the impact of individual suicidality on the use of online health forums and support groups. Comput Human Behav 2016 Aug;61:80-88. [doi: 10.1016/j.chb.2016.03.009]
22. Till B, Tran US, Voracek M, Niederkrotenthaler T. Beneficial and harmful effects of educative suicide prevention websites: randomised controlled trial exploring Papageno . Werther effects. Br J Psychiatry 2017 Aug 02;211(2):109-115. [doi: 10.1192/bjp.bp.115.177394] [Medline: 28522433]
23. Cheng Q, Yom-Tov E. Do search engine helpline notices aid in preventing suicide? Analysis of archival data. J Med Internet Res 2019 Mar 26;21(3):e12235 [FREE Full text] [doi: 10.2196/12235] [Medline: 30912753]
24. Holland J, Thomson R. Gaining a perspective on choice and fate: revisiting critical moments. Eur Soc 2009 Jul;11(3):451-469. [doi: 10.1080/14616690902764799]
25. Smit D, Vrijsen JN, Broekman T, Groeneweg B, Spijker J. User engagement within an online peer support community (depression connect) and recovery-related changes in empowerment: longitudinal user survey. JMIR Form Res 2022 Nov 02;6(11):e39912 [FREE Full text] [doi: 10.2196/39912] [Medline: 36322110]
26. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
27. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov 10;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
28. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
29. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
30. De Beurs D, Cleare S, Wetherall K, Eschle-Byrne S, Ferguson E, O'Connor DB, et al. Entrapment scale--short-form. APA PsycTests. URL: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft79989-000 [accessed 2024-02-21]
31. Bass M, Dawkin M, Muncer S, Vigurs S, Bostock J. Validation of Warwick Edinburgh mental well-being scale (WEMWBS) in a population of people using secondary care mental health services. J Ment Health 2016 Aug 28;25(4):323-329. [doi: 10.3109/09638237.2015.1124401] [Medline: 26821730]
32. Grossoehme D, Lipstein E. Analyzing longitudinal qualitative data: the application of trajectory and recurrent cross-sectional approaches. BMC Res Notes 2016 Mar 02;9:136 [FREE Full text] [doi: 10.1186/s13104-016-1954-1] [Medline: 26936266]
33. Dajani DR, Uddin LQ. Demystifying cognitive flexibility: implications for clinical and developmental neuroscience. Trends Neurosci 2015 Sep;38(9):571-578 [FREE Full text] [doi: 10.1016/j.tins.2015.07.003] [Medline: 26343956]
34. Miranda R, Gallagher M, Bauchner B, Vaysman R, Marroquín B. Cognitive inflexibility as a prospective predictor of suicidal ideation among young adults with a suicide attempt history. Depress Anxiety 2012 Mar 06;29(3):180-186. [doi: 10.1002/da.20915] [Medline: 22147587]

35. Nilsson M, Lundh L, Westrin Å, Westling S. Executive functioning in psychiatric patients with deliberate self-harm, as compared with a psychiatric and a healthy comparison group. J Clin Exp Neuropsychol 2021 Apr 05;43(3):225-237. [doi: 10.1080/13803395.2021.1894094] [Medline: 33949907]
36. Park Y, Ammerman BA. For better or worse?: the role of cognitive flexibility in the association between nonsuicidal self-injury and suicide attempt. J Psychiatr Res 2023 Feb;158:157-164. [doi: 10.1016/j.jpsychires.2022.12.040] [Medline: 36586214]
37. Cañas JJ, Fajardo I, Salmeron L. Cognitive flexibility. In: Karwowski W, editor. International Encyclopedia of Ergonomics and Human Factors. Boca Raton, FL: CRC Press; 2006:297-301.
38. Miranda R, Valderrama J, Tsypes A, Gadol E, Gallagher M. Cognitive inflexibility and suicidal ideation: mediating role of brooding and hopelessness. Psychiatry Res 2013 Nov 30;210(1):174-181 [FREE Full text] [doi: 10.1016/j.psychres.2013.02.033] [Medline: 23528518]
39. Kakoschke N, Aarts E, Verdejo-García A. The cognitive drivers of compulsive eating behavior. Front Behav Neurosci 2018 Jan 17;12:338 [FREE Full text] [doi: 10.3389/fnbeh.2018.00338] [Medline: 30705625]
40. Wu M, Brockmeyer T, Hartmann M, Skunde M, Herzog W, Friederich HC. Set-shifting ability across the spectrum of eating disorders and in overweight and obesity: a systematic review and meta-analysis. Psychol Med 2014 Dec;44(16):3365-3385. [doi: 10.1017/S0033291714000294] [Medline: 25066267]
41. Liu C, Rotaru K, Lee RS, Tiego J, Suo C, Yücel M, et al. Distress-driven impulsivity interacts with cognitive inflexibility to determine addiction-like eating. J Behav Addict 2021 Apr 27;10(3):534-539 [FREE Full text] [doi: 10.1556/2006.2021.00027] [Medline: 33909594]
42. Aditi B, Arunjyoti B. Addiction severity, social functioning, and life satisfaction of patients diagnosed with substance use disorders. Indian J Psy Nsg 2018;15(2):13. [doi: 10.4103/2231-1505.262434]
43. Dennis JP, Vander Wal JS. The cognitive flexibility inventory: instrument development and estimates of reliability and validity. Cogn Ther Res 2009 Nov 11;34(3):241-253. [doi: 10.1007/s10608-009-9276-4]
44. Eshet-Alkali Y, Amichai-Hamburger Y. Experiments in digital literacy. Cyberpsychol Behav 2004 Aug;7(4):421-429. [doi: 10.1089/cpb.2004.7.421] [Medline: 15331029]
45. Lai YJ, Tan HC, Wang CT, Wu WC, Wang LY, Shen YC. Difference in cognitive flexibility between passive and active suicidal ideation in patients with depression. Neuropsychiatry 2018;08(04):1182-1185. [doi: 10.4172/neuropsychiatry.1000446]
46. Eshet-Alkalai Y, Chajut E. You can teach old dogs new tricks: the factors that affect changes over time in digital literacy. J Inf Technol Educ Res 2010;9:173-181. [doi: 10.28945/1186]
47. Martinez ME. What is metacognition? Phi Delta Kappan 2016 Jun 24;87(9):696-699. [doi: 10.1177/003172170608700916]
48. Merrett F. Reflections on the Hawthorne effect. Educ Psychol 2006 Jan;26(1):143-146. [doi: 10.1080/01443410500341080]
49. Derks YP, Klaassen R, Westerhof GJ, Bohlmeijer ET, Noordzij ML. Development of an ambulatory biofeedback app to enhance emotional awareness in patients with borderline personality disorder: multicycle usability testing study. JMIR Mhealth Uhealth 2019 Oct 15;7(10):e13479 [FREE Full text] [doi: 10.2196/13479] [Medline: 31617851]
50. Kurt M, Kurt S. Improving design understandings and skills through enhanced metacognition: reflective design journals. Int J Art Design Ed 2016 Jul 09;36(2):226-238. [doi: 10.1111/jade.12094]
51. Abdelraheem M, McAloon J, Shand F. Mediating and moderating variables in the prediction of self-harm in young people: a systematic review of prospective longitudinal studies. J Affect Disord 2019 Mar 01;246:14-28. [doi: 10.1016/j.jad.2018.12.004] [Medline: 30572208]
52. Vogl S, Zartler U. Interviewing adolescents through time: balancing continuity and flexibility in a qualitative longitudinal study. Longit Life Course Stud 2021;12:83-97. [doi: 10.1332/175795920x15986464938219]

Edited by E Lee; submitted 04.04.23; peer-reviewed by T Officer, R Plackett; comments to author 20.09.23; revised version received 10.10.23; accepted 15.02.24; published 28.03.24.

Please cite as:
Haime Z, Kennedy L, Grace L, Cohen R, Derges J, Biddle L
The Journey of Engaging With Web-Based Self-Harm and Suicide Content: Longitudinal Qualitative Study
JMIR Infodemiology 2024;4:e47699
URL: https://infodemiology.jmir.org/2024/1/e47699
doi:10.2196/47699
PMID:38546718

Original Paper

Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences

Masaru Kamba[1], PhD; Wan Jou She[1], PhD; Kiki Ferawati[1], MStat; Shoko Wakamiya[1,2], PhD; Eiji Aramaki[1], PhD

[1]Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

[2]Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Eiji Aramaki, PhD
Division of Information Science
Graduate School of Science and Technology
Nara Institute of Science and Technology
8916-5 Takayama-cho
Ikoma, 630-0192
Japan
Phone: 81 0743 72 5250
Email: aramaki@is.naist.jp

Related Article:

This is a corrected version. See correction statement: https://infodemiology.jmir.org/2024/1/e67434

Abstract

Background: Despite being a pandemic, the impact of the spread of COVID-19 extends beyond public health, influencing areas such as the economy, education, work style, and social relationships. Research studies that document public opinions and estimate the long-term potential impact after the pandemic can be of value to the field.

Objective: This study aims to uncover and track concerns in Japan throughout the COVID-19 pandemic by analyzing Japanese individuals' self-disclosure of disruptions to their life plans on social media. This approach offers alternative evidence for identifying concerns that may require further attention for individuals living in Japan.

Methods: We extracted 300,778 tweets using the query phrase *Corona-no-sei* ("due to COVID-19," "because of COVID-19," or "considering COVID-19"), enabling us to identify the activities and life plans disrupted by the pandemic. The correlation between the number of tweets and COVID-19 cases was analyzed, along with an examination of frequently co-occurring words.

Results: The top 20 nouns, verbs, and noun plus verb pairs co-occurring with *Corona no-sei* were extracted. The top 5 keywords were *graduation ceremony*, *cancel*, *school*, *work*, and *event*. The top 5 verbs were *disappear*, *go*, *rest*, *can go*, and *end*. Our findings indicate that education emerged as the top concern when the Japanese government announced the first state of emergency. We also observed a sudden surge in anxiety about material shortages such as toilet paper. As the pandemic persisted and more states of emergency were declared, we noticed a shift toward long-term concerns, including careers, social relationships, and education.

Conclusions: Our study incorporated machine learning techniques for disease monitoring through the use of tweet data, allowing the identification of underlying concerns (eg, disrupted education and work conditions) throughout the 3 stages of Japanese government emergency announcements. The comparison with COVID-19 case numbers provides valuable insights into the short- and long-term societal impacts, emphasizing the importance of considering citizens' perspectives in policy-making and supporting those affected by the pandemic, particularly in the context of Japanese government decision-making.

(JMIR Infodemiology 2024;4:e49699) doi:10.2196/49699

KEYWORDS

COVID-19; natural language processing; NLP; Twitter; disrupted plans; concerns

Introduction

Background

The spread of COVID-19 has become a global pandemic, significantly affecting social and economic sectors worldwide [1]. In the early stages of the pandemic, health authorities recommended social distancing to control the spread of the virus, reduce cases, and avoid overwhelming health care facilities [2-4]. Each country had its own strategy for dealing with COVID-19. A survey conducted across 6 countries illustrated the public's perception of measures taken in response to COVID-19 [5]. Other surveys have been conducted in the United Kingdom and European countries to aid interdisciplinary research on public health, particularly regarding COVID-19 [6]. Different results were observed owing to social distancing policies, which affected several aspects of life, including economic activities [7] and consumer behavior, such as drops in mobility [8]. Concerns about cybersecurity risks were also raised, as companies might not have been prepared for adequate work-from-home options for employees [9]. The association between implementing some mitigation policies in response to COVID-19 and the outcomes regarding public mobility were noted [10], one of which was also observed in Japan.

After the government confirmed the first COVID-19 case in Japan on January 16, 2020, the number of cases quickly escalated within 3 months, leading to the declaration of a state of emergency to prevent the further spread of the infection. This measure significantly impacted the daily routines and social lives of Japanese residents, forcing them to refrain from going out, close schools, work from home, and be restricted from visiting crowded locations such as department stores and movie theaters. The first state of emergency effectively reduced the number of COVID-19 cases [11], albeit at a high cost to public mental well-being, education quality, and the economy. Furthermore, the number of cases quickly bounced back, peaking at 1762 new daily cases after the state of emergency was lifted, an increase from the peak of 701 new daily cases during the first wave [11]. These numbers suggest that the government was confronted with the dilemma of mitigating the social and economic impact of the lockdown and stopping the spread of COVID-19 [12]. Due to the fluctuations in COVID-19 cases, the government declared other states of emergency, recognizing the profound and deeply rooted impact the COVID-19 lockdown could have on societal and economic levels.

There have been various investigations into the states of emergency. For instance, studies have predicted SARS-CoV-2 infections using state-space models [13] and examined their impact on mental health [14]. In the aspect of mobility, studies have shown the suppression of social activities of the masses [15]. The tourism industry was among the hardest hit sectors, and the arrival of visitors decreased by 93% by March 2020 [16]. Statistics also show that Japan's gross domestic product in 2020 decreased by 4.28%, indicating a substantial impact on the economy [17]. Interestingly, the unemployment rate only slightly increased to 2.8% in 2020, but started declining by 2022 (2.64%), following the gradual recovery of the gross domestic product (2.14% growth by 2021 and 1.03% by 2022) [17,18]. This trend of recovery indicates the strong resilience of the Japanese economy.

Furthermore, it also changed people's behavior, such as following the advisory to stay at home, as confirmed by cell phone location data [19-21]. Such large-scale societal and behavioral changes warrant further investigation through various means to offer a chance to monitor and reflect the short- and long-term impacts of COVID-19 in the future.

Literature Review

The disruption caused by pandemic-related restrictions may be seen as a failure to perform planned activities, but detecting such disruptions was challenging. For example, it is difficult to obtain behavioral data on trips that individuals could not take or events they could not attend owing to the restriction. Social media, which people use to share their activities, proved to be a great source of information in such cases. Twitter (currently X) is a widely used social media platform in many countries and has a sufficiently large population for social data analysis in health care contexts [22,23]. Japan has a particularly high population density of Twitter users, even when compared to the major countries that use Twitter, such as the United States. Furthermore, owing to language exclusivity, it is easier to filter comments related to Japanese society using Japanese keywords [24]. Twitter has also been frequently used to help summarize peoples' responses about the pandemic and its measures, showing the challenges experienced throughout [25]. Prior studies in Korea and Japan used Twitter to determine public opinion, showing popular words during the pandemic [26]. Because people actively share their daily lives on Twitter, the site has the potential to be a data source for investigating the impact of restrictions on the public. Using Twitter as a resource, this study aims to explore and visualize plans disrupted in Japan due to COVID-19 pandemic measures.

There are many studies on COVID-19 that investigate social media platforms, such as Twitter. Chen et al [27] investigated the levels of anxiety during the COVID-19 pandemic. The adverse effects on the mental health of the public was also one of the impacts of the pandemic, as explained in the research by Li et al [28], who analyzed COVID-19–related tweets into different emotions and investigated the mental health aspects and how they recovered from the COVID-19 crisis. Lyu et al [29] investigated the topics and sentiments in public COVID-19 vaccine–related discussions, whereas Krittanawong et al [30] investigated misinformation dissemination related to COVID-19 on Twitter. Aside from studies focused on the pandemic itself, COVID-19 vaccines have also been highly researched topics on Twitter. Ansari and Khan assessed public responses through sentiment analysis of COVID-19 vaccines using Twitter, revealing an overall negative tone in the tweets [31]. Ferawati et al [32] explored how Twitter reported vaccine-related side effects by comparing the side effects of 2 types of messenger RNA vaccines developed by Pfizer and Moderna in Japan and Indonesia, respectively. Gao et al [33] examined COVID-19 concerns in each Japanese prefecture, and Uehara et al [34] investigated the attitudes toward vaccines or vaccination during the COVID-19 pandemic in different Japanese prefectures using search queries from Yahoo! JAPAN. Our study adopts a unique approach to examine how the COVID-19 pandemic has disrupted everyday activities. Our main focus is on understanding the direct impact of the pandemic on society through the observation of expressions, life disruptions, and plans.

For research on citizen feedback, Ishida et al [35] proposed a method that uses social media data. They implemented a multitask learning framework to estimate the associated viewpoints using bidirectional encoder representations from the transformer model. However, this method requires considerable effort to label the data. This study uses search queries and validates word co-occurrence to infer the themes of topics discussed during the COVID-19 pandemic in Japan, proposing an efficient and low-resource method for social media analysis.

Objectives and Approach

We aimed to report on the impact of COVID-19 on Japanese society by analyzing public opinions extracted from social media data. Specifically, we focused on the popular term *Corona no-sei* (in Japanese コロナのせい, meaning "due to COVID-19," "because of COVID-19," or "considering COVID-19"), which clearly conveyed complaints or concerns about life event disruptions caused by the COVID-19 pandemic. Our study used 2 types of data: the daily COVID-19 case count and Japanese tweets containing the Japanese phrase *Corona no-sei* posted on Twitter between February 1, 2020, and November 30, 2021. We analyzed the trends in the number of tweets and COVID-19 cases to quantitatively explore their relationship and the words frequently used in the tweets to qualitatively explore social needs in the first 2 years of the COVID-19 pandemic.

In conclusion, we critically compared our findings with those identified in other similar studies to provide an alternative evidence base for the impact of COVID-19 on Japanese society.

Methods

COVID-19 Cases

To track the daily rise in COVID-19 cases, we gathered the number of new positive cases in Japan by manually downloading data from a dedicated COVID-19 site maintained by the NHK, Japan's national broadcaster [36]. Our aim was to investigate the correlation between the number of positive cases and the volume of tweets. A total of 1,726,943 COVID-19–positive cases were recorded between February 1, 2020, and November 30, 2021.

Tweets and Keywords Extraction

Another data set for this study includes 300,778 tweets containing the Japanese phrase *Corona no-sei* during the same period as the recorded COVID-19 cases (between February 1, 2020, and November 30, 2021). We chose this period because by the end of January 2020, the Japanese government had officially established the Japan Anti-Coronavirus National Task Force to actively address the pandemic. In addition, we aimed to include the maximum possible data until the initiation of this study in mid-November 2021. Furthermore, this period also included 3 emergency announcements by the Japanese government, making it a representative period for studying the impact of COVID-19 on Japanese society.

We counted the number of tweets per month and found that the maximum number of tweets was 517,688 in April 2020; the minimum number of tweets was 24,625 in November 2021; and the average number of tweets was 136,717.6. The *Corona no-sei* phrase is frequently used by the public in social media and everyday conversation to express the (often negative) feelings when Japanese individuals' activities or life plans were interrupted by the COVID-19 outbreak. Although there are several expressions synonymous with *Corona no-sei* (eg, "because of the new coronavirus" and "because of COVID-19"), we chose *Corona no-sei* as a casual expression used by the public in social media and colloquial speech. The tweet data were provided by the NTT DATA Corporation, which has a real-time backup of Japanese firehose data from X Corporation (formerly known as Twitter). Data access was granted to a few collaborative research institutes, including the University of Tokyo, and one of the authors was granted permission to use the self-adaptive classification system to extract the data and keywords [37].

Although applying a clustering approach for topic modeling can be useful in grasping the topics discussed in the tweets, it does not apply to our context wherein we were targeting *COVID-19* as the main subject and aiming to identify the co-occurrence of events. Instead, we extracted co-occurrence nouns and verbs from the obtained *Corona no-sei* tweets by applying dependency analysis implemented in the system developed by Yoshinaga et al [37-39]. We used the base-phrase chunker to extract all tweets containing the *Corona no-sei* keyword ("keyword" is *bunsetsu* in Japanese). The built-in classifier then extracted the relevant verbs, nouns, and verb-noun-pairs for users based on the nonstack dependency parser, which achieved 99.01% accuracy in base-phrase chunking and 92.23% accuracy in dependency parsing [37]. Researchers who did not use the system and database maintained by the University of Tokyo could use the same tool published by the laboratory Pecco and DepP [37-40]. To avoid overinterpretation, we omitted tweets that described a disruption of plans but did not include COVID-19–related keywords.

Analysis of the Keywords and its Correlation to the COVID-19 Pandemic Trends

The contexts following *Corona no-sei*, which indicate a high level of negative concern about COVID-19, frequently contain verbs in the negative form and nouns associated with them. By aggregating these nouns and verbs, we extracted information on the restrictions imposed and the events or plans canceled owing to the COVID-19 pandemic. This information enabled us to capture the potential social and psychological impact of disrupted life plans. Note that, by events or plans, we refer to the specific *type of occurrences* (eg, university entrance exam) rather than a certain event (eg, a pop singer's concert in 2019). The frequency of nouns and verbs in tweets containing *Corona no-sei* was counted to identify the restrictions placed on people's lives.

To investigate the correlation between tweet volumes and COVID-19 cases, we constructed transition diagrams for each.

In addition, Pearson correlation coefficients were also calculated. Next, we examined the nouns and verbs co-occurring with *Corona no-sei* over the entire study period and specifically on the day with the highest tweet activity.

The cross-validation of the keywords and tweet contents was performed by randomly extracting 20 tweets from the top 5 verb and noun pairs and other keyword pairs that were deemed worthy of discussion by the researchers. The tweet contents were further annotated to ensure that they were aligned with the researchers' interpretations of keywords. We then discussed the themes extracted by analyzing and cross-validating the themes and noteworthy keywords.

Ethical Considerations

This study used publicly available data and did not handle identifiable private information, meaning that it was exempt from Institutional Review Board approval according to the Ethical Guidelines for Research of the Japanese National Government [41]. The NTT DATA Corporation obtained tweets according to Twitter terms of service and approved the use of the data for this study.

Results

Figure 1 shows the time trend of *Corona no-sei* tweets (blue line) compared to the trend of positive cases (red line). There were 3 states of emergency announcements within our targeted period between February 1, 2020, and November 30, 2021, which are highlighted in gray in Figure 1. The number of areas under the state of emergency is indicated by the bar graph in the upper part of the figure because the target areas were changed during each state of emergency. The periods during which the states of emergency were imposed roughly corresponded to an increase in case numbers. Interestingly, the announcement of a state of emergency was highly effective in suppressing the number of cases. Regarding the spike caused by the Tokyo Olympics (which took place between July 23, 2021, and August 8, 2021), the case number quickly dropped to below 5000 per day within 3 months.

As the blue line indicates, the *Corona no-sei* tweets peaked in March 2020, roughly before the first state of emergency was announced and reached the second highest number when the first state of emergency was imposed. After the first announcement of the state of emergency, the number of tweets using *Corona no-sei* showed a downward trend until the end of our data collection period. There were a few instances of small increases in *Corona no-sei* tweets before the second and third states of emergency announcements, but overall, the number of reported plan disruptions never reached the level observed before the first state of emergency announcement. The scatter plot for case numbers and the numbers of *Corona no-sei* tweets is shown in Figure 2, with Pearson correlation coefficients of 0.86, 0.93, and 0.61, respectively, for the first, second, and third states of emergency.

When compared with the number of the *Corona no-sei* tweets during the entire period, the correlation between COVID-19 daily cases and the *Corona no-sei* was not very evident. We were able to observe a slight increase of *Corona no-sei* tweets before the case number started rising, but the extent of increase in case numbers was disproportional to the extent of increase in *Corona no-sei* tweets. Even though the number of cases peaked in September 2021 during the third state of emergency, there was only a slight increase in *Corona no-sei* tweets compared to the high number of complaints at the very beginning of the COVID-19 pandemic. This indicates that Japanese residents might have adapted to the restrictions or disruptions caused by the COVID-19 pandemic lockdown.

Figure 1. Trends in the number of *Corona no-sei* tweets and the number of patients with a COVID-19–positive result. The blue line indicates the number of *Corona no-sei* tweets, and the red line indicates the number of positive COVID-19 cases.

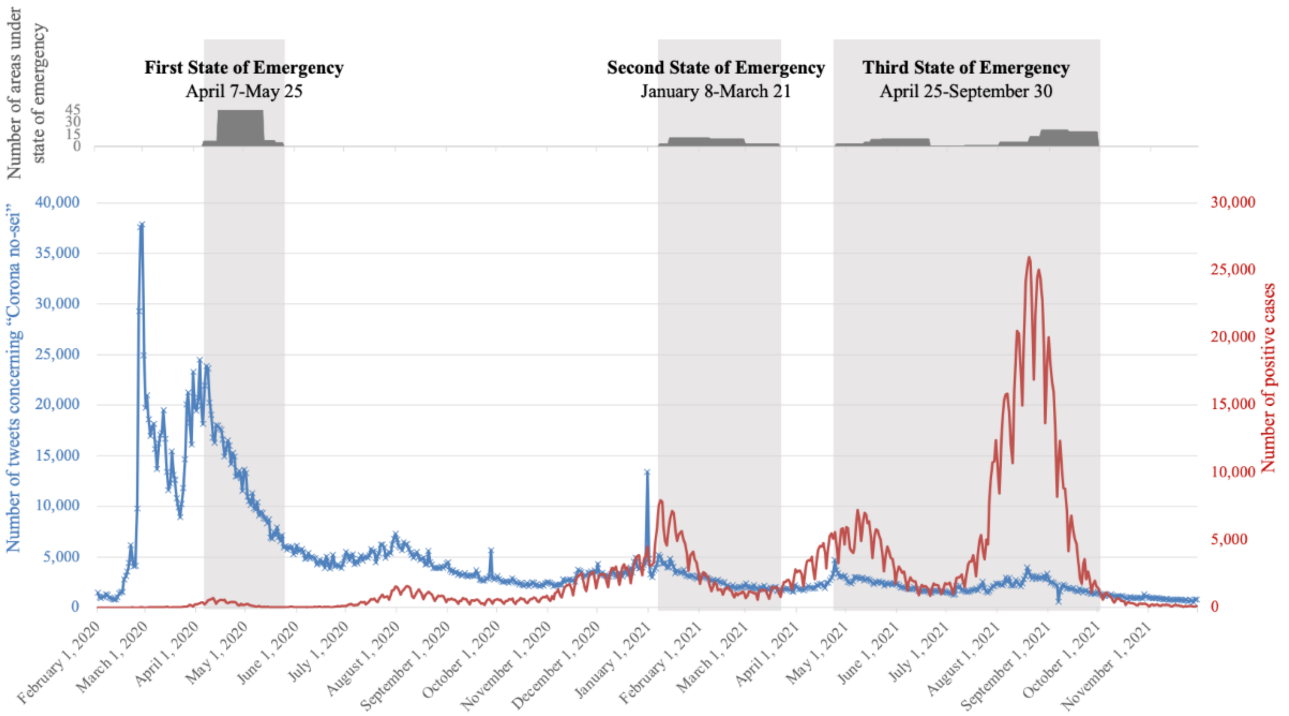
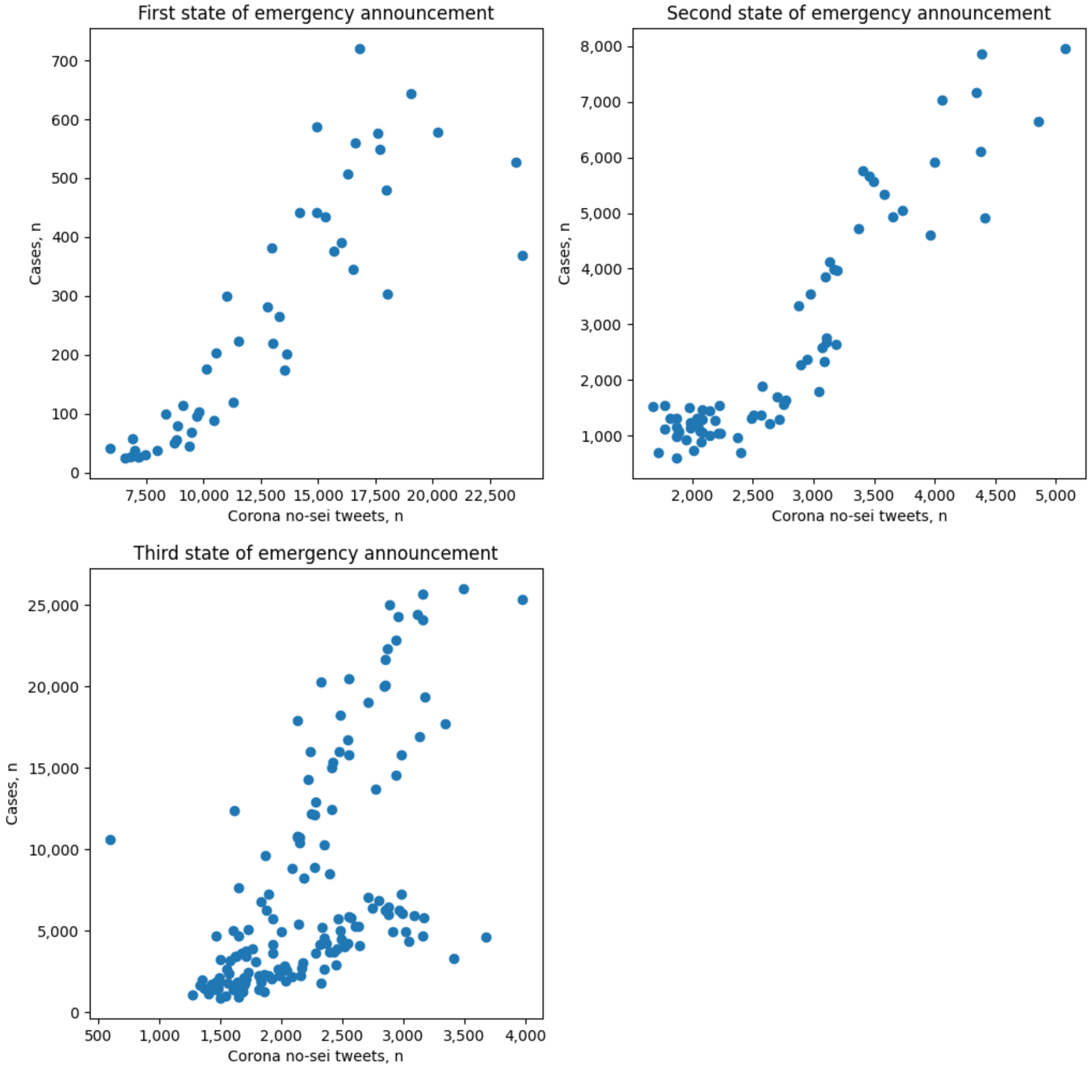


Figure 2. Scatter plot for COVID-19 case numbers and the number of *Corona no-sei* tweets for the first, second, and third state of emergency announcements.



We further investigated the nouns and verbs in the tweets that we sampled. Tables 1 and 2 show the number of tweets for the top 20 nouns and verbs tweeted on February 28, 2020, when the tweet number reached the highest level. Tables 3 and 4 show the top 20 words (nouns and verbs) that co-occurred with *Corona no-sei* tweets in descending order to highlight the most disrupted activities or plans during our data aggregation period. For nouns, here, *Corona* was excluded because it was a word included in the query and was clearly the most frequently detected. For nouns, the top 5 most frequently mentioned words were *work*, *abort*, *home*, *live*, and *friends* after excluding the words that indicate the grammatical tense. These keywords indicated that, over a longer period, Japanese individuals started developing concerns over their disrupted work and social life. For verbs, *go* was the most frequent, but in the actual tweets, it was sometimes used in the negative, and in the context, the verb was unlikely to be used in the affirmative, so the verb was likely used to indicate they *cannot go* even if it is in the affirmative in this paper (Textbox 1). Hence, the top 5 mean *go*, *can go* <negation>, *look*, *meet* <negation>, and *get out*. The results show that there are restrictions on the actions of going, seeing, and meeting as verbs. Compared with the single-day result on February 28, 2020, the concern about work appeared as the top 5 in Tables 1 and 2, suggesting that Japanese individuals placed clear emphasis on their work routines. In addition, the desire for live concerts increased over the long run, making *live concert* the fourth most frequently mentioned keyword in Table 1. Coincidentally, concerns related to friends and missing opportunities to meet up were also observed in both tables, showing the disruption of social relationships and recreational occasions. Both studies indicated that people regarded the COVID-19 pandemic as the main cause of their disrupted plans to hang out with friends or attend large public events. In addition to activities, the keyword finding also reflected the concern of resource shortage, such as toilet paper, masks, and even money, which were critical in supporting daily lives or normal health care practices.

Because the keywords indicated both long-term and short-term concerns, we cross-validated the tweet contents by selecting keyword pairs based on the top 5 keywords related to long-term concerns and those related to short-term anxiety on material shortage. A total of 160 tweets were randomly sampled based on the following keyword pairs (20 tweets each): ライブ＋行く／行けない (*live concert*+*go*/*go*<negation>); 家＋行く／行けない (home+go/go<negation>); 友達＋行く／行けない

(*friends*+*go*/*go*<negation>); 中止＋行く／行けない (*cancel*+*go*/*go*<negation>); 友達＋会う／会える (friends+meet/can meet); 仕事＋行く／行けない (*work*+*go*/*go*<negation>); トイレットペーパー＋なくなる (toilet paper+vanish); and マスク＋なくなる (*mask*+*vanish*). One of the authors annotated the tweets according to the themes reflected by the keywords. Key findings are discussed in the following section.

Table 1. The number of tweets with co-occurring nouns on February 28, 2020.

Noun	Tweets, n (%)
graduation ceremony	2154 (14.55)
cancel	1813 (12.25)
school	1498 (10.12)
work	1210 (8.17)
event	775 (5.24)
live concert	694 (4.69)
toilet paper	667 (4.51)
part-time job	639 (4.32)
school holiday	636 (4.3)
friends	616 (4.16)
disney	605 (4.09)
postponement	562 (3.8)
home	553 (3.74)
mask	531 (3.59)
new corona	330 (2.23)
test	328 (2.22)
tissue	321 (2.17)
hoax	318 (2.15)
company	304 (2.05)
next month	250 (1.69)

Table 2. The number of tweets with co-occurring verbs on February 28, 2020.

Verb	Tweets, n (%)
disappear	3265 (20.15)
go	1835 (11.33)
rest	1169 (7.22)
can go <negation>	1044 (6.44)
end	862 (5.32)
go out	754 (4.65)
look	734 (4.53)
make effort	711 (4.39)
buy	602 (3.72)
cry	576 (3.56)
vanish	556 (3.43)
can meet <negation>	524 (3.23)
play	505 (3.12)
spring rest	501 (3.09)
dies	499 (3.08)
be told	491 (3.03)
come	441 (2.72)
think	436 (2.69)
return	378 (2.33)
crush	319 (1.97)

Table 3. Noun words co-occurring with *Corona no-sei* in descending order.

Order	Noun
1	work
2	cancel
3	home
4	live concert
5	friends
6	event
7	postponement
8	stress
9	school
10	part-time job
11	company
12	new corona
13	mask
14	graduation ceremony
15	hospital
16	one person
17	opportunity
18	university
19	family
20	money

Table 4. Verb words co-occurring with *Corona no-sei* in descending order.

Order	Verb
1	go
2	can go <negation>
3	look
4	meet <negation>
5	get out
6	lose
7	make effort
8	increase
9	buy
10	lose
11	end
12	come
13	think
14	meet
15	rest
16	can go
17	decrease
18	be told
19	meet
20	play

Textbox 1. Examples of tweets posted on Twitter (Japanese tweets were translated into English).

Verb and example

- Go: “Due to COVID-19, the day I've been looking forward to going out with the guy I love has been postponed... I can't help it now and I'll accept it, but I was looking forward to it.”
- Meet: “It doesn't feel like April at all due to COVID-19, but I can't wait for it to end so that we can all meet, eat, and shop together comfortably. Six years already... I want to quit my job lol.”

Discussion

Principal Findings

Our findings revealed that the COVID-19 pandemic significantly disrupted daily routines in Japan, particularly in terms of work, education, social activities, and material shortages (with regard to the temporary spike of anxiety). The findings from our study correspond with numerous studies conducted in diverse countries, highlighting the extensive impact of the COVID-19 pandemic on social life, economy, public mental health, and education [5]. In this section, we discuss key findings across a temporal spectrum, focusing on 4 crucial aspects: disruption of work routines, public anxiety stemming from perceived resource shortages, concerns regarding social relationships, and interference with the curriculum.

Top Concerns

The impact of the COVID-19 pandemic on the labor market in Japan is unequivocal, mirroring the challenges faced by numerous countries. The pandemic necessitated a shift in work dynamics with the unintended pilot of remote collaboration. Notably, certain categories of Japanese workers, contingent on their employment contracts, exhibited heightened susceptibility to these alterations in work patterns. In our findings, the keyword *work* demonstrated associations with *part-time*, *abort*, and *money*, indicating that individuals expressing concerns about their work conditions may grapple with job uncertainty, stemming either from the part-time nature of their employment or an abrupt reduction in income. This discovery aligns seamlessly with prior research examining the repercussions of the COVID-19 pandemic on Japan’s labor sector. As described by Kikuchi et al [42] in their study, individuals in contingent employment, along with women and those with lower income, were notably susceptible. The shift toward teleworking and the accompanying uncertainty about long-term income during the COVID-19 pandemic had a disproportionately adverse impact on these specific demographic groups [42]. Fukai et al [43] endorsed these findings through extensive government statistical analysis. According to their research, Japanese individuals employed part-time in service industries or compelled to take leave or face job loss following the declaration of a state of emergency were identified as particularly vulnerable groups significantly affected by the COVID-19 pandemic [43]. Although the use of part-time or contingent workers has traditionally been a standard practice for Japanese companies seeking to optimize budget and resource allocation, the advent of the COVID-19 pandemic has pushed issues related to *work* to the forefront of public concern. Researchers caution that this could potentially exacerbate inequality for susceptible individuals unless actively addressed by government support [44].

In summary, our findings provide substantial evidence for concerns among Japanese internet users regarding job disruption, employment disparities, and inadequate financial resilience. Failing to address these issues during multiple states of emergency, the Japanese government risks compromising the equality within Japan's labor market significantly. Interestingly, a study conducted by Chen et al [27], who sampled 6535 Reddit posts, identified strikingly identical subjects that propelled nationwide anxiety in the United States. Notably, concerns about career, finance, and the future were prevalent. However, our research suggests that health and death concerns were not as prominent in Japan, as observed in the study by Chen et al [27]. We hypothesized that the emphasis on collectivism and harmony in Japanese society could shape individuals' concerns during crises (particularly in the case of a national crisis). For example, apprehensions about not being perceived as "useful" or causing "inconvenience" to others, possibly even relying on government subsidies, were more pronounced than concerns related to health and mortality.

Sudden and Perhaps Excessive Anxiety About Material Shortage

The scarcity of certain items, including toilet paper, masks, and tissues, as outlined in Table 1, emerged as a significant issue in Japan. Our findings closely parallel earlier Twitter studies investigating hoarding behaviors, particularly concerning toilet paper [45]. Although initially observed in the United States, panic buying for household goods rapidly became a global phenomenon. Notably, toilet paper has emerged as a frequently hoarded item, often signaling a surge in demand during natural disasters [46,47]. Although the act of stockpiling toilet paper may seem irrational and has been widely ridiculed on social media, the adverse effects of bulk purchasing have not been as severe. Social scientists may view this behavior as a coping mechanism during a natural disaster [48]. Contrary to the commonly perceived overhoarding of toilet paper, the mask shortage was deemed a more severe public health crisis and a direct threat to well-being. A 2020 agent-based simulation conducted by Tatapudi et al [48] illustrated that universal mask use could potentially reduce infections by 20% [49]. At the time of the study, the total number of people infected by COVID-19 was 541 million, indicating that implementing universal mask use could potentially spare 108 million cases. Numerous studies have indicated a negative correlation between mask use and the COVID-19 infection rate [44,50].

However, the situation in Japan presents a slightly different scenario. The Japanese government faced criticism for a perceived slow response to the awareness of mask shortages, as the pandemic was considered relatively "under control" in its early stages. As the mask crisis unfolded, many Japanese citizens became concerned about their reliance on masks manufactured abroad, prompting the government to take actions to boost domestic mask production. Unfortunately, heightened anxiety also led to the "Abenomasks" incident, wherein the government faced backlash for stockpiling over 82 million unused masks [51]. A crucial lesson learned from this incident is that although social media serves as a critical channel for the dissemination of news and raising public awareness, the emotional contagion and overpromotion of a particular disaster can backfire, impeding the rational coping mechanisms of citizens and the decision-making processes of the government. Our findings, along with those of numerous other studies, indicate that further efforts are needed to develop effective protocols for addressing the widely contagious anxiety stemming from the dissemination of information about natural disasters on social media.

Concerns About Social Relationships

Keywords pertaining to relationships, social life, and collective events were prevalent in our analysis. For instance, the top 20 frequently occurring nouns associated with *Corona no-sei* included friends, family, live, events, and one person. The most frequent verbs in the context of *Corona no-sei* were *go*, *can go* <negation>, *meet* <negation>, *buy*, *meet*, *can go*, and *play*. The example in Table 4 illustrates how Japanese individuals linked *go* and *meet* to their social events. While it may appear that many tweets express concerns about social relationships, these keywords actually reflect people venting their frustration about being unable to meet and engage in activities together, rather than indicating an actual loss of relationships. Interestingly, a study by Goodwin and Takahashi [52] also yielded similar findings. Most Japanese respondents in their survey gauging perceptions of relationship quality during the COVID-19 pandemic indicated that there were no discernible changes in their perceived relationship quality. Only a few reported that their trust and relationship with communities had declined compared to the prepandemic era [52]. There was also a report indicating that students, due to reduced communication with friends, face an increased risk of mental health problems [53].

These findings suggest that events, such as the COVID-19 pandemic, may lead individuals to experience heightened anxiety and stress. While this emotional response could temporarily disrupt their social activities and coping mechanisms against trauma, it may not have a lasting impact on their perceived relationship quality. In fact, the example tweets we analyzed illustrated how individuals, despite feeling frustrated, expressed eagerness to resume their social activities after the pandemic. Hence, we argue that concerns about relationship disruption may be transient and serve as a positive signal prompting individuals in Japan to actively nurture their relationships. As suggested by the study conducted by Goodwin and Takahashi [52], dedicating additional time to communication, particularly in the context of romantic relationships, could further enhance the quality of these connections [42].

Concern for Education Discontinuation

The peak volume of tweets was recorded on February 28, 2020, coinciding with the government's announcement of the simultaneous closure of all elementary, junior high, and senior high schools in Japan. In fact, in the most frequent nouns and verbs shown in Tables 1 and 2, the top words related to the simultaneous closure of schools were *graduation ceremony*, *cancel*, *lose*, *rest*, and *go* <negation>, all of which reflected Japanese citizens' concerns about the discontinuation of education, the cancelation of the graduation ceremony, and missing school classes. It is essential to note that in Japan, the graduation ceremony typically takes place in March and the new school and work year commences in April. Despite the Japanese government's earnest efforts to mitigate the spread of COVID-19, as scrutinized by scientists, the decision to close schools in Japan did not yield a substantial impact on preventing the spread of COVID-19. Instead, it deprived children of valuable learning and developmental opportunities [54]. Moreover, with the closure of schools, there was a surge in the demand for digital education or internet-based learning platforms. However, many schools and student households were ill-equipped to handle this impromptu shift to an internet-based education system. As discussed in detail by Iwabuchi et al [55], the unequal distribution of resources among schools in Japan further intensified the digital learning disparities brought about by the COVID-19 pandemic–induced school closures. The more well-funded private and prefecture-sponsored schools had often already implemented or could quickly set up the necessary e-learning system to cope with the lack of face-to-face lecturing. However, most public schools were forced to send learning materials to students by mail, risking a huge learning gap between students in private and public schools. The long-term impact on students' physical and mental development remains uncertain, given that most schools were able to resume normalcy after the lifting of the state of emergency. A study conducted by Nishimura et al [56] on medical students clearly indicated a deterioration in subjective mental well-being.

Concerns were also observed regarding web-based alternatives, with growing apprehensions that they fail to adequately substitute the essential in-person learning and hands-on field practice integral to medical education. The diverse concerns reflected in education-related keywords in Table 1 suggest that many Japanese individuals transitioned their focus from one-time events, such as *graduation ceremony* and *school holiday*, to longer-term mental and societal impacts, such as *opportunity*, *stress*, and *university*. This shift implies that the long-term effects would take time to manifest compared to short-term disruptions of specific incidents, such as a graduation ceremony. Further studies are crucial to monitor and unveil a complete picture of this disruption.

Long- and Short-Term Concerns and the Impact on the Society

Following the World Health Organization's official declaration that COVID-19 was no longer considered a global health emergency on May 4, 2023, individuals who survived now faced a familiar daily life with some changes that were difficult to imagine in the prepandemic era. However, there is still an impact on society that can be challenging to trace and monitor. The economic repercussion, such as inflation and tumbling currency values in Japan, are gradually occurring. Schoolchildren who have lost education for almost 1 year are bracing for their future growth. An increasing number of companies are eager to get talent to opt in for remote working styles to attract employees who were reluctant to return to city offices. Individuals are probably no longer worried about toilet paper but will gradually sense the subtle shifting of their workstyles, social styles, and even learning styles. However, due to the limitations of our data, we were not able to speculate about the postpandemic future. Our discussion offers possible clues to further trace the causes of societal changes. The profound effects of the COVID-19 pandemic on society and public health require further investigation and monitoring.

Limitations and Future Work

It should be noted that our study had some limitations in extracting data from social platforms such as Twitter. One limitation is the lack of geolocation metadata. Although we capitalized on the language exclusivity of Japanese tweets and the domestic majority to extract representative samples of Japanese citizens, it is important to note that there may be some minor contributions from Japanese speakers residing outside Japan. This limitation should be considered when interpreting the findings of this study. Another limitation arises from the bias present on Twitter, as its use is lower among older adults compared to the younger population. To mitigate this bias, stratified analysis is necessary to account for the effects of age. However, the current system lacks age data. Consequently, the results should be interpreted with the awareness that the perspectives of the older adults are underrepresented.

Because the purpose of this study was to derive an interruption schedule, we specifically targeted verbs and nouns to better represent social connections (families and friends), locations, events, subjects, and actions, rather than using adjectives or phrases that might focus on emotional descriptions or concrete situations. This approach limited our options for sentiment-related analysis methods or topic modeling, which could reveal emotional reactions instead of generic events and the involvement of close connections. Although people's sentiments were deemed beyond the scope of this study, in future studies, we would like to analyze how people's sentiments have changed through sentiment analysis [57]. With the introduction of transformer-based large language models, such as bidirectional encoder representations from transformers and text-to-text transfer transformers, more contextual and in-depth understanding and analysis might be made available for researchers in social media data. This should be considered in future studies.

We also did not address concerns regarding the safety of cybersecurity during the work-from-home period during the pandemic. We noticed that in the United States, data breaches and the security of the work environment were one of the top concerns [58]; however, based on our current results, there was no direct implication on this aspect in Japan during the

COVID-19 pandemic. This will be considered in our future work.

Conclusions

Overall, by adding the analysis on *Corona no-sei* to the conventional symptom-based monitoring, we were able to identify the underlying concerns at the peak of the disruption and across the whole-time span of the 3 announcements of state of emergency. Our findings and a comparison of the tweets against COVID-19 case numbers yielded rich insights into people's short- and long-term concerns and potential aspects of societal impact caused by the announcements of the state of emergency. Although more studies from different fields would help to reveal the whole landscape of social and psychological impact caused by COVID-19, we believed that the keywords reflected in *Corona no-sei* tweets provided more nuanced descriptions of real-life problems Japanese individuals faced during the COVID-19 pandemic and revealed the development of different concerns in response to the change of policies.

Timely communication of analysis results is crucial, especially when dealing with issues of significant social impact, such as a global pandemic. A delay in delivering results can hinder decision-making processes and require substantial resources to recover from the initial losses caused by poor decisions. For policy makers, especially the Japanese government, this study reflects the opinions of citizens and should be considered when reviewing the effectiveness and suitability of a policy as well as assessing further measures to support those impacted during the pandemic.

Acknowledgments

This research was conducted as part of the COVID-19 AI & Simulation Project run by Mitsubishi Research Institute commissioned by the Cabinet Secretariat. The methodology used in this study was developed through the support from Japan Science and Technology Agency (JST) Strategic International Collaborative Research Program (SICORP) (grant JPMJSC2107) and Japan Society for the Promotion of Science (JSPS) KAKENHI (grant JP 22K12041) in Japan. The authors acknowledge Dr Naoki Yoshinaga, Dr Masashi Toyoda, and Dr Masaru Kitsuregawa for providing the tweet-analysis system.

Conflicts of Interest

None declared.

References

1. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): a review. Int J Surg 2020 Jun;78:185-193 [FREE Full text] [doi: 10.1016/j.ijsu.2020.04.018] [Medline: 32305533]
2. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. Lancet Infect Dis 2020 Jun;20(6):631-633 [FREE Full text] [doi: 10.1016/S1473-3099(20)30190-0] [Medline: 32213329]
3. Qian M, Jiang J. COVID-19 and social distancing. Z Gesundh Wiss 2022 May 25;30(1):259-261 [FREE Full text] [doi: 10.1007/s10389-020-01321-z] [Medline: 32837835]
4. Sen-Crowe B, McKenney M, Elkbuli A. Social distancing during the COVID-19 pandemic: staying home save lives. Am J Emerg Med 2020 Jul;38(7):1519-1520 [FREE Full text] [doi: 10.1016/j.ajem.2020.03.063] [Medline: 32305155]
5. Belot M, Choi S, Jamison JC, Papageorge NW, Tripodi E, van den Broek-Altenburg E. Six-country survey on COVID-19. SSRN J 2020. [doi: 10.2139/ssrn.3596697]
6. McBride O, Murphy J, Shevlin M, Gibson-Miller J, Hartman TK, Hyland P, et al. Monitoring the psychological, social, and economic impact of the COVID-19 pandemic in the population: context, design and conduct of the longitudinal COVID-19 psychological research consortium (C19PRC) study. Int J Methods Psychiatr Res 2021 Mar;30(1):e1861 [FREE Full text] [doi: 10.1002/mpr.1861] [Medline: 33166018]
7. Andersen L, Hansen T, Johannesen N, Sheridan A. Pandemic, shutdown and consumer spending: lessons from scandinavian policy responses to COVID-19. arXiv Preprint posted online May 10, 2020 [FREE Full text] [doi: 10.48550/arXiv.2005.04630]
8. Alexander D, Karger E. Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior (working paper). Federal Reserve Bank of Chicago. 2020. URL: https://ideas.repec.org/p/fip/fedhwp/87999.html [accessed 2024-03-21]
9. Glorin S. A descriptive study on cybersecurity challenges of working from home during COVID-19 pandemic and a proposed 8 step WFH cyber-attack mitigation plan. Commun CIBIMA 2021 Feb 17;2021:1-7 [FREE Full text] [doi: 10.5171/2021.589235]
10. Hakim AJ, Victory KR, Chevinsky JR, Hast MA, Weikum D, Kazazian L, et al. Mitigation policies, community mobility, and COVID-19 case counts in Australia, Japan, Hong Kong, and Singapore. Public Health 2021 May;194:238-244 [FREE Full text] [doi: 10.1016/j.puhe.2021.02.001] [Medline: 33965795]
11. Kuniya T. Evaluation of the effect of the state of emergency for the first wave of COVID-19 in Japan. Infect Dis Model 2020;5:580-587 [FREE Full text] [doi: 10.1016/j.idm.2020.08.004] [Medline: 32844135]

12. Karako K, Song P, Chen Y, Tang W, Kokudo N. Overview of the characteristics of and responses to the three waves of COVID-19 in Japan during 2020-2021. Biosci Trends 2021 Mar 15;15(1):1-8 [FREE Full text] [doi: 10.5582/bst.2021.01019] [Medline: 33518668]
13. Kobayashi G, Sugasawa S, Tamae H, Ozu T. Predicting intervention effect for COVID-19 in Japan: state space modeling approach. Biosci Trends 2020 Jul 17;14(3):174-181 [FREE Full text] [doi: 10.5582/bst.2020.03133] [Medline: 32461511]
14. Yamamura E, Tsutsui Y. How does the impact of the COVID-19 state of emergency change? An analysis of preventive behaviors and mental health using panel data in Japan. J Jpn Int Econ 2022 Jun;64:101194 [FREE Full text] [doi: 10.1016/j.jjie.2022.101194] [Medline: 35125647]
15. Katafuchi KY, Kurita K, Managi S. COVID-19 with stigma: theory and evidence from mobility data. Econ Disaster Clim Chang 2021;5(1):71-95 [FREE Full text] [doi: 10.1007/s41885-020-00077-w] [Medline: 32984755]
16. Kitamura Y, Karkour S, Ichisugi Y, Itsubo N. Evaluation of the economic, environmental, and social impacts of the COVID-19 pandemic on the Japanese tourism industry. Sustainability 2020 Dec 09;12(24):10302. [doi: 10.3390/su122410302]
17. Japan GDP 1960-2024 by World Bank. Macrotrends LLC. URL: https://tinyurl.com/4pdrtsuh [accessed 2024-02-15]
18. Japan unemployment rate 1991-2024 by World Bank. Macrotrends LLC. URL: https://tinyurl.com/bjv6ku2v [accessed 2024-02-15]
19. Watanabe T, Yabu T. Japan's voluntary lockdown. PLoS One 2021 Jun 10;16(6):e0252468 [FREE Full text] [doi: 10.1371/journal.pone.0252468] [Medline: 34111163]
20. Watanabe T, Yabu T. Japan's voluntary lockdown: further evidence based on age-specific mobile location data. Jpn Econ Rev (Oxf) 2021;72(3):333-370 [FREE Full text] [doi: 10.1007/s42973-021-00077-9] [Medline: 34177343]
21. Mizuno T, Ohnishi T, Watanabe T. Visualizing social and behavior change due to the outbreak of COVID-19 using mobile phone location data. New Gener Comput 2021 Nov 02;39(3-4):453-468 [FREE Full text] [doi: 10.1007/s00354-021-00139-x] [Medline: 34744249]
22. Aramaki E, Wakamiya S, Yada S, Nakamura Y. Natural language processing: from bedside to everywhere. Yearb Med Inform 2022 Aug 02;31(1):243-253 [FREE Full text] [doi: 10.1055/s-0042-1742510] [Medline: 35654422]
23. Mahmoud E, Émilien 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 URL: https://www.scitepress.org/PublishedPapers/2021/104145/104145.pdf [doi: 10.5220/0010414508250832]
24. Izutsu MN, Izutsu K. Why is Twitter so popular in Japan? Internet Pragmat 2019 Jul 16;2(2):260-289 [FREE Full text] [doi: 10.1075/ip.00030.izu]
25. Shanthakumar SG, Seetharam A, Ramesh A. Understanding the societal disruption due to COVID-19 via user tweets. In: Proceedings of the 2021 IEEE International Conference on Smart Computing. 2021 Presented at: SMARTCOMP '21; August 23-27, 2021; Irvine, CA p. 137-144 URL: https://tinyurl.com/5ye73y5j [doi: 10.1109/smartcomp52413.2021.00039]
26. Lee H, Noh EB, Choi SH, Zhao B, Nam EW. Determining public opinion of the COVID-19 pandemic in South Korea and Japan: social network mining on Twitter. Healthc Inform Res 2020 Oct;26(4):335-343 [FREE Full text] [doi: 10.4258/hir.2020.26.4.335] [Medline: 33190468]
27. Chen LL, Wilson SR, Lohmann S, Negraia DV. What are you anxious about? Examining subjects of anxiety during the COVID-19 pandemic. Proc Int AAAI Conf Web Soc Media 2023 Jun 02;17:137-148. [doi: 10.1609/icwsm.v17i1.22133]
28. Li I, Li Y, Li T. What are we depressed about when we talk about COVID-19: mental health analysis on tweets using natural language processing. In: Proceedings of the 40th SGAI International Conference on Artificial Intelligence. 2020 Presented at: SGAI '20; December 15-17, 2020; Cambridge, UK URL: https://doi.org/10.1007/978-3-030-63799-6_27 [doi: 10.1007/978-3-030-63799-6_27]
29. 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]
30. Krittanawong C, Narasimhan B, Virk HU, Narasimhan H, Hahn J, Wang Z, et al. Misinformation dissemination in Twitter in the COVID-19 era. Am J Med 2020 Dec;133(12):1367-1369 [FREE Full text] [doi: 10.1016/j.amjmed.2020.07.012] [Medline: 32805227]
31. 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]
32. Ferawati K, Liew K, Aramaki E, Wakamiya S. Monitoring mentions of COVID-19 vaccine side effects on Japanese and Indonesian Twitter: infodemiological study. JMIR Infodemiology 2022;2(2):e39504 [FREE Full text] [doi: 10.2196/39504] [Medline: 36277140]
33. Gao Z, Fujita S, Shimizu N, Liew K, Murayama T, Yada S, et al. Measuring public concern about COVID-19 in Japanese internet users through search queries: infodemiological study. JMIR Public Health Surveill 2021 Jul 20;7(7):e29865 [FREE Full text] [doi: 10.2196/29865] [Medline: 34174781]
34. Uehara M, Fujita S, Shimizu N, Liew K, Wakamiya S, Aramaki E. Measuring concerns about the COVID-19 vaccine among Japanese internet users through search queries. Sci Rep 2022 Sep 03;12(1):15037 [FREE Full text] [doi: 10.1038/s41598-022-18307-4] [Medline: 36057657]

35. Ishida T, Seki Y, Kashino W, Kando N. Extracting citizen feedback from social media by appraisal opinion type viewpoint. J Nat Lang Process 2022;29(2):416-442. [doi: 10.5715/jnlp.29.416]
36. Special site for coronavirus in Japan. NHK. URL: https://www3.nhk.or.jp/news/special/coronavirus/data/ [accessed 2023-04-19]
37. Yoshinaga N, Kitsuregawa M. A self-adaptive classifier for efficient text-stream processing. In: Proceedings of the 25th International Conference on Computational Linguistics: Technical Papers. 2014 Presented at: COLING '14; August 23-29, 2014; Dublin, Ireland p. 1091-1102 URL: https://aclanthology.org/C14-1103.pdf
38. Yoshinaga N, Kitsuregawa M. Polynomial to linear: Efficient classification with conjunctive features. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. 2009 Presented at: EMNLP '09; August 6-7, 2009; Singapore p. 1542-1551 URL: https://aclanthology.org/D09-1160.pdf [doi: 10.3115/1699648.1699701]
39. Yoshinaga N, Kitsuregawa M. Kernel slicing: scalable online training with conjunctive features. In: Proceedings of the 23rd International Conference on Computational Linguistics. 2010 Presented at: COLING '10; August 23-27, 2010; Beijing, China p. 1245-1253 URL: https://aclanthology.org/C10-1140.pdf
40. Oh JH, Torisawa K, Hashimoto C, Iida R, Tanaka M, Kloetzer J. A semi-supervised learning approach to why-question answering. Proc Int AAAI Conf Web Soc Media 2016 Mar 05;30(1):10388. [doi: 10.1609/aaai.v30i1.10388]
41. Ethical guidelines for life sciences and medical research involving human subjects. NHK. URL: https://www.meti.go.jp/press/2021/03/2022031 [accessed 2022-06-28]
42. Kikuchi S, Kitao S, Mikoshiba M. Who suffers from the COVID-19 shocks? Labor market heterogeneity and welfare consequences in Japan. J Jpn Int Econ 2021 Mar;59:101117. [doi: 10.1016/j.jjie.2020.101117]
43. Fukai T, Ichimura H, Kawata K. Describing the impacts of COVID-19 on the labor market in Japan until June 2020. Jpn Econ Rev (Oxf) 2021 Jul 19;72(3):439-470 [FREE Full text] [doi: 10.1007/s42973-021-00081-z] [Medline: 34305434]
44. Kikuchi S, Kitao S, Mikoshiba M. Heterogeneous vulnerability to the COVID-19 crisis and implications for inequality in Japan. Research Institute of Economy, Trade, and Industry. 2020 Apr. URL: https://www.rieti.go.jp/jp/publications/dp/20e039.pdf [accessed 2024-03-05]
45. Leung J, Chung JY, Tisdale C, Chiu V, Lim CC, Chan G. Anxiety and panic buying behaviour during COVID-19 pandemic-a qualitative analysis of toilet paper hoarding contents on Twitter. Int J Environ Res Public Health 2021 Jan 27;18(3):1127 [FREE Full text] [doi: 10.3390/ijerph18031127] [Medline: 33514049]
46. Kaigo M. Social media usage during disasters and social capital: Twitter and the Great East Japan earthquake. Keio Commun Rev 2012;34(1):19-35 [FREE Full text]
47. Kirk CP, Rifkin LS. I'll trade you diamonds for toilet paper: consumer reacting, coping and adapting behaviors in the COVID-19 pandemic. J Bus Res 2020 Sep;117:124-131 [FREE Full text] [doi: 10.1016/j.jbusres.2020.05.028] [Medline: 32834208]
48. Tatapudi H, Das R, Das TK. Impact assessment of full and partial stay-at-home orders, face mask usage, and contact tracing: an agent-based simulation study of COVID-19 for an urban region. Glob Epidemiol 2020 Nov;2:100036 [FREE Full text] [doi: 10.1016/j.gloepi.2020.100036] [Medline: 33103108]
49. Catching A, Capponi S, Yeh MT, Bianco S, Andino R. Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of COVID-19. Sci Rep 2021 Aug 06;11(1):15998 [FREE Full text] [doi: 10.1038/s41598-021-94960-5] [Medline: 34362936]
50. Adjodah D, Dinakar K, Chinazzi M, Fraiberger SP, Pentland A, Bates S, et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. PLoS One 2021 Jun 23;16(6):e0252315 [FREE Full text] [doi: 10.1371/journal.pone.0252315] [Medline: 34161332]
51. Osaki T. Abenomask? Prime minister's 'two masks per household' policy spawns memes on social media. The Japan Times. 2020 Apr 2. URL: https://www.japantimes.co.jp/news/2020/04/02/national/ [accessed 2024-03-05]
52. Goodwin R, Takahashi M. Anxiety, past trauma and changes in relationships in Japan during COVID-19. J Psychiatr Res 2022 Jul;151:377-381 [FREE Full text] [doi: 10.1016/j.jpsychires.2022.04.032] [Medline: 35567802]
53. Tahara M, Mashizume Y, Takahashi K. Mental health crisis and stress coping among healthcare college students momentarily displaced from their campus community because of COVID-19 restrictions in Japan. Int J Environ Res Public Health 2021 Jul 06;18(14):7245 [FREE Full text] [doi: 10.3390/ijerph18147245] [Medline: 34299694]
54. Fukumoto K, McClean CT, Nakagawa K. No causal effect of school closures in Japan on the spread of COVID-19 in spring 2020. Nat Med 2021 Dec 27;27(12):2111-2119 [FREE Full text] [doi: 10.1038/s41591-021-01571-8] [Medline: 34707318]
55. Iwabuchi K, Hodama K, Onishi Y, Miyazaki S, Nakae S, Suzuki KH. COVID-19 and education on the front lines in Japan: what caused learning disparities and how did the government and schools take initiative? In: Reimers FM, editor. Primary and Secondary Education During Covid-19: Disruptions to Educational Opportunity During a Pandemic. Cham, Switzerland: Springer; 2022:125-151.
56. Nishimura Y, Ochi K, Tokumasu K, Obika M, Hagiya H, Kataoka H, et al. Impact of the COVID-19 pandemic on the psychological distress of medical students in Japan: cross-sectional survey study. J Med Internet Res 2021 Feb 18;23(2):e25232 [FREE Full text] [doi: 10.2196/25232] [Medline: 33556033]

57. Ainley E, Witwicki C, Tallett A, Graham C. Using Twitter comments to understand people's experiences of UK health care during the COVID-19 pandemic: thematic and sentiment analysis. J Med Internet Res 2021 Oct 25;23(10):e31101 [FREE Full text] [doi: 10.2196/31101] [Medline: 34469327]
58. Glorin S. Cyber kill chain analysis of five major US data breaches: lessons learnt and prevention plan. Int J Cyber Warf Terror 2022;12(1):1-15 [FREE Full text] [doi: 10.4018/ijcwt.315651]

Edited by T Mackey; submitted 06.06.23; peer-reviewed by L Chen, Y Pachankis, G Sebastian, M Elbattah; comments to author 29.06.23; revised version received 11.08.23; accepted 06.03.24; published 01.04.24.

Please cite as:
Kamba M, She WJ, Ferawati K, Wakamiya S, Aramaki E
Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences
JMIR Infodemiology 2024;4:e49699
URL: https://infodemiology.jmir.org/2024/1/e49699
doi:10.2196/49699
PMID:38557446

Original Paper

Collective Intelligence–Based Participatory COVID-19 Surveillance in Accra, Ghana: Pilot Mixed Methods Study

Gifty Marley[1*], BSc, MSc, PhD; Phyllis Dako-Gyeke[2*], PhD; Prajwol Nepal[1], PhD; Rohini Rajgopal[1], MScPH, MBA; Evelyn Koko[2], MBA; Elizabeth Chen[3], PhD; Kwabena Nuamah[4], PhD; Kingsley Osei[4], MSc; Hubertus Hofkirchner[5], PhD; Michael Marks[6,7], PhD; Joseph D Tucker[6,8], PhD; Rosalind Eggo[6], PhD; William Ampofo[9], PhD; Sean Sylvia[1], PhD

[1]Department of Health Policy and Management, University of North Carolina, Chapel Hill, NC, United States
[2]Department of Social and Behavioral Sciences, School of Public Health, University of Ghana, Accra, Ghana
[3]Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States
[4]Cognate Systems Company Limited, Accra, Ghana
[5]Prediki Prediction Markets GmbH, Vienna, Austria
[6]Clinical Research Department, London School of Hygiene and Tropical Medicine, London, United Kingdom
[7]Division of Infection and Immunity, University College London, London, United Kingdom
[8]Institute for Global Health and Infectious Diseases, University of North Carolina, Chapel Hill, NC, United States
[9]Noguchi Memorial Institute of Medical Research, University of Ghana, Accra, Ghana
[*]these authors contributed equally

Corresponding Author:
Sean Sylvia, PhD
Department of Health Policy and Management, University of North Carolina
1101D McGavran-Greenberg Hall
CB #7411 Chapel Hill, NC 27599-7411
Chapel Hill, NC, 27599
United States
Phone: 1 919 966 6328
Email: sysylvia@email.unc.edu

Abstract

Background: Infectious disease surveillance is difficult in many low- and middle-income countries. Information market (IM)–based participatory surveillance is a crowdsourcing method that encourages individuals to actively report health symptoms and observed trends by trading web-based virtual “stocks” with payoffs tied to a future event.

Objective: This study aims to assess the feasibility and acceptability of a tailored IM surveillance system to monitor population-level COVID-19 outcomes in Accra, Ghana.

Methods: We designed and evaluated a prediction markets IM system from October to December 2021 using a mixed methods study approach. Health care workers and community volunteers aged ≥18 years living in Accra participated in the pilot trading. Participants received 10,000 virtual credits to trade on 12 questions on COVID-19–related outcomes. Payoffs were tied to the cost estimation of new and cumulative cases in the region (Greater Accra) and nationwide (Ghana) at specified future time points. Questions included the number of new COVID-19 cases, the number of people likely to get the COVID-19 vaccination, and the total number of COVID-19 cases in Ghana by the end of the year. Phone credits were awarded based on the tally of virtual credits left and the participant’s percentile ranking. Data collected included age, occupation, and trading frequency. In-depth interviews explored the reasons and factors associated with participants’ user journey experience, barriers to system use, and willingness to use IM systems in the future. Trading frequency was assessed using trend analysis, and ordinary least squares regression analysis was conducted to determine the factors associated with trading at least once.

Results: Of the 105 eligible participants invited, 21 (84%) traded at least once on the platform. Questions estimating the national-level number of COVID-19 cases received 13 to 19 trades, and obtaining COVID-19–related information mainly from television and radio was associated with less likelihood of trading (marginal effect: −0.184). Individuals aged <30 years traded 7.5 times more and earned GH ¢134.1 (US $11.7) more in rewards than those aged >30 years (marginal effect: 0.0135). Implementing the IM surveillance was feasible; all 21 participants who traded found using IM for COVID-19 surveillance acceptable. Active trading by friends with communal discussion and a strong onboarding process facilitated participation. The lack of bidirectional communication on social media and technical difficulties were key barriers.

Conclusions: Using an IM system for disease surveillance is feasible and acceptable in Ghana. This approach shows promise as a cost-effective source of information on disease trends in low- and middle-income countries where surveillance is underdeveloped, but further studies are needed to optimize its use.

(JMIR Infodemiology 2024;4:e50125) doi:10.2196/50125

KEYWORDS

information markets; participatory disease surveillance; collective intelligence; community engagement; the wisdom of the crowds; Ghana; mobile phone

Introduction

Background

Emerging infectious diseases pose substantial and persistent risks to human and animal health globally. Yet, our ability to monitor these threats is limited, especially in low- and middle-income countries (LMICs) where many infectious disease outbreaks initially emerge [1,2]. Traditional, test-based infectious disease surveillance is expensive and prone to selection bias and involves significant time lags in contexts with weak public health infrastructure. Test-based systems relying solely on recorded case information from individuals seeking medical care tend to underestimate the true disease burden [3]. Underestimation is likely to be exacerbated in settings common in LMICs, where access to formal care is limited. More affordable and reliable methods are required to deliver timely data on disease patterns and inform the response to outbreaks in LMICs as either a supplement or stopgap for test-based methods.

Several low-cost supplementary surveillance approaches have been deployed in recent years [2,4-7]. Key among these are participatory syndromic surveillance systems, such as Influenzanet and Flu Near You. These systems rely on volunteers to regularly report symptoms using brief web-based surveys, mobile apps, interactive voice response, and SMS text messages [8-13]. Participatory disease surveillance systems can, in theory, collect information from a more representative population (including those who do not seek formal care) and can provide near–real-time data on disease trends to enable rapid health response [14,15]. Challenges in outbreak control because of delays in timely data capturing during the COVID-19 pandemic demonstrated the need for participatory disease surveillance systems that can be used in LMIC settings [16,17]. Key challenges documented include low and selective participation, which introduces bias because of a lack of population representativeness [18]. Therefore, data from information market (IM)–based participatory surveillance could be used to complement traditional participatory disease surveillance systems, such as Influenzanet, and reduce the consequential impact of the lack of population representativeness [11].

Participatory disease surveillance systems that incorporate features of IMs could be an approach to circumventing some of these challenges simultaneously. IMs are a crowdsourcing approach that relies on fundamental insights from economics to encourage accurate reporting by participants and efficiently aggregate community beliefs into a single interpretable “now-cast” or forecast. In the simplest design, participants trade (buy and sell) shares in a given outcome (eg, whether or not an event will occur by a specified date), typically in a virtual marketplace. The supply and demand for shares at a given time yield a market-clearing price, which is interpretable as an aggregate index of participants’ perceptions of the likelihood of the event [19]. With active participation, this index fluctuates over time, similar to a stock price, increasing and decreasing as community perceptions evolve. Prediction markets, one form of IMs, have been widely used in industry and government to forecast product sales and make predictions regarding geopolitical events as well as in various other contexts where reliable data are scarce or difficult to collect [20-22].

IM surveillance uses this mechanism to elicit real-time information on community participants’ perceptions of disease trends. In contrast to prediction markets, the objective is not to forecast disease trends into the future but to efficiently elicit reliable perceptions from the community in real time. When designed effectively and with sufficient engagement, IM surveillance systems may prove more reliable than syndromic participatory disease surveillance systems by (1) improving scalability by motivating wide participation, (2) eliminating the need for participants to be representative of a population, (3) automatically encouraging more reporting from key demographic groups or geographic areas, and (4) efficiently aggregating participant information into indices useful for decision makers without additional modeling or confounder adjustment.

IM disease surveillance has not been evaluated in an LMIC context. Two previous studies describe case studies in higher-income countries (the United States and Taiwan) [23,24]. These studies suggest that IM surveillance is feasible and accurate for influenza and dengue fever in these settings [19,23-25].

Objectives

The objective of this study was to evaluate the approach’s feasibility for a lower-income context. We deployed an IM-based surveillance platform to monitor disease trends and associated socioeconomic indicators in Accra, Ghana, during the COVID-19 pandemic.

Methods

Study Design

This pilot study assessed the use of a tailored IM system for disease surveillance in Ghana. In IMs, participants report by trading contracts specifying payoffs tied to a future event [26]. Each participant can buy or sell shares in these contracts based on their expectations. For example, consider a contract that pays US $100 if candidate X wins an election. If the market price of an X contract is currently US $53 and the price of a Y contract (for candidate Y) is currently US $21, we would interpret this to mean that the market "believes" candidate X has a 53% chance of winning and candidate Y has a 21% chance of winning [5]. This approach functions as an information collection and aggregation mechanism: each participant knows and is incentivized to use it to make trades. The market for each contract yields prices that reflect the trader's expectations of future events [27].

Study Setting

This pilot study was conducted among residents in Accra City, Greater Accra Region, Ghana. Accra is Ghana's capital and the largest city and was one of the main COVID-19 epicenters in 2021 [28]. As Ghana's commercial and political capital, Accra receives visitors from other regions of Ghana and neighboring countries [29]. There are 16.99 million internet users in Ghana as of February 2022, of which 99.3% use smartphones [30].

Sample Size

The study included individuals aged ≥18 years who had access to a personal desktop or computer and self-reported as frequent internet users (>1 time/wk). Previous studies that used IMs to forecast infectious diseases recruited 40 to 130 participants but observed high attrition rates. Considering the potential for a low participation rate, the study assumed that 105 participants would be sufficient to achieve the pilot study goals of assessing the feasibility and acceptability of the approach. Thus, total enrollment was completed for 50 health workers and 55 community members in Accra.

Recruitment

An eligibility survey was designed based on gender; age; occupation; device ownership; and digital experience, which this study defined as proficiency in using modern technological devices, especially computers. Two types of respondents were targeted for this study: health workers and volunteer community members. A total of 5 recruiters were trained on the recruitment questionnaire, and participants were purposively recruited until a point of saturation was reached on major issues in the criteria. Health workers, including medical doctors, nurses, pharmacists, community health workers, local health officials, and laboratory technicians, were purposively recruited from clinics and hospitals in Accra. Community members were purposively recruited in different localities across Accra. Recruiters approached potential respondents in their homes and workplaces. This target included students, teachers, drivers, IT experts, web developers, and computer system engineers. Eligible health workers and community volunteers willing to participate were encouraged to promote the study among their colleagues, friends, peers, and family members. They could refer interested individuals for eligibility screening.

Data Collection

Overview

A total of 105 recruits, including 50 health workers and 55 community members, were invited to participate. The study was pretested among 4 community members and 2 health workers over 3 weeks; learnings were used to improve the questions and some platform features. The main pilot was conducted over 3 months, from October 1 to December 31, 2021. The Prediki platform was used to coordinate trading and interactions with the participants regarding expectations of future COVID-19 cases, deaths, and COVID-19 vaccine uptake in the Greater Accra region and nationwide. The buying and selling of contracts was tied to the future realizations of the outcomes based on regional statistics, as reported by the Ghana Health Service and an independent surveillance system (FluNet). Prices in these markets were used to track expectations of trends over time.

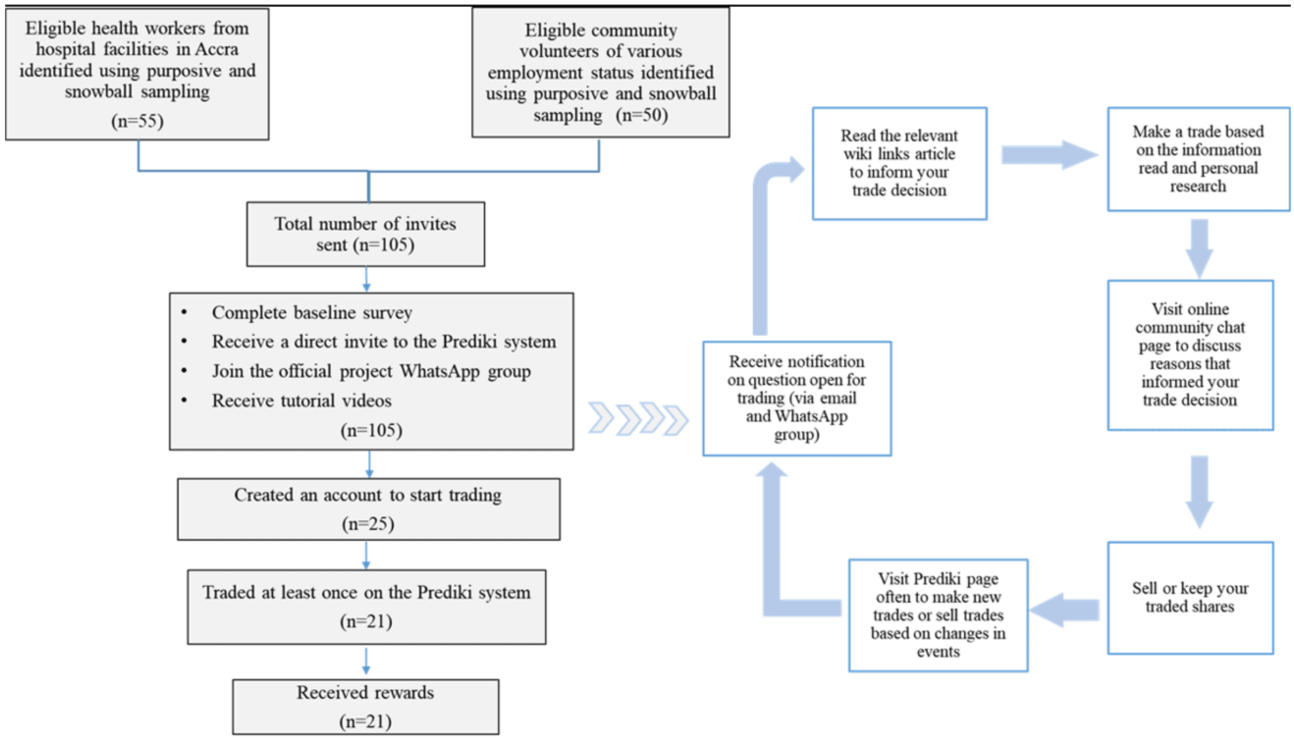
Participant Engagement

Between September 15, 2021, and October 1, 2021, all eligible participants (health care workers and community members) willing to participate in the study were recruited until the desired sample size was attained. Recruits completed a web-based baseline survey collecting sociodemographic characteristics, including age, occupation, and a major news source (such as television, radio, and social media). Participants were enrolled in a research group on WhatsApp Messenger (WhatsApp Inc) administered by the study team, where instructional videos and project-related announcements were shared to increase engagement.

Prediction Market Implementation

Figure 1 shows the step-by-step recruitment process and the stages of participating in the prediction markets survey.

Figure 1. Flowchart showing the step-by-step recruitment process and participation cycle in the COVID-19 prediction markets survey in Ghana in 2021.

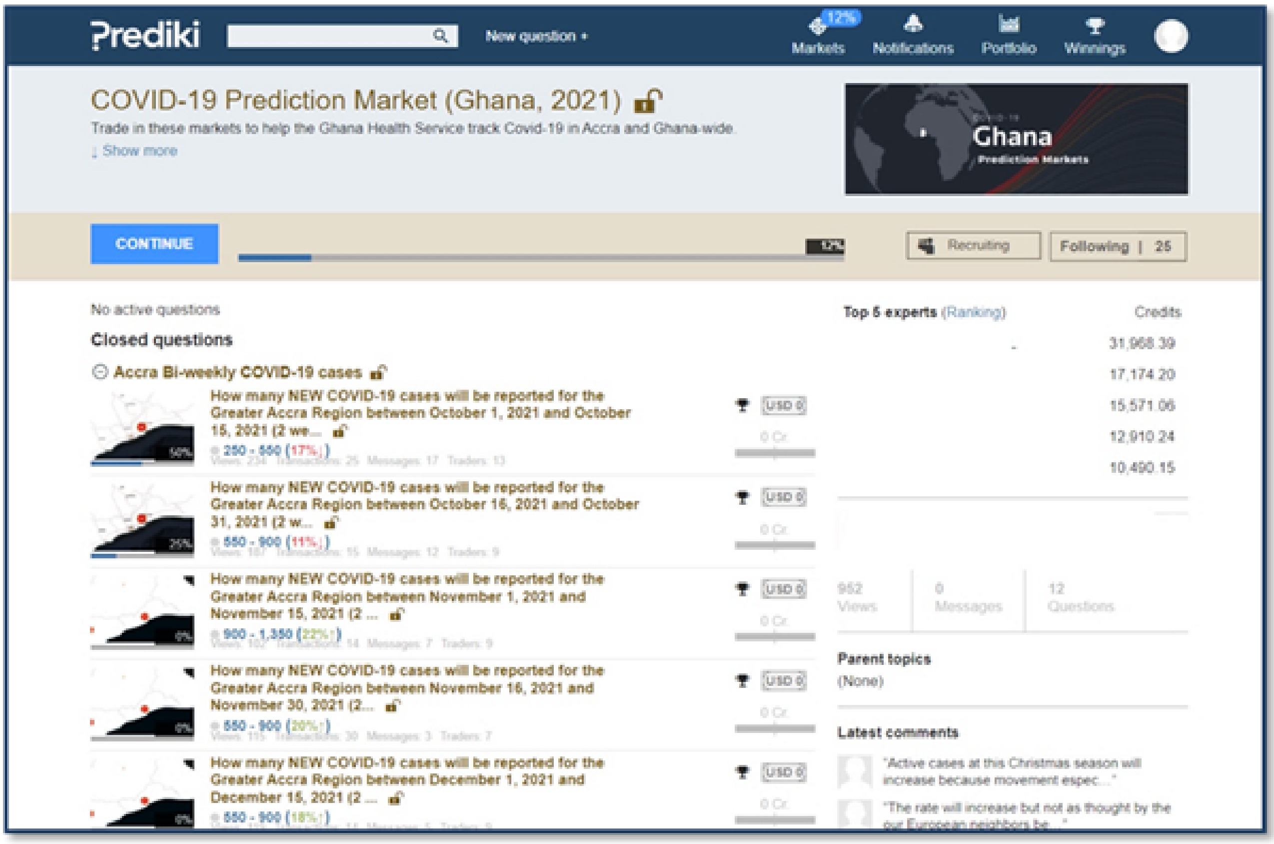


Preparation

We used the Prediki Prediction Markets 2.0 system (Prediki Prediction Markets GmbH) to design a COVID-19 prediction market for Ghana (Figure 2). Participants received individual invite links to the Prediki platform and 3 instructional videos on using prediction markets before the prediction markets opened for trading (Figure 3). One video informed participants of the need to help the Ghana Health Service monitor outbreaks and the potential of using prediction markets in crowdsourcing forecasts on COVID-19 (Multimedia Appendix 1). The other videos guided participants on how to make trades on the Prediki platform (Multimedia Appendix 2) and explained the reward mechanism to be used in the study (Multimedia Appendix 3).

Figure 2. The Prediki prediction markets interface showing the user landing page for the COVID-19 Ghana prediction markets survey.

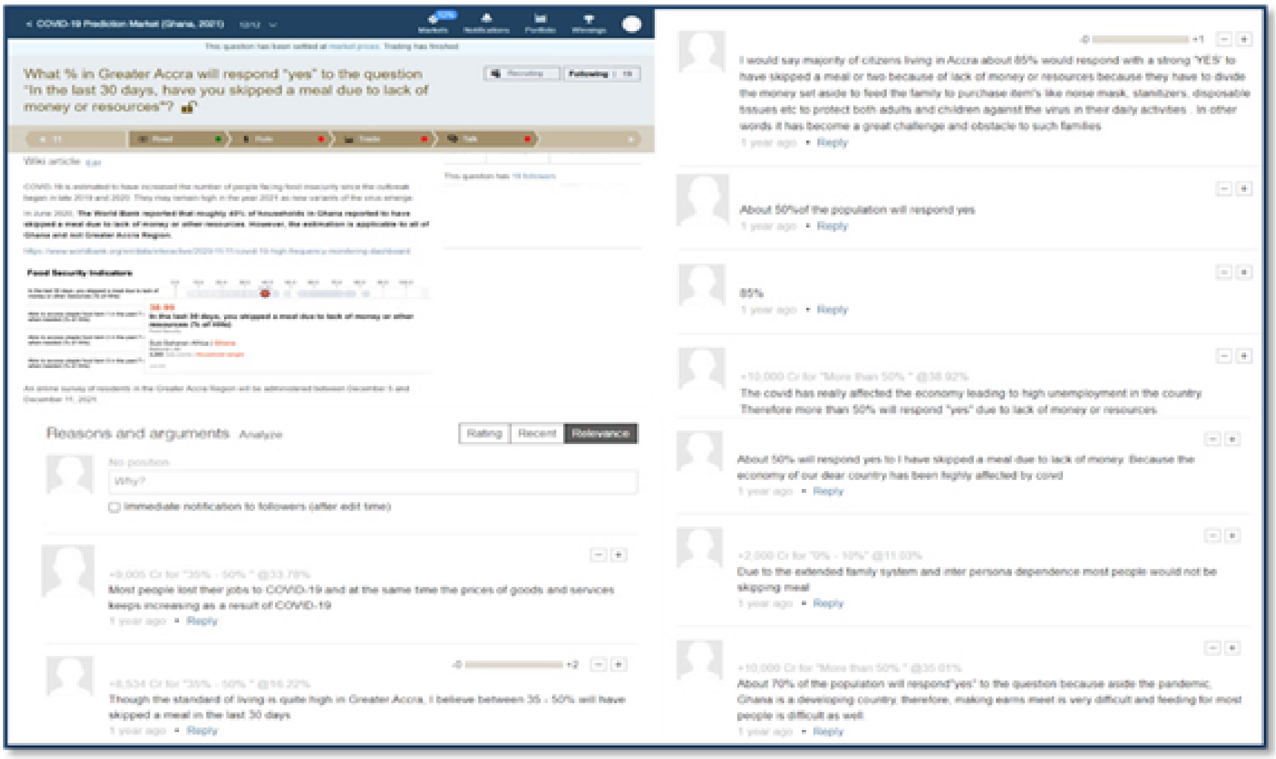
Figure 3. How questions opened for trading are depicted on the user interface of the Prediki markets system. Participants can click on each question to be directed to the trading options.



Trading

The trading on the platform began on October 1, 2021, with 2 sets of question types (Figure 3). Five questions had an end date of December 31, 2021, and were open for trading from October 1 to December 31, 2021. The second set asked 6 questions, which were made available periodically for a trading period of 2 weeks each. At the trading onset, each participant received 10,000 virtual credits for each question, with options to buy or sell the contracts. Each question had 6 prediction options and useful links to wiki articles and websites to inform trading choices (Figure 4). Each question had a judgment rule—how the result was determined—predetermined before trading was opened. In addition, participants received weekly texts and reminders to encourage participation.

Figure 4. A picture example showing some website links to relevant wiki articles and the community discussions that helped inform participants' trading decisions and choices based on ongoing events.



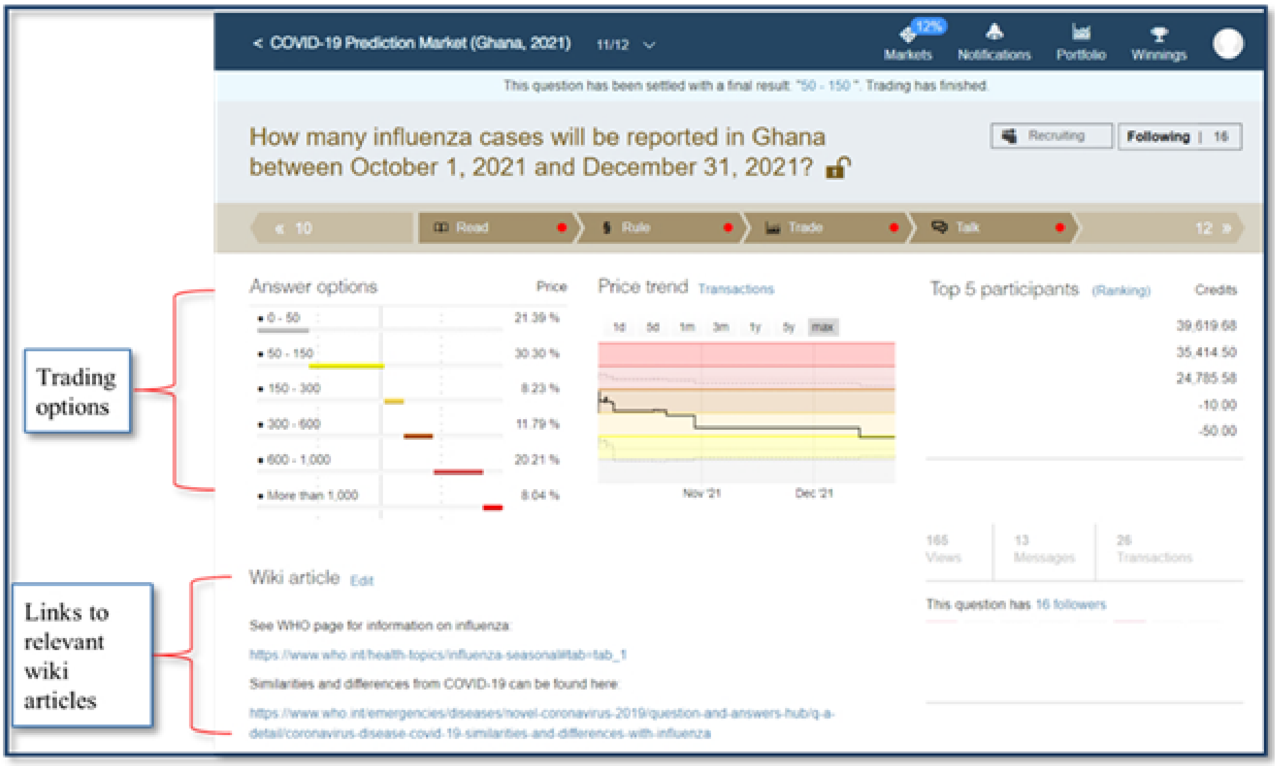
System Design

Questions on COVID-19 outcomes were judged against numbers reported in the Ghana Health Service dashboard. Rewards were assessed at the end of the trading period based on the virtual credits left for each participant. The total virtual credits at the end of the trading period were tallied, and rewards were distributed per percentile placement, that is, those in the highest 10 percentile earned the highest rewards, and the rewards decreased as placement in the percentile was further from the top 10 percentile.

Rewards

Participants received rewards in phone credits ranging from GH ¢2.7 to GH ¢134.1 (US $0.23 to US $11.5) worth. Trade results are compared to the official reports for the period after the trading period to determine reward distributions (Figure 5). After buying and selling, participants could post the reasons for their trade choices in an interactive feature called "market talk." This ensured higher participation and allowed participants to make a case for their trade (prices rose as more participants made the same trade).

Figure 5. An example of how trading results are compared to the official reports for the period after the trading period to determine reward distributions.



Data Collection

Overview

Participants completed a self-administered questionnaire before trading onset. The survey collected data on participant age, category of participants (health worker or community volunteer), employment status, and main source of information (radio, television, social media, other web-based platforms, and word of mouth), and COVID-19 knowledge. Data on trading frequencies were also extracted directly from the Prediki system.

A semistructured interview guide was used to explore the concepts of feasibility, user acceptability and engagement, changes in their health behaviors, and perceived COVID-19 risks related to the prediction markets used at the study's exit. The interviews followed a chronological format, with questions centered around the participants' experience before, during, and after participating in the activity. Participation in the in-depth interviews was open to anyone who had traded on the platform.

Process Evaluation

A total of 3 research team members were simultaneously interviewed to understand their perspectives on the implementation process of the IM activity. The interview guide was tailored to speak to the roles of these individuals. Rather than focusing on a chronological format—before, during, and after participation in the activity—the guide focused on stages in the implementation process: interest in the implementation process, planning, recruitment, execution, and reflection. The complete interview guide for this implementation team can be found in Multimedia Appendix 4. Key insights were added to an existing memo of the participant interviews.

Statistical Analysis

Descriptive statistics were used to summarize the baseline characteristics of recruited participants. We also conducted a trend analysis to assess if there was a day-of-the-week effect per time of the day (average by day) on trading frequency. Ordinary least squares regression analysis was conducted to determine the factors associated with trading at least once. All statistical analyses were conducted using STATA (StataCorp LLC) software.

Qualitative data were thematically analyzed. A codebook developed for these interviews included 3 parent codes: 1 of a topical nature, 1 focused on the participant experience, and 1 focused on a participant's journey. The journey parent code had 5 subcodes at the following levels within a participant's journey: awareness, consideration, conversion, loyalty, and advocacy (Table S1 in Multimedia Appendix 4). The codes and transcripts were added to ATLAS.ti v3.16.1 (ATLAS.ti Scientific Software Development GmbH) for the coding process. Journey maps are a human-centered approach to capturing and sharing a person's experience of learning about something (awareness), trying a solution (consideration or convert), and championing a solution (loyalty or advocacy) [31]. After the initial compilation of interview results, a high-level summary of each stage in the participant journey map as well as an ideal platform journey were shared with a few select participants in a focus group setting. This web-based process was to complete a "return of results," explaining the results of the initial findings of the qualitative interviews back to the participants and adding any additional insights to these results to contribute to this iterative process and study design.

Measures

The primary outcome was feasibility and participant acceptability. This was assessed using platform use metrics (market-specific hit rates) and qualitative interviews after the study's conclusion. The secondary outcome was prediction market acceptability, which was qualitatively explored based on participants' opinions and perceptions of the prediction market system from experience.

Ethical Considerations

The Ghana Health Service Ethics Review Committee reviewed and the Institutional Review Board at The University of North Carolina at Chapel Hill approved this study (IRB 21- 3117). An informed consent form and intake survey were administered on the web through Qualtrics. Links were sent to all eligible participants. Participants were rewarded with their payoffs in the market via virtual credits on the Prediki system. These credits were converted to phone minutes and credited to participants' accounts. Participants were rewarded as each contract was resolved (determined by how closely the prediction matched actual outcomes, as reported by the Ghana Health Service on the target date of the prediction).

Results

User Statistics

Overview

Of the 105 participants enrolled and sent links to trade on the IM platform (Prediki), 36 (34.3%) completed the baseline survey. The average age of the participants recruited was 28 (SD 5.2; range 20-45) years. Half of the participants (19/36, 53%) had received at least 1 COVID-19 vaccine shot, 75% (27/36) were sure they did not have COVID-19 at the time of the study, and all participants (36/36, 100%) had high COVID-19 knowledge (Table 1).

Table 1. Descriptive summary showing the baseline characteristics and COVID-19–related knowledge of participants who completed the baseline survey in Ghana in 2021 (n=36).

Factors	Values
Age (y), mean (SD)	29 (5)
Health workers[a], n (%)	18 (50)
Has a doctor or another health care professional diagnosed you with COVID-19?, n (%)	
No	33 (92)
Yes	1 (3)
Prefer not to answer	2 (6)
Do you think you have been infected with COVID-19?, n (%)	
No	27 (75)
Unsure	7 (19)
Yes	2 (6)
Have you received at least 1 dose of a COVID-19 vaccine?, n (%)	
Yes	19 (53)
No	17 (47)
COVID-19 can spread through mosquitoes, n (%)	
True	4 (11)
False	32 (89)
SARS-CoV-2 can live on certain surfaces for days, n (%)	
True	29 (81)
False	5 (14)
Unsure	2 (6)
Antibiotics are effective at preventing COVID-19 infection, n (%)	
True	10 (27.8)
False	18 (50.0)
Unsure	8 (22.2)
COVID-19 can spread through droplets coughed out by an infected person, n (%)	
True	36 (100)
False	—[b]
Blisters or sores on the throat are not symptoms of COVID-19, n (%)	
True	10 (28)
False	18 (50)
Unsure	8 (22)
Washed hands with soap or used hand sanitizer several times a day, n (%)	
Yes	36 (100)
No	0 (0)
Avoided public spaces, gatherings, or crowds, n (%)	
Yes	29 (81)
No	7 (19)
Avoided contact with people who could be high risk, n (%)	
Yes	30 (83)
No	6 (17)
Canceled or postponed travel for work or pleasure, n (%)	

Factors	Values
Yes	17 (47)
No	19 (53)

[a]Health workers include nurses, doctors, laboratory technicians, physician assistants, midwives, and orderlies.

[b]Not applicable.

Traders Versus Nontraders

Most trading and nontrading participants were aged ≥30 years (13/19, 68% vs 53/86, 62%; *P*=.80) and health care workers (11/19, 58% vs 44/86, 51%; *P*=.59). Most participants who traded mainly obtained their COVID-19–related information from web-based platforms and social media (12/19, 63%), whereas most nontraders mainly obtained their information from multiple sources (41/86, 48%; Table 2).

Table 2. Differences in characteristics between the recruited participants who traded at least once on the Prediki system during the 3-month active trading period in 2021.

Factors	Traded at least once, n (%)		*P* value
	No (n=86)	Yes (n=19)	
Age (y)			.58
<30	33 (38)	6 (32)	
≥30	53 (62)	13 (68)	
Occupation			.59
Health worker[a]	44 (51)	11 (58)	
Nonhealth worker[b]	42 (49)	8 (42)	
Main sources of COVID-19–related information			.09
Internet[c] or social media	33 (38)	12 (63)	
Television or radio	12 (14)	3 (16)	
Multiple sources[d]	41 (48)	4 (21)	

[a]Health workers include nurses, doctors, laboratory technicians, physician assistants, midwives, and orderlies.

[b]Nonhealth workers include students, unemployed persons, market traders, software engineers, bankers, drivers, and computer programmers.

[c]Internet includes news websites, search engines, blogs, and scientific publications.

[d]Multiple sources include web-based and offline sources, paper graphics, and word of mouth from others.

IM Uptake

Of the 105 participants invited to trade on the platform, 21 (20%) traded at least once. A total of 321 trades were made by the 21 participants across 12 questions in the platform. Questions estimating the number of COVID-19 cases at the regional level (for the Greater Accra region) received the least number of trade activities (range 6-13 trades). Comparatively, questions were used to estimate the national-level number of COVID-19 cases, and the final bonus question received the highest number of trades (range 13-19 trades). Multimedia Appendix 5 presents a detailed summary of trading activities for each question and the number of rewards paid out per question.

Factors Associated With Trading

Participants who obtained their COVID-19–related information mainly from television or radio were less likely to participate (marginal effect: −0.184) than those who obtained information from social media or internet. However, if they did trade, those who got information from television or radio made more trades and earned more rewards. Among those who traded, those aged <30 years made 7.5 times more trades than those aged >30 years (marginal effect: 0.0135) and earned GH ¢134.1 (approximately US $11.7) more in rewards. There was no substantial relationship between being a health worker and participating in any trade or among those who traded regarding the number of questions traded and rewards earned. Table 3 presents the details of the ordinary least squares regression estimations with robust SEs.

Table 3. Ordinary least squares (linear probability model) regression results showing factors associated with COVID-19 prediction markets trading uptake and frequency among Ghanaian participants in 2021 (n=21).

Variables	Any trade made, OLS[a]	Number of questions, OLS	Rewards, OLS
Age (y)			
<30	0.0135	7.548	134.1
≥30	1	1	1
Occupation			
Health worker[b]	−0.0602	0.918	11.95
Nonhealth worker[c]	1	1	1
The main source of COVID-19 information			
Multiple sources[d]	−0.0912	3.688	81.30[e]
Television or radio	−0.184[e]	4.671[f]	85.54[e]
Internet[g] or social media	1	1	1

[a]OLS: ordinary least squares.

[b]Health workers include nurses, doctors, laboratory technicians, physician assistants, midwives, and orderlies.

[c]Nonhealth workers include students, unemployed persons, market traders, software engineers, bankers, drivers, and computer programmers.

[d]Multiple sources include web-based and offline sources, paper graphics, and word of mouth from others.

[e]*P*<.05.

[f]*P*<.10.

[g]Internet includes news websites, search engines, blogs, and scientific publications.

Qualitative Results

Overall, 10 (48%) of the 21 trading participants were engaged in our in-depth interview over Zoom (Zoom Video Communications Inc). Many factors influencing the feasibility of implementing a viable IM-based disease surveillance program were identified in this study. Key facilitators identified included having (1) a strong onboarding process through the WhatsApp platform, (2) a genuine curiosity and motivation to experiment with the platform, (3) interactive features within the platform, (4) peer influence, and (5) rewards. The main barriers identified included (1) difficulty understanding the trading process and (2) various IT issues and considerations. We found little evidence that participants changed their long-term COVID-19 risk mitigating behaviors.

Facilitators

Strong Onboarding Process Through WhatsApp

A project-specific WhatsApp group was an important tool that immediately created a positive onboarding experience for participants and facilitated communication between participants and the research team. Participants could communicate directly with the designated support project staff if they had any specific questions. WhatsApp messages served as reminders for participants to check the digital platform. In addition, the introductory videos showing how to use the platform, trade, and earn more rewards were shared on the WhatsApp group, allowing participants to understand how to operate and navigate the platform, given that the videos were short and visual. Participants could also ask for technical support from designated staff on WhatsApp during the trading process and receive notifications about new open questions and trading reminders on the group page. This helped keep participants updated on questions open for trades and actively engaged in the IM process:

> *...The toggle video is the one with the cartoon. It was very helpful. Because I had my hands on many things, taking time to read was difficult, so I think the videos were helpful. You watch it in about 2 minutes and get the whole idea.* [Participant 6]

Curiosity About IMs

All participants had genuine curiosity, interest, and motivation to participate in this pilot. Of 10 participants, 6 (60%) heard about the activity from friends and chose to participate because their friends participated. Half of the participants (5/10, 50%) interviewed were curious about how IMs worked and wanted to learn more about the platform. In total, 20% (2/10) of participants wanted to understand the frequency of how COVID-19 was evolving in Ghana, and others wanted to share their ideas to give back to their community and help their country. Participating health care workers also expressed that this was a good way to share their voices on this topic as health professionals.

Furthermore, seeing other participants actively engaged on the platform motivated participants to continue trading. A total of 20% (2/10) of interviewees and some focus group discussion participants indicated that their motivation did not mainly stem from the novelty of platform features or the direct reminder messages from the administrator but rather from seeing the activeness of their peers on the platform. Moreover, understanding one's genuine interest in this material was a precursor for their acceptability of the platform when assessing their willingness to participate in future IM surveillance in their daily life. Moreover, participants cited that persons who are genuinely curious about IMs and want to explore data (perhaps a student or someone interested in research) are the ideal participants to consider when deciding whether to share this platform with a friend:

> *I just wanted to do something different from what I used to do, to learn more and to share ideas, and as a health professional, I was thinking that maybe this was right for me to put out my voice concerning this particular new disease that has come either to stay or will go away in a short while.* [Participant 3]

Interactive Features Within the Platform

Besides the trading process, the platform (Prediki) had an interactive chat feature that participants enjoyed. Many participants did not see the chat capabilities and initially felt there needed to be more interaction on the platform. However, once they discovered this feature, it allowed for an interactive experience that encouraged sharing knowledge and diverse opinions. Many participants valued the diversity of thoughts shared in the community section on the platform and deemed it a key attractor to this type of trading activity. The competitive nature of the activity attracted some participants but deterred others. While some enjoyed the back-and-forth in the comments section and partly based their predictions on others' opinions, others wished the platform was more collaborative and had participants working together more explicitly. The platform also had a ranking system that ranked participants based on how many points they had accrued from their predictions. Many participants appreciated this feature and saw it as a motivator to try and improve their predictions to increase their rank:

> *I found it interesting because, at some point, you will have conflicting ideas, you're your colleagues, and at some point, you will realize that your other colleague will come in to support your explanation that you have given to the reason why probably you chose a particular range...it was a nice way of sharing knowledge.* [Participant 1]

Peer Promotion and Rewards

Some participants (4/10, 40%) heard about the activity through a friend or had a friend participate. This was a key facilitator in getting people to sign on to the pilot and for sharing information in the future. To scale up the platform use, all participants said they would recommend it to friends and family as those people would be more trusting and willing to try and iterate how they would describe the platform to others. Other participants proposed promoting the platform's social and fun angle, while some advocated promoting the educational appeal. All participants agreed that social media should be explored to recruit more diverse participants. However, friends and family would be the starting recruitment point for future platform iterations.

In addition, rewards were deemed a strong motivator for participation in prediction markets. Several participants appreciated the phone credits but mentioned that a monetary reward would be more valuable than the phone credits. This is because some participants received phone credits through work; other participants already had beneficial phone credit bundle subscriptions. Therefore, cash rewards allowed for more flexibility in spending on other things they may need besides phone credit. Moreover, nearly all participants agreed that more active participants should receive more compensation for their time and efforts and there should be rewards for the referral of new participants:

> *...basically, I was contacted by a friend, my mate in school, who told me about it, and I did not have to think twice. I told him I was in, and that is where it started.* [Participant 8]

Barriers

Trading and Technical Difficulties

Although all participants cited a willingness to participate again if offered, understanding the trading process could have helped active engagement and enjoyment of the experience. Some participants needed clarification on the trading process even after repeatedly watching the tutorial videos. This difficulty remained across every step of the trading process, from understanding the terminology of "confirming a trade" to the concept of "altering an existing trade." Although the videos were helpful, 3 (30%) out of the 10 participants interviewed suggested having a "classroom training" session using Zoom or another platform to answer participant questions live as they watched the explanatory videos. Not only would this help answer any questions that participants may have, but as 1 (10%) participant pointed out, it would also help increase engagement in the overall activity. In addition, receiving a successful trade confirmation notification was recommended by both participants and research team members to resolve some of the confusion surrounding trading. Of 10 participants, 5 (50%) also suggested that providing detailed education and training on trading and how trading markets work earlier in the onboarding process would help people overcome these difficulties (Multimedia Appendix 4).

It was recommended that allowing for a dynamic onboarding approach where participants may communicate via oral language versus written in WhatsApp rather than static videos would provide a more educational experience. In addition, highlighting some difficulty-related questions in a frequently asked questions section on the platform for future reference was recommended. Some participants also mentioned the need for more visuals ("Maybe some videos, 2 minutes, 1-minute video" [Participant 8] and "colored and boldened so that when you get to the platform, you know that this is what you are talking about..." [Participant 7]), short excerpts, or short voice notes with detailed explanations on trading to offer a broad set of learning options. The discrepancy between points and the actual monetary value of the points earned also confused some participants as they "had no idea what they would earn" and ended up discouraged when their points gained largely exceeded the financial reward received. Implementation team members also mentioned that disseminating the rewards was cumbersome and could be streamlined in the future.

Unidirectional Communication

Besides the technical specifications of the onboarding process, participants noted that it would have been nice if members were lively and more interactive with each other from the start. Initially, the WhatsApp platform was only set up for 1-way communication (ie, from the administrator to the participants as a group). Although participants could directly message the administrator for help, many participants would have liked 2-way communication with other participants on WhatsApp. Many participants also emphasized the need for a more mobile-friendly platform in the future, as many participants engaged with the website through their laptops. Only 1 (10%) participant was technology savvy enough to interact with the site on their mobile device by keeping a window tab open in the web browser. However, not everyone would understand how to do this. Therefore, creating an app would allow for a smoother participant and referral experience (Multimedia Appendix 4).

Low Participation

Although 105 participants were recruited, only 21 (20%) traded on the platform. Of the 10 participants interviewed, 2 (20%) mentioned that their busy schedules hindered greater participation. Others may need constant reminders to remember to trade, as many Ghanaians do not check their email regularly. Therefore, 3 (30%) of the 10 participants suggested a "notifications" feature to keep people engaged daily, besides the WhatsApp reminders for a future iteration of the platform. More specifically, 1 (10%) of the 10 participants highlighted the need for notifications sent via a message (such as SMS text message or push notifications), call, or app notification, instead of directing people to their email, recognizing the effects of limited data use, particularly in settings with poor telecommunication services. The implementation team members also noted that addressing password recovery as a high-return but low-effort task would remarkably improve a participant's experience with the platform:

> *It was cool. I was very active from the beginning, but right after the end, I was not that active because I was doing my internship and helped my mom take care of me in school. So, I was working two jobs at the same time.* [Participant 4]

Competition With Other Web-Based Spaces and Pilot Delays

Although we thought participation rates would have been higher considering COVID-19's impact and global acceleration into the virtual world, this acceleration also increased the competition for people's time in other spaces (such as social media and dating apps). Another implementation team member also mentioned the long delay between the initial participant recruitment and onboarding to the platform. Therefore, the curiosity about IMs and interest in how they work would have withered when the platform was finally opened for trading. This may have contributed to lower participation rates. One of the research team members narrated that "several individuals had forgotten why the administrator was reaching out to them as a length of time had passed between being recruited and the actual launch of the platform." Although this delay is not unusual when developing and piloting new systems, the duration between recruitment and opening of the platform for trading can be reduced by ensuring that recruitment and training only commence when the platform is actively ready for use. In addition, establishing a social media group to keep participants engaged through demonstrations and bidirectional communication would have kept participants actively engaged during such waiting periods.

Discussion

Overview

Traditional infectious disease surveillance is challenging and prohibitively costly in many LMICs, increasing the importance of developing new infectious disease surveillance methods. Participatory disease surveillance offers a cheap and flexible option but is vulnerable to multiple issues that limit their ability to inform disease response. Designing participatory disease surveillance systems using IMs can address many of these challenges, and IM-based participatory disease surveillance has been successfully piloted in high-income contexts [19,24]. Our data suggest that prediction markets were feasible for engaging local Ghanaian communities in COVID-19 control responses. We also found that user feedback could be used to improve the prediction market platform. This study extends the literature by focusing on an LMIC setting and iteratively developing the prediction market with end users using participatory methods.

Principal Findings

Our results demonstrate that using IMs for infectious disease surveillance is feasible in Ghana. This finding is consistent with a small amount of literature on using IMs to enhance infectious disease surveillance [19,23-25]. Most prediction markets for infectious diseases have focused on the United States and other high-income countries, neglecting resource-limited settings [23]. This may result from more familiarity with information or prediction markets in high-income settings or fewer digital health resources in LMICs. Designing an IM system as an ad hoc structure to complement disease surveillance or integrating it into existing health monitoring structures could expand public health surveillance in resource-limited countries. IMs could effectively and efficiently capture real-time data to accelerate policy and planning. Further research is needed to scale up this approach and assess its effectiveness in tracking disease trends.

The qualitative findings from our study showed how social innovation approaches could be used to improve prototypes based on local feedback iteratively. Social innovation is a community-engaged process that links social change and health improvement [32]. Our findings corroborate this, which suggested that prediction market systems could facilitate community participation in public health monitoring, but the system design is crucial to increasing community engagement. Thus, the ability of the target population to use digital tools at varying levels, the cost of internet services, and internet coverage must be critically considered when determining the design and features of a prediction market system for LMICs. In addition, researchers should consider designing IM systems as applications that can be supported by multiple mobile phone operating software (eg, Androids, iOS, and Windows) and web browsers to cater to users of varying levels of technology savviness.

Our data also confirm observations of other studies demonstrating that IM surveys could effectively enable community participation in real-time infectious disease monitoring. IMs provide a formal mechanism to aggregate views of diverse individuals to generate forecasts that could improve response to infectious disease emergencies [19,33]. This approach may allow individuals with low income, ethnic and racial minority groups, and others to contribute to infectious disease control efforts. However, our data also highlights the need for further platform refinement. Many participants found some aspects of the user interface duplicative and not intuitive, making navigation difficult. In addition, other methods to decrease the digital divide and ensure broad access are needed. For example, an unstructured supplementary service data interface could allow people to participate without a smartphone.

Low participation could hinder the successful use of IMs for infectious disease surveillance in LMICs. The participation rate that we observed (21/105, 20%) is similar to the findings of another study, which reported that only 20% of their study participants accounted for 80% of trades made during the season under review [24]. Despite this, relatively low take-up may be less consequential for IM-based systems than for more traditional participatory disease surveillance systems [18]. Therefore, IM surveillance data could complement traditional participatory disease surveillance and reduce the impact of the lack of population representativeness. Moreover, there are multiple ways in which participation can be increased, such as tweaking platform design, enabling more engagement, and offering optional notifications and trading updates.

Beyond implementation in an LMIC, our study built on the existing literature by expanding the sample population to a broader set of participants (including general community members in addition to health care professionals) and included human-centered design thinking principles in the methods to better capture participant experience. Although this study was not designed to evaluate the quantitative results of the IM as they related to existing surveillance methods, which is left for future studies, our pilot data have implications for research and policy. From a research perspective, this study shows that IMs for infectious disease surveillance are feasible in LMIC settings. On the basis of the feedback from the participant and implementation team interviews and focus group discussions, we have compiled the ideal participant journey map for a future version of an IM-based platform (Figure S1 in Multimedia Appendix 4). Further research is needed to ensure broader engagement, especially among people with limited technology experience. In addition, randomized controlled trials are needed to assess the effectiveness of prediction markets compared to conventional case-based surveillance systems. From a policy perspective, there is a need for technological and regulatory support to create IMs, and the availability of public data on reported infectious diseases is necessary.

Limitations

Our study has limitations. First, our sample size is small, and therefore, the views and experiences of the study participants are not representative of the general population. However, our pilot study provides preliminary data and lessons to inform subsequent research into using predictive markets for infectious disease surveillance in LMICs. Second, specific characteristics of the prediction market system may have limited engagement and limited our ability to use the data to estimate COVID-19 cases. Although the pilot was implemented with a computer default, several users accessed the platform on their smartphones. Subsequent studies should consider adopting smartphone-based designs for user convenience to mitigate this limitation and encourage community engagement. Subsequent studies should also consider having social media–based chat groups where participants can directly interact with each other to make their experience more enjoyable. Third, IMs are limited in the context of capturing the syndromic profiles of reporting individuals. Therefore, IMs should be used to complement data from traditional surveillance and passive disease surveillance systems but not as replacements for those systems.

Conclusions

This study contributes to the growing literature around participatory disease surveillance approaches by demonstrating the feasibility of tailoring existing systems for use in LMIC settings, where they are urgently needed. Our findings showed that using IM surveillance to supplement traditional laboratory and syndromic monitoring systems is feasible and acceptable in LMICs. Our qualitative findings also better explained the barriers and facilitators of implementing an IM-based program. Moreover, the ideal journey framework we have developed could inform future implementation designs. However, adequate education on trading and a more user-friendly technological interface with offline notification systems are needed to ensure optimum participation in future IMs. In addition, ensuring a comprehensive onboarding process with bidirectional communication and social media promotion platforms is essential to keep participants continuously engaged and interested in prediction markets. Future research should incorporate greater equity in the sample population and explore variations in geographic diversity to fully understand the potential of IMs in daily national disease surveillance systems of other LMICs.

Acknowledgments

This study was financially supported by the European & Developing Countries Clinical Trial Partnership (grant RIA2020EF-2983). The funding body, European & Developing Countries Clinical Trial Partnership, had no role in the design, methods, subject recruitment, data collection, analysis, and preparation of the paper. The authors wish to thank the study staff and participants from the University of Ghana, Cognate Systems, SIHI Ghana, and community residents of Accra, who made this work possible.

The authors would also like to thank Prediki for providing access to their platform and providing guidance on market implementation. SS is grateful for support from the National Institute of Allergy and Infectious Diseases (award K01AI59233).

Authors' Contributions

SS, JDT, PDG, and MM conceived the study concept; SS, KN, KO, WA, and HH designed the system; PN, RR, EC, and EK collected the data; PN, RR, and EC analyzed the data; GM, SS, JDT, and PDG wrote the manuscript; and MM, JDT, RE, RR, EK, and HH critically revised the manuscript. All authors contributed to the final draft for submission. The corresponding author had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Video introduction on COVID-19 prediction markets surveillance to complement efforts of the Ghana Health Service in monitoring COVID-19 outbreaks during the pandemic.
[MP4 File (MP4 Video), 101652 KB - infodemiology_v4i1e50125_app1.mp4]

Multimedia Appendix 2
Video showing the step-by-step process of trading participation in the COVID-19 prediction markets survey on the Prediki platform.
[MP4 File (MP4 Video), 12459 KB - infodemiology_v4i1e50125_app2.mp4]

Multimedia Appendix 3
Video showing the trading rewarding process and how rewards are determined in the prediction markets survey.
[MP4 File (MP4 Video), 5221 KB - infodemiology_v4i1e50125_app3.mp4]

Multimedia Appendix 4
Supplementary tables, explanatory texts, participants excerpt quotes, thematic coding tree, and a figure showing the ideal user journey.
[PDF File (Adobe PDF File), 121 KB - infodemiology_v4i1e50125_app4.pdf]

Multimedia Appendix 5
The questions used in the Prediki prediction markets surveillance for Ghana on the Prediki platform and the frequency of trades made by participants per question for 3 months in 2021.
[DOCX File , 16 KB - infodemiology_v4i1e50125_app5.docx]

References

1. Arthur RR, LeDuc JW, Hughes JM. Global surveillance for emerging infectious diseases. In: Guerrant RL, Walker DH, Weller PF, editors. Tropical Infectious Diseases. 3rd edition. Amsterdam, the Netherlands: Elsevier Inc; Sep 2011:105-109.
2. 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]
3. Chunara R, Smolinski MS, Brownstein JS. Why we need crowdsourced data in infectious disease surveillance. Curr Infect Dis Rep 2013 Aug;15(4):316-319 [FREE Full text] [doi: 10.1007/s11908-013-0341-5] [Medline: 23689991]
4. Althouse BM, Ng YY, Cummings DA. Prediction of dengue incidence using search query surveillance. PLoS Negl Trop Dis 2011 Aug;5(8):e1258 [FREE Full text] [doi: 10.1371/journal.pntd.0001258] [Medline: 21829744]
5. Arrow KJ, Forsythe R, Gorham M, Hahn R, Hanson R, Ledyard JO, et al. Economics. The promise of prediction markets. Science 2008 May 16;320(5878):877-878. [doi: 10.1126/science.1157679] [Medline: 18487176]
6. Babakazo P, Kabamba-Tshilobo J, Wemakoy EO, Lubula L, Manya LK, Ilunga BK, et al. Evaluation of the influenza sentinel surveillance system in the Democratic Republic of Congo, 2012-2015. BMC Public Health 2019 Dec 10;19(1):1652 [FREE Full text] [doi: 10.1186/s12889-019-8008-2] [Medline: 31823763]
7. Mtema Z, Changalucha J, Cleaveland S, Elias M, Ferguson HM, Halliday JE, et al. Mobile phones as surveillance tools: implementing and evaluating a large-scale intersectoral surveillance system for rabies in Tanzania. PLoS Med 2016 Apr 12;13(4):e1002002 [FREE Full text] [doi: 10.1371/journal.pmed.1002002] [Medline: 27070315]
8. Bayer C, Remschmidt C, an der Heiden M, Tolksdorf K, Herzhoff M, Kaersten S, et al. Internet-based syndromic monitoring of acute respiratory illness in the general population of Germany, weeks 35/2011 to 34/2012. Euro Surveill 2014 Jan 30;19(4):20684 [FREE Full text] [doi: 10.2807/1560-7917.es2014.19.4.20684] [Medline: 24507468]

9. Gupta A, Katarya R. Social media based surveillance systems for healthcare using machine learning: a systematic review. J Biomed Inform 2020 Aug;108:103500 [FREE Full text] [doi: 10.1016/j.jbi.2020.103500] [Medline: 32622833]
10. Charbonneau DH, James LN. FluView and FluNet: tools for influenza activity and surveillance. Med Ref Serv Q 2019 Nov 05;38(4):358-368. [doi: 10.1080/02763869.2019.1657734] [Medline: 31687905]
11. Lee L, Desroches M, Mukhi S, Bancej C. FluWatchers: evaluation of a crowdsourced influenza-like illness surveillance application for Canadian influenza seasons 2015-2016 to 2018-2019. Can Commun Dis Rep 2021 Sep 10;47(9):357-363 [FREE Full text] [doi: 10.14745/ccdr.v47i09a02] [Medline: 34650332]
12. Leal Neto O, Haenni S, Phuka J, Ozella L, Paolotti D, Cattuto C, et al. Combining wearable devices and mobile surveys to study child and youth development in Malawi: implementation study of a multimodal approach. JMIR Public Health Surveill 2021 Mar 05;7(3):e23154 [FREE Full text] [doi: 10.2196/23154] [Medline: 33536159]
13. Leal-Neto O, Egger T, Schlegel M, Flury D, Sumer J, Albrich W, et al. Digital SARS-CoV-2 detection among hospital employees: participatory surveillance study. JMIR Public Health Surveill 2021 Nov 22;7(11):e33576 [FREE Full text] [doi: 10.2196/33576] [Medline: 34727046]
14. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats. JMIR Public Health Surveill 2017 Oct 11;3(4):e62 [FREE Full text] [doi: 10.2196/publichealth.7540] [Medline: 29021131]
15. Ruebush TK, Zeissig R, Koplan JP, Klein RE, Godoy HA. Community participation in malaria surveillance and treatment. III. An evaluation of modifications in the Volunteer Collaborator Network of Guatemala. Am J Trop Med Hyg 1994 Jan;50(1):85-98. [doi: 10.4269/ajtmh.1994.50.85] [Medline: 8304577]
16. Worsley-Tonks KE, Bender JB, Deem SL, Ferguson AW, Fèvre EM, Martins DJ, et al. Strengthening global health security by improving disease surveillance in remote rural areas of low-income and middle-income countries. Lancet Glob Health 2022 Apr;10(4):e579-e584. [doi: 10.1016/s2214-109x(22)00031-6]
17. Olufadewa II, Adesina MA, Ekpo MD, Akinloye SJ, Iyanda TO, Nwachukwu P, et al. Lessons from the coronavirus disease 2019 (COVID-19) pandemic response in China, Italy, and the U.S.: a guide for Africa and low- and middle-income countries. Glob Health J 2021 Mar;5(1):56-61 [FREE Full text] [doi: 10.1016/j.glohj.2021.02.003] [Medline: 33585052]
18. Cantarelli P, Debin M, Turbelin C, Poletto C, Blanchon T, Falchi A, et al. The representativeness of a European multi-center network for influenza-like-illness participatory surveillance. BMC Public Health 2014 Sep 20;14(1):984 [FREE Full text] [doi: 10.1186/1471-2458-14-984] [Medline: 25240865]
19. Li EY, Tung CY, Chang SH. The wisdom of crowds in action: forecasting epidemic diseases with a web-based prediction market system. Int J Med Inform 2016 Aug;92:35-43. [doi: 10.1016/j.ijmedinf.2016.04.014] [Medline: 27318069]
20. Burnham K. How Motorola uses prediction markets to choose innovations. ComputerWorld. 2009. URL: https://www.computerworld.com/article/1598880/how-motorola-uses-prediction-markets-to-choose-innovations.html [accessed 2024-04-29]
21. Guerrisi C, Turbelin C, Souty C, Poletto C, Blanchon T, Hanslik T, et al. The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France. Euro Surveill 2018 Jun;23(25):1700337 [FREE Full text] [doi: 10.2807/1560-7917.ES.2018.23.25.1700337] [Medline: 29945696]
22. Chunara R, Chhaya V, Bane S, Mekaru SR, Chan EH, Freifeld CC, et al. Online reporting for malaria surveillance using micro-monetary incentives, in urban India 2010-2011. Malar J 2012 Feb 13;11:43 [FREE Full text] [doi: 10.1186/1475-2875-11-43] [Medline: 22330227]
23. Tung CY, Chou TC, Lin JW. Using prediction markets of market scoring rule to forecast infectious diseases: a case study in Taiwan. BMC Public Health 2015 Aug 11;15:766 [FREE Full text] [doi: 10.1186/s12889-015-2121-7] [Medline: 26259612]
24. Polgreen PM, Nelson FD, Neumann GR. Use of prediction markets to forecast infectious disease activity. Clin Infect Dis 2007 Jan 15;44(2):272-279. [doi: 10.1086/510427] [Medline: 17173231]
25. Sell TK, Warmbrod KL, Watson C, Trotochaud M, Martin E, Ravi SJ, et al. Using prediction polling to harness collective intelligence for disease forecasting. BMC Public Health 2021 Nov 20;21(1):2132 [FREE Full text] [doi: 10.1186/s12889-021-12083-y] [Medline: 34801014]
26. Jennifer HW. Prediction markets as an aggregation mechanism for collective intelligence. In: Proceedings of the 2007 International Astronomical Consortium for High-Energy Calibration. 2007 Presented at: IACHEC '07; May 9-11, 2007; Lake Arrowhead, CA p. 2007 URL: https://escholarship.org/uc/item/8mg0p0zc#main [doi: 10.1145/1250910.1250968]
27. Wolfers J, Zitzewitz E. Prediction Markets. J Econ Perspect 2004 May 01;18(2):107-126. [doi: 10.1257/0895330041371321]
28. Kenu E, Odikro MA, Malm KL, Asiedu-Bekoe F, Noora CL, Frimpong JA, et al. Epidemiology of COVID-19 outbreak in Ghana, 2020. Ghana Med J 2020 Dec;54(4 Suppl):5-15 [FREE Full text] [doi: 10.4314/gmj.v54i4s.3] [Medline: 33976436]
29. Ghana SS. Population projection for Ghana. Ghana Statistical Service. 2020. URL: https://statsghana.gov.gh/nationalaccount_macros.php?Stats=MTA1NTY1NjgxLjUwNg==/webstats/s679n2sn87 [accessed 2023-01-10]
30. Global digital reports series: digital Ghana 2022. DataPortal. URL: https://datareportal.com/reports/digital-2022-ghana [accessed 2023-02-04]
31. IDEO. The field guide to human-centered design: a step-by-step guide that will get you solving problems like a designer. IDEO.org. URL: https://tinyurl.com/5x56dyry [accessed 2022-12-22]

32. Halpaap BM, Tucker JD, Mathanga D, Juban N, Awor P, Saravia NG, et al. Social innovation in global health: sparking location action. Lancet Glob Health 2020 May;8(5):e633-e634. [doi: 10.1016/s2214-109x(20)30070-x]
33. Pfeiffer T, Almenberg J. Prediction markets and their potential role in biomedical research--a review. Biosystems 2010;102(2-3):71-76. [doi: 10.1016/j.biosystems.2010.09.005] [Medline: 20837097]

Abbreviations

IM: information market
LMIC: low- and middle-income country

Edited by T Purnat; submitted 20.06.23; peer-reviewed by O Leal Neto, P Washington; comments to author 08.02.24; revised version received 11.04.24; accepted 04.05.24; published 12.08.24.

Please cite as:
Marley G, Dako-Gyeke P, Nepal P, Rajgopal R, Koko E, Chen E, Nuamah K, Osei K, Hofkirchner H, Marks M, Tucker JD, Eggo R, Ampofo W, Sylvia S
Collective Intelligence–Based Participatory COVID-19 Surveillance in Accra, Ghana: Pilot Mixed Methods Study
JMIR Infodemiology 2024;4:e50125
URL: https://infodemiology.jmir.org/2024/1/e50125
doi:10.2196/50125
PMID:

Original Paper

Perceptions of Health Misinformation on Social Media: Cross-Sectional Survey Study

Anna Gaysynsky[1,2], MPH; Nicole Senft Everson[1], PhD; Kathryn Heley[1], PhD, MPH; Wen-Ying Sylvia Chou[1], PhD, MPH

[1]Health Communication and Informatics Research Branch, Behavioral Research Program, National Cancer Institute, Rockville, MD, United States
[2]ICF Next, ICF, Rockville, MD, United States

Corresponding Author:
Anna Gaysynsky, MPH
Health Communication and Informatics Research Branch
Behavioral Research Program
National Cancer Institute
9609 Medical Center Drive
Rockville, MD, 20850
United States
Phone: 1 240 276 5284
Email: anna.gaysynsky@nih.gov

Abstract

Background: Health misinformation on social media can negatively affect knowledge, attitudes, and behaviors, undermining clinical care and public health efforts. Therefore, it is vital to better understand the public's experience with health misinformation on social media.

Objective: The goal of this analysis was to examine perceptions of the social media information environment and identify associations between health misinformation perceptions and health communication behaviors among US adults.

Methods: Analyses used data from the 2022 Health Information National Trends Survey (N=6252). Weighted unadjusted proportions described respondents' perceptions of the amount of false or misleading health information on social media ("perceived misinformation amount") and how difficult it is to discern true from false information on social media ("perceived discernment difficulty"). Weighted multivariable logistic regressions examined (1) associations of sociodemographic characteristics and subjective literacy measures with misinformation perceptions and (2) relationships between misinformation perceptions and health communication behaviors (ie, sharing personal or general health information on social media and using social media information in health decisions or in discussions with health care providers).

Results: Over one-third of social media users (35.61%) perceived high levels of health misinformation, and approximately two-thirds (66.56%) reported high perceived discernment difficulty. Odds of perceiving high amounts of misinformation were lower among non-Hispanic Black/African American (adjusted odds ratio [aOR] 0.407, 95% CI 0.282-0.587) and Hispanic (aOR 0.610, 95% CI 0.449-0.831) individuals compared to White individuals. Those with lower subjective health literacy were less likely to report high perceived misinformation amount (aOR 0.602, 95% CI 0.374-0.970), whereas those with lower subjective digital literacy were more likely to report high perceived misinformation amount (aOR 1.775, 95% CI 1.400-2.251). Compared to White individuals, Hispanic individuals had lower odds of reporting high discernment difficulty (aOR 0.620, 95% CI 0.462-0.831). Those with lower subjective digital literacy (aOR 1.873, 95% CI 1.478-2.374) or numeracy (aOR 1.465, 95% CI 1.047-2.049) were more likely to report high discernment difficulty. High perceived misinformation amount was associated with lower odds of sharing general health information on social media (aOR 0.742, 95% CI 0.568-0.968), using social media information to make health decisions (aOR 0.273, 95% CI 0.156-0.479), and using social media information in discussions with health care providers (aOR 0.460, 95% CI 0.323-0.655). High perceived discernment difficulty was associated with higher odds of using social media information in health decisions (aOR 1.724, 95% CI 1.208-2.460) and health care provider discussions (aOR 1.389, 95% CI 1.035-1.864).

Conclusions: Perceptions of high health misinformation prevalence and discernment difficulty are widespread among social media users, and each has unique associations with sociodemographic characteristics, literacy, and health communication behaviors. These insights can help inform future health communication interventions.

(JMIR Infodemiology 2024;4:e51127) doi:[10.2196/51127](https://doi.org/10.2196/51127)

KEYWORDS

social media; misinformation; health communication; health literacy; patient-provider communication

Introduction

Background

The Pew Research Center estimates that approximately 72% of Americans use social media [1], and research suggests that social media is widely used for health-related purposes specifically [2]. Social media has become an important venue for the exchange of health-related information and advice [3]. In 2019, 41% of internet users in the United States reported watching health-related YouTube videos, and 17% reported sharing health information on social networking sites [4]. Social media can help people find and access more useful and personally relevant information, facilitate the exchange of social support, and aid with disease management efforts [3]. However, while social media can make health information more accessible, the use of social media for health information seeking can also create the risk of harm through exposure to misinformation.

Defined as "health-related information that is false, inaccurate, or misleading according to the best available evidence at the time," health misinformation is increasingly recognized as a threat to public health [5,6] (note that this definition includes *disinformation*, or false information that is created and spread with the intent to deceive, as a subset of misinformation [7]). Although health misinformation is not a new phenomenon, social media facilitates the rapid spread of falsehoods [6], thereby exacerbating the potential negative impact of misinformation on both individual and population health. Certain features of social media platforms, such as incentives that reward the sharing of content that receives more engagement, can result in a focus on sharing emotionally charged or provocative content rather than accurate content [6,8]. Additionally, algorithms that suggest content to users are often based on past engagement behavior, which can reinforce echo chambers, whereby users who engage with misinformation increasingly encounter further misinformation [6,8].

Many studies have documented substantial health-related misinformation on social media across a range of topics (including tobacco products, drugs, and vaccines) [9], and research increasingly suggests that social media misinformation can have a negative impact on health-related attitudes, behaviors, and outcomes. For example, Pierri et al [10] found that the prevalence of COVID-19 vaccine misinformation on Twitter (now rebranded as X) was related to higher levels of vaccine hesitancy and lower vaccination uptake rates in the United States even after accounting for political and sociodemographic factors. Furthermore, their causality analysis suggested a directional relationship between social media misinformation and vaccine hesitancy, with a lag of approximately 2 to 6 days from misinformation being posted in a county to a corresponding increase in vaccine hesitancy in that county [10]. Further evidence of a causal relationship between exposure to misinformation and health-related attitudes and intentions is provided by a randomized controlled trial conducted in the United States and the United Kingdom, which showed that exposure to misinformation in the form of social media posts decreased the number of respondents who said that they would "definitely" take the COVID-19 vaccine by approximately 6 percentage points relative to the control group [11].

While there is a growing body of research examining the prevalence of misinformation on social media, as well as the association between social media misinformation and health-related outcomes, little work to date has focused on understanding how individuals perceive misinformation on social media or how these perceptions impact behavior. Surveys conducted over the past few years show that many US adults believe that much of the news they see on social media is false or inaccurate [12,13]. This is significant because studies have shown that misinformation perceptions impact communication behaviors. For example, a study conducted in Germany found higher self-perceived exposure to "fake news" to be associated with more frequent engagement in information verification behaviors on Facebook [14]. Meanwhile, a study conducted on Amazon Mechanical Turk found that perceiving fake news to have a greater influence on others than on oneself (ie, the third-person effect) was associated with lower intent to share news obtained from social media (either online or offline) [15].

Studies also suggest that perceptions of misinformation prevalence are associated with attitudes toward health issues—for example, one cross-sectional study found perceptions of high misinformation prevalence to be correlated with worry about COVID-19 [16]. Notably, the study found neither a significant association between actual misinformation prevalence (as measured using the "Infodemic Risk Index," which produces frequency estimates of misinformation on Twitter by country) and worry about COVID-19 nor an interaction between actual misinformation prevalence and perceived prevalence in explaining pandemic worry [16]. This suggests critical psychological and cognitive effects of misinformation perceptions independent of actual misinformation prevalence (and separate from misinformation endorsement or belief) [16].

In addition to assessing perceptions of misinformation prevalence, assessing people's confidence in their ability to detect misinformation is important because confidence can affect the way people make subsequent judgments [17]. For example, confidence can determine whether an individual acts on their initial judgment or seeks out additional information [18]. Confidence levels also affect a person's willingness and ability to defend their assessments such that individuals who are able to discern true from false information—and are confident about their judgments—are more resistant to misinformation [18]. A person's perceptions about their ability to accurately detect misinformation can also influence their perceptions about their capacity to manage health issues or make health decisions. For example, Park et al [19] found that people who had higher confidence in their ability to distinguish between true and false COVID-19 information also had higher COVID-19 risk readiness perceptions (ie, felt that they had a “handle on the issues and developments surrounding the coronavirus outbreak”). Unfortunately, confidence can also be easily undermined, particularly when an individual is unsure about the validity of the material they are considering or lacks the necessary skills or literacy competencies to feel secure in their assessment [17].

Beyond obtaining a better understanding of misinformation perceptions and how they impact cognitive and behavioral processes and outcomes, it is also important to assess whether these perceptions vary by sociodemographic or other characteristics to identify groups that may be more vulnerable to misinformation and in need of more targeted efforts. For example, the trial conducted by Loomba et al [11] showed that some groups were differentially affected by exposure to misinformation—in the United States, female individuals were found to be less resistant to misinformation than male individuals, whereas those with lower incomes were found to be more resistant. Additionally, a survey conducted by the Pew Research Center in 2016 found that White individuals were more likely than Black and Hispanic individuals to say that they often saw fake political news online, and those with annual incomes of at least US $75,000 were more likely to report seeing fake news compared to those who made less than US $75,000 per year [20]. Findings regarding demographic differences in perceived ability to discern true from false information are more mixed. The Pew survey found that confidence in detecting fake political news did not differ significantly by sociodemographic characteristics (such as age, gender, income, or race) [20], whereas the study conducted by Park et al [19] found education and income to be significant predictors of confidence in distinguishing true from false information about COVID-19. Furthermore, a large survey experiment conducted by Sirlin et al [21] found digital literacy (as measured by familiarity with internet-related terms and attitudes toward technology as well as understanding of social media algorithms) to be an important predictor of the ability to discern truths from falsehoods when judging headline accuracy for both political and COVID-19 articles. Taken together, these findings demonstrate the importance of assessing how perceptions of health misinformation and misinformation discernment vary by sociodemographic characteristics and literacy. Although research regarding vulnerability to misinformation remains mixed (eg, with regard to age, sex, and income) [22], the potential impact of social media misinformation on health disparities is an important issue that requires ongoing attention.

Study Aims

Because misinformation perceptions can affect attitudes and behaviors, a better understanding of the public’s perceptions of health misinformation on social media and their ability to detect it, as well as possible subgroup differences in such perceptions, is needed. Toward that end, this study analyzed data from the National Cancer Institute’s 2022 Health Information National Trends Survey (HINTS 6) to (1) assess the prevalence of 2 distinct misinformation-related perceptions—perceived amount of health misinformation on social media and perceived ability to distinguish true from false health information on social media—(2) identify sociodemographic factors associated with these health misinformation perceptions; and (3) explore associations between these misinformation perceptions and health communication behaviors, including information sharing, health decision-making, and communicating with health care providers.

Methods

Data and Sample Selection

HINTS is a nationally representative, cross-sectional, self-administered survey of civilian, noninstitutionalized US adults aged ≥18 years. Data for HINTS 6 (N=6252) were collected between March 7, 2022, and November 8, 2022, using questionnaires administered via mailed paper or web-based surveys. The overall response rate for HINTS 6 was 28.1%. Respondents who reported that they did not use social media (1211/6252, 19.37%) were excluded from the analyses, resulting in a starting analytic sample of 5041. Details regarding the design of HINTS 6, including methodology, sampling, and weighting procedures, have been published elsewhere [23].

Ethical Considerations

HINTS 6 received approval from the Westat Institutional Review Board on May 10, 2021 (6632.03.51), and was designated as non–human subjects research by the National Institutes of Health Office of Human Subjects Research on August 16, 2021 (000626). Respondents’ return of the completed survey indicated consent to participate.

Measures

Social Media Health Misinformation Perceptions

A total of 2 social media misinformation–related perceptions were measured. Perceived amount of misinformation on social media (“perceived misinformation amount”) was assessed with the following item: “How much of the health information that you see on social media do you think is false or misleading?” Response options were *none*, *a little*, *some*, *a lot*, and *I do not use social media* (as noted previously, those who selected “I do not use social media” in response to this item were excluded from the analyses).

Perceived difficulty distinguishing true from false information on social media (“perceived discernment difficulty”) was measured by assessing agreement with the following statement—“I find it hard to tell whether health information on social media is true or false”—among respondents who reported social media use. Response options were *strongly agree*, *somewhat agree*, *somewhat disagree*, and *strongly disagree*.

Health Communication Behaviors Related to Social Media Use

Information sharing on social media was assessed using two items that asked how often in the previous 12 months respondents (1) “share[d] personal health information on social media” and (2) “share[d] general health-related information on social media (for example, a news article).” Response options were *almost every day*, *at least once a week*, *a few times a month*, *less than once a month*, and *never*.

Respondents’ use of information encountered on social media was assessed through reported agreement with 2 items: “I use information from social media to make decisions about my health” and “I use information from social media in discussions with my healthcare provider.” Response options were *strongly agree*, *somewhat agree*, *somewhat disagree*, and *strongly disagree*.

Sociodemographic Characteristics

Sociodemographic variables included (1) educational level (categorized as high school degree or lower, some college or vocational training, and college graduate or higher), (2) sex (male or female), (3) age (18-24 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, and ≥65 years), (4) race or ethnicity (non-Hispanic White; non-Hispanic Black/African American; Hispanic; and non-Hispanic other, which included non-Hispanic American Indian or Alaska Native, non-Hispanic Asian, non-Hispanic Native Hawaiian or other Pacific Islander, and non-Hispanic multiple races), (5) annual household income (<US $20,000, US $20,000-<$35,000, US $35,000-<$50,000, US $50,000-<$75,000, US $75,000-<$100,000, and ≥US $100,000), and (6) geographic residence (urban or rural based on the 2013 Rural-Urban Continuum Codes).

Literacy Measures

Subjective health literacy was assessed using the following item: “How confident are you filling out medical forms by yourself?” Response options were *very* [confident], *somewhat* [confident], *a little* [confident], and *not at all* [confident]. This measure is one of the brief screening questions identified by Chew et al [24] for detecting inadequate or marginal health literacy among adults. Subjective digital literacy was assessed using the following item: “How confident are you that you can find helpful health resources on the Internet?” Response options were *completely confident*, *very confident*, *somewhat confident*, *a little confident*, and *not at all confident*. This measure was adapted from the eHealth Literacy Scale [25]. Subjective numeracy was assessed using the following item: “In general, how easy or hard do you find it to understand medical statistics?” Response options were *very easy*, *easy*, *hard*, and *very hard*. This item, which is part of the STAT-Confidence scale developed by Woloshin et al [26], has been shown to be a strong predictor of scores on the Newest Vital Sign measure (an objective measure of health literacy and numeracy) [27].

Statistical Analysis

To account for the complex sampling design of HINTS, analyses were conducted in SAS (version 9.4; SAS Institute) using final sample weights to obtain population-level point estimates and a set of 50 replicate weights to compute accurate variance estimates [23]. Frequencies and survey-weighted unadjusted proportions were used to describe the distributions of perceived misinformation amount and perceived discernment difficulty.

In total, 2 weighted multivariable logistic regression models examined associations of sociodemographic characteristics and literacy measures with perceived misinformation amount and perceived discernment difficulty. For these analyses, perceived misinformation amount was dichotomized to reflect high perceived misinformation amount (*a lot*) versus low perceived misinformation amount (*none*, *a little*, or *some*) to facilitate comparison between those who perceived misinformation to be a significant problem in the information environment and those who did not. Furthermore, only a relatively small proportion of respondents felt that “none” or only “a little” of the information they saw on social media was false or misleading, whereas over a third of the sample reported that “a lot” of the information they saw was false or misleading. Perceived discernment difficulty was dichotomized as high (*strongly agree* or *somewhat agree*) versus low (*somewhat disagree* or *strongly disagree*). Additionally, subjective health literacy was dichotomized as high (*very* [confident] or *somewhat* [confident]) versus low (*a little* [confident] or *not at all* [confident]), digital literacy was dichotomized as high (*completely confident* or *very confident*) versus low (*somewhat confident*, *a little confident*, or *not at all confident*), and numeracy was dichotomized as high (*very easy* or *easy*) versus low (*hard* or *very hard*).

A total of 4 additional weighted multivariable logistic regression models tested associations of high versus low perceived misinformation amount and high versus low perceived discernment difficulty with communication behaviors related to social media use (ie, sharing personal health information on social media, sharing general health information on social media, using information from social media to make health decisions, and using information from social media in discussions with health care providers) adjusted for sociodemographic characteristics and dichotomized literacy measures. The 2 information-sharing behavior measures were dichotomized as ever shared (*almost every day*, *at least once a week*, *a few times a month*, or *less than once a month*) versus never shared, whereas the 2 social media information use items were dichotomized as agreement (*strongly agree* or *somewhat agree*) versus disagreement (*somewhat disagree* or *strongly disagree*). Sensitivity analyses tested the interaction of perceived misinformation amount and discernment difficulty in predicting these communication outcomes, but this interaction was not statistically significant in any of the models.

Adjusted analyses used complete case analysis, with valid analytic samples reported in tables corresponding to each analysis. Descriptive information on missing data for each variable is publicly available on the HINTS website [28]. Tests of significance were conducted at the P<.05 level.

Results

Prevalence of Social Media Health Misinformation Perceptions

As shown in Figure 1, over one-third of American social media users (35.61%) perceived “a lot” of misinformation on social media (ie, expressed high perceived misinformation amount), whereas only a very small percentage (1.54%) perceived that “none” of the health information they see is false or misleading. Figure 2 shows that approximately two-thirds of American social media users agreed that they find it hard to tell whether health information on social media is true or false (ie, endorsed high discernment difficulty).

Figure 1. Weighted proportions of perceived health misinformation amount among American social media users.

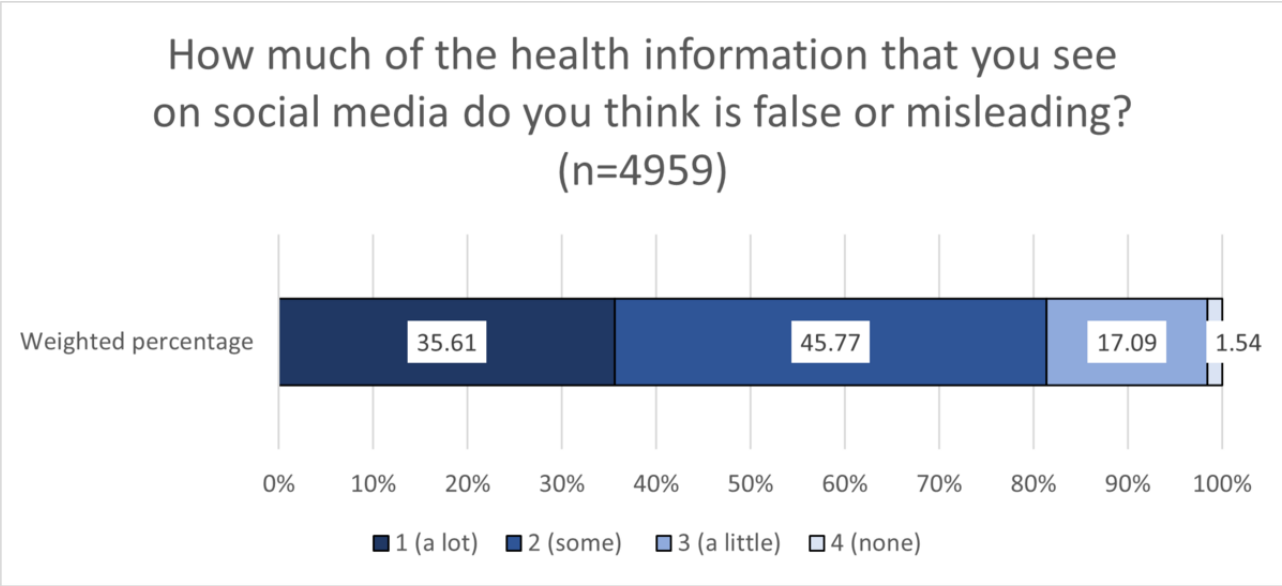
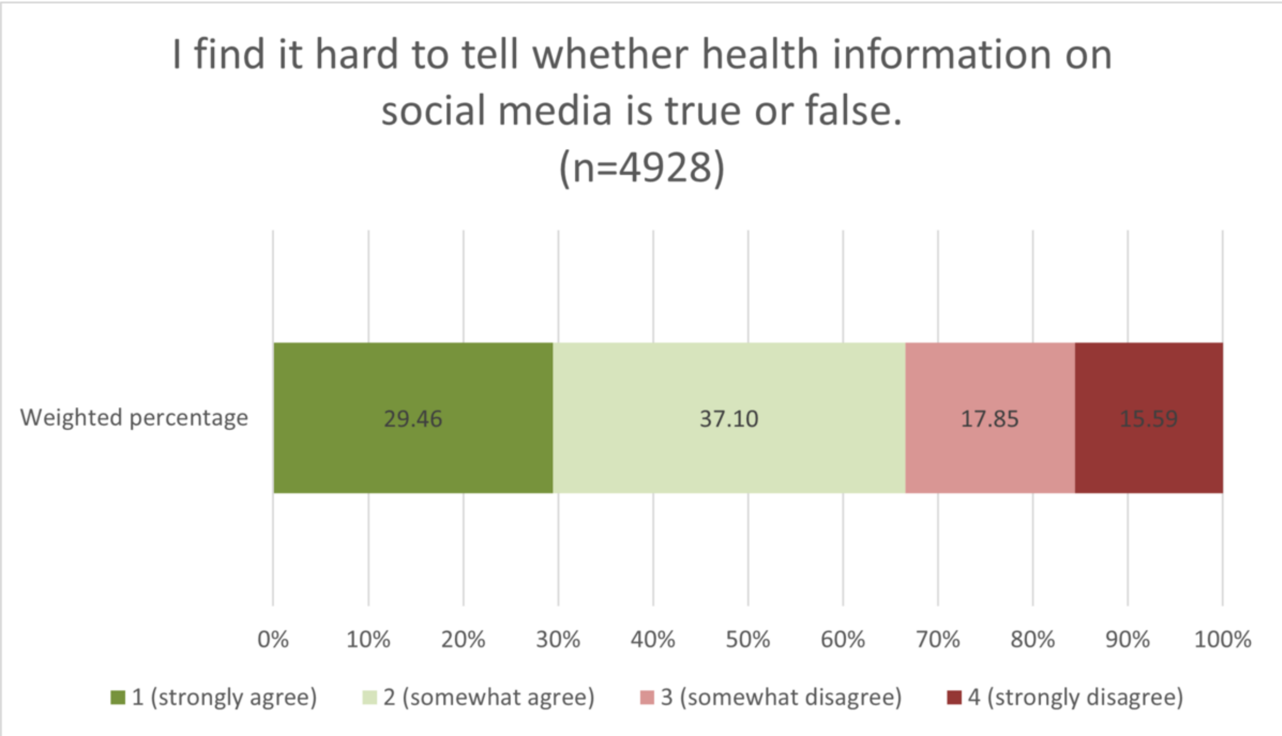


Figure 2. Weighted proportions of perceived discernment difficulty among American social media users.



Predictors of Social Media Health Misinformation Perceptions

Perceived Misinformation Amount

As shown in Table 1, individuals who were non-Hispanic Black/African American (compared to non-Hispanic White individuals; adjusted odds ratio [aOR] 0.407, 95% CI 0.282-0.587) or Hispanic (compared non-Hispanic White individuals; aOR 0.610, 95% CI 0.449-0.831) or who had lower subjective health literacy (vs those with higher health literacy; aOR 0.602, 95% CI 0.374-0.970) were less likely to report high perceived misinformation amount. Comparatively, respondents with lower subjective digital literacy were more likely to report high misinformation amount (vs those with higher digital literacy; aOR 1.775, 95% CI 1.400-2.251). Age, sex, educational level, income, geographic residence, and numeracy were not statistically significantly related to perceived amount of misinformation.

Table 1. Predictors of social media health misinformation perceptions.

Independent variable	Perceived misinformation amount[a] (n=4218), adjusted odds ratio (95% CI)	Perceived discernment difficulty[b] (n=4205), adjusted odds ratio (95% CI)
Age (y; reference: 18-24)		
25-34	0.745 (0.403-1.376)	*0.485* (0.289-0.816)[c]
35-44	0.859 (0.495-1.491)	*0.487* (0.311-0.763)
45-54	0.772 (0.410-1.454)	0.624 (0.384-1.014)
55-64	0.790 (0.425-1.468)	*0.605* (0.369-0.990)
≥65	0.688 (0.384-1.232)	0.841 (0.525-1.346)
Sex (reference: male)		
Female	1.037 (0.791-1.358)	1.077 (0.871-1.331)
Educational level (reference: high school or lower)		
Some college or vocational training	1.206 (0.867-1.677)	1.210 (0.882-1.660)
College graduate or higher	1.144 (0.809-1.618)	0.871 (0.632-1.200)
Race or ethnicity (reference: non-Hispanic White)		
Hispanic	*0.610* (0.449-0.831)	*0.620* (0.462-0.831)
Non-Hispanic Black/African American	*0.407* (0.282-0.587)	0.830 (0.596-1.156)
Non-Hispanic other	0.977 (0.662-1.442)	1.126 (0.788-1.607)
Annual household income (reference: <US $20,000)		
US $20,000-$34,999	1.283 (0.747-2.202)	1.016 (0.640-1.614)
US $35,000-$49,999	1.039 (0.613-1.760)	1.061 (0.678-1.659)
US $50,000-$74,999	1.619 (0.968-2.709)	1.058 (0.740-1.512)
US $75,000-$99,999	1.693 (0.996-2.880)	1.459 (0.932-2.283)
≥US $100,000	1.469 (0.910-2.369)	1.245 (0.870-1.780)
Geographic residence (reference: urban)		
Rural	1.012 (0.770-1.331)	1.109 (0.769-1.600)
Health literacy (reference: high health literacy)		
Low health literacy	*0.602* (0.374-0.970)	1.230 (0.829-1.824)
Digital literacy (reference: high digital literacy)		
Low digital literacy	*1.775* (1.400-2.251)	*1.873* (1.478-2.374)
Numeracy (reference: high numeracy)		
Low numeracy	1.030 (0.771-1.376)	*1.465* (1.047-2.049)

[a]The probability modeled was odds of high perceived misinformation amount (*a lot*) in reference to low perceived misinformation amount (*none*, *a little*, or *some*).

[b]The probability modeled was odds of reporting high perceived discernment difficulty (*strongly agree* or *somewhat agree*) in reference to low perceived discernment difficulty (*somewhat disagree* or *strongly disagree*).

[c]Italicized values are statistically significant (P<.05).

Perceived Discernment Difficulty

As shown in Table 1, there were differences in perceived discernment difficulty by age—adults aged 25 to 34 years, 35 to 44 years, and 55 to 64 years were less likely to report high discernment difficulty compared to those aged 18 to 24 years, whereas adults aged 45 to 54 years and those aged ≥65 years did not differ significantly from the youngest age group. Hispanic individuals (vs non-Hispanic White individuals; aOR 0.620, 95% CI 0.462-0.831) were less likely to report high discernment difficulty. Those with lower (vs higher) subjective digital literacy (aOR 1.873, 95% CI 1.478-2.374) or lower (vs higher) subjective numeracy (aOR 1.465, 95% CI 1.047-2.049) were more likely to report high discernment difficulty. The associations between perceived discernment difficulty and sex, educational level, income, geographic residence, and health literacy were not statistically significant.

Associations Between Social Media Health Misinformation Perceptions and Communication Behaviors

After adjusting for sociodemographic characteristics and literacy measures, individuals who perceived high (vs low) levels of social media misinformation were less likely to report sharing general health information on social media (aOR 0.742, 95% CI 0.568-0.968), using social media information to make health decisions (aOR 0.273, 95% CI 0.156-0.479), and using social media information in discussions with health care providers (aOR 0.460, 95% CI 0.323-0.655). Perceived misinformation amount was not significantly associated with sharing personal health information on social media (Table 2).

Individuals with high (vs low) perceived discernment difficulty were more likely to report using information from social media to make health decisions (aOR 1.724, 95% CI 1.208-2.460) and in discussions with health care providers (aOR 1.389, 95% CI 1.035-1.864). Perceived discernment difficulty was not significantly associated with sharing personal or general health information on social media.

Table 2. Adjusted odds ratio (aOR) and 95% CI of health information sharing and social media information use by social media health misinformation perceptions[a].

Social media misinformation perception	Sharing personal health information[b] (n=4136), aOR (95% CI)	Sharing general health information[c] (n=4159), aOR (95% CI)	Using social media information to make health decisions[d] (n=4177), aOR (95% CI)	Using social media information in discussions with health care providers[e] (n=4174), aOR (95% CI)
High perceived misinformation amount[f]	0.803 (0.591-1.092)	*0.742* (0.568-0.968)[g]	*0.273* (0.156-0.479)	*0.460* (0.323-0.655)
High perceived discernment difficulty[h]	1.163 (0.862-1.570)	1.100 (0.878-1.379)	*1.724* (1.208-2.460)	*1.389* (1.035-1.864)

[a]Analyses were adjusted for age, sex, educational level, race or ethnicity, income, geographic residence, health literacy, digital literacy, and numeracy.

[b]The probability modeled was odds of having ever shared personal information on social media (*shared almost every day, at least once a week, a few times a month*, or *less than once a month in the past 12 months*) in reference to having never shared.

[c]The probability modeled was odds of having ever shared general information on social media (*shared almost every day, at least once a week, a few times a month*, or *less than once a month in the past 12 months*) in reference to having never shared.

[d]The probability modeled was odds of using social media information for making health decisions (*strongly agree* or *somewhat agree*) in reference to not using social media information for making health decisions (*strongly disagree* or *somewhat disagree*).

[e]The probability modeled was odds of using social media information in discussions with health care providers (*strongly agree* or *somewhat agree*) in reference to not using social media information in discussions with health care providers (*strongly disagree* or *somewhat disagree*).

[f]High perceived misinformation amount=thinking that *a lot* of the health information on social media is false or misleading; low perceived misinformation amount=thinking that *none*, *a little*, or *some* of the health information on social media is false or misleading.

[g]Italicized values are statistically significant (*P*<.05).

[h]High perceived discernment difficulty=strongly or somewhat agreeing that it is hard to tell whether health information on social media is true or false; low perceived discernment difficulty=strongly or somewhat disagreeing that it is hard to tell whether health information on social media is true or false.

Discussion

Principal Findings

This study examined 2 misinformation-related perceptions among social media users (perception of the amount of health misinformation on social media and perceived ability to distinguish true from false health information on social media) to better understand the prevalence of these perceptions, subgroup differences in these perceptions, and how these perceptions are related to health communication behaviors. The study found that over one-third of social media users perceived their information environment to contain "a lot" of misleading or false content, and two-thirds expressed difficulty discerning true from false information on social media, with significant variation in these perceptions by sociodemographic characteristics and self-reported literacy skills. The analysis also showed that perceiving a high amount of misinformation on social media was related to lower information sharing on social media and lower use of social media information in discussions with providers and in health decisions, whereas difficulty distinguishing true from false information was associated with higher use of social media information in discussions with providers and health decisions. These results suggest that understanding misinformation perceptions could help inform health communication interventions and efforts to mitigate the impact of web-based misinformation, and that different approaches may be needed in response to each of these misinformation perceptions.

A substantial proportion of American social media users reported that "a lot" of the health information they see on social media is false or misleading, and this perception varied by race and ethnicity, as well as subjective measures of literacy. Non-Hispanic Black/African American and Hispanic individuals were less likely to say that "a lot" of the health information they see on social media is false or misleading. Because this analysis relied on self-report measures, it is not possible to ascertain whether minority groups are actually less exposed to social media misinformation (eg, due to the nature of their web-based networks) or if they are less aware that the information they are seeing is, in fact, false. However, the reasons behind these differences in misinformation perceptions and the potential for these differences to exacerbate health disparities deserve further attention given that Black and Hispanic individuals use social media at higher rates than White individuals [1] and substantial proportions of individuals in these groups report regularly obtaining their news from social media platforms [29]. Additional research that attempts to triangulate user perceptions with the social media content they encounter [30] could help shed light on the unique impact of objective and subjective social media experiences.

The analysis also found that individuals with lower digital literacy were more likely to report that “a lot” of the health information they see on social media is false or misleading, whereas those with lower health literacy were less likely to do so. This may be because individuals who self-report low confidence in their ability to find helpful resources on the web are more aware of content quality issues on the internet (including on social media), whereas reporting low subjective health literacy (eg, expressing difficulty filling out medical forms) may not be similarly related to concerns about the online information environment. In fact, a small study conducted in Europe found that participants with low health literacy (as measured using the Newest Vital Sign) had higher scores on the eHealth Literacy Scale, suggesting that they perceived themselves to have higher digital literacy than those in the high health literacy group [31]. The authors hypothesized that this finding might reflect differences in awareness of the issue of web-based health information quality between those with high versus low health literacy as well as differences in knowledge and use of established information evaluation criteria [31].

This study also revealed that approximately two-thirds of American social media users find it hard to tell whether health information on social media is true or false. High levels of discernment difficulty among the public are concerning. Low confidence in one’s ability to distinguish true from false information could result in lower motivation to seek additional information [19], apathy, and confusion, which could lead to negative health outcomes not just because people might act on misinformation but also because they may fail to act on accurate information or adhere to public health recommendations. Research suggests that self-efficacy (ie, judgments regarding how well one can execute a course of action required to deal with a prospective situation) plays an important role in how people select and evaluate information in web-based environments [32]. Individuals with higher self-efficacy may be better able to make accurate credibility assessments because they are more motivated to engage in deep cognitive processing and critical thinking [32], whereas those with lower self-efficacy may avoid engaging in extensive evaluations of information credibility, especially in contexts characterized by uncertainty and ambiguity, as they may not feel that they have a high likelihood of achieving desirable outcomes and, therefore, may experience negative affect (eg, anxiety, frustration, and confusion) in response to these situations [32]. However, although some research suggests that confidence in one’s ability to spot misinformation is associated with better performance in accurately distinguishing false from accurate news [32], the evidence is somewhat limited, and further research combining both subjective perceptions of ability and objective measures of ability is needed in order to investigate the impact of confidence on the way in which individuals navigate health information on social media.

Beyond generally high rates of discernment difficulty, this analysis also identified differences in perceived discernment ability in certain demographic subgroups. Specifically, adults aged 25 to 34 years, 35 to 44 years, and 55 to 64 years were less likely than those in the youngest age group to report discernment difficulty, and Hispanic individuals reported less discernment difficulty compared to non-Hispanic White individuals. Higher confidence in discernment ability among these groups could be justified (eg, slightly older adults may be just as technologically savvy as young adults but also have more experience and therefore may be better equipped to make accurate credibility assessments); however, it is also possible that discernment confidence in these groups is misplaced, which would be a cause for concern as it might mean that individuals in these groups are less likely to verify information that might be false (eg, through additional research or by speaking to a health care provider), potentially putting them at greater risk of acting on false information. Additional research is needed to better understand why these groups express higher levels of confidence in their discernment ability.

In contrast, those with lower digital literacy and those with lower numeracy were more likely to report high discernment difficulty. This is perhaps not surprising as research has shown lack of digital literacy to be associated with lower objective ability to successfully judge the accuracy of news stories [21], suggesting that individuals with lower digital literacy may be aware of their limitations in this area. Therefore, digital literacy skills as well as health information evaluation abilities may be important targets for interventions seeking to increase resiliency against misinformation—particularly among more susceptible groups.

The results of this analysis also indicated an association between health misinformation perceptions and distinct behavioral patterns. For example, individuals who perceived high levels of misinformation were less likely to share general health information on social media (perhaps because they are more aware of the problem and are more hesitant to share information that could be false), whereas self-reported discernment difficulty was not similarly associated with sharing behaviors on social media. This finding is in line with the results of previous research showing that confidence in one’s ability to identify factually incorrect information is not significantly associated with likelihood of sharing misinformation [32,33]. Some studies suggest that accuracy may not be the most important factor that people consider when making sharing decisions [34,35], which may help explain why uncertainty about the veracity of information does not significantly influence sharing behavior. Research has shown that interventions that prime individuals to consider accuracy when making sharing decisions on social media could be a promising way to mitigate the spread of misinformation [35], and the results of this analysis suggest that this strategy deserves further attention.

Additionally, the results of this study showed that individuals who perceived high levels of misinformation were less likely to use social media information in making health decisions or in discussions with health care providers. It is possible that, because these individuals perceive high amounts of health misinformation on social media, they are skeptical of the information they encounter on these platforms and, therefore, do not rely on it to inform either their conversations with health care providers or their health decision-making. In contrast, individuals who reported difficulty distinguishing between true and false information on social media were more likely to use information from these platforms in making health decisions and in discussions with health care providers, perhaps because they seek assistance from their health care providers in assessing the credibility of the information. These findings were somewhat counterintuitive, and future research exploring *how* and *why* individuals who report high discernment difficulty use the information they encounter on social media in health-related decisions and discussions could help provide important insights that are beyond the scope of this analysis. For example, exploring whether these individuals are asking for clarification about social media information in discussions with providers versus seeking a "second opinion" on social media after speaking to their clinicians would provide important context to these findings and could help inform how providers can best communicate with patients about information obtained from social media.

Significance

This study offers a unique contribution to our understanding of social media misinformation by focusing on perceptions of the issue rather than objective assessments of misinformation prevalence, exposure, endorsement, or discernment. Assessing perceptions is important because perception of widespread misinformation on social media, as well as perceptions of personal ability to navigate misinformation in web-based spaces, can affect attitudes and behaviors—over and above the impact of actual exposure or ability [16]. In fact, individuals who report high perceived misinformation are likely less susceptible to the direct effects of misinformation (as individuals who characterize a claim as "misinformation" are unlikely to accept it or act on it); however, as demonstrated in this study as well as in previous work, misinformation perceptions can still shape their responses and behaviors [16].

Notably, there are limitations to using self-reported measures of perception—for example, it is impossible to know whether people's perceptions are an accurate reflection of "the ground truth" (ie, whether a lot of the social media information they are exposed to really is or is not false and whether they are really capable of discerning the veracity of social media information) [20]. However, there is still value in assessing these perceptions to obtain a high-level understanding of the public's views on the scope of the problem and the extent to which it affects them as well as their judgment of their own capacity to cope with the problem. In the context of political misinformation, individuals who perceived a lot of exposure to misinformation were more likely to believe that misinformation is a serious problem that creates a lot of confusion about the basic facts of current issues and events and were also more confident in their ability to identify misinformation [20]. Perceptions of the information environment can also impact attitudes and behaviors in ways that are important to health [19]—for example, people may feel overwhelmed and discouraged from seeking additional information about a health topic or develop inaccurate risk perceptions. Additionally, the differences in misinformation perceptions by demographics and literacy levels identified in this study are concerning as they threaten to increase disparities among vulnerable populations. However, while perceptions are important in and of themselves, future research could benefit from including both subjective and objective measures of the information environment to better understand the unique contribution of each construct and provide a more comprehensive understanding of how people respond to social media information.

Nonetheless, the results of this study suggest several practical measures that could help mitigate the impact of misinformation on social media. First, they point to specific populations that may benefit from targeted interventions. For example, those who perceived "a lot" of misinformation on social media were less likely to use this information in health decision-making, suggesting that interventions that raise awareness of information quality issues on social media could limit the extent to which individuals rely on questionable information from social media to make health decisions. These efforts may be especially impactful among groups who are less likely to report perceiving high amounts of misinformation on social media (eg, Black/African American and Hispanic individuals and individuals with lower health literacy). Additionally, the finding that those who express high discernment difficulty still use information from social media to make health decisions suggests that these individuals should be targeted for training interventions that can increase their ability to discern misinformation to (1) ensure that they are relying on accurate information to make these decisions and (2) increase their confidence in their ability to navigate the social media information environment. For example, instructional programs that train people to recognize misinformation techniques have been shown to increase their awareness of these tactics as well as confidence in their ability to successfully deal with misinformation [36]. Furthermore, increasing confidence in discernment ability may itself be a viable target for encouraging careful evaluation of information and increasing resilience to misinformation. For example, Ferrucci and Hopp [37] found that a short intervention providing positive verbal persuasion regarding participants' ability to identify false information on social media increased fake news self-efficacy and that higher self-efficacy beliefs were in turn associated with ability to correctly classify both credible and "fake" news headlines in an information accuracy assessment task.

Second, the finding that those who express high discernment difficulty are more likely to have discussions with health care providers regarding social media health information suggests a need for training aimed at providers to support them in effectively helping patients navigate the information they encounter on the web (eg, teaching providers how to invite these conversations, address misinformation with empathy, and empower patients by recommending accurate sources of information) [38,39]. Research suggests that providers rarely initiate conversations about web-based health information seeking with patients [40,41], but the results of this study indicate that asking about patients' perceptions and use of social media health information could be helpful to incorporate into patient-provider conversations.

Although this study looks at individual-level perceptions and has implications for individual-level interventions (eg, increasing digital literacy), the onus should not be solely on individuals (or providers) to address the problem of social media misinformation. Social media platforms could take steps to decrease the amount of misinformation that users are exposed to in the first place and make it easier for them to discern true from false information (eg, through the use of fact-checking labels and account verification). However, in the absence of these types of more systematic changes in the social media environment, individuals will likely be left to navigate the increasingly confusing information landscape on their own and will need to be supported in their efforts, for example, through campaigns to raise awareness of the issue (particularly among vulnerable populations), training on information evaluation strategies and common misinformation techniques, and encouragement to discuss social media health information with providers and others with relevant expertise. These interventions can be deployed in both web-based and offline contexts (eg, through video advertisements on social media platforms [42] or through educational services delivered in health care settings [43]).

Limitations

This study has several limitations. First, the cross-sectional nature of HINTS data precludes causal inferences from being drawn about observed relationships between variables. Second, the misinformation measures included in this analysis are subjective perception items. As such, there is no way to determine the objective truth about a respondent's actual misinformation exposure or their ability to differentiate true from false information. However, even if they do not reflect objective reality, perceptions are valuable to assess because they enable a better understanding of the public's views on the scope of the misinformation problem and their capacity to cope with it and can help shed light on the way in which perceptions of the information environment shape health-related attitudes and behaviors. Third, the lack of information on certain aspects of respondents' social media experiences and behaviors (eg, the specific platforms they use) is a limitation of this analysis—and reflects a disadvantage of using a national health communication survey that includes only a limited number of items regarding social media use due to space constraints. Finally, the response rate for HINTS 6 (28.1%) was relatively low, which may introduce bias into the data [44]. However, methodological research suggests that the impact of low response rates on data quality may be less significant than previously assumed [44]. Despite these limitations, this analysis provides an important contribution to the broader health misinformation literature as there has been limited research to date focusing on perceptions of misinformation, particularly outside the context of COVID-19.

Conclusions

Many social media users in the United States perceive high levels of misinformation on social media and report difficulty discerning true from false information. This is concerning because perceptions of high misinformation prevalence could increase negative affect (eg, anxiety and worry) regarding health issues, whereas low discernment confidence could result in apathy, confusion, and lower motivation to seek additional information. The fact that health misinformation perceptions were found to vary across race, ethnicity, age, and literacy levels may suggest a need to raise awareness about misinformation and provide training for certain populations (eg, those with low health literacy) to ensure that they approach the information environment with sufficient skepticism and are better able to verify the health claims they see on social media. Finally, the associations between misinformation perceptions and social media–related communication behaviors found in this study can help inform future research as well as health communication interventions and misinformation mitigation efforts. For example, the finding that individuals who have low confidence in their discernment ability are more likely to use social media information to make health decisions and in discussions with health care providers suggests that they may benefit from providers assisting them in navigating and verifying web-based information.

Although a growing body of literature focusing on social media misinformation has emerged in recent years, to date, very little work has been done to look at subjective assessments of the problem of misinformation. This study provides initial insights into the prevalence, disparities, and potential impact of social media misinformation perceptions. However, more research is needed to understand how perceptions of misinformation affect the public's health-related cognitions, attitudes, communication behaviors, and outcomes.

Disclaimer

Any opinions expressed by the authors are their own, and this material should not be interpreted as representing the official viewpoint of the US Department of Health and Human Services, the National Institutes of Health, or the National Cancer Institute.

Conflicts of Interest

None declared.

References

1. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2023-12-15]

2. 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]
3. Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. Health Info Libr J 2017 Dec;34(4):268-283 [FREE Full text] [doi: 10.1111/hir.12192] [Medline: 29045011]
4. Chou WYS, Gaysynsky A, Trivedi N, Vanderpool RC. Using social media for health: national data from HINTS 2019. J Health Commun 2021 Mar 04;26(3):184-193. [doi: 10.1080/10810730.2021.1903627] [Medline: 33856286]
5. Chou WYS, 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]
6. Confronting health misinformation: the U.S. surgeon general's advisory on building a healthy information environment. US Department of Health and Human Services. 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK572169/pdf/Bookshelf_NBK572169.pdf [accessed 2023-12-15]
7. Treen KM, Williams HT, O'Neill SJ. Online misinformation about climate change. Wiley Interdiscip Rev Clim Change 2020 Jun 18;11(5):e665. [doi: 10.1002/wcc.665]
8. Chou WY, Gaysynsky A, Lama Y. The health misinformation ecosystem on social media: emerging evidence and research gaps. In: Keselman A, Smith CA, Wilson AJ, editors. Combating Online Health Misinformation: A Professional's Guide to Helping the Public. Lanham, MD: Rowman & Littlefield; Apr 01, 2023:31-44.
9. 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]
10. Pierri F, Perry BL, DeVerna MR, Yang KC, Flammini A, Menczer F, et al. Online misinformation is linked to early COVID-19 vaccination hesitancy and refusal. Sci Rep 2022 Apr 26;12(1):5966. [doi: 10.1038/s41598-022-10070-w] [Medline: 35474313]
11. 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]
12. Americans' views of misinformation in the news and how to counteract it. Gallup Inc. 2018. URL: https://knightfoundation.org/reports/americans-views-of-misinformation-in-the-news-and-how-to-counteract-it/ [accessed 2023-12-15]
13. Shearer E, Matsa KE. News use across social media platforms 2018. Pew Research Center. 2018 Sep 10. URL: https://www.pewresearch.org/journalism/2018/09/10/news-use-across-social-media-platforms-2018/ [accessed 2023-12-15]
14. Müller P, Schulz A. Facebook or Fakebook? How users' perceptions of 'fake news' are related to their evaluation and verification of news on Facebook. Stud Commun Media 2019;8(4):547-559. [doi: 10.5771/2192-4007-2019-4-547]
15. Yang F, Horning M. Reluctant to share: how third person perceptions of fake news discourage news readers from sharing "real news" on social media. Soc Media Soc 2020 Sep 28;6(3):205630512095517. [doi: 10.1177/2056305120955173]
16. Matthes J, Corbu N, Jin S, Theocharis Y, Schemer C, van Aelst P, et al. Perceived prevalence of misinformation fuels worries about COVID-19: a cross-country, multi-method investigation. Inf Commun Soc 2022 Nov 29;26(16):3133-3156. [doi: 10.1080/1369118x.2022.2146983]
17. Hinsley A, Ju I, Park T, Ohs J. Credibility in the time of COVID-19: cues that audiences look for when assessing information on social media and building confidence in identifying 'fake news' about the virus. Open Inf Sci 2022;6(1):61-73. [doi: 10.1515/opis-2022-0132]
18. Roozenbeek J, Maertens R, Herzog SM, Geers M, Kurvers R, Sultan M, et al. Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. Judgm Decis Mak 2023 Jan 01;17(3):547-573. [doi: 10.1017/s1930297500003570]
19. Park YJ, Chung JE, Kim JN. Social media, misinformation, and cultivation of informational mistrust: cultivating COVID-19 mistrust. Journalism 2022 Apr 08;23(12):2571-2590. [doi: 10.1177/14648849221085050]
20. Barthel M, Mitchell A, Holcomb J. Many Americans believe fake news is sowing confusion. Pew Research Center. 2016 Dec 15. URL: https://www.pewresearch.org/journalism/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/ [accessed 2023-12-15]
21. Sirlin N, Epstein Z, Arechar AA, Rand DG. Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. Misinformation Review 2021 Dec 06;2(6):1-13 [FREE Full text] [doi: 10.37016/mr-2020-83]
22. Nan X, Wang Y, Thier K. Why do people believe health misinformation and who is at risk? A systematic review of individual differences in susceptibility to health misinformation. Soc Sci Med 2022 Dec;314:115398. [doi: 10.1016/j.socscimed.2022.115398] [Medline: 36327631]
23. HINTS 6 methodology report. Westat. 2023. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS_6_MethodologyReport.pdf [accessed 2023-12-15]
24. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
25. 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]
26. Woloshin S, Schwartz LM, Welch HG. Patients and medical statistics: interest, confidence, and ability. J Gen Intern Med 2005 Nov;20(11):996-1000 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.00179.x] [Medline: 16307623]

27. Stagliano V, Wallace LS. Brief health literacy screening items predict newest vital sign scores. J Am Board Fam Med 2013;26(5):558-565 [FREE Full text] [doi: 10.3122/jabfm.2013.05.130096] [Medline: 24004707]
28. Data. National Cancer Institute: Health Information National Trends Survey. 2023. URL: https://hints.cancer.gov/data/Default.aspx [accessed 2023-12-15]
29. Appendix: demographic profiles and party identification of regular social media news consumers in the United States. Pew Research Center. 2023 Nov 15. URL: https://www.pewresearch.org/journalism/2023/11/15/appendix-demographic-and-party-identification-of-regular-social-media-news-consumers-in-the-u-s/ [accessed 2023-12-15]
30. Rivera YM, Moran MB, Thrul J, Joshu C, Smith KC. Contextualizing engagement with health information on Facebook: using the social media content and context elicitation method. J Med Internet Res 2022 Mar 04;24(3):e25243 [FREE Full text] [doi: 10.2196/25243] [Medline: 35254266]
31. Diviani N, van den Putte B, Meppelink CS, van Weert JC. Exploring the role of health literacy in the evaluation of online health information: insights from a mixed-methods study. Patient Educ Couns 2016 Jun;99(6):1017-1025. [doi: 10.1016/j.pec.2016.01.007] [Medline: 26817407]
32. Hopp T. Fake news self-efficacy, fake news identification, and content sharing on Facebook. J Inf Technol Politics 2021 Aug 12;19(2):229-252. [doi: 10.1080/19331681.2021.1962778]
33. Paciello M, Corbelli G, D'Errico F. The role of self-efficacy beliefs in dealing with misinformation among adolescents. Front Psychol 2023;14:1155280 [FREE Full text] [doi: 10.3389/fpsyg.2023.1155280] [Medline: 37275715]
34. Pennycook G, Epstein Z, Mosleh M, Arechar AA, Eckles D, Rand DG. Shifting attention to accuracy can reduce misinformation online. Nature 2021 Apr;592(7855):590-595. [doi: 10.1038/s41586-021-03344-2] [Medline: 33731933]
35. 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 [FREE Full text] [doi: 10.1177/0956797620939054] [Medline: 32603243]
36. Rapp DN, Withall MM. Confidence as a metacognitive contributor to and consequence of misinformation experiences. Curr Opin Psychol 2024 Feb;55:101735. [doi: 10.1016/j.copsyc.2023.101735] [Medline: 38041918]
37. Ferrucci P, Hopp T. Let's intervene: how platforms can combine media literacy and self-efficacy to fight fake news. Commun Public 2023 Oct 31;8(4):367-389. [doi: 10.1177/20570473231203081]
38. Wood JL, Lee GY, Stinnett SS, Southwell BG. A pilot study of medical misinformation perceptions and training among practitioners in North Carolina (USA). Inquiry 2021;58:469580211035742 [FREE Full text] [doi: 10.1177/00469580211035742] [Medline: 34399597]
39. Southwell BG, Wood JL, Navar AM. Roles for health care professionals in addressing patient-held misinformation beyond fact correction. Am J Public Health 2020 Oct;110(S3):S288-S289. [doi: 10.2105/AJPH.2020.305729] [Medline: 33001723]
40. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther 2020 Nov;33(6):e14098. [doi: 10.1111/dth.14098] [Medline: 32725746]
41. Shen MJ, Dyson RC, D'Agostino TA, Ostroff JS, Dickler MN, Heerdt AS, et al. Cancer-related internet information communication between oncologists and patients with breast cancer: a qualitative study. Psychooncology 2015 Nov;24(11):1439-1447 [FREE Full text] [doi: 10.1002/pon.3752] [Medline: 25631285]
42. 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 [FREE Full text] [doi: 10.1126/sciadv.abo6254] [Medline: 36001675]
43. Smith C, Martin-Lillie C, Higano JD, Turner L, Phu S, Arthurs J, et al. Challenging misinformation and engaging patients: characterizing a regenerative medicine consult service. Regen Med 2020 Mar;15(3):1427-1440 [FREE Full text] [doi: 10.2217/rme-2020-0018] [Medline: 32319855]
44. Blake KD, Ottenbacher AJ, Finney Rutten LJ, Grady MA, Kobrin SC, Jacobson RM, et al. Predictors of human papillomavirus awareness and knowledge in 2013: gaps and opportunities for targeted communication strategies. Am J Prev Med 2015 Apr;48(4):402-410 [FREE Full text] [doi: 10.1016/j.amepre.2014.10.024] [Medline: 25700651]

Abbreviations

aOR: adjusted odds ratio
HINTS: Health Information National Trends Survey

Edited by T Mackey; submitted 21.07.23; peer-reviewed by S Ganesh, B Hoffman; comments to author 28.11.23; revised version received 15.12.23; accepted 20.03.24; published 30.04.24.

Please cite as:
Gaysynsky A, Senft Everson N, Heley K, Chou WYS
Perceptions of Health Misinformation on Social Media: Cross-Sectional Survey Study
JMIR Infodemiology 2024;4:e51127
URL: https://infodemiology.jmir.org/2024/1/e51127
doi:10.2196/51127
PMID:38687591

Original Paper

Effects of COVID-19 Illness and Vaccination Infodemic Through Mobile Health, Social Media, and Electronic Media on the Attitudes of Caregivers and Health Care Providers in Pakistan: Qualitative Exploratory Study

Abdul Momin Kazi[1*], MBBS, MPH; Nazia Ahsan[1*], MSc; Rawshan Jabeen[1], MSc; Raheel Allana[1], BDS, MSBE; Saima Jamal[1], MSc; Muhammad Ayub Khan Mughal[1], Grad Dip; Kathryn L Hopkins[2], MPH, PhD; Fauzia Aman Malik[3], MSc, PhD

[1]Aga Khan University, Karachi, Pakistan
[2]Sabin Vaccine Institute, Karachi, Pakistan
[3]University of Texas Southwestern Medical Center, Dallas, TX, United States
[*]these authors contributed equally

Corresponding Author:
Abdul Momin Kazi, MBBS, MPH
Aga Khan University
Stadium Road
Karachi
Pakistan
Phone: 92 3002139610
Email: momin.kazi@aku.edu

Abstract

Background: The COVID-19 pandemic has had a significant impact on different countries because of which various health and safety measures were implemented, with digital media playing a pivotal role. However, digital media also pose significant concerns such as misinformation and lack of direction.

Objective: We aimed to explore the effects of COVID-19–related infodemics through digital, social, and electronic media on the vaccine-related attitudes of caregivers and health care providers in Pakistan.

Methods: This study employs a qualitative exploratory study design with purposive sampling strategies, and it was conducted at 3 primary health care facilities in the province of Sindh, Pakistan. Seven focus group discussions with health care providers and 60 in-depth interviews with caregivers were conducted using semistructured interviews through virtual platforms (ConnectOnCall and Zoom). Transcripts were analyzed through thematic analysis.

Results: Our study reveals the pivotal role of electronic media, mobile health (mHealth), and social media during the COVID-19 pandemic. Four major themes were identified: (1) sources of information on COVID-19 and its vaccination, (2) electronic media value and misleading communication, (3) mHealth leveraging and limitations during COVID-19, and (4) social media influence and barriers during COVID-19. Health care providers and caregivers reported that the common sources of information were electronic media and mHealth, followed by social media. Some participants also used global media for more reliable information related to COVID-19. mHealth solutions such as public awareness messages, videos, call ringtones, and helplines promoted COVID-19 prevention techniques and vaccine registration. However, the overwhelming influx of news and sociobehavioral narratives, including misinformation/disinformation through social media such as WhatsApp, Facebook, and Twitter, were found to be the primary enablers of vaccine-related infodemics. Electronic media and mHealth were utilized more widely to promote information and communication on the COVID-19 pandemic and vaccination. However, social media and electronic media–driven infodemics were identified as the major factors for misinformation related to COVID-19 and vaccine hesitancy. Further, we found a digital divide between the urban and rural populations, with the use of electronic media in rural settings and social media in urban settings.

Conclusions: In a resource-constrained setting like Pakistan, the usage of mHealth, social media, and electronic media for information spread (both factual and mis/disinformation) related to COVID-19 and its vaccination had a significant impact on attitudes toward COVID-19 vaccination. Based on the qualitative findings, we generated a model of digital communications and information dissemination to increase knowledge about COVID-19 and its prevention measures, including vaccination, which can be replicated in similar settings for other disease burdens and related infodemics. Further, to mitigate the infodemics, both digital and nondigital interventions are needed at a larger scale.

(JMIR Infodemiology 2024;4:e49366) doi:10.2196/49366

KEYWORDS

infodemics; mHealth; social media; electronic media; Pakistan; vaccination; misinformation; infodemiology; mobile phone

Introduction

The COVID-19 pandemic has had a deleterious impact on health care systems, economies, and societies globally [1]. This impact has been exacerbated by the resulting COVID-19 infodemic or accompaniment of too much information, including false or misleading information, also occurring in this new digital age of information sharing. With more than 3 billion digital media users globally, digital media has become the key source of information and communication, particularly during a crisis, creating a digital pandemic of information disseminated in multiple forms, regardless of the legitimacy of the sources [2-4]. The downstream effects are confusion, low COVID-19 vaccine confidence, vaccination refusal, and other poorly informed health behavior–related decision-making [5,6]. The most severe consequences are evident in low- and middle-income countries (LMICs) and among marginalized communities [7], where trust in and exposure to official health information sources are comparatively low [8] and there are low health literacy levels, poor health care infrastructure, and less resources [9].

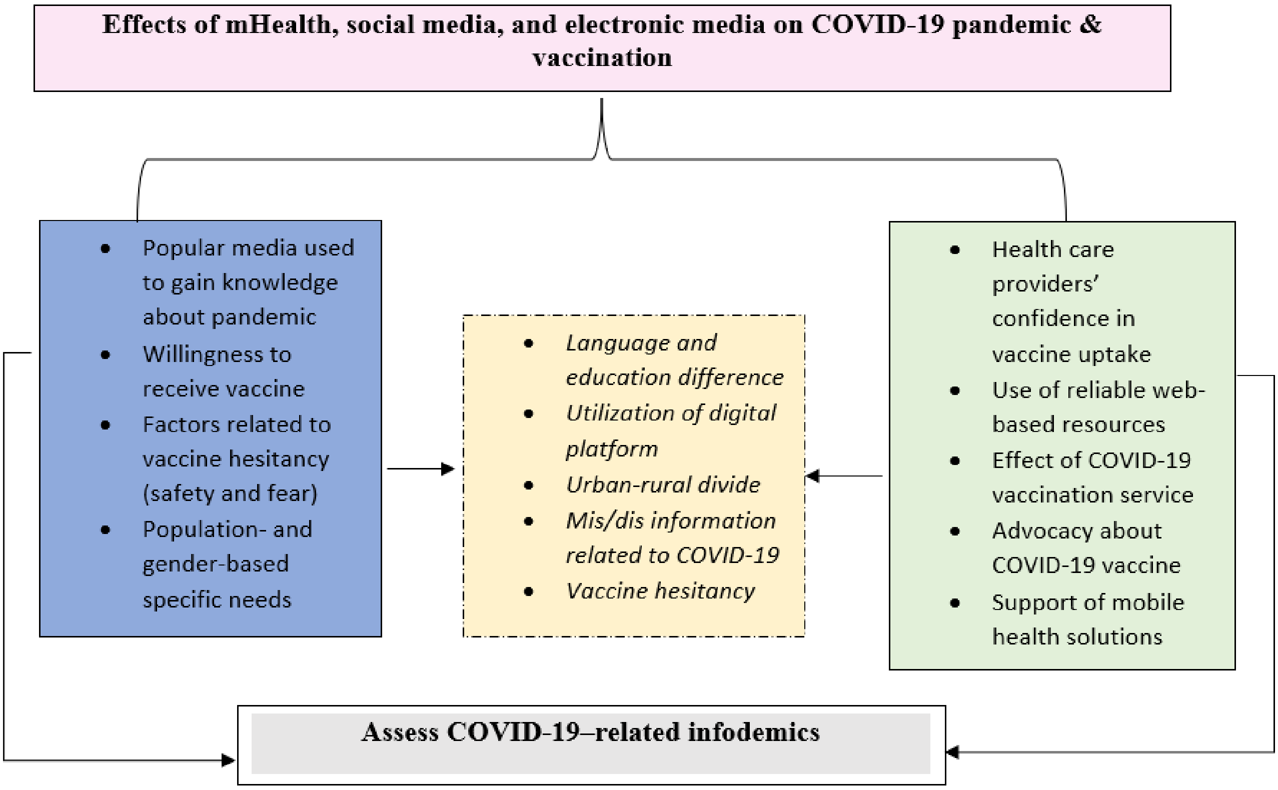
Amid experienced disruptions to health care associated with the COVID-19 pandemic and infodemic, an opportunity also arose. With the advancement of information and technology, digital health played a critical role in COVID-19 response, advocacy, and mobilization [10]. Specifically, digital media assisted in disseminating correct information through mobile health (mHealth)—mobile wireless technologies for public health. These innovations are an integral part of eHealth, which include the cost-effective and secure use of information and communication technologies in support of health and health-related fields, social media–promoted public health initiatives, and electronic media–raised awareness, and these encouraged preventative measures (eg, hand hygiene, vaccine uptake) [11].

During the COVID-19 pandemic, many countries developed mHealth apps to assist with the identification of prevalent symptoms for self-assessment, implementing contact tracing, disseminating information, minimizing exposure, and reducing face-to-face interaction between patients and health workers [12]. A WhatsApp chatbot app in South Africa used machine learning technology to provide free, automated responses to user queries on COVID-19, relating to travel advice, recent statistics, symptoms, and debunking of myths and misinformation [13]. Facebook groups were utilized by most health care professionals during the pandemic to discuss and integrate real-time experiences in COVID-19 treatment [14]. Data visualization dashboards enabled data-driven infographics representing global-to-local pandemic-related statistics, which allowed for the public and researchers to comprehend and track the pandemic in real time [15]. Additionally, social media channels were used to inform citizens about pandemic-related government response efforts and updates such as the national WhatsApp channel established in Singapore during the COVID-19 surge [16].

Based on the described initiatives above during the COVID-19 pandemic, we have developed a comprehensive framework highlighting the key roles of various digital platforms. This framework highlights how digital tools played diverse and complementary roles in pandemic management—from disseminating information and assessing symptoms to real-time data tracking and communication. These tools were pivotal in proactive intervention, personalized guidance, knowledge exchange among experts, data-driven decision-making, and fostering community resilience through amplified public health messaging and grassroots initiatives (Figure 1).

Pakistan has faced challenges related to vaccine hesitancy and low vaccine acceptance in the past. Pakistan has a history of vaccine-preventable disease outbreaks due to various factors, including misinformation, cultural beliefs, lack of awareness, and mistrust in vaccines. One notable example is the polio eradication efforts in Pakistan. Despite considerable progress made globally, Pakistan has remained one of the few countries where polio cases continue to be reported [17]. Further, Pakistan has diverse cultural and linguistic backgrounds, where different regions may have unique sociodemographic factors that influence vaccine acceptance and hesitancy [18]. Moreover, the COVID-19 pandemic has brought about new challenges and concerns regarding vaccine acceptance globally. It is essential to examine the role of digital media and mHealth interventions in vaccine acceptance, as uptake differs across regions in Pakistan. Factors such as regional beliefs, levels of trust in digital media sources, and access to health care information may vary, leading to different outcomes in vaccine-related decision-making. By exploring these regional differences, we can provide valuable insights into the role of digital health interventions in addressing vaccine hesitancy in diverse settings. It is vital that these digital health interventions continue to be developed to harness social media for the public good and to increase trust in vaccines and vaccination, especially within LMICs like Pakistan. Thus, we explored the effects of mHealth, social media (eg, Facebook, Twitter, Instagram), and electronic media (ie, television and radio) during the COVID-19 pandemic and its association with COVID-19 vaccination and childhood routine immunization acceptance and uptake among parents and child caregivers and health care providers (HCPs) in a resource-constrained setting like Pakistan.

Figure 1. Effects of mobile health, social media, and electronic media during the COVID-19 pandemic. mHealth: mobile health.

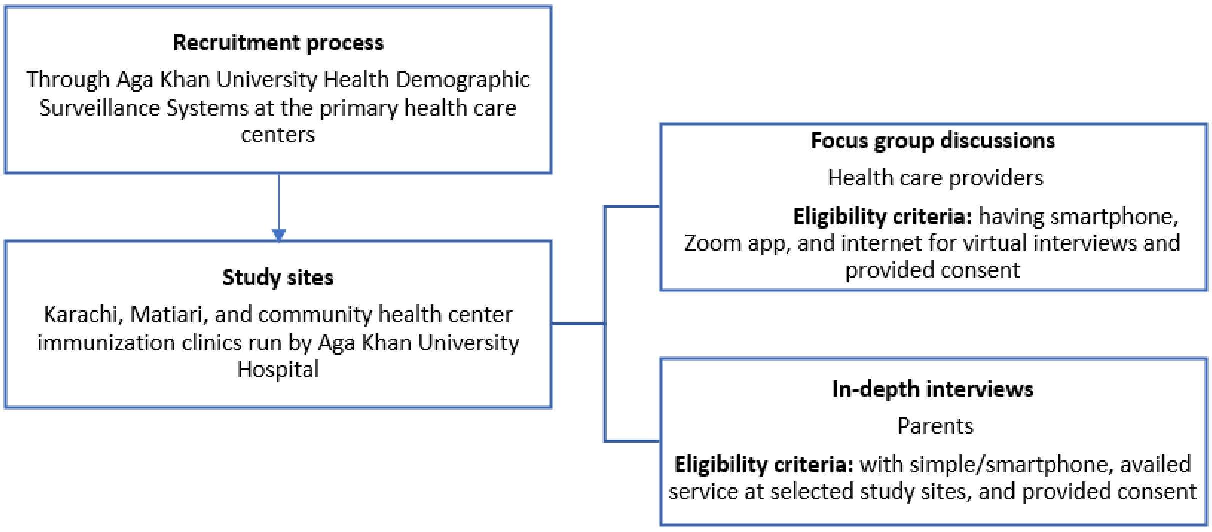


Methods

Study Design, Setting, and Population

This exploratory qualitative research was employed to assess the role of infodemics through mHealth, electronic media, and social media in Pakistan to explore the unique experiences and insights of populations, which may be overlooked in quantitative studies [19,20]. This study was implemented between May 2020 and August 2021 at 3 sites in Pakistan: the periurban Aga Khan University (AKU) Health Demographic Surveillance System at primary health care centers of Ali Akbar Shah Goth and Bhains Colony, the rural Sindhi district of Matiari, and the community health center (CHC) vaccination clinic at the AKU hospital in urban Karachi. The 3 selected sites in Pakistan were chosen to represent diverse populations and to capture unique socioeconomic and cultural factors influencing vaccine acceptance and hesitancy. The vaccination center at the periurban sites targeted low-to-middle and low-income populations. The rural area represented a low-income population, and the vaccination center at AKU hospital (urban site) represented both high-income and low-to-middle-income populations. This approach ensured a comprehensive understanding of vaccination behaviors by assessing the role and impact of mHealth, social media, and electronic media on vaccine-related attitudes and behaviors across different socioeconomic contexts. Focus group discussions (FGDs) were implemented among health care workers (doctors, nurses, pharmacists, lady health visitors, vaccinators, lady health workers, and community health workers) in Karachi and Matiari who worked in selected centers, and in-depth interviews (IDIs) were conducted with parents/caregivers of children aged <1 year. Figure 2 shows the flowchart of participant recruitment.

Figure 2. Flowchart depicting participant enrollment.



Ethics Approval

Ethics approval for this study was obtained from the ethics review committee of AKU (2020-5316-14620).

Inclusion Criteria and Sampling Approach

We used purposive sampling for recruiting HCPs, and convenient sampling was performed for parent selection. Parents or child caregivers were eligible for IDIs if (1) they had at least one child younger than 1 year and (2) their telephonic contact number was listed in either the registry maintained by the AKU hospital and its associated centers at the study sites or provided by community health workers affiliated with the study sites. The IDI was conducted with 1 caregiver at a time. HCPs such as doctors, nurses, pharmacists, lady health visitors, vaccinators, and community health workers at each of the 3 study sites were eligible as FGD participants; 6-8 participants were included in the FGDs. The parents/caregivers and HCPs who did not provide consent were excluded from this study.

Study Framework and Tool

Semistructured qualitative interview guides were developed for both IDI and FGD data collection in English and local Urdu language with the help of a literature review [21-23]. The main topics discussed were access to electronic and social media among diverse populations, barriers and perceived challenges, caregivers/parental concerns about vaccine safety, understanding of content available on digital media, and word-of-mouth communication within the community. Moreover, HCPs were more focused on factors associated with COVID-19 fear and dis/misinformation in the data, which may lead to vaccine hesitancy. We used the process defined in Figure 2 to gather data.

Data Collection and Management

A team of researchers from AKU designed and piloted the research tools and trained qualitative research staff for conducting FGDs and IDIs. Due to the ongoing pandemic, all data collection was conducted remotely as per COVID-19 standard operating procedures (SOPs). Seven FGDs with HCPs and 60 IDIs with caregivers were conducted using semistructured interviews through virtual platforms (ConnectOnCall and Zoom) till the point of saturation. Two alternative methods were employed to acquire data: IDIs of 40-60 minutes were conducted telephonically, while FGDs of 60-90 minutes with HCPs were conducted using the Zoom videoconferencing platform. After obtaining verbal informed consent for participation, the respondents affiliated with all 3 study sites were given an overview of the study objectives, and parents/caregivers were provided training or guidance on how to attend virtual interviews on phone, which was helpful for individuals who were not tech-savvy or had limited experience with virtual communication. HCPs were debriefed about the Zoom platform and its related features to avoid any hindrance while interviews were being conducted. Data were recorded on the phone and a Zoom password-protected device.

Data Analysis

Thematic analyses were adopted as part of the qualitative study. Each audio recorded Sindhi and Urdu-spoken IDIs, and FGDs were transcribed and translated to English. The English transcripts were then assessed and coded separately and connected to text fragments that reflected crucial user perspectives. The data were subsequently organized into thematic categories by looking for topics and then reviewed, defined, and named. We ensured trustworthiness of the data analysis by using Lincoln and Guba guidelines [24] to reduce researcher bias. Further, all discrepancies were resolved after team discussion; codes were finalized to generate the major themes emerging through FGDs and IDIs. The team then contrasted the themes before more targeted coding, focusing on ideas linked to mHealth, social media, and electronic media.

Results

COVID-19–Related Infodemics

The key findings with regard to the COVID-19–related infodemic are listed in Table 1.

Table 1. COVID-19–related infodemic.

Subtheme	Key findings
Electronic media infodemic	• Overselling of news • Both factual and false information • Fabricated news of COVID-19 and vaccination • Influence of negative political narratives • Anxiety due to exaggerated breaking news
Social media infodemic	• Negative impact of videos (burial of dead due to COVID-19) • Overhype of content, creating fear of contracting infection • Sharing of conflicting posts (audios and videos) • False information through nonmedical professionals • Myths about adverse effects of COVID-19 vaccine: brain damage, infertility, and death
Mobile health infodemic	• Lack of pandemic-related information in rural populations • Limited access to mobile phone • Parents were unable to understand information about COVID-19 information through ringtones • Reluctant about sharing personal identification number required for vaccination registration

Access to Information Through Print, Electronic, and Digital Media

HCPs and caregivers shared that COVID-19–related information was commonly disseminated through print media, mHealth, and social and electronic media. The HCPs of the urban site reflected their beliefs about the use of the medium as authentic, as mobile phones (WhatsApp) were easily accessible and conveniently used. Moreover, HCPs also used Facebook to access information from the official pages of World Health Organization, Centers for Disease Control, and other international websites. Although a few caregivers also used international media for more reliable information related to COVID-19, caregivers in rural areas showed concern about the lack of media access and limited services on mobile networks.

> *…I get credible information from Google. We also access COVID-related information from Facebook pages of WHO, CDC and other international health organizations and they (HCPs) received official updated information through WhatsApp by senior management.* [FGD, HCP from Karachi]

> *…Women do not own mobile phones. Usually, they are used by males in our community.* [FGD, HCP from Matiari]

> *…COVID-19–related SMS is important in those areas where there are not many ways of information like small cities and villages, where there is no internet available.* [IDI, parent/caregiver from CHC]

mHealth: Perceived mHealth Solution Usefulness and Challenges During COVID-19

During the pandemic, SMS text message–based mobile interventions were used to promote COVID-19 prevention techniques and register the general population for COVID-19 vaccination at both urban and rural sites. These messages tried to influence public behaviors such as SOPs related to COVID-19 prevention, follow-up appointments, and immunization advantages. Caregivers endorsed these government services. Caregivers shared that they received messages for the registration for COVID-19 vaccination of people older than 60 years. HCPs stated that conducting vaccination registration with national identity card numbers was a good initiative and felt an automated system that shared reminder messages would help reduce the defaulters.

> *…Public service-SMS are beneficial in spreading awareness about COVID-19 pandemic, SOPs, and getting vaccinated against COVID-19.* [FGD, HCP from Matiari]

> *…The public service-SMS is about COVID-19 vaccine registration, if your age is above 60 then you can get your vaccines, for registration send your CNIC number to 1166.* [IDI, parent/caregiver from CHC]

Urban caregivers were suspicious whenever the national identity card number was asked for vaccination or other treatments, as a circulating rumor via Twitter stated that these local data were being gathered for use by foreign agencies such as the United States or China.

> *…They ask for your NIC number when you go to get your COVID-19 test done. There were such gossips making rounds that your data will be sold to China or USA or on Twitter, like China and USA will control the world.* [IDI, parent/caregiver from CHC]

Social Media

Community Access to Social Media and Information

HCPs perceived that social media, including WhatsApp, YouTube, and Facebook, were used as a positive channel for circulating updated information during the COVID-19 pandemic. However, they also shared that vaccine hesitancy increased among the general population because of the social media infodemic. Caregivers in urban areas reported using social media such as Facebook and the internet to learn about COVID-19 prevalence and news updates, but those in rural areas reported that they never used social media.

> *…Social media (Facebook, WhatsApp, Instagram, YouTube) is the first source of information about COVID.* [FGD, HCP from CHC Karachi]

Conflicting News Related to COVID-19 and Its Spread Through Social Media

Caregivers shared that the news circulating via social media were largely negative and fabricated, creating fear among the general population in both settings. Participants felt social media was full of opinionated people who deliberately post to impose their personal beliefs on the public.

> *…Things shared on social media are completely meaningless. Everyone became a doctor and began to demonstrate their competence. Social media is rife with stories about people becoming unwell after receiving their first dose of vaccination. People are skeptical of the COVID-19 vaccine.* [IDI, parent/caregiver from Matiari]

> *…We get awareness through these messages, but sometimes rumors are being spread through social media posts, so we need to check the information being shared and look at its sources and determine if the messages are credible and true or false* [FGD, HCP from Karachi]

COVID-19 Pandemic and Vaccine Rumors Spread Through Social Media

According to these caregivers, the news about vaccine-related adverse events were shared by trusted community members such as friends and relatives on platforms such as WhatsApp and Facebook. This information spread quickly, leading to skepticism about the safety and efficacy of COVID-19 vaccines. A few caregivers reported that the news on social media shared that COVID-19 immunizations caused brain stroke/clotting and that this news led to vaccine hesitancy.

> *…People got hesitant to receive the COVID-19 vaccination after reading on social media that it causes brain coagulation.* [IDI, parent/caregiver from CHC]

> *…Then there were rumors that so many people died or got infected after getting the vaccine.* [FGD, HCP from Matiari]

One of the significant challenges regarding news was the uncertainty of the delivery of accurate information and its understanding among the population. Caregivers shared multiple rumors regarding COVID-19 immunization such as it caused infertility in unmarried people and congenital malformations in children of vaccinated pregnant women and increased mortality rate among vaccinated people.

> *…After getting COVID-19 vaccine, you might become infertile and unable to conceive. There were rumors that many people died or got infected after COVID-19 vaccine.* [FGD, HCP from Matiari]

Electronic Media

Role of Electronic Media Platforms During the COVID-19 Pandemic

Electronic media was the main source of information for the rural population. The caregivers reported that electronic media had a significant impact on communities and that international media such as the British Broadcasting Corporation (BBC) never fabricated information as compared to the other national news channels. Further, COVID-19 news about morbidity and mortality was telecast as a sensational issue. According to HCPs, there were many erroneous calls and SMS text messages with incorrect information circulating among the caregivers. Nonetheless, the information was validated by HCPs before passing it to others. Further, they double-checked every news item from official sources before sending it to others.

> *…In my opinion, news that is broadcasted on BBC is not fabricated. So, media should encourage more programs and talk shows on COVID-19 prevention and its vaccination rather than sensational news on morbidities and mortalities.* [IDI, parent/caregiver from CHC]

> *…Public service messages are important in creating awareness about COVID-19 pandemic only if they are sent from a relevant source like official government websites or numbers, then you know that this is a credible bit of information.* [FGD, HCPs from CHC Karachi]

COVID-19 Vaccine Hesitancy and Misleading Information Spread Through Electronic Media

HCPs expressed that sometimes nonmedical personalities shared their opinions on the news, which confused the population and negatively influenced other individuals' vaccination behavior. One of the caregivers in the urban site expressed that he felt worried and stressed after receiving COVID-19–related news. The rural population expressed comparable concerns about news-related anxiety. Caregivers believed that both positive and negative news broadcasted on television and other media influenced the population's mental health.

> *…I heard Provincial Government's statement on television that if you want to get this vaccine then get it at your own risk. If government officials continue to talk like this, it will create a negative impact regarding vaccines in the minds of the people. Educational people will also have uncertainties and concerns because of it.* [FGD, HCP from Matiari]

Discussion

Principal Findings

The COVID-19 pandemic has had both health and societal implications, which were not comparable to any recent event in global history. Social distancing during COVID-19 had a great impact, which led to an increase in the widespread use of digital platforms for acquiring the latest updates and COVID-19–related communication. Our qualitative study highlights many critical issues based on the usage of mHealth and social and electronic media related to the COVID-19 pandemic and vaccine coverage. The primary findings of our study indicate that according to HCPs and caregivers, electronic media and mHealth were used more broadly to promote COVID-19 pandemic-related information and communication. However, the social and electronic media–driven infodemics and urban-rural divide posed major obstacles. Further, digital media not only enabled individuals to quickly access and share information on disease spread and emerging health policy changes but also provided access to resources, community engagement, and misinformation.

Our study shows electronic media as the most effective tool for communicating health-related messages such as awareness of COVID-19 SOPs and vaccination to reduce the general population's fear of acquiring infection. Electronic media provided the most reliable information related to the pandemic and hence became the most dependable medium to receive information, especially for the rural population. Electronic media was the primary source for filling the information gap among the general populations in Pakistan. However, there was an element of misinformation and disinformation in the news broadcasts related to the COVID-19 pandemic, which created confusion and misled the general population. This finding was similar to findings in Korea [25], wherein using electronic media to access vaccine-related information had a positive impact on the population's vaccination-related decision-making as well as higher perceived benefits of the COVID-19 vaccine and greater trust in the government to address vaccine hesitancy. Further, an Italian study highlighted the importance of health information transmitted to raise awareness among individuals through radio, television, and journalistic communications [26]. However, despite motivating the population to reduce their fear and stress related to the COVID-19 pandemic, television broadcasts commonly used complicated language, technical jargon, and disseminated debates, thereby creating more confusion about the situation [27].

Our study explores the usefulness of mHealth as a communication media to disseminate information and support the population and HCPs during the pandemic. Participants in our study reported that mHealth strategies such as ringtone messages disseminated COVID-related awareness and preventive strategies and were actively used as a communication medium about public health awareness messages during different phases of the pandemic in the overall population. Most people could understand the ringtone because it was played in local languages. A similar strategy has been recommended to replace entertainment and religious ringtones with health promotion, particularly for LMICs [28]. Our findings also revealed that HCPs approached the COVID-19 dashboard and national and international resources for credible sources of information, which motivated them to deliver services and enabled vaccination uptake. SMS text messaging also played a critical role in the web-based registration for the COVID-19 vaccine. Moreover, reminders were automatically forwarded to the general population for missing doses of COVID-19 vaccine. Another study revealed that sending SMS text messages, emails, and postal messages enhanced influenza vaccination appointments [29]. According to our study participants, SMS text message–based vaccine registration was significantly accepted by the population and could potentially be further expanded to other adult and child vaccinations in Pakistan and other LMICs.

During the pandemic, the role of social media was widespread in providing information about COVID-19 and its vaccines through Facebook, Twitter, and WhatsApp. However, social media also propagated misinformation and disinformation related to COVID-19, hence introducing the phenomena of infodemics. In Pakistan, WhatsApp was widely used for disseminating pandemic-related information. Similarly, a study conducted in the United Arab Emirates reported WhatsApp as being mostly used to acquire COVID-19–related information [30].

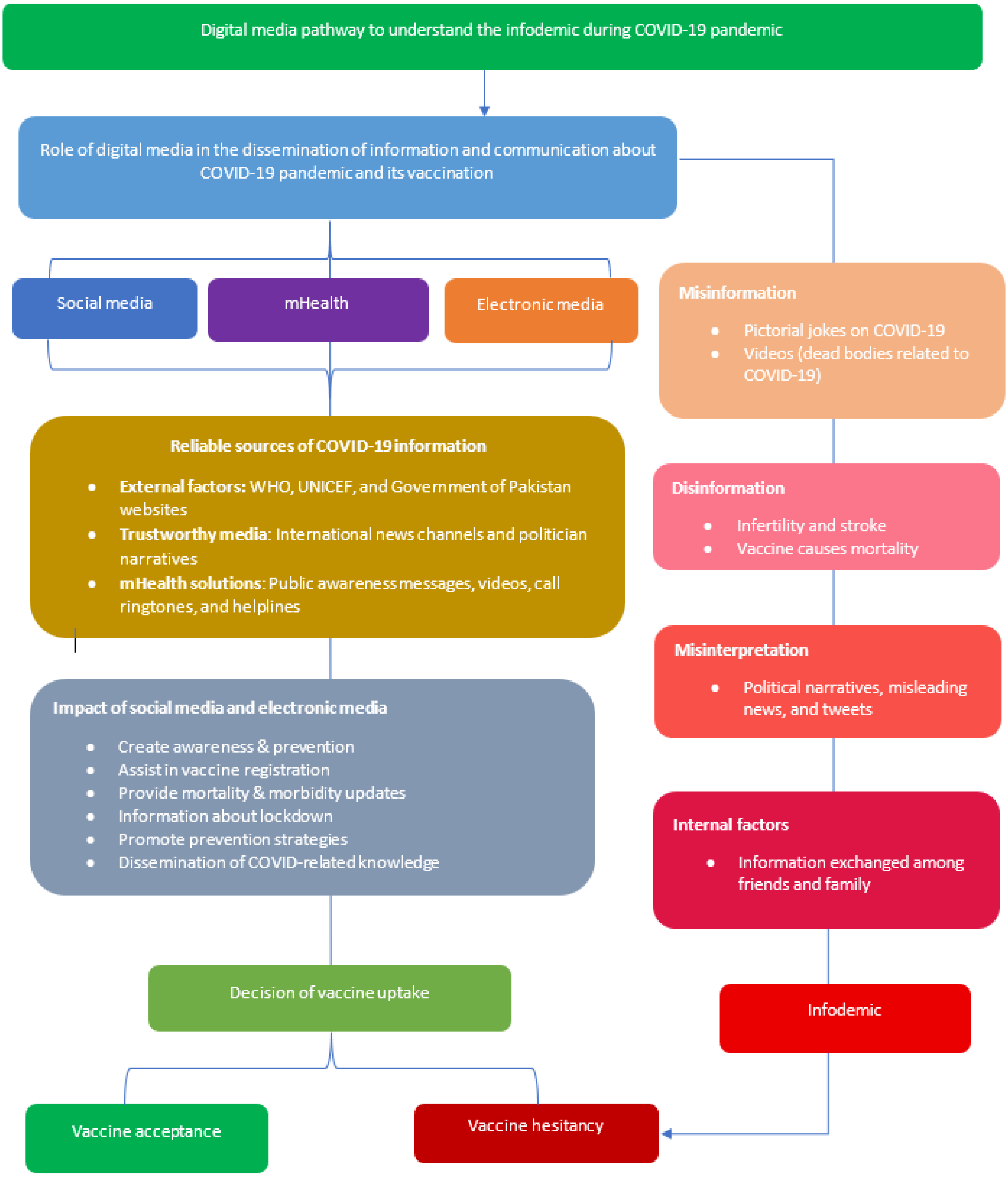
We found that social media news and its content contributed to increasing COVID-19–related anxiety and stress among the population, as stated by HCPs. In our study, political tweets received more engagement, as leaders' tweets had an enormous influence on the public, creating confusion and misconception. Another study showed that leaders' tweets had higher interactions in high-income nations like the United Kingdom as well, and the leader's tweets had a stronger impact on the general population, creating misinterpretation, which ultimately impacted vaccine uptake [31].

During the pandemic, information regarding vaccine adverse effects and misinformation was circulating on television and social media, adding fear and uncertainty among the population, and leading to a decline in vaccine acceptance. Our study participants also reported infertility, brain stroke, and death-related misinformation related to COVID-19 vaccines. Our findings are consistent with those of a US-based study, which reported similar misinformation falsely correlating COVID-19 vaccination with infertility and population growth control, electronic tattooing or microchipping individuals for global surveillance, and autism, which resulted in low vaccine uptake by the general population overall [32]. Thus, our HCP participants emphasized that lack of news verification and monitoring and incorrect information dissemination resulted in confusion and vaccine hesitancy among caregivers. This qualitative study provides insights that lack of technological literacy in the rural population was a hindrance in adopting mHealth interventions in the rural setting. Our study participants were not aware of or did not report about the official COVID-19 Government of Pakistan website. Besides that, the rural population, especially women, did not own personal mobile phones and were unfamiliar with its usage due to lack of technological literacy. However, in a rural setting in India, where the literacy rate of women was 40.35%, nearly 85% of the rural illiterate women were found to be using a mobile phone without necessarily owning it. It was their quickest means of communication and receiving information [33].

The digital pathway model that we conceptualized is based on national policy interventions (SMS text message–based interventions, caller tunes, vaccine registration through SMS text messages) and the responses of our research participants, which enabled them to communicate and disseminate information during the COVID-19 pandemic and assisted them in the fight against pandemic-related infodemics (Figure 3). We have discussed the diversity of mHealth interventions and their usage along this pathway. This framework is required to understand how digital methods might be integrated into COVID-19 control efforts and to aid in future pandemic preparedness (Textbox 1).

We examined platforms commonly used in misinformation campaigns in our settings, such as mHealth and social and electronic media. The flowchart's development revealed a critical insight into how different media play distinct roles in supporting the spread of misinformation via different paths. Our model has crucial information for policy makers who want to combat the phenomena of infodemics and the digital divide. Policy makers at the government level across the world need to strategize social and legal regulatory frameworks to curb the spread of misinformation and disinformation in digital media.

Figure 3. Infodemic pathway of digital media. mHealth: mobile health; UNICEF: United Nations International Children's Emergency Fund; WHO: World Health Organization.



Textbox 1. Digital methods that were integrated into COVID-19 control efforts and that can aid in future pandemic preparedness.

- Electronic media provided coverage to the overall population through advertisements, news, and broadcasts, minimizing the digital communication gap between the urban and rural populations.
- Mobile health–based initiatives such as SMS text messages, calls, helplines, and caller ringtones were the sources of public awareness and information, which were enhanced using local languages. Other external media such as World Health Organization websites and news channels were used as reliable resources.
- Social media enabled wide access to information, particularly in the urban population, and dissemination of misinformation and disinformation about vaccines and pandemics.
- Electronic media, mobile health, and social media were the major modes of communication and dissemination during the pandemic, which promoted COVID-related standard operating procedures and vaccine acceptance. However, these same media also became the source of infodemics, leading to misguided information and vaccine hesitancy.

Study Limitations

Our results were based on participants visiting primary health care centers operated by private health entities with limited opening hours, staffing, and financial resources, constituting a significant study limitation. Further, the findings related to COVID-19 vaccine hesitancy cannot be generalized to other groups with different social and cultural backgrounds of other provinces. Thus, quantitative research is required to mitigate infodemics in larger populations.

Implications and Recommendations

The COVID-19 pandemic has necessitated the use of digital platforms for communication and information dissemination, with electronic media, mHealth, and social media playing crucial roles. Electronic media emerged as an effective tool for promoting COVID-19–related information and reducing fear among the general population. However, it also became a source of confusion causing mis/disinformation. Further, to improve public health interventions, precise and accurate information that pertains to World Health Organization or Centers for Disease Control guidelines should be allowed to be aired on televisions and media so that they can function as a bridge for people to connect with health officials and local governments for assistance, and collaborations between government and media outlets can establish guidelines for responsible reporting [34]. mHealth strategies such as SMS text message notifications, ringtone messages, and mobile apps proved useful in raising public awareness, especially in rural areas. Policy makers should invest in user-friendly mHealth platforms to ensure that timely and accurate information reaches individuals, particularly in low-income settings. Lastly, social media platforms also played a pivotal role in providing pandemic-related information, which was not previously seen on that scale, but they unfortunately also contributed toward infodemics. Regulatory bodies should develop frameworks to mitigate misinformation on social media, including collaboration with social media companies, monitoring mechanisms, and awareness campaigns. These frameworks should include collaboration with platforms for fact-checking and prioritizing authoritative sources, real-time monitoring using artificial intelligence, public awareness campaigns promoting critical thinking, and regulatory oversight ensuring transparency and accountability [35]. Addressing the challenges posed by infodemics and the digital divide requires enhancing information verification, promoting digital media literacy in rural areas, and strengthening public health communication through partnerships with influencers and HCPs. By implementing these strategies, policy makers can improve information dissemination, mitigate misinformation, and enhance future pandemic preparedness efforts.

Conclusion

This study proposes the implementation of a communication pathway focused on diseases and pandemics to be integrated into a national digital policy in resource-constraint setups, including Pakistan. This would enhance the country's readiness to respond to health crises and boost public awareness and comprehension of these matters. Although digital media emerged as the primary source of information during the pandemic, it also contributed to misinformation and disinformation, causing infodemics. It is therefore essential to comprehend the sources and content of information within each digital medium element that triggers infodemics. LMICs are more vulnerable to infodemics because they have limited access, awareness, and lower literacy levels to comprehend and evaluate health-related information. Our research findings further revealed a digital divide between urban and rural populations, resulting in digital inequalities. To address these challenges, both digital and nondigital solutions must play a vital role. Moreover, credible information must be widely disseminated from trustworthy sources, verified by subject matter experts, and tailored to fit the local context. Lastly, training programs on the usage of digital media, dissemination strategies, and information reliability need to be conducted, particularly for HCPs and different settings, especially to reach rural communities.

Acknowledgments

This study was funded by the Sabin Vaccine Institute (agreement: 050110-00).

Authors' Contributions

AMK, FAM, and NA conceptualized this study. AMK, NA, FAM, and RJ contributed to the methodology. NA and RJ performed the thematic analysis. NA and SJ performed all the investigations. SJ provided the resources. NA and RJ curated the data. AK contributed to technology use. NA, RJ, and RA contributed to writing and preparing the original draft. KLH, NA, RJ, and RA contributed to the writing—review and editing. AMK and FAM contributed to supervision. AMK and NA administered the project. KLH acquired the funds. All authors have read and agreed to the published version of the paper.

Conflicts of Interest

None declared.

References

1. Khattab MF, Abou-Madawi AM. Current effect of COVID-19 global pandemic on the professional and life profiles of the Egyptian spine surgeons. SICOT J 2020;6:31 [FREE Full text] [doi: 10.1051/sicotj/2020029] [Medline: 32819456]
2. Nasajpour M, Pouriyeh S, Parizi RM, Dorodchi M, Valero M, Arabnia HR. Internet of things for current COVID-19 and future pandemics: an exploratory study. J Healthc Inform Res 2020;4(4):325-364 [FREE Full text] [doi: 10.1007/s41666-020-00080-6] [Medline: 33204938]
3. 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]
4. Appel G, Grewal L, Hadi R, Stephen AT. The future of social media in marketing. J Acad Mark Sci 2020;48(1):79-95 [FREE Full text] [doi: 10.1007/s11747-019-00695-1] [Medline: 32431463]
5. Purnat TD, Vacca P, Czerniak C, Ball S, Burzo S, Zecchin T, 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 Jul 28;1(1):e30971. [doi: 10.2196/30971]
6. Garett R, Young S. Online misinformation and vaccine hesitancy. Transl Behav Med 2021 Dec 14;11(12):2194-2199 [FREE Full text] [doi: 10.1093/tbm/ibab128] [Medline: 34529080]
7. Ehrlich H, McKenney M, Elkbuli A. Protecting our healthcare workers during the COVID-19 pandemic. Am J Emerg Med 2020 Jul;38(7):1527-1528 [FREE Full text] [doi: 10.1016/j.ajem.2020.04.024] [Medline: 32336585]
8. Shet A, Carr K, Danovaro-Holliday MC, Sodha SV, Prosperi C, Wunderlich J, et al. Impact of the SARS-CoV-2 pandemic on routine immunisation services: evidence of disruption and recovery from 170 countries and territories. Lancet Glob Health 2022 Feb;10(2):e186-e194 [FREE Full text] [doi: 10.1016/S2214-109X(21)00512-X] [Medline: 34951973]
9. Marzo RR, Su TT, Ismail R, Htay MNN, Essar MY, Chauhan S, et al. Digital health literacy for COVID-19 vaccination and intention to be immunized: A cross sectional multi-country study among the general adult population. Front Public Health 2022;10:998234 [FREE Full text] [doi: 10.3389/fpubh.2022.998234] [Medline: 36187686]
10. Saher R, Anjum M. Role of technology in COVID-19 pandemic. Researches and Applications of Artificial Intelligence to Mitigate Pandemics 2021 Apr 30:109-138. [doi: 10.1016/B978-0-323-90959-4.00005-5]
11. Bao H, Cao B, Xiong Y, Tang W. Digital media's role in the COVID-19 pandemic. JMIR Mhealth Uhealth 2020 Sep 18;8(9):e20156 [FREE Full text] [doi: 10.2196/20156] [Medline: 32530817]
12. Asadzadeh A, Kalankesh LR. A scope of mobile health solutions in COVID-19 pandemics. Inform Med Unlocked 2021;23:100558 [FREE Full text] [doi: 10.1016/j.imu.2021.100558] [Medline: 33842688]
13. Matiashe F. WHO is raising coronavirus awareness globally using a WhatsApp bot developed in South Africa. Quartz. 2020 Mar 26. URL: https://qz.com/africa/1826415/coronavirus-who-adopts-south-african-whatsapp-health-alert [accessed 2024-08-23]
14. Xun H, He W, Chen J, Sylvester S, Lerman SF, Caffrey J. Characterization and comparison of the utilization of Facebook groups between public medical professionals and technical communities to facilitate idea sharing and crowdsourcing during the COVID-19 pandemic: cross-sectional observational study. JMIR Form Res 2021 Apr 30;5(4):e22983 [FREE Full text] [doi: 10.2196/22983] [Medline: 33878013]
15. Ivanković D, Barbazza E, Bos V, Brito Fernandes, Jamieson Gilmore K, Jansen T, et al. Features constituting actionable COVID-19 dashboards: descriptive assessment and expert appraisal of 158 public web-based covid-19 dashboards. J Med Internet Res 2021 Feb 24;23(2):e25682 [FREE Full text] [doi: 10.2196/25682] [Medline: 33577467]
16. Liu JCJ, Tong EMW. The relation between official WhatsApp-distributed COVID-19 news exposure and psychological symptoms: cross-sectional survey study. J Med Internet Res 2020 Sep 25;22(9):e22142 [FREE Full text] [doi: 10.2196/22142] [Medline: 32877349]
17. Abbas B, Abbas S, Rafique A, Aslam S. Fallacy of COVID-19 vaccine coverage: discovering vaccine hesitancy and compliance in private dental practitioners and hospital-based dentists to improve immunization levels in Pakistan. EJDENT 2021 Sep 30;2(5):13-18. [doi: 10.24018/ejdent.2021.2.5.94]
18. Hussain SF, Boyle P, Patel P, Sullivan R. Eradicating polio in Pakistan: an analysis of the challenges and solutions to this security and health issue. Global Health 2016 Oct 12;12(1):63 [FREE Full text] [doi: 10.1186/s12992-016-0195-3] [Medline: 27729081]

19. Kennedy J. Populist politics and vaccine hesitancy in Western Europe: an analysis of national-level data. Eur J Public Health 2019 Jun 01;29(3):512-516. [doi: 10.1093/eurpub/ckz004] [Medline: 30801109]
20. Eller NM, Henrikson NB, Opel DJ. Vaccine information sources and parental trust in their child's health care provider. Health Educ Behav 2019 Jun;46(3):445-453 [FREE Full text] [doi: 10.1177/1090198118819716] [Medline: 30616381]
21. Facciolà A, Visalli G, Orlando A, Bertuccio MP, Spataro P, Squeri R, et al. Vaccine hesitancy: An overview on parents' opinions about vaccination and possible reasons of vaccine refusal. J Public Health Res 2019 Mar 11;8(1):1436 [FREE Full text] [doi: 10.4081/jphr.2019.1436] [Medline: 30997357]
22. Signorelli C, Odone A, Ricciardi W, Lorenzin B. The social responsibility of public health: Italy's lesson on vaccine hesitancy. Eur J Public Health 2019 Dec 01;29(6):1003-1004. [doi: 10.1093/eurpub/ckz135] [Medline: 31808818]
23. Benin AL, Wisler-Scher DJ, Colson E, Shapiro ED, Holmboe ES. Qualitative analysis of mothers' decision-making about vaccines for infants: the importance of trust. Pediatrics 2006 May;117(5):1532-1541. [doi: 10.1542/peds.2005-1728] [Medline: 16651306]
24. Schwandt TA, Lincoln YS, Guba EG. Judging interpretations: But is it rigorous? trustworthiness and authenticity in naturalistic evaluation. New Drctns Evaluation 2007 May 29;2007(114):11-25. [doi: 10.1002/ev.223]
25. Lee M, You M. Direct and indirect associations of media use with COVID-19 vaccine hesitancy in South Korea: cross-sectional web-based survey. J Med Internet Res 2022 Jan 06;24(1):e32329 [FREE Full text] [doi: 10.2196/32329] [Medline: 34870605]
26. Giansanti D. The role of the mHealth in the fight against the COVID-19: successes and failures. Healthcare 2021 Jan 08;9(1):58. [doi: 10.3390/healthcare9010058]
27. Bilal, Latif F, Bashir MF, Komal B, Tan D. Role of electronic media in mitigating the psychological impacts of novel coronavirus (COVID-19). Psychiatry Res 2020 Jul;289:113041 [FREE Full text] [doi: 10.1016/j.psychres.2020.113041] [Medline: 32388417]
28. Appiah B, Asamoah-Akuoko L, France C, Rene A, Amanquah N, Bates I. Pharmacists and COVID-19 vaccination - considering mobile phone caller tunes as a novel approach to promote vaccine uptake in low- and middle-income countries. Res Social Adm Pharm 2022 May;18(5):2898-2903 [FREE Full text] [doi: 10.1016/j.sapharm.2021.07.022] [Medline: 34362688]
29. Herrett E, Williamson E, van Staa T, Ranopa M, Free C, Chadborn T, et al. Text messaging reminders for influenza vaccine in primary care: a cluster randomised controlled trial (TXT4FLUJAB). BMJ Open 2016 Feb 19;6(2):e010069 [FREE Full text] [doi: 10.1136/bmjopen-2015-010069] [Medline: 26895984]
30. Abdel-Razig S, Anglade P, Ibrahim H. Impact of the COVID-19 pandemic on a physician group's WhatsApp chat: qualitative content analysis. JMIR Form Res 2021 Dec 07;5(12):e31791 [FREE Full text] [doi: 10.2196/31791] [Medline: 34784291]
31. Fazel S, Zhang L, Javid B, Brikell I, Chang Z. Harnessing Twitter data to survey public attention and attitudes towards COVID-19 vaccines in the UK. Sci Rep 2021 Dec 14;11(1):23402 [FREE Full text] [doi: 10.1038/s41598-021-02710-4] [Medline: 34907201]
32. Mannan KA, Farhana KM. Knowledge, attitude and acceptance of a COVID-19 vaccine: A global cross-sectional study. SSRN Journal 2021 Jan 14:1-23. [doi: 10.2139/ssrn.3763373]
33. Garg C. Is mobile phone use invading multiple boundaries? A study of rural illiterate women in India. Indian Journal of Gender Studies 2021 Jan 20;28(1):29-45. [doi: 10.1177/0971521520974845]
34. 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]
35. Wardle C, Derakhshan H. Information disorder: Toward an interdisciplinary framework for research and policymaking. Council of Europe. 2017 Sep 27. URL: http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf [accessed 2024-08-15]

Abbreviations

AKU: Aga Khan University
BBC: British Broadcasting Corporation
CHC: community health center
FGD: focus group discussion
HCP: health care provider
IDI: in-depth interview
LMIC: low- and middle-income country
mHealth: mobile health
SOP: standard operating procedure

Edited by E Lee; submitted 28.05.23; peer-reviewed by H Ibrahim, S Mahdi; comments to author 17.07.23; revised version received 01.08.23; accepted 26.05.24; published 04.09.24.

Please cite as:
Kazi AM, Ahsan N, Jabeen R, Allana R, Jamal S, Mughal MAK, Hopkins KL, Malik FA
Effects of COVID-19 Illness and Vaccination Infodemic Through Mobile Health, Social Media, and Electronic Media on the Attitudes of Caregivers and Health Care Providers in Pakistan: Qualitative Exploratory Study
JMIR Infodemiology 2024;4:e49366
URL: https://infodemiology.jmir.org/2024/1/e49366
doi:10.2196/49366
PMID:

Original Paper

TikTok as a Source of Health Information and Misinformation for Young Women in the United States: Survey Study

Ciera E Kirkpatrick[1], PhD; LaRissa L Lawrie[2], MA

[1]College of Journalism & Mass Communications, University of Nebraska-Lincoln, Lincoln, NE, United States

[2]School of Journalism, University of Missouri, Columbia, MO, United States

Corresponding Author:
Ciera E Kirkpatrick, PhD
College of Journalism & Mass Communications
University of Nebraska-Lincoln
331 Andersen Hall
Lincoln, NE, 68588
United States
Phone: 1 3165594397
Email: ciera.kirkpatrick@unl.edu

Abstract

Background: TikTok is one of the most-used and fastest-growing social media platforms in the world, and recent reports indicate that it has become an increasingly popular source of news and information in the United States. These trends have important implications for public health because an abundance of health information exists on the platform. Women are among the largest group of TikTok users in the United States and may be especially affected by the dissemination of health information on TikTok. Prior research has shown that women are not only more likely to look for information on the internet but are also more likely to have their health-related behaviors and perceptions affected by their involvement with social media.

Objective: We conducted a survey of young women in the United States to better understand their use of TikTok for health information as well as their perceptions of TikTok's health information and health communication sources.

Methods: A web-based survey of US women aged 18 to 29 years (N=1172) was conducted in April-May 2023. The sample was recruited from a Qualtrics research panel and 2 public universities in the United States.

Results: The results indicate that the majority of young women in the United States who have used TikTok have obtained health information from the platform either intentionally (672/1026, 65.5%) or unintentionally (948/1026, 92.4%). Age (959/1026, 93.47%; r=0.30; P<.001), education (959/1026, 93.47%; ρ=0.10; P=.001), and TikTok intensity (ie, participants' emotional connectedness to TikTok and TikTok's integration into their daily lives; 959/1026, 93.47%; r=0.32; P<.001) were positively correlated with overall credibility perceptions of the health information. Nearly the entire sample reported that they think that misinformation is prevalent on TikTok to at least some extent (1007/1026, 98.15%), but a third-person effect was found because the young women reported that they believe that other people are more susceptible to health misinformation on TikTok than they personally are (t_{1025}=21.16; P<.001). Both health professionals and general users were common sources of health information on TikTok: 93.08% (955/1026) of the participants indicated that they had obtained health information from a health professional, and 93.86% (963/1026) indicated that they had obtained health information from a general user. The respondents showed greater preference for health information from health professionals (vs general users; t_{1025}=23.75; P<.001); the respondents also reported obtaining health information from health professionals more often than from general users (t_{1025}=8.13; P<.001), and they were more likely to act on health information from health professionals (vs general users; t_{1025}=12.74; P<.001).

Conclusions: The findings suggest that health professionals and health communication scholars need to proactively consider using TikTok as a platform for disseminating health information to young women because young women are obtaining health information from TikTok and prefer information from health professionals.

(JMIR Infodemiol 2024;4:e54663) doi:10.2196/54663

KEYWORDS

credibility perceptions; health information; health misinformation; information seeking; misinformation perceptions; public health; social media; strategic communication; third-person effect; TikTok

Introduction

Background

As one of the most-used and fastest-growing social media platforms in the world, TikTok has drastically changed American culture [1,2]. The social media platform, which allows users to create and watch short-form videos ranging in length from 15 seconds to 10 minutes, has >150 million active users in the United States and is expected to reach 955 million users worldwide by 2025 [3,4]. This rapid popularity has caught the attention of health communication scholars and practitioners because the platform is a vehicle for finding and disseminating information, including health-related content [5,6]. According to the Pew Research Center, the number of adults in the United States who regularly get news from TikTok has more than tripled (from 3% in 2020 to 10% in 2022). Adults aged <30 years are the most likely group, with a third (32%) of adults aged 18 to 29 years saying that they regularly get their news from TikTok [7].

Americans turned to the internet to find health information during the COVID-19 pandemic, and medical professionals and health institutions met them on TikTok, delivering pandemic-related information and recommendations [5]. The presence of health-related content has extended beyond COVID-19–related information and includes a wide range of health topics such as cervical cancer screening, chronic pulmonary obstructive disease, diabetes, mental health, and more [8-11]. Studies examining the engagement with this content [10,11] suggest that users like finding health information on the platform. A recent survey of 2000 Americans conducted by the prescription savings company CharityRx found that 1 in 5 Americans turns to TikTok for advice before going to their physician. Of the participants belonging to Generation Z, specifically, 1 in 3 reported obtaining health information from TikTok [12].

TikTok's concise video format encourages users to convey their message in a brief yet engaging way, while the relaxed atmosphere and the capacity to engage with viewers facilitate a more intimate and authentic form of communication [13]. These videos may be especially memorable (and thus influential) because viewers can retrieve both visual and verbal information that they have processed and stored while watching the videos [14]. Social media platforms have the power to spread credible, useful health information [8]. However, recent research has indicated that 1 in 5 TikTok videos likely contains misinformation [15], and fact-checking has been found to be uncommon on social media [16]. Consequently, as young women turn to TikTok for health information, they may encounter both the beneficial aspects and drawbacks of this accessible platform. For these reasons, we suggest the need for a better empirical understanding of the extent to which young women are obtaining health information on TikTok and their associated perceptions and behaviors related to the information they encounter.

Objectives

In light of the popularity of TikTok, we conducted a survey of young women (assigned female sex at birth) in the United States to better understand their use of the social media platform for health information as well as their perceptions of the platform's health information and related communication sources. We focus on women aged 18 to 29 years for this study because women make up the majority of TikTok users in the United States and because survey data have revealed that users aged 18 to 19 years and 20 to 29 years were the 2 largest age groups using TikTok during the time period that this study was conducted [17,18]. Furthermore, it has been suggested that women may be especially affected by health information on TikTok. Mainstream news programs (eg, Good Morning America on the ABC network) have anecdotally reported that many young women turn to TikTok for health information and that obstetrician-gynecologists and other physicians have developed TikTok brands specifically to reach these users [6]. According to prior research, young women's involvement with social media significantly influences their perceptions and behaviors concerning their health [19]; women are more likely to look for health information, including via the internet [20,21]; and women tend to have a leading role in the majority of decisions for their families' health [22].

Specifically, we first sought to explore how often young women are intentionally and unintentionally obtaining health information from TikTok and their top reasons for obtaining health information from the platform. Second, we explored their perceptions of credibility (ie, the perceived credibility of TikTok health information overall) and misinformation in relation to the health information they see on TikTok, as well as their frequency of verifying the health information they see. Within this, we asked questions about perceived susceptibility to health misinformation on TikTok to see if a third-person effect might exist. The third-person effect is a communication theory that suggests that people tend to perceive that messages in the media have a greater effect on other people than on themselves [23], which, in the context of misinformation on TikTok, could cause young women to underestimate the potential impact of misinformation on their own health-related decisions and behaviors. Third, we explored perceptions and behaviors related to the top 2 types of sources that share health information on TikTok (*health professionals* and *general users*). In terms of perceptions, we examined how often young women obtain information from these sources, how much they prefer to obtain information from these sources, and how credible they perceive the information from these sources to be (ie, the perceived credibility of TikTok health information from health professionals and the perceived credibility of TikTok health information from general users). For behaviors related to these source types, we examined whether the young women have acted on health information they obtained from these sources, their likelihood of acting on health information from these sources in the future, and their likelihood of fact-checking information from these sources. In exploring each of these 3 areas, we also examined whether the women's age, highest level of education, and level of TikTok intensity (defined as their emotional connectedness to TikTok and TikTok's integration into their daily lives) had a relationship with their use of TikTok as a source of health information.

Altogether, this study seeks to help both health communication researchers and practitioners by illuminating the role that TikTok plays in young women's acquisition of health information in the United States.

Methods

Design and Sample

A web-based survey of US women (assigned female sex at birth) aged 18 to 29 years (N=1172) was conducted between April and May 2023. While the sample was focused on individuals who were assigned female sex at birth, we also asked about their gender identity (refer to the Results section). The sample was recruited using a Qualtrics research panel as well as convenience sampling at 2 public universities in the United States.

Ethical Considerations

The institutional review boards at the University of Nebraska-Lincoln (IRB 20230122526EX) and the University of Missouri (IRB 2095651) approved the study. Respondents recruited via the Qualtrics panel were compensated in agreement with their Qualtrics contract, and the respondents recruited at the universities were compensated with course credit.

Procedure

The study began with a web-based informed consent form that briefly explained the purpose of the study and gave the survey respondents information about the study's investigator, the expected length of the survey, and how their data would be used (for the reporting of aggregate data) and stored (in a password-protected electronic format). Respondents then answered screening questions, and respondents who were not assigned female sex at birth and not aged 18 to 29 years were excluded from the survey. After passing the eligibility criteria, respondents were asked whether they had ever used TikTok to either watch or post videos. The respondents who had used TikTok were then asked about their average amount of use, whether they had ever intentionally used TikTok to look for advice or information about their health or health care (and whether they had done so in the past 3 months), and whether they had ever unintentionally been exposed to health information on TikTok. From here, the questions they saw depended upon whether they had ever seen health information (intentionally or unintentionally) on TikTok.

All respondents who had ever used TikTok (n=1026) responded to items measuring reasons for health-related TikTok use; the perceptions of health misinformation on TikTok; the use of, and preference for, particular sources (health professionals and general users) of health information on TikTok; the perceived credibility of health information from health professionals and general users on TikTok; the likelihood of acting on health information obtained from health professionals and general users on TikTok; and the likelihood of fact-checking health information from health professionals and general users on TikTok.

Respondents who had seen health information on TikTok (n=959) additionally responded to items measuring their perceived credibility of the health information they have seen overall on TikTok and their verification of the health information they have seen on TikTok. Respondents who had ever received health information on TikTok from either source of interest (health professionals or general users) were also asked about whether they had acted on health information from these sources.

Finally, all participants (N=1172) responded to items measuring TikTok intensity (ie, their emotional connectedness to TikTok and TikTok's integration into their daily lives) and answered demographic questions, including their highest level of education, race, and ethnicity.

Measures

Using TikTok as a Health Information Source

Frequency of Use

Respondents were asked whether they had ever used TikTok (either to watch or to post videos). The respondents who had used TikTok were asked to indicate their average amount of use, using the following options: less than once a month, once a month, once a week, a few times a week, once a day, more often than once a day. The respondents who had ever used TikTok were also asked whether they had ever used TikTok to look for advice or information about their health or health care (selecting *yes* or *no*). Those who selected *yes* were asked how often they *intentionally* use TikTok to obtain health information, and all respondents who had ever used TikTok were asked how often they unintentionally obtain health information on TikTok (hourly, daily, weekly, monthly, less often, or not at all).

Reasons for Health-Related TikTok Use

Respondents indicated their reasons for health-related TikTok use by indicating their level of agreement (ranging from 1=*strongly disagree* to 7=*strongly agree*) with 10 different statements (eg, "I like to get health information from TikTok because it can help me to maintain a healthy lifestyle") adapted from prior research [24-26] (refer to the Results section).

Credibility, Misinformation, and Verification of Health Information on TikTok

Perceived Credibility of TikTok Health Information Overall

Using the 7-point scale (ranging from 1=*not at all* to 7=*extremely*) for media credibility developed by Flanagin and Metzger [27], respondents rated how believable, accurate, trustworthy, biased (reverse coded), and complete they perceive health information on TikTok, overall, to be. Specifically, the respondents were asked, "To what degree do you rate the health information provided on TikTok?" The 5 items were averaged to create a perceived credibility score for each respondent (mean 4.48, SD 1.28; Cronbach α=0.90).

Misinformation Perceptions

For the perceptions of misinformation, respondents were asked to indicate on a 7-point scale how prevalent they think health misinformation is on TikTok (ranging from 1=*not prevalent at all* to 7=*very prevalent*; mean 5.14, SD 1.42), how serious they think the impact of health misinformation on TikTok is (ranging from 1=*not serious at all* to 7=*very serious*; mean 5.57, SD

1.46), and how susceptible they think they are to the influence of health misinformation on TikTok (ranging from 1=*not susceptible at all* to 7=*very susceptible*; mean 4.07, SD 1.75; adapted from the study by Chang [28]). In addition, to explore the possibility of a third-person effect, the respondents were asked to indicate how susceptible (ranging from 1=*not susceptible at all* to 7=*very susceptible*) they think others (eg, the public; mean 5.26, SD 1.46) are to the influence of health misinformation on TikTok (adapted from the study by van der Meer et al [23]). Participants were also asked about their self-perceived direct experience with health information on TikTok with the following item (adapted from the study by Chang [28]): "Have you encountered health misinformation on TikTok in the past? (*yes, no, or unsure*)."

Verification of Health Information on TikTok

Using a scale adapted from the study by Flanagin and Metzger [27], respondents indicated on a 7-point scale (ranging from 1=*never* to 7=*always*) how often they performed 6 different verification behaviors (eg, "Check to see if the information is current") when seeing health information on TikTok. The scores for the 6 items were averaged to create a *verification of TikTok health information* score for each respondent (mean 4.83, SD 1.53; Cronbach α=0.92).

Health Professionals and General Users as Sources of Health Information on TikTok

Source Preferences

Using a 7-point scale (ranging from 1=*not at all* to 7=*very often*), respondents indicated how often they obtain health information from health professionals on TikTok (mean 5.04, SD 1.83) and how often they obtain health information from general users on TikTok (mean 4.55, SD 1.89). In addition, using a 7-point scale (ranging from 1=*don't prefer them at all* to 7=*prefer them a lot*), respondents were asked to indicate how much they prefer to obtain health information from health professionals on TikTok (mean 5.65, SD 1.75) and how much they prefer to obtain health information from general users on TikTok (mean 4.08, SD 1.96).

Perceived Credibility of TikTok Health Information From Health Professionals and General Users

In addition to measuring the respondents' perceived credibility of TikTok health information *overall*, we also measured the respondents' perceived credibility of the 2 sources of interest (health professionals and general users) using the 7-point scale for media credibility developed by Flanagin and Metzger [27]. Specifically, we asked, "To what degree do you rate the health information provided by *health professionals* (eg, a doctor or nurse) on TikTok?" and "To what degree do you rate the health information provided by *general users* (someone like you) on TikTok?" Respondents rated how believable, accurate, trustworthy, biased (reverse coded), and complete they believe the health information on TikTok to be from each of these sources. The scores for each of the 5 items were averaged for each of the source types such that each respondent had a score for the perceived credibility of TikTok health information from health professionals (mean 5.16, SD 1.18; Cronbach α=0.90) and the perceived credibility of TikTok health information from general users (mean 3.95, SD 1.54; Cronbach α=0.94).

Acting on Health Information

Respondents' likelihood of acting on health information was measured with items adapted from the study by Hu and Shyam Sundar [29]. Using a 7-point scale (ranging from 1=*not at all likely* to 7=*extremely likely*), respondents indicated how likely they are to act on health information from a health professional on TikTok (mean 4.50, SD 1.79) and from a general user on TikTok (mean 3.96, SD 1.89). Respondents were also asked whether they ever have acted on health information provided on TikTok by a health professional or general user.

Fact-Checking Information

Respondents were asked, on a scale ranging from 1=*not at all likely* to 7=*very likely*, to rate how likely they are to fact-check health information on TikTok from a health professional (mean 4.88, SD 1.80) and a general user (mean 5.37, SD 1.83).

Audience Characteristics

TikTok Intensity

Scores for TikTok intensity were created using an adapted form of the scale for Facebook intensity developed by Ellison et al [30]. The scale was created to measure how emotionally connected participants are to the social media platform as well as the extent to which the platform is part of their everyday lives. This TikTok-modified version of the scale (ranging from 1=*strongly disagree* to 7=*strongly agree*) asked respondents to rate their agreement with the following six items: (1) "TikTok is part of my everyday life," (2) "I am proud to tell people I'm on TikTok," (3) "TikTok has become part of my daily routine," (4) "I feel out of touch when I haven't logged into TikTok for a while," (5) "I feel I am part of the TikTok community," and (6) "I would be sorry if TikTok shut down." The scores of the 6 items were averaged to create TikTok intensity scores for each respondent who had reported ever having used TikTok (mean 4.91, SD 1.49; Cronbach α=0.90).

Survey Questionnaire and Descriptive Statistics

The full survey questionnaire and descriptive statistics for each variable across the student and Qualtrics samples can be found in Multimedia Appendix 1 and Multimedia Appendix 2, respectively.

Data Analysis

Statistical analyses were performed using SPSS software (version 29.0; IBM Corp). Descriptive analyses were conducted to describe the respondents' frequency of TikTok use, frequency of intentional and unintentional exposure to health information on TikTok, reasons for health-related TikTok use, beliefs about encountering health misinformation, perceptions of misinformation on TikTok, frequency of performing verification behaviors on TikTok, frequency and preferences related to obtaining health information from health professionals and general users on TikTok, and frequency of acting on TikTok health information.

Bivariate correlational analyses were conducted to examine the relationships that credibility perceptions (the perceived credibility of TikTok health information overall, the perceived credibility of TikTok health information from health professionals, and the perceived credibility of TikTok health information from general users), misinformation perceptions, verification behaviors, the likelihood of acting on TikTok health information, and the likelihood of fact-checking TikTok health information had with the respondents' age, highest level of education, and level of TikTok intensity. In addition, bivariate correlational analyses were conducted to examine the relationship between respondents' perceived credibility of TikTok health information overall and their likelihood of acting on the health information. For these correlational analyses, a correlation was considered weak if the correlation coefficient was between −0.4 and 0.4. A correlation was considered moderate if the correlation coefficient was between −0.8 and −0.4 or between 0.4 and 0.8. A correlation was considered strong if the correlation coefficient was between −1 and −0.8 or between 0.8 and 1.

Finally, paired samples *t* tests (2-tailed) were conducted to observe the statistical differences between respondents' perceived susceptibility of health misinformation on TikTok for themselves versus for others, frequency of obtaining health information on TikTok from health professionals versus general users, preference for obtaining health information on TikTok from health professionals versus general users, perceived credibility of TikTok health information from health professionals versus general users, likelihood of acting on TikTok health information from health professionals versus general users, and likelihood of fact-checking TikTok health information from a health professional versus a general user.

Results

Overview

A total of 1172 qualified responses were collected, with the average age of the sample being 22.82 (SD 3.15) years. A little more than half of the participants came from the Qualtrics panel (636/1172, 54.27%), and the rest were recruited through the universities (536/1172, 45.73%). The majority of the sample identified as White (910/1172, 77.65%), and most of the sample reported having used TikTok (1026/1172, 87.54%). Approximately half of the participants reported using TikTok more often than once a day (615/1172, 52.47%). Further demographic information is included in Table 1, and the full list of demographic questions can be found in Multimedia Appendix 1.

Table 1. Respondent demographics (N=1172).

Characteristics	Participants, n (%)
Race	
American Indian or Alaska Native	3 (0.26)
Asian	16 (1.37)
Black or African American	117 (9.98)
Native Hawaiian or Pacific Islander	57 (4.86)
White	910 (77.65)
Other	53 (4.52)
Prefer not to answer	16 (1.37)
Ethnicity	
Cuban	13 (1.11)
Mexican, Mexican American, or Chicana	109 (9.3)
Puerto Rican	11 (9.39)
Other Spanish, Hispanic, or Latina	98 (8.36)
Not Spanish, Hispanic, or Latina	919 (78.41)
Prefer not to answer	22 (1.88)
Education	
Less than high school	17 (1.45)
High school graduate or equivalent (eg, GED[a])	207 (17.66)
Some college	509 (43.43)
2-year degree	90 (7.68)
4-year degree	224 (19.11)
Professional or master’s degree	68 (5.8)
Doctorate	57 (4.86)
Self-identified gender	
Woman	1154 (98.46)
Transgender	4 (0.34)
Nonbinary	11 (0.94)
Gender fluid	2 (0.17)
Other	1 (0.09)
Frequency of TikTok use	
Never	146 (12.46)
Less than once a month	58 (4.95)
Once a month	50 (4.27)
Once a week	60 (5.12)
A few times a week	108 (9.22)
Once a day	135 (11.52)
More often than once a day	615 (52.47)

[a]GED: General Educational Development test.

Using TikTok as a Health Information Source

Frequency of Use

Of the 1026 respondents who had used TikTok before, 672 (65.5%) reported that they had *intentionally* used TikTok to look for advice or information about their health or health care, while 948 (92.4%) reported that they had *unintentionally* received health information or advice on TikTok. Of the 1026 respondents who had ever used TikTok, 582 (56.73%) reported having intentionally used TikTok to look for advice or information about their health or health care in the last 3 months. A breakdown of the frequency of intentional and unintentional exposure to health information on TikTok is provided in Table 2.

Table 2. Frequency of using TikTok intentionally and unintentionally as a source of health information among respondents who had ever used TikTok (n=1026).

Questions and responses	Participants, n (%)
Frequency of intentional use of TikTok to obtain health information	
Hourly	29 (2.83)
Daily	108 (10.53)
Weekly	143 (13.94)
Monthly	168 (16.37)
Less often	190 (18.52)
Not at all	388 (37.82)
Frequency of unintentional exposure to health information on TikTok	
Hourly	38 (3.7)
Daily	232 (22.61)
Weekly	363 (35.38)
Monthly	172 (16.76)
Less often	143 (13.94)
Not at all	78 (7.6)

Reasons for Health-Related TikTok Use

Of the 10 reasons presented for health-related TikTok use, obtaining advice from others with the same disease or health condition (mean 5.29, SD 1.54), receiving social support from others (mean 5.29, SD 1.57), and gaining knowledge about a disease they had been diagnosed with (mean 5.01, SD 1.59) were the most agreed upon reasons. The least agreed upon reason for health-related TikTok use was communicating with physicians (mean 4.06, SD 1.86). Table 3 shows the mean and SD of the level of agreement for each of the 10 reasons.

Table 3. Reasons for health-related TikTok use among respondents who had ever used TikTok (n=1026).

I like to get health information from TikTok because...	Level of agreement[a], mean (SD)
It can help me to maintain a healthy lifestyle.	4.92 (1.60)
It can help me determine whether I need to see a doctor.	4.81 (1.67)
It can provide me with more information after I've seen my doctor.	4.65 (1.72)
It can help me find different options for treatment or maintenance of my health condition(s).	4.79 (1.65)
I can gain knowledge about a disease I've been diagnosed with.	5.01 (1.59)
I can obtain advice from other patients with the same disease or health condition as me.	5.29 (1.54)
I can receive social support from others.	5.29 (1.57)
I can communicate with physicians.	4.06 (1.86)
I can interact in real time with TikTok users.	4.59 (1.81)
I can obtain immediate health information and make use of it.	4.53 (1.71)

[a]Respondents indicated their level of agreement using a 7-point scale (ranging from 1=*strongly disagree* to 7=*strongly agree*).

Credibility, Misinformation, and Verification of Health Information on TikTok

Perceived Credibility of TikTok Health Information Overall

With a mean of 4.48 (SD 1.28) on a 7-point scale, the credibility perceptions were moderate. A positive correlation was found (959/1026, 93.47%; *r*=0.30; *P*<.001) between the perceived credibility of health information on TikTok overall and the respondents' age. Older participants tended to perceive the content as more credible. A positive correlation was also found (959/1026, 93.47%; ρ=0.10; *P*=.001) between the perceived credibility of health information on TikTok overall and the respondents' highest level of education, with respondents with greater education tending to rate the health information on TikTok as more credible. There was a positive correlation (959/1026, 93.47%; r=0.32; *P*<.001) between TikTok intensity and the perceived credibility of TikTok health information overall. Respondents with higher TikTok intensity scores tended to have greater perceived credibility of health information on TikTok overall.

Misinformation Perceptions

Approximately half (563/1026, 54.87%) of the respondents who had used TikTok indicated that they believe that they have personally encountered health misinformation on the platform at some point, and only 1.85% (19/1026) of the sample stated that they think that health misinformation is not prevalent at all on TikTok. Table 4 shows the levels of perceived prevalence, severity, and susceptibility of health misinformation on TikTok reported by the respondents who had used TikTok.

There was a weak but significant negative correlation between age and the perceived seriousness of health misinformation on social media (1026/1172, 87.54%; *r*=−0.07; *P*=.04). There was not a significant relationship between age and perceived prevalence (1026/1172, 87.54%; *r*=−0.06; *P*=.05) or perceived susceptibility (1026/1172, 87.54%; *r*=−0.02; *P*=.50). There was not a significant relationship between the respondents' highest level of education and the perceived prevalence (1026/1172, 87.54%; ρ=−0.008; *P*=.81), the perceived seriousness (1026/1172, 87.54%; ρ=0.01; *P*=.66), or the perceived susceptibility of health misinformation on TikTok (1026/1172, 87.54%; ρ=0.04; *P*=.17).

There was a weak but significant positive relationship between respondents' TikTok intensity scores and perceived seriousness (1026/1172, 87.54%; ρ=0.07; *P*=.02) and perceived susceptibility (1026/1172, 87.54%; ρ=0.07, *P*=.02) such that respondents with high TikTok intensity scores tended to perceive greater seriousness and susceptibility of health misinformation on TikTok. There was not a significant relationship between TikTok intensity scores and the perceived prevalence of health misinformation on TikTok (1026/1172, 87.54%; ρ=0.03; *P*=.29).

The results also showed that respondents perceive other people (mean 5.26, SD 1.46) as more susceptible to health misinformation on TikTok than they personally are (mean 4.07, SD 1.75; t_{1025}=21.16; *P*<.001).

Table 4. Perceived prevalence, seriousness, and susceptibility of health misinformation on TikTok among respondents who had ever used TikTok (n=1026).

Questions	Level of agreement[a], mean (SD)
How prevalent is health misinformation on TikTok?	5.14 (1.42)
How serious do you think the impact of health misinformation on TikTok is?	5.57 (1.46)
How susceptible are you to the influence of health misinformation on TikTok?	4.07 (1.75)
How susceptible are other people to the influence of health misinformation on TikTok?	5.26 (1.46)

[a]Respondents indicated their level of agreement using a 7-point scale (ranging from 1=*not at all prevalent, serious, or susceptible to 7=very prevalent, serious, or susceptible*).

Verification of Health Information on TikTok

The most frequently used form of verification was considering whether the information represented was opinion or fact and the least frequently used form of verification was verifying the TikTok users' qualifications or credentials. Table 5 shows the respondents' frequency of each verification behavior.

There was a weak but significant positive correlation found between the respondents' age and their likelihood of verifying TikTok health information (959/1026, 93.47%; *r*=0.20; *P*<.001) as well as between the respondents' highest level of education and their likelihood of verifying TikTok health information (959/1026, 93.47%; ρ=0.09; *P*<.001). A weak but significant positive relationship existed between respondents' TikTok intensity scores and their likelihood of verifying TikTok health information (959/1026, 93.47%; *r*=0.21; *P*<.001). Participants with high TikTok intensity scores were more likely to verify the health information.

Table 5. Verification of TikTok health information among participants who had seen health information on TikTok (n=959).

Verification behavior	Participant response[a]						
	Never, n (%)	Almost never, n (%)	Rarely, n (%)	About half of the time, n (%)	Most of the time, n (%)	Almost always, n (%)	Always, n (%)
Verify the TikTok users' qualifications or credentials	121 (12.6)	68 (7.1)	89 (9.3)	131 (13.7)	193 (20.1)	178 (18.6)	179 (18.7)
Consider the TikTok users' goals and objectives for posting information on the web	74 (7.7)	87 (9.1)	70 (7.3)	165 (17.2)	208 (21.7)	199 (20.8)	156 (16.3)
Check to see whether the information is current	72 (7.5)	61 (6.4)	66 (6.9)	143 (14.9)	223 (23.3)	198 (20.7)	196 (20.4)
Seek out other sources to validate the information	63 (6.6)	52 (5.4)	76 (7.9)	121 (12.6)	200 (20.9)	221 (23)	226 (23.6)
Consider whether the information represented is opinion or fact	53 (5.5)	50 (5.2)	49 (5.1)	125 (13)	199 (20.8)	251 (26.2)	232 (24.2)
Check to see that the information is complete and comprehensive	73 (7.6)	55 (5.7)	65 (6.8)	148 (15.4)	193 (20.1)	225 (23.5)	200 (20.9)

[a]Participants were asked to indicate how often they perform each of the 6 different behaviors when seeing health information on TikTok.

Health Professionals and General Users as Sources of Health Information on TikTok

Source Preferences

Of the respondents who had ever used TikTok, 93.08% (955/1026) indicated that they had obtained health information from a health professional on the platform, while 93.86% (963/1026) indicated that they had obtained health information from a general user on the platform. That said, respondents reported obtaining health information from health professionals on TikTok (mean 5.04, SD 1.83) significantly more often than they obtain health information from general users on TikTok (mean 4.55, SD 1.89; t_{1025}=8.13; *P*<.001). This was in line with their preferences for health information sources because the respondents' preference for obtaining health information from health professionals (mean 5.65, SD 1.75) was significantly greater than their preference for obtaining health information from general users (mean 4.08, SD 1.96; t_{1025}=23.75; *P*<.001).

Perceived Credibility of TikTok Health Information From Health Professionals and General Users

Respondents perceived health information from health professionals on TikTok (mean 5.16, SD 1.18) to be significantly more credible than health information provided by general users on TikTok (mean 3.95, SD 1.54; t_{958}=26.737; *P*<.001).

Acting on Health Information

Of the respondents who had received health information from a health professional on TikTok, 43.35% (414/955) reported that they had acted on health information they obtained from a health professional on TikTok. In comparison, 37.8% (364/963) of the respondents who had received health information from a general user on TikTok reported that they had acted on health information they obtained on TikTok from a general user. When asked about their likelihood of acting on health information on TikTok in the future, the respondents' likelihood of acting on health information from a health professional on TikTok (mean 4.50, SD 1.79) was significantly higher than their likelihood of acting on health information from a general user on TikTok (mean 3.96, SD 1.89; t_{1025}=12.74; *P*<.001).

Likewise, the respondents' perceived credibility of TikTok health information overall was positively correlated with their likelihood of acting on the health information. Respondents who perceived health information on TikTok overall as credible were more likely to act on health information they obtained from a health professional on TikTok (959/1026, 93.47%; *r*=0.47; *P*<.001) than from a general user on TikTok (959/1026, 93.47%; *r*=0.53; *P*<.001).

Age was weakly positively correlated with the respondents' likelihood of acting on health advice found on TikTok both from a health professional (1026/1172, 87.54%; *r*=0.11; *P*<.001) and from a general user (1026/1172, 87.54%; *r*=0.23; *P*<.001). Likewise, education was weakly positively correlated with their likelihood of acting on health information found on TikTok both from a health professional (1026/1172, 87.54%; ρ=0.09; *P*=.003) and from a general user (1026/1172, 87.54%; ρ=0.08; *P*=.01).

There was a significant positive relationship between TikTok intensity and the respondents' likelihood of acting on health information from health professionals on TikTok (1026/1172, 87.54%; *r*=0.41; *P*<.001). Likewise, there was a significant positive relationship between TikTok intensity and their likelihood of acting on health information from general users on TikTok (1026/1172, 87.54%; *r*=0.39; *P*<.001). Respondents with higher TikTok intensity scores tended to be more likely to act on health information from both health professionals and general users on TikTok.

Fact-Checking Information

There was a statistically significant difference in the respondents' likelihood of fact-checking health information on TikTok from a health professional versus a general user. The likelihood of fact-checking health information from a general user (mean 5.37, SD 1.83) was higher than the likelihood of fact-checking health information from a health professional (mean 4.88, SD 1.80; t_{1025}=9.71; P<.001).

Age was weakly positively correlated with the respondents' likelihood of fact-checking TikTok health information both from a health professional (1026/1172, 87.54%; r=0.18; P<.001) and from a general user (1026/1172, 87.54%; r=0.07; P=.02). Education was also weakly positively correlated with their likelihood of fact-checking TikTok health information from a health professional (1026/1172, 87.54%; ρ=0.11; P<.001), but education was not correlated with their likelihood of fact-checking TikTok health information for a general user (1026/1172, 87.54%; ρ=0.04; P=.26).

There was a significant but weak positive relationship between TikTok intensity and the respondents' likelihood of fact-checking TikTok health information from both health professionals (1026/1172, 87.54%; r=0.07; P=.03) and general users on TikTok (1026/1172, 87.54%; r=0.10; P=.002). Respondents with higher TikTok intensity scores tended to be more likely to fact-check health information from both health professionals and general users on TikTok.

Discussion

Principal Findings

TikTok has generated substantial attention due to recent reports suggesting its emergence as a significant source of information for many Americans [7]. For some users, TikTok has replaced traditional news networks as well as widely used search engines such as Google [31]. Given this emergence of TikTok as an information source and the presence of health information available on the platform [5,6], we surveyed 1172 women aged 18 to 29 years to understand their use of TikTok as a source of health information. Of the 1172 respondents, 1026 (87.54%) had used TikTok in some capacity.

The findings provide evidence that TikTok has become a source of health information for young women in the United States. More than half of the respondents who had ever used TikTok (672/1026, 65.5%) reported that they had intentionally used TikTok to look for advice or information about their health or health care, and nearly the entire sample of TikTok users (948/1026, 92.4%) reported having unintentionally obtained health information on TikTok. The popularity of health-related hashtags on TikTok (eg, as of November 2023, #medicaltiktok and #healthtok had 7.6 billion and 2.4 billion views, respectively) has illuminated some of TikTok's popularity as a commonly searched platform for information related to health, but the findings of this study provide a greater empirical understanding of the extent to which young women actually obtain health information from the platform.

TikTok's popularity as a source of health information may, in part, be the result of how technology has influenced human beings' desire for immediate information. Rather than having to wait for a physician's appointment to ask about one's symptoms or health condition, one can take to the internet (eg, TikTok) and find related information in a matter of minutes [12]. This phenomenon of individuals seeking immediate information has important implications for health professionals. By knowing that individuals turn to platforms such as TikTok to find health information, health professionals can proactively create content so that credible health information is available when users go to find it. Social media platforms were heavily relied upon for health information during the COVID-19 pandemic [16], and since then, social media, and TikTok specifically, have been recommended as a tool for health promotion [5,32,33]. Given that TikTok is easily accessible and allows anyone to consume information without judgment, it may especially be helpful for populations with barriers to care and for communicating about taboo or stigmatized topics that users may be less comfortable asking about in a traditional setting [6,34].

To better understand why young women are using TikTok as a source of health information, we asked our respondents about their agreement with various reasons for health-related TikTok use. Our findings showed that the most agreed upon reasons were obtaining advice from others with the same disease or health condition, receiving social support from others, and gaining knowledge about a disease they had been diagnosed with. In an examination of how the current digital landscape has affected Americans' consumer behavior, CharityRx found "relatability to a shared personal experience" to be a top reason why people go to health influencers for information [12]. In addition to TikTok having the ability to provide immediate information related to users' health inquiries, it has the capacity to help users locate other people who are similar to them. This may be especially relevant for women. As women have experienced gender bias and poorer treatment in health care settings [35,36], it is possible that they may be especially motivated to seek social support and health information from others like them. Prior research has also indicated that social support is especially beneficial for women [37]. When individuals perceive similarity to a source of information, this can cause the message recipient to feel a stronger sense of connection with the message, which can have important implications in terms of the effects of a message [38]. Perceiving similarity to a source of information can also enhance a user's perceptions of the message and overall acceptance [39]. In terms of obtaining a further understanding of young women's motivations for using TikTok as a source of health information, future research could examine what topics (eg, health conditions) young women are most interested in and likely to search for.

Given the potential for misinformation to rapidly pervade the social media landscape, it has been recommended that experts in medical science, public health, and social sciences collaborate to better understand health misinformation on social media, including its reach and influence [40]. Our findings help show the degree to which young women on TikTok perceive an issue of misinformation and to what extent they try to verify or fact-check the information they consume. In first examining the perceived credibility of TikTok health information overall, we found that credibility perceptions were moderate among the respondents who had reported ever having obtained health-related advice or information (intentionally or unintentionally) from TikTok. Notably, health information on TikTok was perceived to be more credible by participants who were older, more educated, or had higher TikTok intensity scores (ie, were more emotionally connected to TikTok and had greater integration of TikTok into their daily lives). While our findings are able to show these positive correlations, we do not know whether the information they are seeing on TikTok (and thus reflecting on when indicating their credibility perceptions) truly is credible. It may be the case that older age, higher levels of education, and more experience with TikTok lead to following more credible users and being delivered more credible content via the TikTok algorithm. In this case, the content may truly be more credible for these users (leading to their greater perceptions of credibility). Future research could explore whether this is the case.

Nearly all participants (1007/1026, 98.15%) indicated that they believe that misinformation is prevalent on TikTok to at least some extent. This may be the result of mainstream news commonly communicating that misinformation is a problem on social media platforms, including TikTok [15]. A great deal of health misinformation reached social media users during the COVID-19 pandemic [8]. It is possible that respondents in this study were among these users or that they heard about this problematic phenomenon. However, despite nearly all participants (1007/1026, 98.15%) stating that they think that misinformation is prevalent on TikTok to at least some extent, only approximately half of the participants (563/1026, 54.87%) indicated that they believe that they have personally encountered health misinformation on TikTok at some point. This discrepancy could stem from a few factors. It could be that the users know that there is a misinformation epidemic but have not been exposed to misinformation because of their commitment to only following credible users (thus leading the TikTok algorithm to feed them more credible content). However, given the large amounts of misinformation that have been identified on the platform [15,41,42] and Americans' inability to identify most forms of misinformation [43-45], it is more likely that this discrepancy is the result of some respondents not having recognized that they have been exposed to health misinformation. This possibility is further supported by our results that showed that respondents perceive other people as more susceptible to health misinformation on TikTok than they personally are. This finding demonstrates what seems to be a third-person effect, in which the young women perceive that media messages have a greater influence on others than on themselves [23]. As evidence of a third-person effect was provided by the results (with respondents perceiving that other people are more susceptible to health misinformation on TikTok than they personally are), it might be that some of the respondents are naive about their susceptibility to misinformation.

The most common form of verification (to verify the accuracy of the health information found on TikTok) was considering whether the information presented was opinion or fact. The least frequently reported form of verification was verifying the TikTok users' qualifications or credentials. Age, education, and TikTok intensity were each found to have a weak positive correlation with the likelihood of verifying TikTok health information. Participants who were older, more educated, or had higher TikTok intensity scores were more likely to verify health information on TikTok. As discussed in the Results section, these demographics (age, education, and TikTok intensity) were also found to each be positively correlated with the perceived credibility of TikTok health information overall. Thus, as both credibility perceptions and the verification of the information increase with age, education, and TikTok intensity, perhaps these users truly are seeing more credible content and appropriately perceiving it to be credible. Future research should further explore these relationships because this is beyond the scope of this study. In prior research, fact-checking has typically been found to not be very common on social media. In a survey by Neely et al [16] of Americans' reliance on social media during the COVID-19 pandemic, it was found that three-fourths of those surveyed had relied on social media to some extent to stay informed about the pandemic, but the majority of them were unlikely to fact-check the information they found.

Finally, this study investigated young women's perceptions and behaviors related to the top 2 types of sources that share health information on TikTok: health professionals and general users. Prior research has identified that health professionals and general users are 2 of the most prevalent sources on TikTok communicating health information [10,11], and our findings indicate that young women are obtaining health information on TikTok from each of these 2 source types. Of the respondents who had used TikTok, a majority indicated that they had obtained health information both from a health professional on the platform (955/1026, 93.08%) and from a general user on the platform (963/1026, 93.86%). Given that the respondents' top reasons for health-related TikTok use were obtaining advice from others with the same disease or health condition, receiving social support from others, and gaining knowledge about a disease they had been diagnosed with, it is reasonable that the respondents would perceive both health professionals and general users as valuable sources of information. Prior research exploring the effects of communication sources on social media [9] has explained that both expert-type and peer-type sources provide value. While health professionals have formal training and credentialed experience, general users (eg, peers) can have a form of "experiential credibility" from their own personal experiences (such as that of living with a particular health condition) [9,27]. While most of the respondents had obtained health information from each of the 2 source types, their preference for obtaining health information from a health professional was significantly greater than their preference for obtaining health information from general users (t_{1025}=23.75; P<.001), and, in line with their preferences, the young women reported obtaining health information from health professionals on TikTok significantly more often than they reported obtaining health information from general users on TikTok (t_{1025}=8.13; P<.001). As medical professionals and health institutions delivered COVID-19–related information during the pandemic [5], it was found that Americans who used social media as a source of COVID-19–related information expanded their social media networks to include credible sources (eg, medical institutions and scientific sources) [16]. In addition, CharityRx's survey found "medical accreditation and certification" to be the top reported reason why people go to influencers for health information [12]. Together, the prior and current findings seem to indicate a preference for obtaining health information from health professionals on platforms such as TikTok. However, it is important to recognize that both of the common source types—health professionals and general users—are providing health information to these users.

Our findings showed that the perceived credibility of TikTok health information from health professionals was significantly higher than the perceived credibility of TikTok health information from general users (t_{958}=26.737; *P*<.001). This is promising for health professionals who choose to invest in creating a TikTok presence because studies have shown that web-based information is more likely to be attended to when it is perceived as credible [46]. There was also a significant difference in the young women's likelihood of acting on health information from a health professional versus a general user on TikTok, with their likelihood of acting on the information being greater when the information was from a health professional (t_{1025}=12.74; *P*<.001). As the respondents had greater perceived credibility of TikTok health information from health professionals (vs general users), it is logical that they would also be more likely to act on health information from health professionals, given that credibility perceptions enhance the likelihood of persuasion [47,48], including in the context of social media messaging [49]. We found that, of the respondents who had received health information from a health professional on TikTok, 43.35% (414/955) reported that they had acted on health information they obtained from a health professional on TikTok. In comparison, 37.8% (364/963) of the respondents who had received health information from a general user on TikTok reported that they had acted on health information they obtained on TikTok from a general user. The respondents' perceived credibility of health information on TikTok overall was found to be positively correlated with their likelihood of acting on health information from both of the source types, which is in alignment with the relationship between credibility perceptions and persuasive effects [47,48]. On the one hand, these findings are promising in the sense that young women perceive health professionals on TikTok to be more credible and are more influenced by them, further suggesting that it is worthwhile for health professionals to use TikTok as a strategic communication tool. On the other hand, this could mean that young women are more susceptible to being influenced by individuals who give the impression of being qualified health professionals. Medical professionals sometimes provide information that is outside of their scope of expertise [50], and uncredentialed users are often confused for credentialed health professionals [51]. Furthermore, it is important to note that medical credentials and titles vary. A search of the hashtag #womenshealth on TikTok results in videos from a number of different types of health professionals, including nurses, nurse practitioners, obstetrician-gynecologists, medical doctors (with MD or DO credentials), and midwives, and because prior research has shown that many individuals do not understand medical roles and titles or how to differentiate between them [52], this could have profound implications. It is also important to note that anyone on TikTok can present themselves as though they have the necessary credentials for the information they are sharing (eg, adding credentials to their username, presenting themselves with a formal title, wearing a laboratory coat or surgical scrubs, and communicating information in a persuasive manner). An authoritative title, on its own, can be enough to capture an individual's attention and generate respect [53]. Therefore, with anyone being able to add credentials to their TikTok username, this could be problematic, especially given that credibility is hard to distinguish on social media. Users are more likely to rely on heuristic cues (such as the titles included in a username) to determine a user's credibility [9,54]. The aforementioned findings also showed that the least frequently reported form of information verification was verifying the TikTok users' qualifications or credentials, further illustrating that this could be vastly problematic.

The study's findings showed that the young women in our study are more likely to fact-check information from a general user than fact-check information from a health professional. Again, this is promising in terms of the fact that content from general users may be more likely to include misinformation, but this could be problematic if the credibility cues of a health professional lead users to automatically assume that they can believe and trust any of the information. It would be worthwhile for future research to investigate this further, uncovering whether users trust misinformation from health professionals on TikTok simply because of the creators' credentials.

We also found that both age and education are positively correlated with young women's likelihood of acting on health information found on TikTok—both from a health professional and from a general user. However, these correlations were quite weak. A stronger positive correlation was found between the users' emotional connection to TikTok (ie, TikTok intensity) and their likelihood of acting on health information from both health professionals and general users. It may be logical to assume that users who heavily engage more with TikTok will have a greater propensity to act on the information they receive. Social influence theory suggests that individuals are influenced by those around them [55]. This may extend to the web-based environment, such that as TikTok becomes more integrated into one's life, it is more likely to affect one's behaviors.

Limitations

The findings of this survey research should be interpreted in light of some limitations. First, the sample of survey respondents was recruited through convenience sampling methods. While the sample is only a segment of the total population, we tried to ensure that we had a large sample that was representative of the population of interest (women aged 18-29 years throughout the United States) by using both Qualtrics and 2 large public universities to recruit individuals who fit the sample parameters. This study does not discuss differences between individuals recruited from the universities and those recruited from the Qualtrics sample, but we have provided descriptive statistics for each variable across each sample as a means of allowing for some comparison between the 2 samples (Multimedia Appendix 2). In addition, we made sure to include the respondents' highest level of education as a variable in this study to see how education is associated with the variables of interest in this study, which can help to provide some understanding of how one's educational experience might be related to one's use and perceptions of TikTok health information. Second, because this study relies on self-reporting from the survey respondents, there is a chance that the results do not truly capture the real behaviors of the respondents. As we asked questions about fact-checking web-based information (behaviors that individuals likely know they should engage in), the respondents may have answered some questions in a more socially desirable or acceptable manner (ie, what they expect would be an "appropriate" response) rather than being truthful in their responses. Hopefully, though, because participants knew that their responses would be anonymous, this helped to lessen social desirability bias. Finally, this study focused on young women (assigned female sex at birth) as the population of interest for this study. This narrow focus allows us to better understand the implications of TikTok use for this demographic, but it is also important to explore other populations' behaviors and perceptions.

Conclusions and Practical Implications

This study provides a greater understanding of the extent to which TikTok is serving as a source of health information for young women in the United States. With nearly all young women in this study (who had used TikTok) having been exposed to health information on TikTok (948/1026, 92.4%), and more than half of them (582/1026, 56.73%) having actively sought health-related information on the platform in the last 3 months, it is imperative for health professionals and health communication scholars to prioritize the consideration of TikTok as a platform that is influencing health information acquisition and dissemination in the United States. While the popularity and accessibility of TikTok may change, short-form video social media sites are likely to remain a common form of communication [56].

The findings of this study illustrate the potential value that TikTok can have for disseminating health information to an audience of young women. As the respondents of this survey reported a preference for information from health professionals and were more likely to act on the information from these sources, it is worthwhile for health professionals to use TikTok to disseminate health information to this audience, especially given the large number of women on the platform and prior research illustrating that social media significantly influence women's health-related behaviors and perceptions [17,19]. In doing so, health professionals may want to consider how they can align their content with young women's most common motivations for using TikTok as a source of health information, which we found to be obtaining advice from others with the same disease or health condition, receiving social support from others, and gaining knowledge about a disease they had been diagnosed with. Given that young women want advice and support from others who are experiencing similar health conditions, it may be useful for health professionals to consider working with individuals who are willing to share their personal experience with a health condition. Incorporating the stories of patients and other experienced individuals who have similar characteristics to those searching for information on TikTok could be especially influential for increasing attention to, and engagement with, health information on TikTok.

Furthermore, given that our findings indicate that young women have a preference for obtaining health information on TikTok from health professionals and that they are less likely to fact-check information from these sources, it is imperative that future initiatives address the proliferation of individuals sharing information beyond their scope of expertise and the problem of social media users confusing uncredentialed users as credentialed health professionals [50,51]. Future researchers and practitioners should also work on media literacy and education initiatives, given the third-person effect found in this research. It seems that young women know that misinformation is an issue on TikTok, but it seems that they may not be recognizing that they have been exposed to misinformation and that they perceive themselves as less susceptible. It may be beneficial for future interventions to address this perception and help young women to have better recognition of when they are being exposed to health misinformation.

Acknowledgments

This study was supported by an Emerging Scholars grant from the Association for Education in Journalism and Mass Communication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full survey questionnaire.
[DOCX File , 31 KB - infodemiology_v4i1e54663_app1.docx]

Multimedia Appendix 2
Descriptive statistics for study measures across 2 samples (student sample and Qualtrics sample).
[DOCX File , 41 KB - infodemiology_v4i1e54663_app2.docx]

References

1. Dellatto M. TikTok hits 1 billion monthly active users. Forbes. URL: https://www.forbes.com/sites/marisadellatto/2021/09/27/tiktok-hits-1-billion-monthly-active-users/ [accessed 2023-11-02]

2. Willingham A, Asmelash L, Andrew S. The biggest ways TikTok has changed American culture. CNN. URL: https://www.cnn.com/2023/04/02/us/tiktok-american-culture-effects-cec/index.html [accessed 2023-11-02]
3. Shepardson D. TikTok hits 150 mln U.S. monthly users, up from 100 million in 2020. Reuters. URL: https://www.reuters.com/technology/tiktok-tell-congress-it-has-150-million-monthly-active-us-users-2023-03-20/ [accessed 2023-11-02]
4. Yuen M. TikTok users worldwide (2020-2025). Insider Intelligence. URL: https://www.emarketer.com/chart/c/T11778/tiktok-users-worldwide-2020-2025-millions-change-of-internet-users-1 [accessed 2023-05-10]
5. Comp G, Dyer S, Gottlieb M. Is TikTok the next social media frontier for medicine? AEM Educ Train 2021 Jul 21;5(3):1-4 [FREE Full text] [doi: 10.1002/aet2.10532] [Medline: 34095694]
6. Kindelan K. Women turn to TikTok for health information and OBGYNs are there to meet them. Good Morning America. URL: https://www.goodmorningamerica.com/wellness/story/women-turn-tiktok-health-information-obgyns-meet-68824442 [accessed 2023-11-02]
7. Matsa KE. More Americans are getting news on TikTok, bucking the trend on other social media sites. Pew Research Center. URL: https://tinyurl.com/34avwsnk [accessed 2023-10-30]
8. Basch CH, Hillyer GC, Jaime C. COVID-19 on TikTok: harnessing an emerging social media platform to convey important public health messages. Int J Adolesc Med Health 2022 Oct 01;34(5):367-369 [FREE Full text] [doi: 10.1515/ijamh-2020-0111] [Medline: 32776899]
9. Kirkpatrick CE, Lawrie LL. Can videos on TikTok improve pap smear attitudes and intentions? Effects of source and autonomy support in short-form health videos. Health Commun 2023 Sep 10:1-13. [doi: 10.1080/10410236.2023.2254962] [Medline: 37691164]
10. 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]
11. 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]
12. CharityRx Blog. URL: https://tinyurl.com/yck583pj [accessed 2023-11-02]
13. Montenegro L. The rise of short-form video: TikTok is changing the game. Forbes. URL: https://www.forbes.com/sites/forbesagencycouncil/2021/08/27/the-rise-of-short-form-video-tiktok-is-changing-the-game/?sh=78ccf7c15083 [accessed 2023-11-02]
14. Paivio A. Imagery and Verbal Processes. New York, NY: Psychology Press; 2013.
15. Brewster J, Arvanitis L, Pavilonis V, Wang M. Misinformation monitor: September 2022. NewsGuard. URL: https://www.newsguardtech.com/misinformation-monitor/september-2022 [accessed 2023-11-02]
16. 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]
17. Ceci L. U.S. TikTok audience by gender 2023. Statista. URL: https://www.statista.com/statistics/1398182/tiktok-us-audience-by-gender/ [accessed 2023-11-02]
18. Ceci L. U.S. TikTok users by age 2022. Statista. URL: https://www.statista.com/statistics/1095186/tiktok-us-users-age/ [accessed 2023-11-02]
19. Davies C, Mann A. Factors influencing women to accept diet and exercise messages on social media during COVID-19 lockdowns: a qualitative application of the health belief model. Health Mark Q 2023;40(4):415-433. [doi: 10.1080/07359683.2023.2193076] [Medline: 37021625]
20. Dluhos-Sebesto C, Jethwa TE, Bertasi TG, Bertasi RA, Nishi LY, Pantin SA, et al. Women's health information survey: common health concerns and trusted sources of health information among different populations of female patients. Womens Health Rep (New Rochelle) 2021;2(1):173-181 [FREE Full text] [doi: 10.1089/whr.2020.0118] [Medline: 34235504]
21. Manierre MJ. Gaps in knowledge: tracking and explaining gender differences in health information seeking. Soc Sci Med 2015 Mar;128:151-158. [doi: 10.1016/j.socscimed.2015.01.028] [Medline: 25618604]
22. General facts on women and job based health. U.S. Department of Labor. 2018. URL: https://www.dol.gov/sites/dolgov/files/ebsa/about-ebsa/our-activities/resource-center/fact-sheets/women-and-job-based-health.pdf [accessed 2024-04-29]
23. van der Meer TG, Brosius A, Hameleers M. The role of media use and misinformation perceptions in optimistic bias and third-person perceptions in times of high media dependency: evidence from four countries in the first stage of the COVID-19 pandemic. Mass Commun Soc 2022 Feb 25;26(3):438-462. [doi: 10.1080/15205436.2022.2039202]
24. Caiata-Zufferey M, Abraham A, Sommerhalder K, Schulz PJ. Online health information seeking in the context of the medical consultation in Switzerland. Qual Health Res 2010 Aug;20(8):1050-1061. [doi: 10.1177/1049732310368404] [Medline: 20442347]
25. Huo J, Desai R, Hong YR, Turner K, Mainous AG, Bian J. Use of social media in health communication: findings from the health information national trends survey 2013, 2014, and 2017. Cancer Control 2019 Apr 18;26(1):1073274819841442 [FREE Full text] [doi: 10.1177/1073274819841442] [Medline: 30995864]

26. Jia X, Pang Y, Liu LS. Online health information seeking behavior: a systematic review. Healthcare (Basel) 2021 Dec 16;9(12):1740 [FREE Full text] [doi: 10.3390/healthcare9121740] [Medline: 34946466]
27. Flanagin AJ, Metzger MJ. Perceptions of internet information credibility. Mass Commun Q Journal 2016 Jun 25;77(3):515-540. [doi: 10.1177/107769900007700304]
28. Chang C. Fake news: audience perceptions and concerted coping strategies. Digit Journal 2021 May 25;9(5):636-659. [doi: 10.1080/21670811.2021.1923403]
29. Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. Commun Res 2009 Nov 25;37(1):105-132. [doi: 10.1177/0093650209351512]
30. Ellison N, Steinfield C, Lampe C. The benefits of Facebook “friends:” social capital and college students’ use of online social network sites. J Comput Mediat Commun 2007;12(4):1143-1168. [doi: 10.1111/j.1083-6101.2007.00367.x]
31. Perez S. Google exec suggests Instagram and TikTok are eating into Google's core products, Search and Maps. TechCrunch. URL: https://tinyurl.com/mp8zhsy9 [accessed 2023-11-03]
32. de Vere Hunt I, Linos E. Social media for public health: framework for social media-based public health campaigns. J Med Internet Res 2022 Dec 14;24(12):e42179 [FREE Full text] [doi: 10.2196/42179] [Medline: 36515995]
33. 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 09;23(12):e30315 [FREE Full text] [doi: 10.2196/30315] [Medline: 34889739]
34. Gordon D. 33% Of Gen Zers trust TikTok more than doctors, new survey shows. Forbes. URL: https://www.forbes.com/sites/debgordon/2022/12/20/33-of-gen-zers-trust-tiktok-more-than-doctors-new-survey-shows/?sh=2f5770e6c7b2 [accessed 2023-04-29]
35. Samulowitz A, Gremyr I, Eriksson E, Hensing G. "Brave men" and "emotional women": a theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. Pain Res Manag 2018;2018:6358624-6358614 [FREE Full text] [doi: 10.1155/2018/6358624] [Medline: 29682130]
36. Govender V, Penn-Kekana L. Gender biases and discrimination: a review of health care interpersonal interactions. Glob Public Health 2008;3 Suppl 1:90-103. [doi: 10.1080/17441690801892208] [Medline: 19288345]
37. Milner A, Krnjacki L, LaMontagne AD. Age and gender differences in the influence of social support on mental health: a longitudinal fixed-effects analysis using 13 annual waves of the HILDA cohort. Public Health 2016 Nov;140:172-178. [doi: 10.1016/j.puhe.2016.06.029] [Medline: 27527844]
38. Kelman HC. Processes of opinion change. Public Opin Q 1961;25(1):78 [FREE Full text] [doi: 10.1086/266996]
39. Andsager JL, Bemker V, Choi HL, Torwel V. Perceived similarity of exemplar traits and behavior: effects on message evaluation. Commun Res 2016 Jun 30;33(1):3-18. [doi: 10.1177/0093650205283099]
40. Chou WY, 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]
41. O'Sullivan NJ, Nason G, Manecksha RP, O'Kelly F. The unintentional spread of misinformation on 'TikTok'; a paediatric urological perspective. J Pediatr Urol 2022 Jun;18(3):371-375. [doi: 10.1016/j.jpurol.2022.03.001] [Medline: 35331640]
42. 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 03;17(8):2373-2377 [FREE Full text] [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
43. Poll: most Americans encounter health misinformation, and most aren’t sure whether it’s true or false. KFF. URL: https://tinyurl.com/ycx6ms58 [accessed 2023-11-03]
44. Lyons BA, Montgomery JM, Guess AM, Nyhan B, Reifler J. Overconfidence in news judgments is associated with false news susceptibility. Proc Natl Acad Sci U S A 2021 Jun 08;118(23):e2019527118 [FREE Full text] [doi: 10.1073/pnas.2019527118] [Medline: 34074757]
45. Scheufele DA, Krause NM. Science audiences, misinformation, and fake news. Proc Natl Acad Sci U S A 2019 Apr 16;116(16):7662-7669 [FREE Full text] [doi: 10.1073/pnas.1805871115] [Medline: 30642953]
46. Metzger MJ, Flanagin AJ. Using Web 2.0 technologies to enhance evidence-based medical information. J Health Commun 2011;16 Suppl 1:45-58. [doi: 10.1080/10810730.2011.589881] [Medline: 21843095]
47. Dong Z. How to persuade adolescents to use nutrition labels: effects of health consciousness, argument quality, and source credibility. Asian J Commun 2015 Feb 19;25(1):84-101. [doi: 10.1080/01292986.2014.989241]
48. Eagly AH, Wood W, Chaiken S. Causal inferences about communicators and their effect on opinion change. J Pers Soc Psychol 1978;36(4):424-435. [doi: 10.1037//0022-3514.36.4.424]
49. Dockter CE, Lee S, Boman CD, Hinnant A, Cameron GT. The impact of retransmission and modality on communicating health research findings via social media. Health Commun 2021 Sep;36(10):1231-1241 [FREE Full text] [doi: 10.1080/10410236.2020.1749354] [Medline: 32268798]
50. Vorias T. Beware of TikTok’s medical “experts”. The Michigan Daily. 2021. URL: http://www.michigandaily.com/opinion/columns/beware-of-tiktoks-medical-experts/ [accessed 2023-11-12]
51. Sung M. On TikTok, mental health creators are confused for therapists. Mashable. URL: https://mashable.com/article/tiktok-mental-health-therapist-psychology [accessed 2023-11-12]

52. Hause E, Praska C, Pitt MB, Hendrickson MA, Charpentier V, Allen KA, et al. What's in a name? Laypeople's understanding of medical roles and titles. J Hosp Med 2022 Dec 29;17(12):956-960 [FREE Full text] [doi: 10.1002/jhm.12971] [Medline: 36173137]
53. Rhoads KV, Cialdini RB. The business of influence: Principles that lead to success in commercial settings. In: Dillard JP, Pfau M, editors. The Persuasion Handbook: Developments in Theory and Practice Sage. Thousand Oaks, CA: SAGE Publications; 2002:513-542.
54. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. In: Berkowitz L, editor. Advances in Experimental Social Psychology. Washington, DC: Academic Press; 1986:123-205.
55. Kelman HC. Compliance, identification, and internalization three processes of attitude change. J Conflict Resolut 2016 Jul 01;2(1):51-60 [FREE Full text] [doi: 10.1177/002200275800200106]
56. Singh S. Why short-form video is here to stay (even if TikTok isn’t). Linkedin. 2023. URL: https://www.linkedin.com/pulse/why-short-form-video-here-stay-even-tiktok-isnt-sunjay-singh/ [accessed 2023-11-02]

Edited by T Mackey; submitted 17.11.23; peer-reviewed by S Chinn, H Song, M Loades; comments to author 10.01.24; revised version received 05.03.24; accepted 11.04.24; published 21.05.24.

Please cite as:
Kirkpatrick CE, Lawrie LL
TikTok as a Source of Health Information and Misinformation for Young Women in the United States: Survey Study
JMIR Infodemiology 2024;4:e54663
URL: https://infodemiology.jmir.org/2024/1/e54663
doi:10.2196/54663
PMID:38772020

Original Paper

Association Between X/Twitter and Prescribing Behavior During the COVID-19 Pandemic: Retrospective Ecological Study

Scott A Helgeson[1], MS, MD; Rohan M Mudgalkar[2], MS; Keith A Jacobs[2], BS, MPH; Augustine S Lee[1], MS, MD; Devang Sanghavi[3], MHA, MD; Pablo Moreno Franco[4], MD; Ian S Brooks[2], PhD; National COVID Cohort Collaborative (N3C)[5]

[1]Department of Pulmonary and Critical Care Medicine, Mayo Clinic, Jacksonville, FL, United States

[2]School of Information Sciences, Center for Health Informatics, University of Illinois at Urbana-Champaign, Ubana-Champaign, IL, United States

[3]Department of Critical Care Medicine, Mayo Clinic, Jacksonville, FL, United States

[4]Department of Transplant Medicine, Mayo Clinic, Jacksonville, FL, United States

[5]see Acknowledgments

Corresponding Author:
Scott A Helgeson, MS, MD
Department of Pulmonary and Critical Care Medicine
Mayo Clinic
4500 San Pablo Road South
Jacksonville, FL, 32224
United States
Phone: 1 9049532000
Email: helgeson.scott@mayo.edu

Abstract

Background: Social media has become a vital tool for health care providers to quickly share information. However, its lack of content curation and expertise poses risks of misinformation and premature dissemination of unvalidated data, potentially leading to widespread harmful effects due to the rapid and large-scale spread of incorrect information.

Objective: We aim to determine whether social media had an undue association with the prescribing behavior of hydroxychloroquine, using the COVID-19 pandemic as the setting.

Methods: In this retrospective study, we gathered the use of hydroxychloroquine in 48 hospitals in the United States between January and December 2020. Social media data from X/Twitter was collected using Brandwatch, a commercial aggregator with access to X/Twitter's data, and focused on mentions of "hydroxychloroquine" and "Plaquenil." Tweets were categorized by sentiment (positive, negative, or neutral) using Brandwatch's sentiment analysis tool, with results classified by date. Hydroxychloroquine prescription data from the National COVID Cohort Collaborative for 2020 was used. Granger causality and linear regression models were used to examine relationships between X/Twitter mentions and prescription trends, using optimum time lags determined via vector auto-regression.

Results: A total of 581,748 patients with confirmed COVID-19 were identified. The median daily number of positive COVID-19 cases was 1318.5 (IQR 1005.75-1940.3). Before the first confirmed COVID-19 case, hydroxychloroquine was prescribed at a median rate of 559 (IQR 339.25-728.25) new prescriptions per day. A day-of-the-week effect was noted in both prescriptions and case counts. During the pandemic in 2020, hydroxychloroquine prescriptions increased significantly, with a median of 685.5 (IQR 459.75-897.25) per day, representing a 22.6% rise from baseline. The peak occurred on April 2, 2020, with 3411 prescriptions, a 397.6% increase. Hydroxychloroquine mentions on X/Twitter peaked at 254,770 per day on April 5, 2020, compared to a baseline of 9124 mentions per day before January 21, 2020. During this study's period, 3,823,595 total tweets were recorded, with 10.09% (n=386,115) positive, 37.87% (n=1,448,030) negative, and 52.03% (n=1,989,450) neutral sentiments. A 1-day lag was identified as the optimal time for causal association between tweets and hydroxychloroquine prescriptions. Univariate analysis showed significant associations across all sentiment types, with the largest impact from positive tweets. Multivariate analysis revealed only neutral and negative tweets significantly affected next-day prescription rates.

Conclusions: During the first year of the COVID-19 pandemic, there was a significant association between X/Twitter mentions and the number of prescriptions of hydroxychloroquine. This study showed that X/Twitter has an association with the prescribing behavior of hydroxychloroquine. Clinicians need to be vigilant about their potential unconscious exposure to social media as a source of medical knowledge, and health systems and organizations need to be more diligent in identifying expertise, source, and quality of evidence when shared on social media platforms.

(JMIR Infodemiology 2024;4:e56675) doi:10.2196/56675

KEYWORDS

social media; infodemic; COVID-19; healthcare utilization; misinformation; disinformation; Twitter; hydroxychloroquine; X; drugs; pharmacy; pharmacology; pharmacotherapy; pharmaceuticals; medication; prescription; sentiment; SARS-CoV-2; pandemic; respiratory; infectious

Introduction

As of September 1, 2024, SARS-CoV-2 has led to 111,820,082 infections and 1,201,061 deaths reported in the United States alone [1]. When the first nontravel case of COVID-19 was reported in the United States on January 20, 2020, clinicians had no significant knowledge of the virus nor the management of the disease resulting from the virus and looked to colleagues from other countries for their experiences [2,3].

Although peer-reviewed observational reports were being published rapidly in record amounts, and clinical trials were ongoing, clinicians also used social media to share and gain knowledge on the care of their patients in hopes of improving their patients' outcomes [4-6]. Even though social media is increasingly used by health care providers and health care systems, its accuracy and role as a medium for sharing critical information has the potential to be corrupted, leading to the use of unproven, potentially harmful, and costly therapies [7].

Multiple studies have shown misinformation as shown by false tweets was widely available and were even retweeted faster [8-11]. Not only were there concerns about true and false information spreading, but there was regional sentiment variation on X/Twitter throughout the United States [12]. This misinformation was not only spread by nonmedical people but also by physicians from a range of specialties [13]. To combat this infodemic, the World Health Organization established a repository of COVID-19 fact-checking groups that verify COVID-19 claims [14].

We hypothesized that the circumstances of the COVID-19 pandemic (ie, lack of evidence-based medicine, overly stressed health care system, and health care providers) may have magnified the potential for social media to have influenced the care of patients affected by COVID-19. To explore this hypothesis, we aimed to see if we could identify a temporal association between the X/Twitter mentions of hydroxychloroquine and the "sentiment" of these X/Twitter mentions, against the daily use of the medication.

Methods

Social Media Data

All social media data was gathered before the rebranding of Twitter to X, so throughout this paper X will be referred to as X/Twitter and posts will be referred to as tweets. The commercial social media aggregator, Brandwatch was leveraged to gather X/Twitter data and perform sentiment analysis [15,16]. Brandwatch is an official partner of X/Twitter allowing better access to X/Twitter data [17]. The accuracy of Brandwatch's sentiment analysis is around 75%, which is comparable to other sentiment analysis tools [18-20]. In the tool, a query for the drug "hydroxychloroquine" and "Plaquenil" was added, and the date range was set from November 30, 2019, to January 29, 2021. Unlike other tools that rely on a dictionary of "good" and "bad" words for performing sentiment analysis, Brandwatch's sentiment analysis models are based on a library of hundreds of rules created based on natural language processing and are updated and audited regularly [16]. Brandwatch classifies tweets as positive or negative only if it is confident about the sentiment classification. If the sentiment cannot be accurately determined, such tweets are categorized as neutral. Tweets were reported as neutral, nonneutral, total, positive, and negative sentiment. Nonneutral tweets combined positive and negative tweets, while total tweets summed up neutral and nonneutral tweets. The tweets were grouped by date, and a CSV file was generated, which included a count of each category of tweets for each day.

COVID-19 and Medication Data

Medication data was obtained from the National COVID Cohort Collaborative (N3C), a national collection of 48 hospitals or data partners with 4.8 million patients [21]. The N3C cohort is comprised of patients diagnosed with COVID-19 by polymerase chain reaction (PCR) and a control group of patients without COVID-19 matched by age, sex, and race at a 2:1 ratio. For this study, the entire cohort from 2020 was used to capture new hydroxychloroquine prescriptions per day either in the inpatient or in the outpatient setting. Restriction to COVID-19 positivity or negativity was not carried out because the total usage of this medication was being assessed and not specifically the treatment of patients who were tested positive for COVID-19, as patients were treated with hydroxychloroquine before their COVID-19 status resulted, and some of these patients were likely to be negative. Data that could identify patients were not included in the data access, such as age, race, and gender, per N3C. New COVID-19 positivity numbers by PCR were collected with the start date being February 1, 2020, as this was the first date in the N3C database of positive tests. The "drug exposure" dataset of the N3C database was specifically queried for hydroxychloroquine, hydroxychloroquine sulfate, Quineprox, and Plaquenil. There were no exclusion criteria.

Timeline Data

Prepandemic basal rates of hydroxychloroquine prescriptions and tweets were defined as the period prior from January 1, 2020, to the first reported case of COVID-19 in the United States on January 20, 2020. Pandemic rates were determined using the timeframe between January 21, 2020, to December 31, 2020.

The first peer-reviewed publications on hydroxychloroquine for COVID-19, and key sentinel media or public events or announcements were identified, including announcements from the FDA or other government and nongovernment authorities. A timeline illustrating the X/Twitter hits (including the sentiment) was overlayed to highlight impacts with key media events or announcements, and key peer-reviewed publications evaluating the efficacy of hydroxychloroquine on COVID-19.

Statistical Analysis

Median and IQRs were reported for descriptive statistics. COVID-19 PCR positivity rates were used to determine the relative use (percent) of hydroxychloroquine prescriptions per patient who tested positive for COVID-19. Estimation of the annual number of hydroxychloroquine prescriptions without the pandemic was carried out by using the median over the prepandemic period and multiplying by 365 (days) to come up with a yearly total. This was used to determine the difference in the number of prescriptions for unnecessary prescriptions. A 10-day prescription and the National Average Drug Acquisition Cost in 2021 were used to calculate the cost of a prescription [22,23].

Statistical analysis was performed using a Granger causality test to determine whether a temporal association exists between X/Twitter mentions and hydroxychloroquine prescriptions [24]. For analysis, the tweets were classified into 3 sentiment categories: neutral, nonneutral, and total as Brandwatch sentiment analysis allows for the most accuracy with this approach.

To obtain the optimum time lag length for Granger causality tests, a vector auto-regressive (VAR) model was fitted to the dataset. The VAR model was run for neutral, nonneutral, and total tweets, and the optimum time lag was noted for each test. Using the optimum time lags calculated from the VAR model, the Granger causality test was then used to determine if a causal relationship exists between X/Twitter mentions of hydroxychloroquine and its prescriptions. The chi-square metric was used to verify the significance of the time lag.

As the Granger causality test does not specify the direction of the temporal association, the optimum lag day was used in a linear regression model to determine whether the number of prescriptions of hydroxychloroquine could be predicted by the number of tweets mentioning hydroxychloroquine in the preceding days (offset by the optimum lag day determined by the above VAR analysis). Similar analyses using a multivariate linear regression model, but using the tweet sentiments as covariates (positive, neutral, and negative sentiment tweets) were carried out to identify their independent impact on predicting the number of new hydroxychloroquine prescriptions. A *P* value of less than .05 was considered significant. Granger analysis was performed using Python (version 3.9.7; Python Software Foundation) using the “stats models” package and the linear regression models were performed in BlueSky (version 10.3; BlueSky Statistics).

Ethical Considerations

No identifiers were used in the collection or analysis of the data, and this study was considered exempt by the Mayo Clinic institutional review board (#21-002787). This study was approved to use a limited dataset (access level 3) by the N3C Data Use Request Committee. All data were deidentified and a waiver of informed consent was approved consistent with social media postings.

Results

COVID-19 and Hydroxychloroquine Prescriptions

A total of 581,748 cases of confirmed patients with COVID-19 were identified. The median number of daily positive case counts of patients with COVID-19 was 1318.5 (IQR 1005.75-1940.3) as there was documented reinfection during this study’s period (Figure 1). At a baseline, the median daily use of hydroxychloroquine before the first confirmed case of COVID-19 on January 20, 2020, was 559 (IQR 339.25-728.25) new prescriptions per day (Figure 2). There appears to be a day-of-the-week effect resulting in oscillations of prescriptions and COVID-19 diagnosis. As the pandemic progressed, there was a significant increase in the use of hydroxychloroquine. During the pandemic in 2020, hydroxychloroquine was prescribed a median of 685.5 (IQR 459.75-897.25) times per day, which was a 22.6% increase above the baseline rate. The peak number of prescriptions during a single day was 3411, which occurred on April 2, 2020, which was a 397.6% increase in prescriptions during that day.

Figure 1. Using the N3C database, the daily census data of new patients with COVID-19 diagnosed through PCR testing was displayed from February 1, 2020, to December 31, 2020. N3C: National COVID Cohort Collaborative; PCR: polymerase chain reaction.

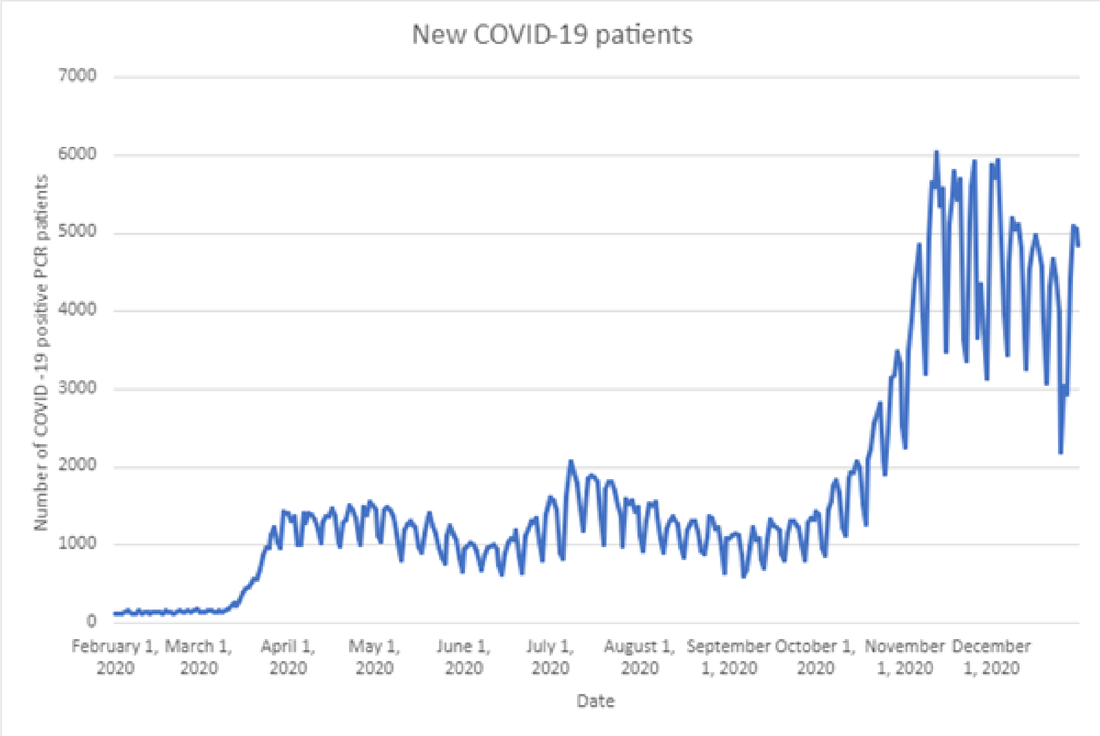
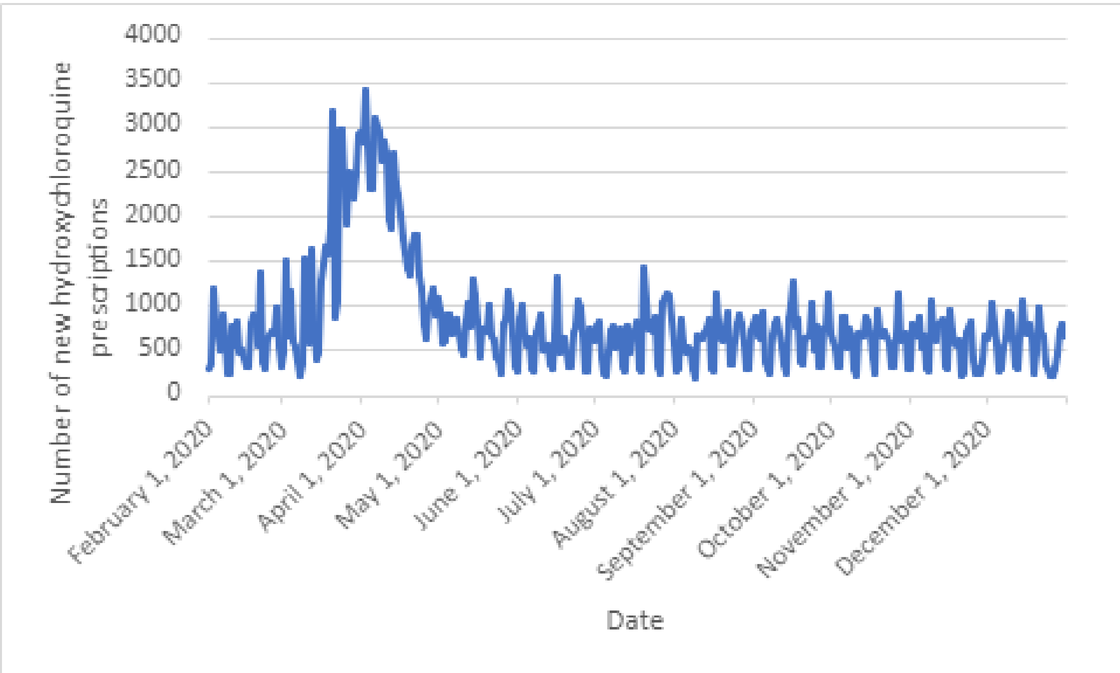


Figure 2. Using the N3C database, daily new hydroxychloroquine prescriptions for any reason were displayed from February 1, 2020, to December 31, 2020. N3C: National COVID Cohort Collaborative.



X/Twitter Data

Hydroxychloroquine mentions on X/Twitter did not increase significantly until March 2020. X/Twitter mentions and retweets on hydroxychloroquine at baseline were 9124 per day before January 21, 2020. It increased from the pre-COVID baseline and peaked at 254,770 tweets per day on April 5, 2020 (Figure 3 [3,25-31]). There were notable associations between several key publications, public events, or announcements, and a rise in X/Twitter mentions and retweets (Figure 3 [3,25-31]). X/Twitter data were separated into total, neutral, positive, and negative sentiments (Figure 4). For the investigational period, the total tweets mentioning hydroxychloroquine were 3,823,595, with 10.09% (n=386,115) positive, 37.87% (n=1,448,030) negative, and 52.03% (n=1,989,450) neutral sentiments.

Figure 3. The total X/Twitter mentions of hydroxychloroquine, obtained by Brandwatch, in a natural log scale was displayed from January 1, 2020, to December 31, 2020, and was marked with some key publications and media moments noted by numbers. (1) The first paper was published describing the use of chloroquine for COVID-19 from expert opinion on March 12, 2020 [25]. (2) On March 17, 2020, an open-label nonrandomized study was published showing success with the use of hydroxychloroquine in patients with COVID-19 [26]. (3) On April 5, 2020, the president of the United States promoted the use of hydroxychloroquine for COVID-19 [27]. (4) Then the Center for Disease Control deleted a report of physicians using hydroxychloroquine [3]. (5) The Food and Drug Administration on April 24, 2020, announced to use hydroxychloroquine with caution outside of the hospital or clinical trial secondary to arrhythmias [28]. (6) On May 18, 2020, the president of the United States reported tweets that he is taking hydroxychloroquine to protect himself from COVID-19 [29]. (7) The World Health Organization stopped a trial with hydroxychloroquine over safety concerns [30]. (8) On June 15, 2020, the Food and Drug Administration revoked the Emergency Use Authorization for chloroquine and hydroxychloroquine on COVID-19 [31].

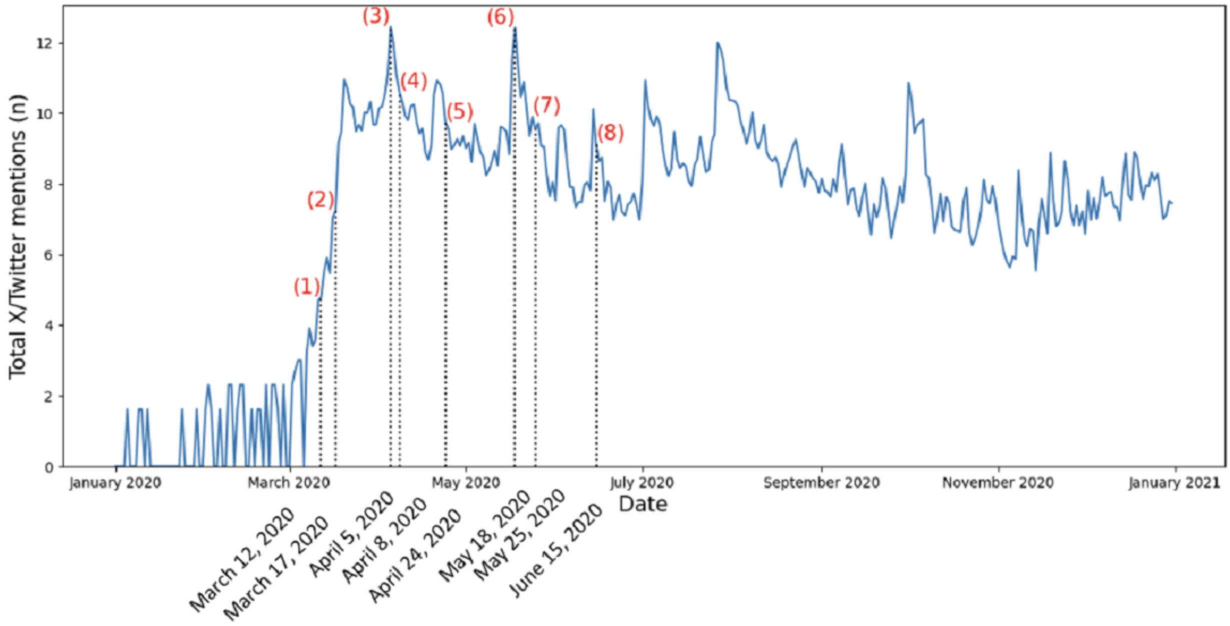
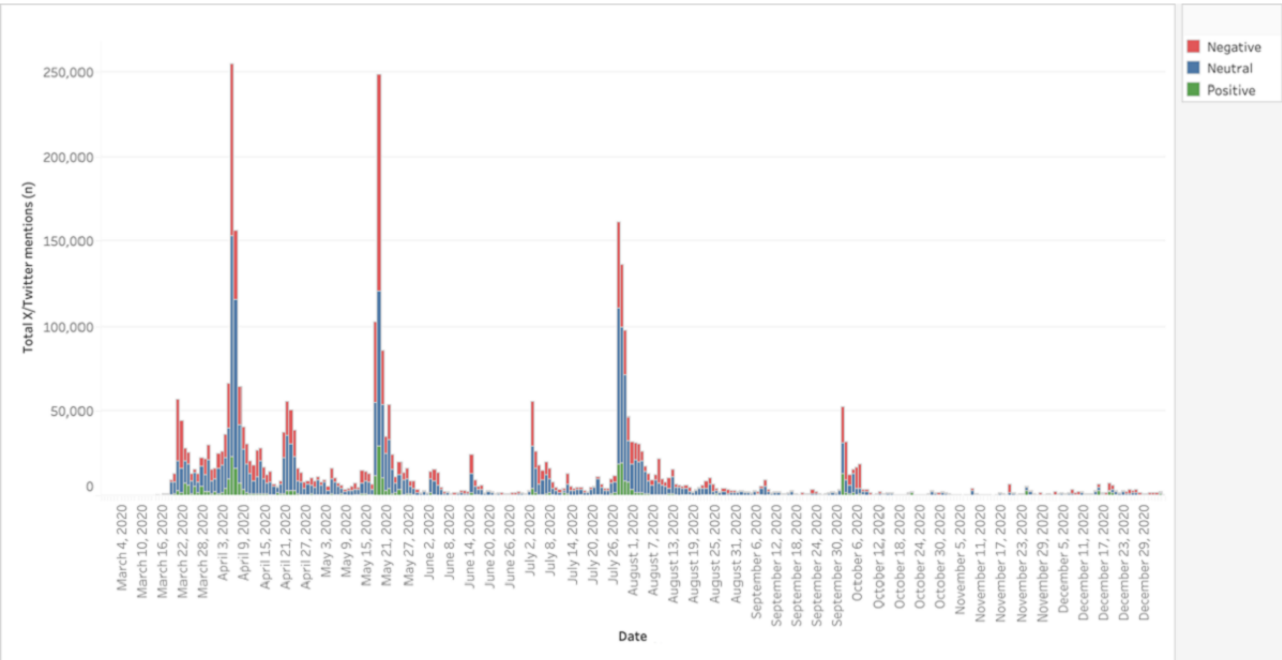


Figure 4. Using Brandwatch, the volume of tweets by total amount and sentiment category (neutral and nonneutral divided out by positive and negative sentiment) was displayed from throughout the year 2020.



Optimal Lag-Day and Temporal Causality Analysis

The optimal lag-day of 1 day was identified using the VAR model, indicating that the tweets mentioning hydroxychloroquine were most strongly associated with hydroxychloroquine prescriptions within 1 day (Table 1). Although Granger analysis identified significant causal relationships between tweets and prescriptions of hydroxychloroquine across multiple lag days, it confirmed that the strongest statistical relationship occurred with a 1-day time lag identified (total tweets, *P*=.005; neutral tweets, *P*=.001; and nonneutral tweets, *P*=.02; Table S1 in Multimedia Appendix 1).

Table 1. Univariate and multivariate analyses for the direction of relationships between tweets and the number of hydroxychloroquine prescriptions with a 1-day time lag through the 2020 year. Tweets were separated by sentiment and total number.

Sentiment	β value	*P* value
Univariate analysis		
Positive	.0641	<.001
Negative	.0155	<.001
Neutral	.0169	<.001
Total	.0078	<.001
Multivariate analysis		
Positive	.0029	.92
Negative	–.0149	.04
Neutral	.0271	<.001

Directional Analysis

The optimal lag day of 1 day demonstrated that an increase in tweets increased hydroxychloroquine prescriptions the following day (Table 1). The biggest impact was with positive tweets, which showed that for every 100 positive sentiment tweets, there would be 6 new hydroxychloroquine prescriptions the next day. Negative tweets, as well as neutral tweets, appeared to be associated with an increase in hydroxychloroquine prescriptions the next day.

However, in multivariate analysis with covariates of positive, negative, and neutral tweets, only neutral and negative tweets seemed to affect hydroxychloroquine prescriptions the next day. On average, the impact of 1000 positive tweets would result in 27.1 new prescriptions the following day, while 1000 negative sentiment tweets would result in 14.9 fewer hydroxychloroquine prescriptions the following day.

Potential Harm and Costs

In 1 meta-analysis of 9 phase 3 randomized controlled clinical trials, the rate of adverse events with hydroxychloroquine was 12%, with a computed number needed to harm of 9 [32]. The total number of hydroxychloroquine prescriptions over the basal prepandemic period was 68,893. This would translate to 7654.8 patients potentially risking harm if all the hydroxychloroquine prescriptions were for COVID-19. Additionally, hydroxychloroquine National Average Drug Acquisition Costs were estimated at US $3.43 for each new prescription, which may have contributed to an excess of US $236,473.85 in unnecessary costs during this study's period.

Discussion

The main findings of our investigation provide temporal evidence that prescribing practices of health care providers were associated with X/Twitter tweets. Specifically, the number of hydroxychloroquine-related tweets was temporally associated with new hydroxychloroquine prescriptions the following day. Supporting the temporal association, multivariate analysis showed that negative sentiment tweets decreased the subsequent prescribing of hydroxychloroquine. Recognizing that hydroxychloroquine has been shown to be ineffective for COVID-19 infection, there may have been 68,893 unnecessary hydroxychloroquine prescriptions resulting in avoidable harm, and US $236,473.85 in excess costs from this study's sample.

Hydroxychloroquine is an antimalarial medication most often used in the United States to treat rheumatologic disorders. Early reports in the COVID-19 pandemic suggested a potential benefit of hydroxychloroquine in treating COVID-19 [26,32-34]. There was biological plausibility to these studies, with prior in vitro studies showing the activity of hydroxychloroquine against SARS-CoV-1 (severe acute respiratory syndrome coronavirus) and MERS-CoV (Middle East respiratory syndrome) [35,36]. None of the COVID-19 studies reported that considered hydroxychloroquine showed a strong enough level of evidence to meet the criteria for use in patients. The deviation from traditional drug approval pathways and its dissemination to the bedside, which normally takes years, because of the need for effective treatments accelerated this process and arguably led to shortcuts in the process [37]. This scenario led to the backdrop for our study.

The internet era has undoubtedly contributed to improved access and sharing of knowledge and information with physicians and patients [38]. Social media platforms promised to facilitate and strengthen relationships between diverse people and opinions worldwide, by enhancing and supporting collaborations and improving knowledge sharing in the professional world [39,40]. The pace at which information can be shared and disseminated by social media is staggering, considering that previously a confirmed best practice may take more than a decade before its widespread implementation in clinical practice. False news spreads have been reported to spread about 6 times faster and to more than 10 times the people than accurate news when evaluating X/Twitter data [41]. However, much like modern-day centralized electronic health systems, a small error could be magnified and applied systematically across a population of patients unknowingly [42,43]. In the case of social media, there is misinformation (incorrect information) and disinformation (deliberately inaccurate information) on the care of patients that could be magnified, particularly under stressed circumstances, such as a devastating pandemic. X/Twitter responded to this concern by flagging potentially concerning information and removing 11,230 accounts because of misinformation [44]. In the specific case of clinical care, there is concern about circumventing the scientific peer-reviewed process to test the safety of interventions and their risks rigorously. Social media may prematurely propagate preliminary unvalidated therapies such as hydroxychloroquine for COVID-19. Examples of clinical findings conveyed to the media before peer review are suspect and should generally be avoided, with recent examples of disinformation through social media being high doses of vitamin C, thiamine, and hydrocortisone in patients who develop sepsis [45].

COVID-19 caused an infodemic with huge releases of medical literature through peer-reviewed and non–peer-reviewed sources with more than 1000 publications monthly [4]. Along with these publications, technology and social media played a massive role in information dissemination, with millions of tweets weekly about COVID-19 [46]. This sheer volume of data could negatively impact patient care, given that a significant amount of all health- and COVID-19–related tweets contained false information [7]. The World Health Organization raised concerns that this infodemic adversely affected global health related to COVID-19 [47]. Poor-quality health care data have historically been the problem paving the way for a misinformation infodemic [48]. Other prior epidemics (Ebola and Zika) suggested the same infodemic concerns with misinformation spread by social media [49-56]. Outside of epidemics and pandemics, there has also been consistent misinformation about vaccinations portrayed on social media platforms [57-62]. With the COVID-19 pandemic, qualitative studies have shown significant misinformation (up to 70% of information was false or lacked evidence) on social media platforms [7,9,63-67]. One promising study on X/Twitter data showed that false information was tweeted more than science-based information, but science-based tweets were retweeted more [68]. Our study, regarding specifically hydroxychloroquine use, showed that X/Twitter was associated with the prescribing practice during the COVID-19 pandemic, which raises significant concern about a misinformation infodemic.

Some potential fixes to this problem consist of education for physicians or clinicians and appropriate tagging of X/Twitter data. X/Twitter tried to combat misinformation with the implementation of “verified” status, which means that the account is authentic [69]. This verification does not mean that someone is an expert in the field or disease for which they are tweeting. On top of that, the verification system has stopped. Around 39,000 physicians were active on X/Twitter during the start of the pandemic, but it is not known how many provided accurate information. The appropriate use of X/Twitter for health care information or education has not been standardized. There are no educational programs about the appropriate use of X/Twitter and a physician’s responsibility for accurate and verified information. Many of the X/Twitter best practices are geared toward physicians on how to set up their accounts, become noticed, and create connections [70,71]. The development of physician best practices on X/Twitter needs to occur with a focus on disseminating information that is accurate and that nonmedical people will understand. This also creates a need for medical journals to enhance their presence on social media platforms so peer-reviewed studies take the forefront and people can gain access to these studies earlier than for printed copies.

Our findings are compelling but have limitations that should be considered in their interpretation. First, this is an ecological study of multiple databases and their temporal trends, but they are not tied to individual prescribers or patients. For example, we do not know for sure that the use of hydroxychloroquine was prescribed specifically to patients who had COVID-19 and we cannot track whether these prescribing health care professionals used the social media platform. We also do not know if the patients who were prescribed hydroxychloroquine were diagnosed with COVID-19, only had symptoms, or were presumed positive. Additionally, although our findings support that social media did affect COVID-19 prescribing habits, this study did not look at physician versus nonphysician handles. The infodemic could have enhanced the public perception of possible therapies, leading to questioning clinicians taking care of their loved ones, as to why such treatments are being withheld, particularly under the pressure of a patient who is rapidly deteriorating or already on life support [72]. Using X/Twitter as the sample platform to gauge whether social media influenced clinician behavior is faulty, and we acknowledge that social media represents multiple ever-changing platforms and media. However, it is one of the most used social media platforms, with over 396 million users worldwide. The sheer number of hits supported that X/Twitter reasonably represents overall social media sentiments. Sentiment analysis can be difficult because distinguishing between someone being happy that hydroxychloroquine was being used and someone being happy that it was stopped cannot be performed. Brandwatch uses a set of proprietary rules to classify and score tweets into positive, negative, and neutral sentiments. As such, there is no disclosure as to how high these rules classify tweets and we were forced to use a “black box” algorithm, albeit a standard throughout this area of study. Lastly, X/Twitter is not the only source of information and is not the only means of creating undue influence on prescribing habits. There are likely additional factors that cannot be controlled for in this study that impacted prescribing patterns. Additionally, The World Health Organization, Center for Disease Control, hospitals, and medical societies, among others, could have actually mediated the association found in this study.

Similarly, our sample of acute care hospitals in the United States may not precisely reflect the behavior of the rest of the country, as participation and providing data were voluntary, but this is the largest cohort of hospitals to evaluate for prescribing practices [21]. There may be regional differences between academic and nonacademic practices and between types of prescribers, which could not be assessed. We must follow up our findings with more systematic data as they become available (such as from multiple payers and the Center for Medicare Services) and numerous social media platforms. Despite these limitations, our results are compelling, which clinicians, payers, administrators, and policy makers must know as additional supportive data arise.

In conclusion, X/Twitter information about hydroxychloroquine was significantly associated with the prescribing habits of clinicians within 1 day of the tweet. More granular data are necessary to evaluate how specific prescriber details and location affect prescribing practices. The impact of charging for verified

X/Twitter accounts may affect future misinformation infodemics, but the primary responsibility for misinformation should be on the person spreading the information and not on the platform or method of dissemination. This study shows a strong analytical case for the dangers of social media and the inappropriate attention and influence on prescribers. Further study is necessary on how to prevent or reduce infodemics in the future as this will likely keep occurring if an intervention is not performed.

Acknowledgments

The authors would also like to thank Michael Wieck-Sosa, BS, and Yiren Liu, BS, for their help in with this project. The analyses described in this publication were conducted with data or tools accessed through the National Center for Advancing Translational Sciences (NCATS) N3C Data Enclave [21] and N3C Attribution & Publication Policy (version 1.2-2020-08-25b; grant NCATS U24 TR002306). This research was possible because of the patients whose information is included in the data and the organizations [73] and scientists who have contributed to the ongoing development of this community resource [74]. We gratefully acknowledge the following core contributors to N3C: Adam B Wilcox, Adam M Lee, Alexis Graves, Alfred (Jerrod) Anzalone, Amin Manna, Amit Saha, Amy Olex, Andrea Zhou, Andrew E Williams, Andrew Southerland, Andrew T Girvin, Anita Walden, Anjali A Sharathkumar, Benjamin Amor, Benjamin Bates, Brian Hendricks, Brijesh Patel, Caleb Alexander, Carolyn Bramante, Cavin Ward-Caviness, Charisse Madlock-Brown, Christine Suver, Christopher Chute, Christopher Dillon, Chunlei Wu, Clare Schmitt, Cliff Takemoto, Dan Housman, Davera Gabriel, David A Eichmann, Diego Mazzotti, Don Brown, Eilis Boudreau, Elaine Hill, Elizabeth Zampino, Emily Carlson Marti, Emily R Pfaff, Evan French, Farrukh M Koraishy, Federico Mariona, Fred Prior, George Sokos, Greg Martin, Harold Lehmann, Heidi Spratt, Hemalkumar Mehta, Hongfang Liu, Hythem Sidky, JW Awori Hayanga, Jami Pincavitch, Jaylyn Clark, Jeremy Richard Harper, Jessica Islam, Jin Ge, Joel Gagnier, Joel H Saltz, Joel Saltz, Johanna Loomba, John Buse, Jomol Mathew, Joni L Rutter, Julie A McMurry, Justin Guinney, Justin Starren, Karen Crowley, Katie Rebecca Bradwell, Kellie M Walters, Ken Wilkins, Kenneth R Gersing, Kenrick Dwain Cato, Kimberly Murray, Kristin Kostka, Lavance Northington, Lee Allan Pyles, Leonie Misquitta, Lesley Cottrell, Lili Portilla, Mariam Deacy, Mark M Bissell, Marshall Clark, Mary Emmett, Mary Morrison Saltz, Matvey B Palchuk, Melissa A Haendel, Meredith Adams, Meredith Temple-O'Connor, Michael G Kurilla, Michele Morris, Nabeel Qureshi, Nasia Safdar, Nicole Garbarini, Noha Sharafeldin, Ofer Sadan, Patricia A Francis, Penny Wung Burgoon, Peter Robinson, Philip RO Payne, Rafael Fuentes, Randeep Jawa, Rebecca Erwin-Cohen, Rena Patel, Richard A Moffitt, Richard L Zhu, Rishi Kamaleswaran, Robert Hurley, Robert T Miller, Saiju Pyarajan, Sam G Michael, Samuel Bozzette, Sandeep Mallipattu, Satyanarayana Vedula, Scott Chapman, Shawn T O'Neil, Soko Setoguchi, Stephanie S Hong, Steve Johnson, Tellen D Bennett, Tiffany Callahan, Umit Topaloglu, Usman Sheikh, Valery Gordon, Vignesh Subbian, Warren A Kibbe, Wenndy Hernandez, Will Beasley, Will Cooper, William Hillegass, and Xiaohan Tanner Zhang. The details of contributions are available online [75].

Data Availability

The prescription and COVID-19 datasets generated and analyzed during this study are not publicly available due to the National COVID Cohort Collaborative data sharing rules, but the X/Twitter data are available from the corresponding author on reasonable request.

Authors' Contributions

SAH, ASL, ISB, DS, and PMF were responsible for study conception and design. SAH, RMM, KAJ, and ISB conducted data acquisition. All authors were responsible for data analysis and interpretation of results. All authors discussed the results and wrote this paper. IAS is the Associate Editor for JMIR Infodemiology.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables on tweet sentiment and tweets per day.
[DOCX File , 57 KB - infodemiology_v4i1e56675_app1.docx]

References

1. Centers for Disease Control and Prevention. COVID Data Tracker. Atlanta, GA: US Department of Health and Human Services, CDC URL: https://covid.cdc.gov/covid-data-tracker [accessed 2024-10-10]
2. CDC COVID-19 Response Team, Jorden MA, Rudman SL, Villarino E, Hoferka S, Patel MT, et al. Evidence for limited early spread of COVID-19 within the United States, January-February 2020. MMWR Morb Mortal Wkly Rep 2020;69(22):680-684 [FREE Full text] [doi: 10.15585/mmwr.mm6922e1] [Medline: 32497028]

3. Centers for Disease Control and Prevention. CDC museum COVID-19 timeline. URL: https://www.cdc.gov/museum/timeline/covid19.html [accessed 2024-10-10]
4. Chen Q, Allot A, Lu Z. Keep up with the latest coronavirus research. Nature 2020;579(7798):193. [doi: 10.1038/d41586-020-00694-1] [Medline: 32157233]
5. Ghosh P, Schwartz G, Narouze S. Twitter as a powerful tool for communication between pain physicians during COVID-19 pandemic. Reg Anesth Pain Med 2021;46(2):187-188. [doi: 10.1136/rapm-2020-101530] [Medline: 32321859]
6. Margus C, Brown N, Hertelendy AJ, Safferman MR, Hart A, Ciottone GR. Emergency physician Twitter use in the COVID-19 pandemic as a potential predictor of impending surge: retrospective observational study. J Med Internet Res 2021;23(7):e28615 [FREE Full text] [doi: 10.2196/28615] [Medline: 34081612]
7. Swetland SB, Rothrock AN, Andris H, Davis B, Nguyen L, Davis P, et al. Accuracy of health-related information regarding COVID-19 on Twitter during a global pandemic. World Med Health Policy 2021;13(3):503-517 [FREE Full text] [doi: 10.1002/wmh3.468] [Medline: 34540337]
8. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
9. Kouzy R, Abi Jaoude J, Kraitem A, El Alam MB, Karam B, Adib E, et al. Coronavirus goes viral: quantifying the COVID-19 misinformation epidemic on Twitter. Cureus 2020;12(3):e7255 [FREE Full text] [doi: 10.7759/cureus.7255] [Medline: 32292669]
10. Shahi GK, Dirkson A, Majchrzak TA. An exploratory study of COVID-19 misinformation on Twitter. Online Soc Netw Media 2021;22:100104 [FREE Full text] [doi: 10.1016/j.osnem.2020.100104] [Medline: 33623836]
11. Tasnim S, Hossain MM, Mazumder H. Impact of rumors and misinformation on COVID-19 in social media. J Prev Med Public Health 2020;53(3):171-174 [FREE Full text] [doi: 10.3961/jpmph.20.094] [Medline: 32498140]
12. Guntuku SC, Buttenheim AM, Sherman G, Merchant RM. Twitter discourse reveals geographical and temporal variation in concerns about COVID-19 vaccines in the United States. Vaccine 2021;39(30):4034-4038 [FREE Full text] [doi: 10.1016/j.vaccine.2021.06.014] [Medline: 34140171]
13. Sule S, DaCosta MC, DeCou E, Gilson C, Wallace K, Goff SL. Communication of COVID-19 misinformation on social media by physicians in the US. JAMA Netw Open 2023;6(8):e2328928 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.28928] [Medline: 37581886]
14. COVID-19 misinformation portal? A rapid response project from the social media lab. 2021. URL: https://covid19misinfo.org/ [accessed 2024-10-10]
15. Brandwatch. 2022. URL: https://www.brandwatch.com/ [accessed 2024-10-10]
16. Sentiment analysis. Brandwatch. 2022. URL: https://www.brandwatch.com/wp-content/uploads/brandwatch/Sentiment-Analysis.pdf [accessed 2024-10-10]
17. Palmer G. News: Brandwatch and Twitter expanded strategic partnership. 2017. URL: https://www.brandwatch.com/blog/brandwatch-and-twitter-expand-strategic-partnership/ [accessed 2024-10-10]
18. AminiMotlagh M, Shahhoseini H, Fatehi N. A reliable sentiment analysis for classification of tweets in social networks. Soc Netw Anal Min 2023;13(1):7 [FREE Full text] [doi: 10.1007/s13278-022-00998-2] [Medline: 36532862]
19. Gonzalez JC. Accuracy measures in sentiment analysis: the precision of Meaningcloud's technology. 2021. URL: https://www.meaningcloud.com/blog/accuracy-in-sentiment-analysis [accessed 2024-10-10]
20. Taylor N. Interview: the data science behind Brandwatch's new sentiment analysis. 2022. URL: https://www.brandwatch.com/blog/data-science-behind-brandwatchs-new-sentiment-analysis [accessed 2024-10-10]
21. National Institute of Health (NIH) and National Center for Advancing Translational Sciences (NCATS). National COVID Cohort Collaborative Data Enclave Respository. Bethesda, MA: US Department of Health and Human Services, National Institutes of Health URL: https://covid.cd2h.org/ [accessed 2024-10-10]
22. Self WH, Semler MW, Leither LM, Casey JD, Angus DC, Brower RG, National Heart, Lung,Blood Institute PETAL Clinical Trials Network, et al. Effect of hydroxychloroquine on clinical status at 14 days in hospitalized patients With COVID-19: a randomized clinical trial. JAMA 2020;324(21):2165-2176 [FREE Full text] [doi: 10.1001/jama.2020.22240] [Medline: 33165621]
23. Centers for Medicare and Medicaid Services. NADAC (National Average Drug Acquisition Cost). 2021. URL: https://data.medicaid.gov/dataset/d5eaf378-dcef-5779-83de-acdd8347d68e/data?conditions[0][property]=ndc_description&conditions[0][value]=hydroxychloroquine&conditions[0][operator]=contains [accessed 2024-10-10]
24. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;37(3):424-438. [doi: 10.2307/1912791]
25. Chen Z, Hu J, Zhang Z, Jiang S, Han S, Yan D. Efficacy of hydroxychloroquine in patients with COVID-19: results of a randomized clinical trial. medRxiv. Preprint posted online on 2020. 2004. URL: https://www.medrxiv.org/content/10.1101/2020.03.22.20040758v3 [accessed 2024-10-10]
26. Gautret P, Lagier J, Parola P, Hoang VT, Meddeb L, Mailhe M, et al. Hydroxychloroquine and azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical trial. Int J Antimicrob Agents 2020;56(1):105949 [FREE Full text] [doi: 10.1016/j.ijantimicag.2020.105949] [Medline: 32205204]

27. Crowley M, Thomas K, Haberman M. Ignoring expert opinion, Trump again promotes hydroxychloroquine. New York City: New York Times; 2020. URL: https://www.nytimes.com/2020/04/05/us/politics/trump-hydroxychloroquine-coronavirus.html [accessed 2024-10-10]
28. US Food and Drug Administration. FDA cautions against use of hydroxychloroquine or chloroquine for COVID-19 outside of the hospital setting or a clinical trial due to risk of heart rhythm problems. URL: https://www.fda.gov/drugs/drug-safety-and-availability/fda-cautions-against-use-hydroxychloroquine-or-chloroquine-covid-19-outside-hospital-setting-or [accessed 2024-10-10]
29. Miller Z, Marchione M, Superville D. Trump says he's taking malaria drug to protect against virus. AP News. URL: https://apnews.com/article/virus-outbreak-malaria-donald-trump-ap-top-news-politics-0fbe485717c9a74e6d6c48aee7d208ae [accessed 2024-10-10]
30. Beaubien J. WHO halts hydroxychloroquine trial over safety concerns. NPR News. URL: https://www.npr.org/sections/coronavirus-live-updates/2020/05/25/861913688/who-halts-hydroxychloroquine-trial-over-safety-concerns [accessed 2024-10-10]
31. US Food and Drug Administration. Coronavirus (COVID-19) update: FDA revokes emergency use authorization for chloroquine and hydroxychloroquine. 2020. 2020. URL: https://tinyurl.com/5cczudeh [accessed 2024-10-10]
32. Tanni SE, Bacha HA, Naime A, Bernardo WM. Use of hydroxychloroquine to prevent SARS-CoV-2 infection and treat mild COVID-19: a systematic review and meta-analysis. J Bras Pneumol 2021;47(5):e20210236 [FREE Full text] [doi: 10.36416/1806-3756/e20210236] [Medline: 34669839]
33. Gao J, Tian Z, Yang X. Breakthrough: chloroquine phosphate has shown apparent efficacy in treatment of COVID-19 associated pneumonia in clinical studies. Biosci Trends 2020;14(1):72-73 [FREE Full text] [doi: 10.5582/bst.2020.01047] [Medline: 32074550]
34. Su Y, Ling Y, Ma Y, Tao L, Miao Q, Shi Q, et al. Efficacy of early hydroxychloroquine treatment in preventing COVID-19 pneumonia aggravation, the experience from Shanghai, China. Biosci Trends 2021;14(6):408-414 [FREE Full text] [doi: 10.5582/bst.2020.03340] [Medline: 33342929]
35. de Wilde AH, Jochmans D, Posthuma CC, Zevenhoven-Dobbe JC, van Nieuwkoop S, Bestebroer TM, et al. Screening of an FDA-approved compound library identifies four small-molecule inhibitors of middle east respiratory syndrome coronavirus replication in cell culture. Antimicrob Agents Chemother 2014;58(8):4875-4884 [FREE Full text] [doi: 10.1128/AAC.03011-14] [Medline: 24841269]
36. Keyaerts E, Vijgen L, Maes P, Neyts J, Van Ranst M. In vitro inhibition of severe acute respiratory syndrome coronavirus by chloroquine. Biochem Biophys Res Commun 2004;323(1):264-268 [FREE Full text] [doi: 10.1016/j.bbrc.2004.08.085] [Medline: 15351731]
37. Rome BN, Avorn J. Drug evaluation during the COVID-19 pandemic. N Engl J Med 2020;382(24):2282-2284. [doi: 10.1056/NEJMp2009457] [Medline: 32289216]
38. Yu J, Meng S. Impacts of the internet on health inequality and healthcare access: a cross-country study. Front Public Health 2022;10:935608 [FREE Full text] [doi: 10.3389/fpubh.2022.935608] [Medline: 35757602]
39. Ansari JAN, Khan NA. Exploring the role of social media in collaborative learning the new domain of learning. Smart Learn Environ 2020;7(1):9. [doi: 10.1186/s40561-020-00118-7]
40. Liu S, Zaigham GHK, Rashid RM, Bilal A. Social media-based collaborative learning effects on student performance/learner performance with moderating role of academic self-efficacy. Front Psychol 2022;13:903919 [FREE Full text] [doi: 10.3389/fpsyg.2022.903919] [Medline: 35899006]
41. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
42. Bowman S. Impact of electronic health record systems on information integrity: quality and safety implications. Perspect Health Inf Manag 2013;10(Fall):1c [FREE Full text] [Medline: 24159271]
43. Graber ML, Siegal D, Riah H, Johnston D, Kenyon K. Electronic health record-related events in medical malpractice claims. J Patient Saf 2019;15(2):77-85 [FREE Full text] [doi: 10.1097/PTS.0000000000000240] [Medline: 26558652]
44. Twitter. COVID-19 Misinformation. URL: https://transparency.twitter.com/en/reports/covid19.html#2021-jul-dec [accessed 2024-10-10]
45. Marik PE, Khangoora V, Rivera R, Hooper MH, Catravas J. Hydrocortisone, vitamin C, and thiamine for the treatment of severe sepsis and septic shock: a retrospective before-after study. Chest 2017;151(6):1229-1238. [doi: 10.1016/j.chest.2016.11.036] [Medline: 27940189]
46. Lamsal R. Coronavirus (COVID-19) tweets dataset. IEEE Dataport. 2020. URL: https://dx.doi.org/10.21227/781w-ef42 [accessed 2024-10-10]
47. Joint statement by WHO U, UNICEF, UNDP, UNESCO, UNAIDS, Itu, UN Global Pulse, IFRC. Managing the COVID-19 infodemic: promoting healthy behaviours and mitigating the harm from misinformation and disinformation. 2020. URL: https://tinyurl.com/2s2sf5js [accessed 2024-10-10]
48. 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]

49. Bora K, Das D, Barman B, Borah P. Are internet videos useful sources of information during global public health emergencies? A case study of YouTube videos during the 2015-16 Zika virus pandemic. Pathog Glob Health 2018;112(6):320-328 [FREE Full text] [doi: 10.1080/20477724.2018.1507784] [Medline: 30156974]
50. Fung IC, Fu K, Chan C, Chan BSB, Cheung C, Abraham T, et al. Social media's initial reaction to information and misinformation on Ebola, August 2014: facts and rumors. Public Health Rep 2016;131(3):461-473 [FREE Full text] [doi: 10.1177/003335491613100312] [Medline: 27252566]
51. Jin F, Wang W, Zhao L, Dougherty E, Cao Y, Lu C, et al. Misinformation propagation in the age of Twitter. Computer 2014;47(12):90-94. [doi: 10.1109/mc.2014.361]
52. Pathak R, Poudel D, Karmacharya P, Pathak A, Aryal MR, Mahmood M, et al. YouTube as a source of information on Ebola virus disease. N Am J Med Sci 2015;7(7):306-309 [FREE Full text] [doi: 10.4103/1947-2714.161244] [Medline: 26258077]
53. Sell TK, Hosangadi D, Trotochaud M. Misinformation and the US Ebola communication crisis: analyzing the veracity and content of social media messages related to a fear-inducing infectious disease outbreak. BMC Public Health 2020;20(1):550 [FREE Full text] [doi: 10.1186/s12889-020-08697-3] [Medline: 32375715]
54. Seltzer E, Horst-Martz E, Lu M, Merchant R. Public sentiment and discourse about Zika virus on Instagram. Public Health 2017;150:170-175. [doi: 10.1016/j.puhe.2017.07.015] [Medline: 28806618]
55. 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]
56. Wood MJ. Propagating and debunking conspiracy theories on twitter during the 2015-2016 Zika virus outbreak. Cyberpsychol Behav Soc Netw 2018;21(8):485-490 [FREE Full text] [doi: 10.1089/cyber.2017.0669] [Medline: 30020821]
57. Bahk CY, Cumming M, Paushter L, Madoff LC, Thomson A, Brownstein JS. Publicly available online tool facilitates real-time monitoring of vaccine conversations and sentiments. Health Aff (Millwood) 2016;35(2):341-347. [doi: 10.1377/hlthaff.2015.1092] [Medline: 26858390]
58. DeVerna MR, Pierri F, Truong BT, Bollenbacher J, Axelrod D, Loynes N, et al. CoVaxxy: a collection of English-language Twitter posts about COVID-19 vaccines. Proc Int AAAI Conf Web Soc Media 2021;15(1):992-999. [doi: 10.1609/icwsm.v15i1.18122]
59. Donzelli G, Palomba G, Federigi I, Aquino F, Cioni L, Verani M, et al. Misinformation on vaccination: a quantitative analysis of YouTube videos. Hum Vaccin Immunother 2018;14(7):1654-1659 [FREE Full text] [doi: 10.1080/21645515.2018.1454572] [Medline: 29553872]
60. Mahoney LM, Tang T, Ji K, Ulrich-Schad J. The digital distribution of public health news surrounding the human papillomavirus vaccination: a longitudinal infodemiology study. JMIR Public Health Surveill 2015;1(1):e2 [FREE Full text] [doi: 10.2196/publichealth.3310] [Medline: 27227125]
61. Panatto D, Amicizia D, Arata L, Lai PL, Gasparini R. A comprehensive analysis of Italian web pages mentioning squalene-based influenza vaccine adjuvants reveals a high prevalence of misinformation. Hum Vaccin Immunother 2018;14(4):969-977 [FREE Full text] [doi: 10.1080/21645515.2017.1407483] [Medline: 29172967]
62. 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]
63. Al-Rakhami MS, Al-Amri AM. Lies kill, facts save: detecting COVID-19 misinformation in Twitter. IEEE Access 2020;8:155961-155970 [FREE Full text] [doi: 10.1109/ACCESS.2020.3019600] [Medline: 34192115]
64. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
65. Dutta A, Beriwal N, Van Breugel LM, Sachdeva S, Barman B, Saikia H, et al. YouTube as a source of medical and epidemiological information during COVID-19 pandemic: a cross-sectional study of content across six languages around the globe. Cureus 2020;12(6):e8622 [FREE Full text] [doi: 10.7759/cureus.8622] [Medline: 32685293]
66. Li HO, 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]
67. Memon S, Carley K. Characterizing COVID-19 misinformation communities using a novel Twitter dataset. arXiv. Preprint posted online on 2008 2020:1-9.
68. Pulido CM, Villarejo-Carballido B, Redondo-Sama G, Gómez A. COVID-19 infodemic: more retweets for science-based information on coronavirus than for false information. Int Sociol 2020;35(4):377-392.
69. Twitter. Legacy verification policy. 2023. URL: https://help.twitter.com/en/managing-your-account/legacy-verification-policy#categories [accessed 2024-10-10]
70. American Medical Association. Social meda guidance for physicians taps timeless principles. 2018. URL: https://www.ama-assn.org/delivering-care/ethics/social-media-guidance-physicians-taps-timeless-principles [accessed 2024-10-10]
71. Nguyen BM, Lu E, Bhuyan N, Lin K, Sevilla M. Social media for doctors: taking professional and patient engagement to the next level. Fam Pract Manag 2020;27(1):19-14 [FREE Full text] [Medline: 31934735]
72. Choo EK, Ranney ML, Chan TM, Trueger NS, Walsh AE, Tegtmeyer K, et al. Twitter as a tool for communication and knowledge exchange in academic medicine: a guide for skeptics and novices. Med Teach 2015;37(5):411-416. [doi: 10.3109/0142159X.2014.993371] [Medline: 25523012]

73. Data Transfer Agreement Signatories. National Institutes of Health. URL: https://ncats.nih.gov/n3c/resources/data-contribution/data-transfer-agreement-signatories [accessed 2024-10-10]
74. Haendel MA, Chute CG, Bennett TD, Eichmann DA, Guinney J, Kibbe WA, N3C Consortium. The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. J Am Med Inform Assoc 2021 Mar 01;28(3):427-443. [doi: 10.1093/jamia/ocaa196] [Medline: 32805036]
75. N3C Core Contributors. N3C. URL: https://covid.cd2h.org/contributors/ [accessed 2024-10-10]

Abbreviations

MERS-CoV: Middle East respiratory syndrome
N3C: National COVID Cohort Collaborative
PCR: polymerase chain reaction
SARS-CoV-1: severe acute respiratory syndrome coronavirus
VAR: vector auto-regressive

Edited by A Mavragani; submitted 23.01.24; peer-reviewed by K Lin; comments to author 06.09.24; revised version received 09.09.24; accepted 12.09.24; published 18.11.24.

Please cite as:
Helgeson SA, Mudgalkar RM, Jacobs KA, Lee AS, Sanghavi D, Moreno Franco P, Brooks IS, National COVID Cohort Collaborative (N3C)
Association Between X/Twitter and Prescribing Behavior During the COVID-19 Pandemic: Retrospective Ecological Study
JMIR Infodemiology 2024;4:e56675
URL: https://infodemiology.jmir.org/2024/1/e56675
doi:10.2196/56675
PMID:39556417

Viewpoint

Understanding and Combating Misinformation: An Evolutionary Perspective

Nicola Luigi Bragazzi[1,2,3], MPH, MD, PhD; Sergio Garbarino[3], MD, PhD

[1]Division of Human Nutrition Unit, Department of Food and Drugs, University of Parma, Parma, Italy

[2]Laboratory for Industrial and Applied Mathematics, Department of Mathematics and Statistics, York University, Toronto, ON, Canada

[3]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics and Maternal/Child Sciences, University of Genoa, Genoa, Italy

Corresponding Author:
Nicola Luigi Bragazzi, MPH, MD, PhD
Division of Human Nutrition Unit
Department of Food and Drugs
University of Parma
Via Volturno 39
Parma, 43125
Italy
Phone: 39 0521 903121
Email: robertobragazzi@gmail.com

Abstract

Misinformation represents an evolutionary paradox: despite its harmful impact on society, it persists and evolves, thriving in the information-rich environment of the digital age. This paradox challenges the conventional expectation that detrimental entities should diminish over time. The persistence of misinformation, despite advancements in fact-checking and verification tools, suggests that it possesses adaptive qualities that enable it to survive and propagate. This paper explores how misinformation, as a blend of truth and fiction, continues to resonate with audiences. The role of narratives in human history, particularly in the evolution of Homo narrans, underscores the enduring influence of storytelling on cultural and social cohesion. Despite the increasing ability of individuals to verify the accuracy of sources, misinformation remains a significant challenge, often spreading rapidly through digital platforms. Current behavioral research tends to treat misinformation as completely irrational, static, finite entities that can be definitively debunked, overlooking their dynamic and evolving nature. This approach limits our understanding of the behavioral and societal factors driving the transformation of misinformation over time. The persistence of misinformation can be attributed to several factors, including its role in fostering social cohesion, its perceived short-term benefits, and its use in strategic deception. Techniques such as extrapolation, intrapolation, deformation, cherry-picking, and fabrication contribute to the production and spread of misinformation. Understanding these processes and the evolutionary advantages they confer is crucial for developing effective strategies to counter misinformation. By promoting transparency, critical thinking, and accurate information, society can begin to address the root causes of misinformation and create a more resilient information environment.

(JMIR Infodemiology 2024;4:e65521) doi:10.2196/65521

KEYWORDS

misinformation; infodemics; evolutionary theory; fake news; spoof news; fact-checking; digital platform; behavioral research; social cohesion; extrapolation; deformation; fabrication; disinformation; evolutionary paradox; adaptive qualities; strategic deception; intrapolation; health information; public health

The Evolutionary Paradox of Misinformation

Some stories are unbelievable, yet they can still convince people because, "in substance," they are true, even though only some details are real and the rest is false. This is the conclusion of "Emma Zunz," a short story by Argentine writer Jorge Luis Borges (1899-1986) [1], which illustrates how truth and falsehood can blend ("true lies" and "false truths") [2] and how human life is a complex admixture of fact and fiction. "Emma Zunz" demonstrates how narratives, even when not entirely factual or even completely fabricated, can be perceived as "essentially true," being powerful, persuasive, and impactful as long as they resonate with the audience [3].

Narratives have played a key role in the story of *Homo narrans* [4] since the earliest days of civilization. From the ancient myths that sought to explain the mysteries of the universe to the epic poems that celebrated heroes and their extraordinary deeds, storytelling has been an integral part of human culture and communication. These narratives not only served to entertain but also to pass down knowledge, moral lessons, and cultural values from one generation to the next. Through storytelling, societies have preserved their histories, shaped their identities, and connected with one another, demonstrating the enduring power of narrative throughout the ages [5].

More specifically, six functions of stories have been proposed: namely, (1) communication, (2) the formation of both individual and collective identities, (3) the cultivation of empathy and theory-of-mind, (4) the development and transmission of knowledge, including social knowledge and tacit sociocultural understanding, (5) the use of stories as tools for simulation and modeling to inform decision-making, and (6) their role in persuasion, particularly in belief change. Notably, the latter 5 functions highlight the role of stories as constitutive forces, shaping realities and understanding, rather than merely reflecting or representing them (constitutive vs representational or descriptive stories) [6].

However, when not completely accurate, some narratives can be detrimental and harmful. As such, the ability of humans to verify the accuracy of sources has evolved and significantly increased over time, especially with the advent of advanced technology and the widespread availability of information. Tools and methods for fact-checking are more accessible than ever, enabling people to cross-reference data, identify misinformation, and scrutinize the credibility of their sources. Despite these advancements, this enhanced capacity for verification has not resulted in the complete elimination or extinction of false news. In fact, misinformation continues to proliferate, often spreading rapidly through digital platforms [7-9], where it can still mislead and influence large audiences. The persistence of false news highlights the ongoing challenges in distinguishing truth from falsehood in the information age, despite the improved tools at our disposal.

However, current behavioral research tends to overlook the dynamic, evolving nature of misinformation, often treating it as an irrational, static, and finite phenomenon. Scholars commonly assume that misinformation is just the product of irrational reasoning, consisting of discrete, fully formed pieces of inconsistent content that can be definitively debunked and falsified. This conventional approach has the advantage of making misinformation more manageable and easier to study in controlled, experimental settings. On the other hand, this irrational, finite perspective comes with a significant drawback: it hinders the comprehension of why it resists scrutiny, obscuring the dynamics and the evolving characteristics of misinformation, and leaving researchers unaware of the behavioral (and societal) factors that drive its transformation over time. As a result, the intricate processes by which misinformation is fabricated, spreads, adapts, and changes are often ignored, limiting our understanding of its impact and persistence in society [10].

Furthermore, misinformation is (apparently) paradoxical from an evolutionary perspective. This paradox lies in the fact that, despite misinformation being harmful and detrimental to society, it persists and continues to spread. According to evolutionary logic, harmful entities should be eliminated or should diminish over time; however, misinformation defies this expectation by thriving and evolving, even in environments where tools and mechanisms exist to detect and remove them. This persistence suggests that misinformation may possess adaptive qualities that enable it to survive and propagate, despite its negative impact on individuals and communities.

Misinformation, along with its counterparts disinformation, malinformation, and deinformation, represents a significant challenge in the modern digital age. Understanding why these forms of information persist and spread requires an examination of their evolutionary advantages, challenging the common views (and prejudices) of misinformation. By analyzing the distinct types of false information and applying an evolutionary framework, we can develop strategies to combat their influence more effectively.

Misinformation as Units of the Evolutionary Process

As stated by Marchetti and Mastrogiorgio, misinformation “can be considered units of the evolutionary process” [10]. The evolutionary theory, originally formulated by Charles Darwin, has been refined and adapted to explain how cultural information spreads and evolves. In the theory of memes, proposed by Richard Dawkins in 1976 [11], memes are units of cultural transmission or ideas that propagate within a society, analogous to the way genes transmit biological information. Memes can include beliefs, behaviors, symbols, or practices that are passed from one individual to another through communication, imitation, and other forms of social interaction. Much like genes, memes undergo processes of variation, competition, and inheritance, evolving over time based on their ability to replicate and spread. Dawkins’ concept of memes has been influential in fields such as cultural evolution, communication studies, and psychology, offering insights into how cultural phenomena—ranging from language and religion to fashion trends and internet virality—emerge and persist. Adding to the conversation on how information spreads and evolves, Eva Jablonka and Marion Lamb's theory of the “four dimensions of evolution” [12] provides a broader and more nuanced framework for understanding inheritance and evolution, both biological and cultural. Jablonka and Lamb argue that evolution operates on multiple levels, not just through genetic changes but also through other forms of inheritance. They propose 4 dimensions of evolutionary processes: genetic inheritance, epigenetic inheritance, behavioral inheritance, and symbolic inheritance.

Genetic inheritance is the traditional form of inheritance, where traits are passed down through DNA. Beyond DNA, epigenetic factors such as chemical modifications to DNA and histones can influence gene expression and can sometimes be passed on to subsequent generations. This form of inheritance allows for environmental factors to affect evolutionary outcomes, offering a more dynamic view of how organisms adapt and evolve. Behavioral inheritance is a dimension including behaviors that are learned and transmitted across generations (for example, animals teaching their offspring how to hunt or humans passing down cultural practices). Behavioral inheritance is crucial in the context of cultural evolution, where learned behaviors can have significant adaptive value. The fourth dimension, symbolic inheritance, refers to the transmission of information through symbols, language, and other forms of communication unique to humans. Symbolic inheritance encompasses the spread of memes and is essential for understanding how complex cultural systems—such as religious beliefs, scientific knowledge, and social norms—develop and persist.

By integrating these 4 dimensions, Jablonka and Lamb's theory offers a comprehensive view of how both biological and cultural evolution are driven by a variety of mechanisms. It also highlights the interconnectedness of these processes, where cultural practices can influence genetic evolution, and vice versa.

In the context of misinformation and cultural transmission, both meme theory and the 4 dimensions of evolution can provide valuable insights. Misinformation can be understood as a type of meme that evolves within the symbolic inheritance system, spreading through communication and social media. Their persistence and adaptation over time can be analyzed through the lens of behavioral inheritance, as individuals and groups learn to create, share, and reinforce these narratives. The interaction between genetic, epigenetic, behavioral, and symbolic dimensions could also explain why some individuals are more prone to misinformation and why certain types of misinformation are more resilient and pervasive than others, pointing to the complex interplay between different forms of inheritance, evolution, and behaviors.

The Distinction Among Misinformation, Disinformation, Malinformation, and Deinformation

According to the literature, the various forms of misleading information can be categorized as follows: deinformation, malinformation, misinformation, and disinformation [10,13,14]. Based on the "ABCDE framework," they vary according to the actors disseminating misinformation (A), their intentions and behaviors (B), the content disseminated and its context (C), the degree of its distribution in terms of audience reached (D), and the effect or impact (harm or threat) caused or potential (E) [15-17].

Deinformation is a less commonly discussed category where specific content is true but the overall message becomes unintentionally false due to a lack of context or competence by the producer. It can be seen as an unintentional form of sensationalism. Malinformation involves the use of true information with harmful intent. While the facts themselves are accurate, they are presented or highlighted in a way designed to cause harm, often through mechanisms like cherry-picking or sensationalism. An example of malinformation is doxing, which consists in the public disclosing and sharing of private information. Misinformation refers to false information that is created without harmful intent. It is often spread by individuals who believe the information to be true, making it a "benign" form that resembles traditional news in its presentation. Disinformation is false information deliberately created with the intent to deceive or cause harm. This is the most insidious type, as it is crafted with malicious purposes, such as manipulating public opinion or discrediting individuals or groups.

These categories help differentiate the various ways in which information can be manipulated or fabricated. Recognizing these distinctions is crucial for developing tailored strategies to counter each type effectively.

Information Along the Truth-Falsehood Continuum

The "deinformation / malinformation / misinformation / disinformation" categories can be thought of along a truth-falsehood *continuum*, where each type of information occupies a different position based on its degree of truthfulness, accuracy, and the intent behind its creation and dissemination.

At one end of the *spectrum*, we have deinformation, which starts with true content but becomes misleading due to a lack of context or competence, resulting in an unintentionally false overall message. This reflects low intent to deceive but still carries a risk of causing misunderstanding. Moving further along the *continuum*, malinformation represents accurate information used with malicious intent. Here, the truthfulness of the content is high, but the intent to harm, deceive, or manipulate is also significant, often leading to negative consequences despite the factual nature of the information. Misinformation, positioned closer to the middle of the *continuum*, involves false information shared without the intent to deceive. The accuracy of the content is low, but the intent behind its dissemination is typically benign, as the individuals spreading it usually believe it to be true. At the opposite end of the *spectrum* lies disinformation, which is entirely false and created with a clear intent to deceive, manipulate, or harm. This represents the most dangerous category, combining low accuracy with high malicious intent, resulting in significant potential for harmful impact.

Together, these categories illustrate how information can vary not only in terms of its factual correctness but also in the motivations behind its creation and the potential harm it can cause. Understanding where a piece of information falls on this *continuum* can help in assessing its reliability and the potential risks associated with its dissemination (Table 1).

Table 1. Information and techniques used to produce and deliver information along the "truth-falsehood continuum," categorized by accuracy, intent, and potential harm.

Category and technique	Description	Accuracy	Intent	Potential harm
Deinformation				
Extrapolation	Extending true data points beyond their original context to draw broader, often misleading, conclusions	High (initial facts)	Low	Moderate (misleading message)
Intrapolation	Inserting false or misleading details within a true context, distorting the original message	Mixed (true context, false details)	Low	Moderate
Malinformation				
Cherry-picking	Selecting specific pieces of true information to support a viewpoint while ignoring contradictory evidence	High (selected facts)	High	High (intentional omission)
Contextualization	Placing true information within a misleading context to alter its interpretation	High (facts)	High	High (manipulative context)
Misrepresentation	Presenting information in a way that unintentionally misleads, often by altering context or tone	Low (false presentation)	Low	Moderate (unintentional)
Disinformation				
Deformation	Altering facts or details to fit a specific narrative, often by twisting or exaggerating the truth	Low (twisted facts)	High	Very high (intentional)
Fabrication	Creating completely false information or events with no basis in reality	None (false information)	High	Very high (intentional)
Echo chamber amplification	Repeating or amplifying false or misleading information within a closed group, reinforcing the false narrative	None or low (false information)	High	Very high (reinforced belief)

The Role of Falsehood in Contributing to Knowledge

Paradoxically, falsehood can contribute to knowledge. As stated by Bernecker [18], "falsehood plays an important role in the inference-based production of knowledge" by serving as a catalyst for clarifying concepts, strengthening arguments, and enhancing understanding. When engaging with falsehoods, thinkers often test assumptions and explore alternative possibilities that might initially seem plausible. This process helps refine and redefine the boundaries of true knowledge.

By critically examining and refuting false ideas, the reasoning process is sharpened, and the credibility of true knowledge is bolstered. In philosophical traditions, particularly in the dialectical method, falsehoods are often introduced deliberately to stimulate discussion and challenge existing beliefs. This confrontation between opposing ideas leads to a synthesis that represents a more refined understanding. In education, falsehoods are also used as a pedagogical tool to provoke critical thinking. When students are challenged with false statements, they must apply their knowledge to refute them, which reinforces their learning and deepens their comprehension.

The role of falsehood in scientific progress is particularly notable in the context of Karl Popper's falsifiability (or refutability) principle [19]. Scientific knowledge advances through conjectures and refutations, where theories must be falsifiable. The process of attempting to falsify a theory, and encountering falsehoods along the way, is crucial for scientific advancement. When a theory withstands these challenges, our confidence in its truth increases.

The discovery and correction of errors within scientific theories also contribute to the self-correcting nature of science, driving it closer to the truth. In hypothetical reasoning, falsehoods are used in counterfactual scenarios, which, although not true, provide valuable insights into causal relationships and decision-making processes. These scenarios help in understanding the underlying principles of the subject matter. Furthermore, falsehoods help identify exceptions to general rules and clarify the limitations of existing knowledge. By recognizing where and why certain inferences do not hold, theories can be refined to accommodate a broader range of phenomena.

In sum, falsehoods are essential in the production of knowledge through inference. They challenge existing beliefs, prompt critical examination, and drive the process of inquiry and discovery, ultimately leading to a more robust and nuanced understanding of the world.

Falsehood as Non-(Completely) Irrational

False beliefs and beliefs in conspiracy theories are generally considered inherently irrational due to their resistance to disconfirming evidence and internal incoherence. However, Poth and Dolega [20] argued that such beliefs can be considered rational within a probabilistic framework. More in detail, belief in conspiracy theories can be rational if it is supported by a network of auxiliary beliefs that protect the core belief from disconfirmation. Within this network of beliefs, auxiliary hypotheses are adjusted to protect a core hypothesis from falsification. Central conspiracy beliefs can be, indeed, preserved by revising or rejecting auxiliary beliefs, which may be more easily discarded when faced with disconfirming evidence. This process aligns with Bayesian norms of rationality, where the updating of beliefs depends on prior probabilities and the likelihood of new evidence. The monological nature of conspiracy beliefs (where beliefs support each other in a self-sustaining network) can be reconciled with their apparent insensitivity to counterevidence by viewing such beliefs as different aspects of the same phenomenon within a Bayesian framework. Inductive biases have a role in the formation and maintenance of conspiracy beliefs, as they influence how individuals weigh the evidence and update their beliefs, potentially leading to the adoption of conspiracy theories even in the face of contradictory evidence. Finally, a distinction should be made between rationality and desperation, between “glorious rescues” of beliefs (where adjustments lead to new, confirmable predictions) and “desperate rescues” (where auxiliary beliefs are adjusted without sufficient justification). The latter is considered a hallmark of irrational belief maintenance. In conclusion, belief in conspiracy theories can be rational in certain contexts, especially when viewed through a Bayesian lens. However, the rationality of these beliefs depends on the broader context, including the strength of prior beliefs and the nature of the evidence encountered. The irrationality of some conspiracy beliefs may stem more from desperate attempts to protect poorly confirmed hypotheses rather than a fundamental flaw in the belief-updating process itself.

Techniques of Misinformation Production

Several techniques can be employed in the production of misinformation, including (1) extrapolation, (2) intrapolation, (3) deformation, (4) cherry-picking, (5) misrepresentation, (6) contextualization, (7) fabrication, and (8) echo chamber amplification. Each of these methods plays a distinct role in distorting, manipulating, and fabricating information to create misleading narratives.

Extrapolation involves taking specific facts or data points and extending them beyond their original context to draw broader, often misleading, conclusions. Intrapolation refers to manipulating information within its context, often by inserting false or misleading details that distort the original message while retaining a semblance of truth. Deformation is altering facts or details to create a narrative that fits specific purposes, such as a particular agenda, often by twisting or exaggerating the truth. Cherry-picking refers to selecting only specific pieces of information that support a particular viewpoint or narrative while ignoring contradictory evidence. Misrepresentation involves presenting information in a way that intentionally misleads, often by changing the context, tone, or framing of the content. Contextualization refers to placing true information within a misleading context to alter its interpretation, often to provoke a particular reaction or belief. Fabrication is the creation of completely false information or events that are presented as truth, with no basis in reality. Finally, echo chamber amplification involves repeating or amplifying false or misleading information within a closed or like-minded group, reinforcing the narrative and making it seem more credible.

Becoming True and Becoming False Versus Being True and Being False

True and false news are not merely static collections of facts; they are dynamic processes [10]. The information we encounter often goes through various stages of verification, interpretation, and dissemination, which can affect its status as true or false. True news is not simply about reporting facts accurately. It involves a rigorous process of gathering, verifying, and contextualizing information. This process includes cross-checking sources, analyzing data, and presenting the information in a way that aligns with objective reality. As new information becomes available, what was once thought to be true can be refined or even corrected, highlighting the ongoing nature of truth. Similarly, false news is not just the presentation of incorrect information; it often involves deliberate or unintentional processes that distort or misrepresent facts. This could include selective reporting, biased framing, or the spread of unverified information. Over time, as this false information circulates, it can be amplified and accepted by some as true, further complicating the distinction between truth and falsehood.

News, whether true or false, is often in a state of “becoming.” A story may begin as a rumor, undergo various degrees of scrutiny, and eventually be confirmed as true or debunked as false. The “becoming” aspect emphasizes that the truthfulness of news is not always immediate or self-evident. It is something that can evolve, depending on how the information is handled, interpreted, and understood by both journalists and the public. In this way, true and false news are processes that involve ongoing evaluation and re-evaluation, making them dynamic rather than fixed entities.

From False Claims and Statements to False Narratives, Conspiracy Theories, and False Belief Systems

The journey from false claims and statements to the development of false narratives, conspiracy theories [21,22], and false belief systems [23] is a complex and often insidious process. It begins with a false claim, which is simply an assertion that is factually incorrect. These false claims can be made either deliberately, as lies, or accidentally, as misinformation. If these claims go unchecked, they can spread, leading people to accept these falsehoods as truth. When multiple false claims are strung together, they can form what is known as a false narrative. A false narrative is more than just a collection of incorrect statements; it is a story or explanation that weaves these claims into a broader context that appears coherent and credible. These narratives often simplify complex situations or manipulate facts to fit a particular agenda, making them persuasive to those who encounter them. The impact of false narratives can be significant, influencing public opinion, driving decision-making, and justifying actions based on incorrect or incomplete information. As false narratives gain traction, they can evolve into conspiracy theories. A conspiracy theory is a specific type of false narrative that suggests a secret, often sinister, plot by a group of people or organizations. These theories are typically based on speculation and lack solid evidence, yet they thrive on the distrust of official explanations or the desire to explain complex phenomena in emotionally charged, simplistic ways. The danger of conspiracy theories lies in their ability to challenge established facts, foster distrust in institutions, create social division or even incite violence. Over time, these narratives and conspiracy theories can harden into false belief systems. A false belief system is a cohesive set of beliefs that are based on incorrect information or flawed reasoning. These systems can be religious, political, or ideological in nature and are often deeply ingrained in the individuals or groups that hold them. As people become increasingly committed to these beliefs, they become resistant to contrary evidence, making these belief systems difficult to challenge. The impact of false belief systems is profound, as they can shape individual and collective identities, influence behavior, and perpetuate misinformation across generations. The progression from false claims to false belief systems often follows a recognizable path: false claims give rise to false narratives, which can then develop into conspiracy theories, and ultimately solidify into false belief systems. This progression can be seen in various real-world examples, such as the Flat Earth theory, which began as a false claim, evolved into a narrative questioning mainstream science, became a conspiracy theory involving governments and scientists, and now exists as a false belief system with a community of believers. Similarly, the antivaccine movement started with false claims about the dangers of vaccines, developed into a narrative suggesting vaccines are part of a harmful plot, led to conspiracy theories about government and pharmaceutical companies, and has now become a widespread false belief system influencing public health [24].

Evolutionary Advantages of Misinformation

From an evolutionary perspective, the prevalence of misinformation over correct information can be understood through the lens of how human cognition and social behavior have evolved: (1) survival value of heuristics, (2) social cohesion and group identity, (3) emotional resonance, (4) information overload and cognitive limits, (5) status and influence, and (6) mimicry and the spread of beliefs.

Throughout human evolution, our ancestors relied on mental shortcuts, or heuristics, to make quick decisions in uncertain environments. These heuristics were often based on incomplete or ambiguous information but were crucial for survival. For example, reacting quickly to a potential threat based on limited information (eg, rustling in the bushes might be a predator) was more important than verifying the accuracy of the threat. This predisposition to favor quick, heuristic-based decisions over slow, analytical reasoning persists in modern humans, making us susceptible to misinformation that triggers these quick judgments.

Humans are inherently social animals, and our evolutionary success has been closely tied to our ability to form cohesive groups. Sharing and believing in the same information (even if incorrect) can strengthen group identity and solidarity. Misinformation that aligns with group norms or beliefs may be more readily accepted because it reinforces social bonds. This can be especially powerful in times of uncertainty or conflict, where group cohesion is critical for survival.

Evolutionarily, emotions have played a critical role in decision-making. Information that elicits strong emotions—fear, anger, or excitement—can trigger faster and more decisive actions, which would have been advantageous in survival contexts. Misinformation often spreads because it is emotionally charged, capturing attention and encouraging rapid dissemination, much like how a warning about a potential danger would have spread quickly in early human communities.

The human brain evolved to process a manageable amount of information in relatively stable environments. In today's world; however, we are bombarded with vast amounts of information daily. Evolutionarily, we are not well-equipped to handle this overload, leading to reliance on simple narratives, even if they are inaccurate. Misinformation often provides these simple, digestible narratives that our brains prefer, especially under conditions of information overload. Moreover, storytelling has been a fundamental way humans have passed down knowledge and information through generations. Evolutionarily, stories that were memorable, engaging, and had clear moral lessons were more likely to be retained and shared. Misinformation often comes in the form of compelling stories that are easier to remember and share, even if they are not true.

In ancestral human societies, individuals who were able to provide information (whether true or not) that influenced group decisions could gain status and power within the group. This dynamic still exists today, where spreading sensational or shocking information can elevate a person's status, especially on social media. The desire for social status can drive the spread of misinformation as individuals seek to gain attention and influence.

From an evolutionary standpoint, humans have developed a tendency to imitate the behavior and beliefs of those around them, especially individuals who are perceived as successful or influential. This mimicry would have been beneficial in many survival contexts but can also lead to the spread of misinformation when influential figures or a majority within a group propagate false information.

In summary, from an evolutionary perspective, the factors that once helped our ancestors survive—such as quick decision-making, social cohesion, emotional responsiveness, and storytelling—can also make modern humans more susceptible to misinformation. These deep-rooted tendencies, shaped by millennia of evolution, create fertile ground for misinformation to thrive in contemporary society.

Combating Misinformation

Appreciating the dynamics and evolutionary trajectories of misinformation is key to combating misinformation.

As previously said, misinformation can play a key role in fostering social cohesion and reinforcing group identity. Shared beliefs, even if inaccurate, can unite communities, creating a strong sense of belonging. This dynamic is evident in how myths, conspiracy theories, and certain forms of disinformation become rallying points for groups, strengthening their internal bonds while distinguishing them from outsiders. To counteract this, it is essential to offer alternative narratives that promote unity without relying on falsehoods. Encouraging communities to embrace values grounded in truth can weaken the social cohesion that misinformation creates. Moreover, fostering inclusivity within communities can reduce the appeal of misinformation-driven identities, promoting dialogue and understanding across social divides.

From an evolutionary standpoint, misinformation may sometimes appear to offer survival or reproductive benefits. For instance, deceptive signals or behaviors can deter competitors or attract mates, providing immediate, though misleading, advantages. However, the long-term risks associated with misinformation—such as poor decision-making or health consequences—often outweigh these short-term gains. By highlighting these risks, individuals can be made aware of the dangers inherent in clinging to false beliefs. Education on the value of accurate information, particularly in matters of health and safety, can help individuals recognize the limitations and potential harms of relying on misinformation.

Misinformation is frequently employed as a tool for strategic deception in various contexts, including social, political, and economic arenas. To mitigate this, promoting transparency and accountability is critical. If the benefits of deception are reduced, misinformation becomes a less attractive option. Strengthening the role of fact-checking organizations and technologies that can quickly identify and correct false information can also diminish the effectiveness of misinformation. When individuals and organizations face consequences for spreading misinformation, the incentive to engage in such practices diminishes.

In complex and uncertain environments, misinformation often provides simplified narratives that help individuals make quick decisions, even if those decisions are based on falsehoods. Combating this requires the provision of clear, concise, and easily digestible, accurate information. Simplified yet truthful narratives can compete with misinformation by reducing the cognitive burden required to understand complex topics. Additionally, ensuring that accurate information is available and disseminated quickly, particularly during crises, limits the space for misinformation to take hold. The rapid presentation of truth can preempt the spread of falsehoods that thrive in the vacuum of uncertainty.

Certain forms of misinformation persist because they can indirectly reinforce socially desirable behaviors, even if the factual basis is incorrect. These false beliefs may be rooted in myths, taboos, or cultural narratives that promote actions beneficial to the community. For example, myths about the dangers of overusing certain natural resources might discourage exploitation, thereby preserving the environment. Similarly, health-related taboos, even when scientifically unfounded, may encourage behaviors that reduce the spread of diseases or promote social harmony. In this way, misinformation can be sustained because it serves a protective or stabilizing function for the group, even when the facts themselves are erroneous. Combatting this requires redirecting these motivations toward truthful narratives that also promote social good. By framing accurate information in a way that appeals to prosocial motivations, communities can maintain the positive functions of these beliefs without relying on misinformation. Creating new cultural narratives grounded in factual information can serve as effective replacements for false beliefs, preserving their social benefits while ensuring they are based on reality.

Misinformation can sometimes act as a catalyst for exploration and innovation by sparking curiosity about the unknown. When people encounter mysterious or unverified claims, it can inspire them to investigate further, seeking answers that might reveal new insights or technological advancements. For instance, historical misconceptions about the natural world have occasionally prompted scientific inquiry, leading to groundbreaking discoveries. However, for this curiosity to be productive, it must be guided by critical thinking and skepticism. Without these tools, the exploration may result in reinforcing false beliefs rather than uncovering truths, making it crucial to question and verify the information during the investigative process. As such, fostering a culture of critical thinking, where individuals are encouraged to question and verify information before accepting it, is crucial in combating misinformation. Moreover, education systems should be adaptable, teaching students how to think critically about information sources and adjust their understanding as new, accurate information becomes available. This approach helps counter the flexibility of misinformation, ensuring that innovation and exploration are informed by facts.

The Taxonomy of Strategies to Tackle Misinformation

There exist several strategies to tackle and combat misinformation [25-28]. These various techniques can be broadly categorized into informational, educational, cognitive, social and community-based, technology-driven, and institutional approaches. Moreover, they can be classified into proactive versus reactive and user-centered versus producer-centered countermisinformation interventions.

Cognitive approaches focus on preparing individuals mentally to resist misinformation. Cognitive inoculation [29], for example, exposes people to weakened forms of misinformation, helping them build resistance when they encounter stronger versions. Similarly, cognitive reflection tasks encourage critical thinking, prompting individuals to reflect before accepting information at face value [30].

Educational approaches aim to empower individuals with the skills and knowledge needed to recognize and reject misinformation. Media and social media literacy education teaches critical evaluation of information sources through classroom-based coursework, short web-based videos, or games, enabling people (especially adolescents and young adults) to identify misinformation [31,32]. Debunking involves directly correcting false claims with evidence-based information, while prebunking preemptively exposes individuals to misinformation tactics, equipping them to spot and resist these tactics in the future [33]. Gamification and interactive tools also fall into this category, using engaging, game-like elements to teach users how to recognize and combat misinformation [34].

Social and community-based approaches leverage the influence of social norms and trusted voices within communities. Social norms messaging works by communicating what behaviors or beliefs are typical or acceptable within a group, encouraging individuals to reject misinformation [35]. Community engagement and peer influence involve enlisting trusted community members to spread accurate information and counter misinformation, relying on the power of peer influence [36].

Technology-driven approaches use digital tools and algorithms to combat misinformation. Algorithmic interventions, for instance, detect and limit the spread of false information on online platforms by demoting or flagging misleading content [37].

When considering the timing and nature of these interventions, they can be further divided into proactive and reactive approaches. Proactive interventions aim to prepare individuals before they encounter misinformation. Cognitive inoculation and prebunking are examples of proactive strategies that build resistance to misinformation in advance. Media literacy education and gamification also serve as proactive tools by equipping individuals with the knowledge and skills to recognize and resist misinformation from the outset.

On the other hand, reactive interventions respond to misinformation after it has been encountered. Fact-checking verifies specific claims and provides corrections after misinformation has been spread, while debunking directly addresses and refutes false information. Algorithmic interventions react to misinformation in real-time by limiting its spread on digital platforms. Narrative correction provides an alternative, truthful narrative to counter misinformation that has already been disseminated.

This classification highlights the diverse strategies employed to combat misinformation, encompassing individual cognitive strategies, community engagement, educational efforts, and technological solutions, and distinguishing between proactive prevention and reactive correction (Table 2).

Table 2. Strategies for combating misinformation.

Technique	Description	Application
Cognitive inoculation	Exposing individuals to weakened forms of misinformation to build resistance	Used in public health campaigns to build resistance against persuasive misinformation
Fact-checking	Verifying specific claims and correcting inaccuracies in public discourse	Employed by organizations or fact-checking sites and reference sources to assess the veracity of claims
Media literacy education	Teaching critical evaluation of information sources to recognize misinformation	Implemented in schools, universities, and online platforms to teach media literacy
Debunking	Directly refuting false claims with evidence-based corrections	Common in journalism and science communication to correct false claims
Prebunking	Exposing misinformation tactics to help individuals spot and resist false information	Used by educational campaigns to inform the public about misinformation tactics
Social norms messaging	Communicating what behaviors or beliefs are typical to encourage rejection of misinformation	Applied in public health to encourage following scientific advice by showing majority behavior
Algorithmic interventions	Using algorithms to detect and limit the spread of misinformation online	Implemented by platforms like Facebook, Twitter, and Google to reduce visibility of false content
Community engagement and peer influence	Engaging communities and leveraging trusted voices to counter misinformation	Used in public health and community campaigns to spread accurate information
Narrative correction	Countering misinformation by telling a compelling and truthful story	Used in documentaries and social media to correct widespread myths
Cognitive reflection tasks	Prompting individuals to think critically about information before accepting it	Integrated into educational tools and quizzes that encourage critical thinking
Gamification and interactive tools	Using game-like elements to teach users how to recognize and resist misinformation	Examples include online games that teach users about misinformation tactics

Tackling Misinformation in a Dynamically Evolving Landscape

As stated by Bateman and Jackson [25], “there is no silver bullet or “best” policy option.” Indeed, none of the interventions previously mentioned are both well-researched, highly effective, and easy to scale at the same time. Instead, the effectiveness of most interventions appears to be quite uncertain and likely hinges on a variety of factors that researchers have only just started to explore.

To effectively implement countermisinformation strategies in a dynamic landscape where false news and beliefs are continuously adapting, a multifaceted approach that evolves alongside the misinformation itself is essential.

One critical aspect is the development of adaptive and continuous learning systems. Technology-driven strategies should include algorithms capable of learning from new patterns of misinformation. By updating machine learning models continuously, these algorithms can detect emerging forms of misinformation through the analysis of content trends, linguistic shifts, and changes in misinformation sources. Platforms can then use these adaptive algorithms not only to flag or demote misleading content but also to adjust their detection criteria as misinformation evolves. Feedback loops are crucial in this context; interventions must be regularly assessed for their effectiveness. If a particular fact-checking or debunking strategy proves less effective over time, adjustments can be made based on user feedback, ensuring that the response to misinformation remains robust. The synergy between proactive and reactive approaches also plays a vital role. Proactive strategies, such as cognitive inoculation and prebunking, need to be regularly updated with the latest misinformation tactics and narratives. This ensures that individuals are prepared to recognize the most recent threats.

On the reactive side, interventions must be agile and quick to deploy. Timely fact-checking or narrative correction should match the speed at which misinformation spreads. This necessitates close collaboration between fact-checkers, social media platforms, and news outlets to ensure a rapid response. Community-driven and peer-influence models offer another layer of defense against misinformation. Engaging community leaders and influencers who can quickly adapt their messaging to counter new forms of misinformation is crucial. These trusted voices can amplify accurate information and address specific community concerns as they arise. Peer-led initiatives can further enhance this effort by encouraging individuals within communities to actively participate in identifying and countering misinformation. Establishing peer-to-peer fact-checking networks can help people share verified information and challenge false claims within their social circles. Continuous education and literacy development are essential components of this adaptive strategy. Media and social media literacy programs should be dynamic, incorporating new case studies and evolving threats. Regular updates to educational content ensure that individuals stay informed about the latest misinformation tactics and how to resist them. Gamified learning tools should also be kept current, with scenarios that reflect the most recent forms of misinformation, maintaining their relevance and effectiveness. A cross-platform and multichannel approach is necessary to address the spread of misinformation across different platforms. Since misinformation spreads differently on various platforms, interventions must be tailored to the specific characteristics of each one. For instance, the tactics required to counter misinformation on Twitter may differ from those on TikTok.

Additionally, a multi-channel communication strategy that includes social media, traditional media, and direct community engagement ensures that accurate information reaches a wide audience, regardless of their preferred information sources. Building resilience through cognitive approaches is another key strategy. Encouraging regular cognitive exercises, such as reflection tasks, helps individuals maintain and enhance their resilience to misinformation over time. Exposure to a wide range of information sources also fosters cognitive flexibility, which is crucial for resisting misinformation. Institutional and policy support further strengthens these efforts. Regulatory frameworks should be flexible and capable of adapting as misinformation evolves. Governments and institutions must collaborate with researchers and technology companies to refine these regulations in response to new challenges. Public-private partnerships can also play a significant role by facilitating the sharing of data, insights, and resources, leading to more coordinated and effective responses to evolving misinformation.

Ongoing research and development are indispensable for staying ahead of misinformation. Continuous research into the dynamics of misinformation—how it evolves, spreads, and impacts different populations—can inform the development of new strategies. Piloting new interventions in controlled environments before scaling them up ensures that these strategies are both effective and scalable.

By integrating these approaches, counter-misinformation efforts can remain effective in a rapidly changing information environment, ensuring that interventions are both proactive and reactive and continuously adapt to the evolving nature of false news and beliefs.

Conclusions

Understanding the dynamics and the evolutionary advantages of misinformation provides valuable insights into how it can be effectively countered. By recognizing the distinct types of misleading information—misinformation, disinformation, malinformation, and deinformation—as non-completely irrational, dynamically evolving entities, we can tailor strategies to address each form. Promoting transparency, simplifying accurate information, replacing false beliefs with truthful narratives, and encouraging critical thinking are essential components of this effort. By addressing the root causes of misinformation and offering viable alternatives, we can create a more resilient information environment where truth prevails over falsehood.

Conflicts of Interest

None declared.

References

1. Borges JL. Emma Zunz. Aleph. URL: https://en.wikipedia.org/wiki/Emma_Zunz [accessed 2024-11-16]
2. Porinsky R. True lies: metaphysical games in Borges' Emma Zunz. WELS and eLS Undergraduate Research Symposium 2002.
3. Priel B. Negative capability and truth in Borges's 'Emma Zunz'. Int J Psychoanal 2004;85(Pt 4):935-949. [doi: 10.1516/0020757041557601] [Medline: 15310429]
4. Lehmann A. In: Brednich RW, editor. 'Homo narrans: Individuelle und kollektive Dimensionen des Erzählens', in Erzählkultur: Beiträge zur kulturwissenschaftlichen Erzählforschung: Hans-Jörg Uther zum 65. Geburtstag. Berlin: de Gruyter; 2009.
5. László J. The Science of Stories An Introduction to Narrative Psychology. UK: Routledge; 2008.
6. Dillon S, Craig C. Storylistening Narrative Evidence and Public Reasoning. UK: Taylor & Francis; 2021.
7. Adams Z, Osman M, Bechlivanidis C, Meder B. (Why) is misinformation a problem? Perspect Psychol Sci 2023;18(6):1436-1463 [FREE Full text] [doi: 10.1177/17456916221141344] [Medline: 36795592]
8. Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, et al. The spreading of misinformation online. Proc Natl Acad Sci USA 2016;113(3):554-559 [FREE Full text] [doi: 10.1073/pnas.1517441113] [Medline: 26729863]
9. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
10. Marchetti J, Mastrogiorgio A. Becoming Fake: An Evolutionary-Behavioral Framework on Fake News. 2023. URL: https://ssrn.com/abstract=4340222 [accessed 2023-01-27]
11. Dawkins R. The Selfish Gene: 40th Anniversary Edition. Oxford, England: Oxford University Press; 2016.
12. Jablonka E, Lamb MJ. Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral and Symbolic Variation in the History of Life. Cambridge, MA: MIT Press; 2005:262.
13. Wardle C, Derakhshan H. Thinking about information disorder: formats of misinformation, disinformation, and mal-information. In: Ireton C, Posetti J, editors. Journalism, Fake News & Disinformation (Paris: Unesco). London: Media Defence; 2018.
14. Balkan E, Ülgen S. A primer on misinformation, malinformation and disinformation. EDAM. URL: https://edam.org.tr/en/cyber-governance-digital-democracy/a-primer-on-misinformation-malinformation-and-disinformation [accessed 2024-11-16]
15. François C. Transatlantic Working Group. URL: https://science.house.gov/imo/media/doc/Francois%20 [accessed 2019-09-20]
16. Pamment J. The EU's role in the fight against disinformation: developing policy interventions for the 2020s. Carnegie Endowment for International Peace. URL: https://carnegieendowment.org/2020/09/30/eu-s-role-in-fight-againstdisinformation-developing-policy-interventions-for-2020s-pub-82821 [accessed 2020-09-30]
17. Hameleers M. Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. Commun Theory 2023:1-10 [FREE Full text] [doi: 10.1093/ct/qtac021]
18. Bernecker S. Knowledge from falsehood and truth-closeness. Philosophia 2022;50:1623-1638 [FREE Full text] [doi: 10.1007/s11406-022-00479-y]
19. Popper K. Logik der Forschung. Vienna, Austria: Verlag von Julius Springer; 1935.
20. Poth N, Dolega K. Bayesian belief protection: a study of belief in conspiracy theories. Philos Psychol 2023;36(6):1182-1207 [FREE Full text] [doi: 10.1080/09515089.2023.2168881]
21. 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 [FREE Full text] [doi: 10.1371/journal.pone.0118093] [Medline: 25706981]
22. van Prooijen J, Douglas KM. Belief in conspiracy theories: basic principles of an emerging research domain. Eur J Soc Psychol 2018;48(7):897-908 [FREE Full text] [doi: 10.1002/ejsp.2530] [Medline: 30555188]
23. Botvinik-Nezer R, Jones M, Wager TD. A belief systems analysis of fraud beliefs following the 2020 US election. Nat Hum Behav 2023;7(7):1106-1119. [doi: 10.1038/s41562-023-01570-4] [Medline: 37037989]
24. 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]
25. Bateman J, Jackson D. Countering disinformation effectively. An evidence-based policy guide. Carnegie Endowment for International Peace. 2024. URL: https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en [accessed 2024-11-16]
26. Blair RA, Gottlieb J, Nyhan B, Paler L, Argote P, Stainfield CJ. Interventions to counter misinformation: lessons from the global north and applications to the global south. Curr Opin Psychol 2024;55:101732. [doi: 10.1016/j.copsyc.2023.101732] [Medline: 38070207]
27. Hartwig K, Doell F, Reuter C. The landscape of user-centered misinformation interventions - a systematic literature review. ACM Comput Surv 2024;56(11):1-36. [doi: 10.1145/3674724]

28. Kozyreva A, Lorenz-Spreen P, Herzog SM, Ecker UKH, Lewandowsky S, Hertwig R, et al. Toolbox of individual-level interventions against online misinformation. Nat Hum Behav 2024;8(6):1044-1052. [doi: 10.1038/s41562-024-01881-0] [Medline: 38740990]
29. Lu C, Hu B, Li Q, Bi C, Ju XD. Psychological inoculation for credibility assessment, sharing intention, and discernment of misinformation: systematic review and meta-analysis. J Med Internet Res 2023;25:e49255 [FREE Full text] [doi: 10.2196/49255] [Medline: 37560816]
30. Orosz G, Faragó L, Paskuj B, Krekó P. Strategies to combat misinformation: enduring effects of a 15-minute online intervention on critical-thinking adolescents. Comput Hum Behav 2024;159:108338. [doi: 10.1016/j.chb.2024.108338]
31. Guess AM, Lerner M, Lyons B, Montgomery JM, Nyhan B, Reifler J, et al. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. Proc Natl Acad Sci USA 2020;117(27):15536-15545 [FREE Full text] [doi: 10.1073/pnas.1920498117] [Medline: 32571950]
32. Mason LE, Krutka D. Media literacy, democracy, and the challenge of fake news. J Media Lit Educ 2018;10(2):1-10 [FREE Full text] [doi: 10.23860/jmle-2018-10-2-1]
33. Schmid P, Betsch C. Benefits and pitfalls of debunking interventions to counter mRNA vaccination misinformation during the COVID-19 pandemic. Sci Commun 2022;44(5):531-558 [FREE Full text] [doi: 10.1177/10755470221129608] [Medline: 38603361]
34. Kiili K, Siuko J, Ninaus M. Tackling misinformation with games: a systematic literature review. Interact Learn Environ 2024:1-16 [FREE Full text] [doi: 10.1080/10494820.2023.2299999]
35. Prike T, Butler LH, Ecker UKH. Source-credibility information and social norms improve truth discernment and reduce engagement with misinformation online. Sci Rep 2024;14(1):6900 [FREE Full text] [doi: 10.1038/s41598-024-57560-7] [Medline: 38519569]
36. Pareek S, Goncalves J. Peer-supplied credibility labels as an online misinformation intervention. Int J Hum Comput Stud 2024;188:103276. [doi: 10.1016/j.ijhcs.2024.103276]
37. Armeen I, Niswanger R, Tian C. Combating fake news using implementation intentions. Inf Syst Front 2024:1-14 [FREE Full text] [doi: 10.1007/s10796-024-10502-0]

Edited by T Purnat; submitted 18.08.24; peer-reviewed by S Gordon, K Koidl; comments to author 15.10.24; revised version received 16.10.24; accepted 22.10.24; published 27.12.24.

Please cite as:
Bragazzi NL, Garbarino S
Understanding and Combating Misinformation: An Evolutionary Perspective
JMIR Infodemiology 2024;4:e65521
URL: https://infodemiology.jmir.org/2024/1/e65521
doi:10.2196/65521
PMID:39466077

Original Paper

Changes in Reproductive Health Information-Seeking Behaviors After the Dobbs Decision: Systematic Search of the Wikimedia Database

Mackenzie Lemieux[1], BSc, MD; Cyrus Zhou[1], BA; Caroline Cary[1], BA; Jeannie Kelly[1], MD

Department of Obstetrics and Gynecology, Washington University School of Medicine in St. Louis, St Louis, MO, United States

Corresponding Author:
Mackenzie Lemieux, BSc, MD
Department of Obstetrics and Gynecology
Washington University School of Medicine in St. Louis
660 S. Euclid Ave St. Louis, MO 63110
St Louis, MO, 63110-1010
United States
Phone: 1 (314) 362 7080
Email: l.mackenzie@wustl.edu

Abstract

Background: After the US Supreme Court overturned *Roe v. Wade*, confusion followed regarding the legality of abortion in different states across the country. Recent studies found increased Google searches for abortion-related terms in restricted states after the *Dobbsv. Jackson Women's Health Organization* decision was leaked. As patients and providers use Wikipedia (Wikimedia Foundation) as a predominant medical information source, we hypothesized that changes in reproductive health information-seeking behavior could be better understood by examining Wikipedia article traffic.

Objective: This study aimed to examine trends in Wikipedia usage for abortion and contraception information before and after the *Dobbs* decision.

Methods: Page views of abortion- and contraception-related Wikipedia pages were scraped. Temporal changes in page views before and after the *Dobbs* decision were then analyzed to explore changes in baseline views, differences in views for abortion-related information in states with restrictive abortion laws versus nonrestrictive states, and viewer trends on contraception-related pages.

Results: Wikipedia articles related to abortion topics had significantly increased page views following the leaked and final *Dobbs* decision. There was a 103-fold increase in the page views for the Wikipedia article *Roe v. Wade* following the *Dobbs* decision leak (mean 372,654, SD 135,478 vs mean 3614, SD 248; P<.001) and a 67-fold increase in page views following the release of the final *Dobbs* decision (mean 8942, SD 402 vs mean 595,871, SD 178,649; P<.001). Articles about abortion in the most restrictive states had a greater increase in page views (mean 40.6, SD 12.7; 18/51, 35% states) than articles about abortion in states with some restrictions or protections (mean 26.8, SD 7.3; 24/51, 47% states; P<.001) and in the most protective states (mean 20.6, SD 5.7; 8/51, 16% states; P<.001). Finally, views to pages about common contraceptive methods significantly increased after the *Dobbs* decision. "Vasectomy" page views increased by 183% (P<.001), "IUD" (intrauterine device) page views increased by 80% (P<.001), "Combined oral contraceptive pill" page views increased by 24% (P<.001), "Emergency Contraception" page views increased by 224% (P<.001), and "Tubal ligation" page views increased by 92% (P<.001).

Conclusions: People sought information on Wikipedia about abortion and contraception at increased rates after the *Dobbs* decision. Increased traffic to abortion-related Wikipedia articles correlated to the restrictiveness of state abortion policies. Increased interest in contraception-related pages reflects the increased demand for contraceptives observed after the *Dobbs* decision. Our work positions Wikipedia as an important source of reproductive health information and demands increased attention to maintain and improve Wikipedia as a reliable source of health information after the *Dobbs* decision.

(JMIR Infodemiology 2024;4:e64577) doi:10.2196/64577

KEYWORDS

abortion; Dobbs; internet; viewer trends; Wikipedia; women's health; contraception; contraceptive; trend; information seeking; page view; reproductive; reproduction

Introduction

In the United States, public interest in reproductive health information has been affected by recent drastic changes to the political landscape. In *Dobbs v Jackson Women's Health Organization*, a landmark case decided on June 24, 2022, the Supreme Court overturned *Roe v. Wade*, ruling that the US Constitution does not protect the right to abortion and leaving the states to decide its legality [1,2]. Since June 2022, the legality and accessibility of abortion have changed rapidly and significantly across the country. As of August 2024, a total of 14 states have outlawed abortion with few exceptions, 4 states have enacted 6-week abortion bans [3], and 14 states have adopted or are working to adopt amendments to enshrine the right to abortion in their constitutions [4,5]. These changes have created considerable confusion among the public about the legality of abortion. One 2023 survey found that 45% of the public was unsure if medication abortion was legal in their state [6].

Recent papers published in *JAMA Health Forum* and *JAMA Internal Medicine* [7,8] have explored health information-seeking behaviors in the post-*Dobbs* era to understand the public health impact of increasingly strict legislation on reproductive health care. These studies revealed that Google searches for abortion, contraception, and reproductive health-related topics reached record levels after the *Dobbs* decision was leaked, especially in states with restrictive abortion policies. Importantly, the increased demand for reproductive health information is complicated by the overwhelming amount of abortion misinformation on the internet. Experts have warned that there is currently an abortion "infodemic," [9] a term the World Health Organization (WHO) defines as the rapid spread of an excessive amount of health information, some of which is false, that ultimately confuses the public and decreases trust in the medical field [10]. Maintaining up-to-date online information about reproductive care is imperative for people to stay informed and understand how policy decisions impact their health.

While Google search trends [7,8] provide a surface-level view of internet user behavior, Wikipedia page views reflect an in-depth interest in a topic [11]. Compared with Google, studying Wikipedia also offers more specificity regarding where internet users obtain their health information. Furthermore, the process of crowdsourcing information on Wikipedia represents an opportunity for intervention in the dissemination of accurate reproductive and public health information, as attempted by the WHO during COVID-19 [12].

Research into who uses Wikipedia and why shows that 90% of medical students used Wikipedia to aid in their studies [13], 70% of junior physicians consulted Wikipedia in a clinical setting [14], and 35% of pharmacists reported using Wikipedia at work in a questionnaire [15]. In 2019, Wikipedia was ranked the second most commonly used resource for health information online [9], potentially due to its presence as a top-10 search result 71% to 85% of the time across common search engines such as Google, Yahoo, MSN, and so on [16]. During the COVID-19 pandemic, there was a significant increase in the number of medical articles on Wikipedia, a finding published in the *Journal of Medical Internet Research* in 2021 [17]. The authors of this article suggest that studying Wikipedia traffic could be a valid approach to epidemiologic surveillance [17]. Given the widespread use of Wikipedia as a source of health information for health professionals, students, and patients alike [18,19], we sought to explore viewing trends on Wikipedia pages about abortion policies, medical and surgical abortion, and contraceptive options after the *Dobbs* decision. We hypothesized that Wikipedia is used to obtain reproductive health information more after the *Dobbs* decision and more in restrictive states.

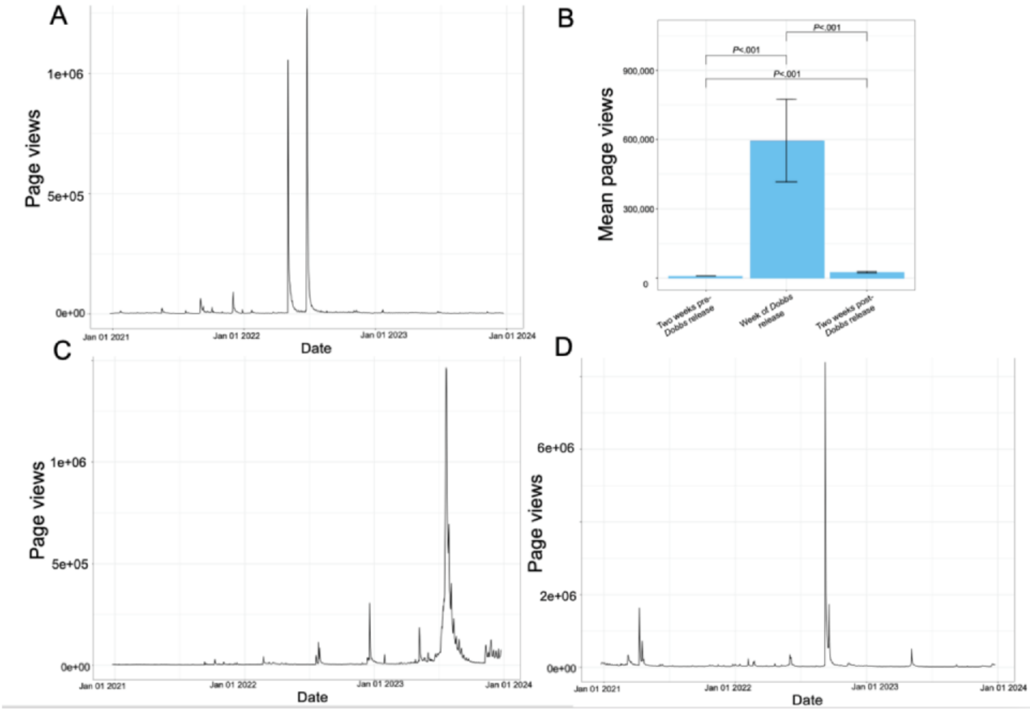
Methods

Data Scraping

A total of 89 Wikipedia articles were examined for this study. The process of choosing these articles and why they were chosen is further detailed below in the *Selection of Reproductive Health-Related Wikipedia Pages* section. A list of all 89 articles and their Wikipedia links can be found in Multimedia Appendix 1. Daily view counts to 87 total Wikipedia articles were obtained for 2 years bookending the June 24, 2022, *Dobbs* decision, from June 24, 2021, to June 24, 2023, using the *pageview* library (version 0.5.0) with the following options: platform "all" and user type "user."

Longitudinal trends in page views of the Wikipedia article *Roe v. Wade* were assessed given the central role of this abortion-related political event to our research question. Longitudinal trends in page views to 2 prominent topics surrounding the *Dobbs* decision with associated Wikipedia articles were selected as controls: Queen Elizabeth II and J. Robert Oppenheimer. Daily view counts for the 3 years bookending the *Dobbs* decision were obtained for these 3 pages and plotted in a histogram format (Figure 1).

Figure 1. (A) A line graph of page view counts over time for the English-language Wikipedia page “Roe v. Wade” for 3 years bookending the Dobbs decision (between January 2021 and January 2024). (B) A bar chart comparing mean page views across a 2-week span before, 1-week span during, and 2-week span after the Dobbs decision was made in June 2022. (C) and (D) Line graphs depicting page view counts over the 3 years bookending the Dobbs decision for 2 control pages: “Elizabeth II” (Queen Elizabeth II) and “J. Robert Oppenheimer.”.



Selection of Reproductive Health-Related Wikipedia Pages

Given that abortion policy varied dramatically by state following the *Dobbs* decision, we sought to determine if Wikipedia page viewing patterns differed by state. Using categories defined by The Guttmacher Institute [3], we grouped individual states into the following categories: restrictive policies, protective policies, and varied restrictive or protective policies. We subsequently quantified views to Wikipedia pages dedicated to each US state’s abortion laws (ie, “Abortion in Missouri” and “Abortion in Florida”). A list of the specific Wikipedia articles is provided in Table 1.

We next examined the percentage change in page views for 3 subsets of pages providing abortion information, including medical, surgical, and self-managed abortion. For medical abortion, we selected pages for the 2 most commonly used abortion medications, “Mifepristone” and “Misoprostol,” as described by the WHO and the American College of Obstetricians and Gynecologists (ACOG) [20,21]. Planned Parenthood (PPFA) and the Guttmacher Institute, 2 well-regarded sources of information within the field of obstetrics and gynecology, also recognize these 2 medications as the most commonly used for medical abortions [22,23]. We then selected blue hyperlinked alternative abortion medications found within the Wikipedia “Mifepristone” and “Misoprostol” articles to simulate Wikipedia user behavior. Similarly, the surgical abortion pages “Vacuum Aspiration” and “Dilation and Extraction” were selected based on PPFA’s web page for procedural abortion [24], and again, we selected blue hyperlinked alternative procedural abortion pages found within the “Vacuum Aspiration” and “Dilation and Extraction” pages. Finally, we examined the Wikipedia article for “Self-induced abortion” and selected the blue-linked alternative on this page, “Menstrual extraction.” A list of all included pages can be found in Table 2.

We next examined percentage changes in page views compared with the baseline for pages about the most commonly used birth control methods as described in the Guttmacher Institute 2021 report [15], including “Vasectomy,” “Tubal ligation,” “IUD” (intrauterine device), “Combined oral contraceptive pill,” “Emergency Contraceptive,” and “Condoms” (Figure 2B).

Examining Wikipedia pages with titles that exhibit more “colloquial language” was a consideration; however, many Wikipedia search attempts using colloquial language “redirect” to one common page that encompasses the colloquial terms. For example, the following Wikipedia search terms redirect to the Wikipedia page “Emergency Contraception” instead of having their own dedicated page: “Morning-after-pill,” “Emergency Contraceptive Pill,” “Emergency Birth Control,” “Postcoital contraception,” “Next day pill,” “Plan B Contraceptive,” and so on. We similarly found many colloquial Wikipedia search terms that redirect to the pages “Medical Abortion,” “Dilation and Curettage,” “Contraceptive Implant,” “Birth Control,” and “Tubal Ligation,” to name a few. Thus, we felt that we were capturing searches in which daily speech or colloquial language was used, given that Wikipedia has streamlined these search terms to one common page in many instances.

Table 1. Change in page views for US state abortion pagesa.

Abortion protectiveness and state	Daily page views, mean (SD)		Fold change
	Before the *Dobbs* decision	After the *Dobbs* decision	
Protective			
Minnesota	27 (7.4)	834 (590.2)	30.9
Maryland	26 (6.6)	758 (529.9)	29.7
Oregon	40 (11.6)	907 (553.4)	22.7
New Jersey	49 (8.0)	963 (649.8)	19.8
Vermont	13 (8.3)	250 (128.5)	19.7
California	114 (31.8)	1929 (1216.7)	16.9
New York	96 (16.4)	1312 (686.2)	13.6
New Mexico	30 (8.4)	350 (277.4)	11.9
Varied			
Ohio	46 (5.7)	2135 (1191.2)	46.6
Utah	23 (4.5)	996 (730.1)	43.6
Delaware	5 (2.4)	221 (162.9)	43.0
Montana	8 (4.6)	325 (246)	43.0
Arizona	29 (8.0)	1184 (446.2)	41.0
Florida	70 (31.9)	2415 (1228.8)	34.3
Wisconsin	37 (7.5)	1156 (1554.5)	30.9
Colorado	83 (8.2)	2395 (1274.8)	28.9
Alaska	22 (5.8)	634 (243.5)	28.6
Connecticut	16 (5.5)	433 (229)	26.6
Nevada	17 (3.7)	436 (342.1)	25.4
Michigan	30 (6)	755 (433.4)	25.3
New Hampshire	19 (7.2)	429 (473.2)	23.1
Pennsylvania	35 (4)	796 (647.3)	22.9
Wyoming	11 (4.8)	244 (206.3)	21.9
Kansas	47 (8.9)	959 (275.5)	20.6
Illinois	60 (16.7)	1222 (716.2)	20.5
Hawaii	16 (3.8)	323 (202.2)	20.4
Iowa	14 (5.2)	288 (165.4)	20.4
Rhode Island	10 (2)	188 (185.1)	18.8
Massachusetts	30 (7.8)	541 (357.7)	17.8
Maine	16 (8.1)	262 (150)	16.1
Washington	23 (5)	302 (430.6)	13.1
Virginia	33 (4.4)	356 (496.6)	10.9
Restrictive			
Kentucky	13 (3.2)	1142 (641.6)	85.9
Missouri	40 (6.3)	3204 (2149.2)	79.5
Idaho	14 (5.1)	846 (581.1)	60.4
Alabama	51 (7.7)	2976 (1581.7)	58.0
North Dakota	8 (4.1)	440 (208.1)	55.0
Arkansas	20 (5.5)	1043 (557.1)	52.2

Abortion protectiveness and state	Daily page views, mean (SD)		Fold change
	Before the *Dobbs* decision	After the *Dobbs* decision	
Mississippi	27 (9.9)	1163 (887.3)	42.4
Tennessee	42 (9.4)	1691 (1252.3)	39.9
Nebraska	9 (2.9)	344 (310.8)	38.2
Louisiana	32 (15.8)	1140 (654.1)	36.1
Georgia	44 (9.2)	1469 (935.9)	33.3
Indiana	25 (6.2)	729 (402.1)	29.1
North Carolina	39 (6.4)	1023 (1037.1)	26.1
Texas	270 (36.6)	6081 (4503.4)	22.6
South Dakota	21 (14.2)	470 (305.2)	22.4
West Virginia	13 (4.2)	224 (498.1)	17.6
South Carolina	26 (6.2)	441 (545.5)	17.2
Oklahoma	108 (16.4)	1581 (1012.3)	14.7

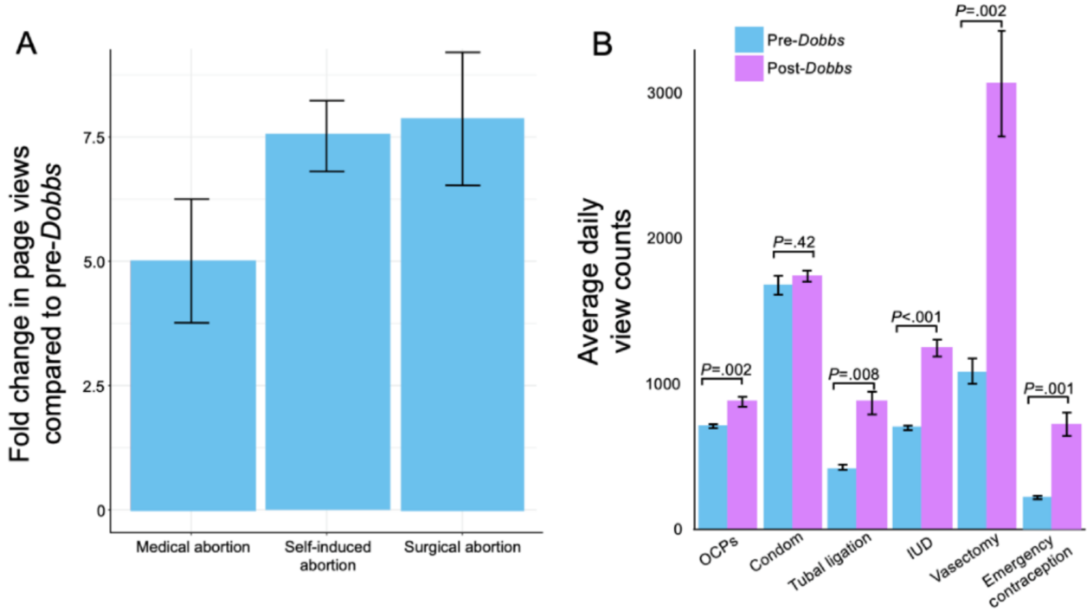
[a]Mean page views and fold change in page views before and after the *Dobbs* decision for each US state's abortion policy Wikipedia page. Within each restrictiveness category, states are ordered from highest to lowest fold change.

Table 2. Changes in page views for abortion-related pages before and after the Dobbs decisiona.

Articles and their categories	Before the *Dobbs* decision, mean (SD)	After the *Dobbs* decision, mean (SD)	Fold change	*P* value
Medical				
Abortifacient	278 (70.9)	2347 (1250.6)	8.5	.01
Gemeprost	10 (3.4)	17.4 (5.3)	1.7	.01
Medical abortion	299 (121.8)	2149 (1306.7)	7.2	.01
Methotrexate	1140 (149.2)	1640 (260.4)	1.4	.001
Mifepristone	491 (54.7)	3575 (1265.9)	7.3	<.001
Misoprostol	749 (69.9)	2953 (1299.7)	3.9	.004
Surgical				
Dilation and curettage	403 (35.0)	2013 (705.0)	5.0	<.001
Dilation and evacuation	169 (45.7)	1144 (379.8)	6.8	<.001
Hysterotomy abortion	30 (6.8)	255 (90.5)	8.5	<.001
Instillation abortion	37 (7.1)	305 (113.2)	8.2	<.001
Intact dilation and extraction	136 (32.1)	1872 (655.0)	13.8	<.001
Vacuum aspiration	146 (28.2)	726 (240.6)	8.2	.01
Self-induced				
Menstrual extraction	64 (11.7)	530 (314.2)	6.8	.004
Self-induced abortion	346 (38.4)	2353 (1152.7)	5.0	<.001

[a]Mean page views and fold change in page views before and after the *Dobbs* decision for all abortion-related pages included in the subcategory analysis of medical, self-induced, and surgical abortion.

Figure 2. Bar charts depicting changes in page views for abortion- and contraception-related pages during the 2 weeks before the Dobbs decision and the week after the Dobbs decision. (A) The fold change in page views compared with the baseline for pages about medication abortions, self-induced abortions, and surgical abortion procedures. (B) Comparison of average daily page view counts before and after the Dobbs decision for the following pages, which are combined: oral contraceptives (OCPs), condoms, tubal ligation, intrauterine device (IUD), vasectomy, and emergency contraception.



Data Cleaning and Statistics

Data were transformed and summarized with the aid of the *dplyr* (v 1.1.3) and *scales* (v 1.2.1; both developed in R Studios) libraries. Data visualization and graphics were generated using *ggplot2* (v 3.4.4) and *ggpubr* (v 0.6.0). All statistic comparisons were performed using Wilcoxon rank-sum (Mann-Whitney *U* test). In cases where multiple comparisons were made, *P* values were adjusted using false discovery rate correction, which was performed using the Benjamini-Hochberg procedure. All data analysis was done using R (R Core Team) (version 4.1.2) using the R studio interface (PositPBC; v 2023.06.2).

Ethical Considerations

This study does not include human subjects research (no human subjects experimentation or intervention was conducted) and so does not require institutional review board approval.

Results

The Public Seeks Information About Current Events on Wikipedia, Including Abortion Legislation

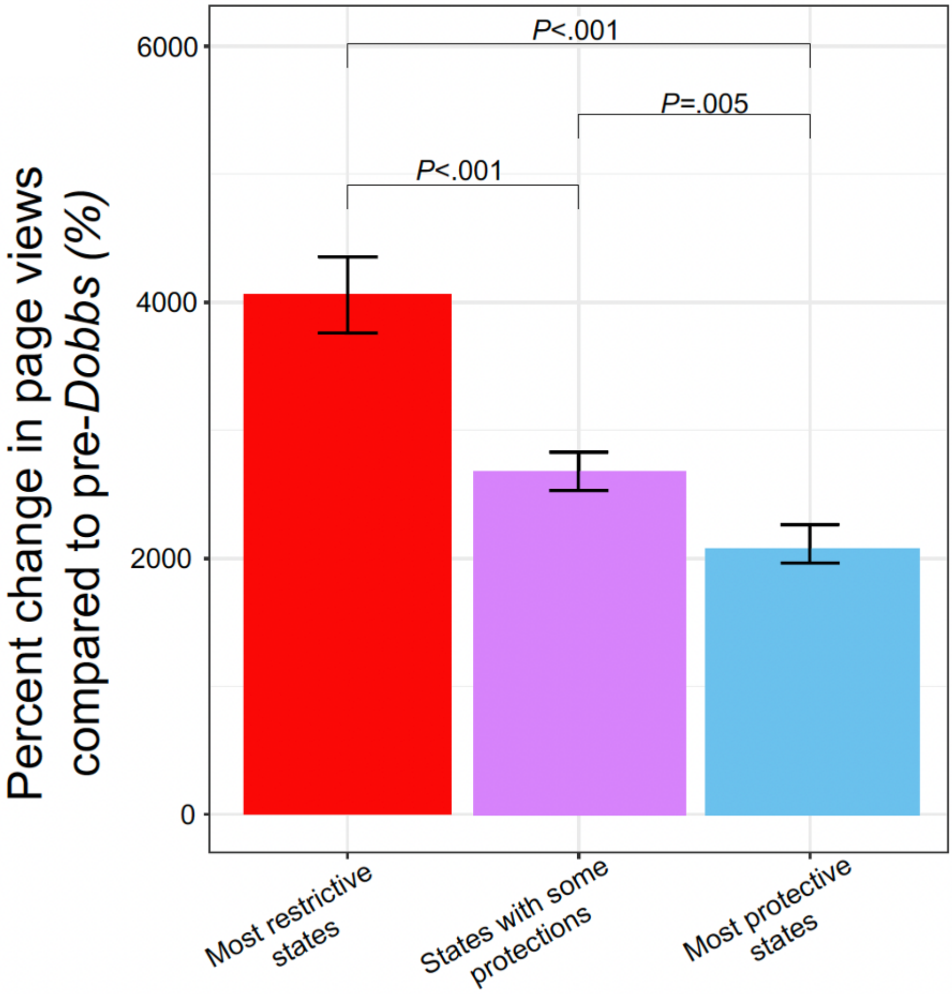
During the study period, 2 prominent spikes in page views were observed for the Wikipedia article *Roe v.Wade* (Figure 1A). These 2 spikes corresponded to the first week of May (when the *Dobbs* decision was leaked to the public on May 2, 2022) and the fourth week of June (when the *Dobbs* decision was formally released).

A comparison of the average daily page views to the Wikipedia article *Roe v.Wade* 2 weeks before the *Dobbs* leak (April 17 to May 1, 2022) and the week of the *Dobbs* leak (May 2–9, 2022) revealed a significant 103-fold increase (mean 372,654, SD 135,478 vs mean 3614, SD 248; *P*<.001) in the number of views after the leak. Similarly, when comparing the average daily page views 2 weeks before the *Dobbs* decision (June 9-23, 2022) to the week of the *Dobbs* decision (June 24 to July 1, 2022), we found a significant 67-fold increase in the number of views the week the *Dobbs* decision was released (Figure 1B; mean 8942, SD 402 vs mean 595,871, SD 178,649; *P*<.001). This pre- versus post-*Dobbs* spike in view counts was not seen in selected Wikipedia control articles for prominent media or news events unrelated to abortion. Wikipedia articles for J. Robert Oppenheimer (Figure 1C) and Queen Elizabeth II (Figure 1D), which saw spikes in other parts of the year, did not have observable spikes in early May and late June as the *Roe v. Wade* article did (Figure 1C and D).

State Abortion Restrictions Impact Information-Seeking Behavior on Wikipedia

Across all states, we observed a significant increase in searches for Wikipedia articles covering state-level abortion policies following the *Dobbs* decision. Similar to what we observed with the view counts in the *Roe v. Wade* article, 2 prominent spikes in viewership correlated to the week of the *Dobbs* leak and the *Dobbs* decision. In aggregate, we saw a roughly 30-fold increase in views for state-level Wikipedia articles following the *Dobbs* decision (eg, the average daily views of the Wikipedia article “Abortion in Alaska” went from 22 views before the *Dobbs* decision to 634 views after the *Dobbs* decision). Although all states had a significant increase following the *Dobbs* release, the relative magnitude of the increase significantly varied by the restrictiveness of state abortion laws (Figure 3 and Table 1). The percentage change in page views was significantly higher for most or very restrictive states (mean 40.6, SD 12.7; 18/51, 35% states) compared with states with some restrictions or protections (mean 26.8, SD 7.3; 24/51, 47% states; *P*<.001) and to states with most protections (mean 20.6, SD 5.7; 8/51, 16% states; *P*<.001).

Figure 3. Bar chart comparing percentage change in page views before and after the Dobbs decision for US state abortion pages (eg, “Abortion in Texas” and “Abortion in California”) on Wikipedia. States were stratified based on abortion restriction status guided by Guttmacher Institute data.



Abortion and Contraceptive Information-Seeking Behavior Increased After the Dobbs Decision on Wikipedia

Views to Wikipedia pages about medical, surgical, and self-managed abortion significantly increased during the week of the *Dobbs* decision compared with baseline (2-week average before the decision), corresponding to a 401% (SD 124%), 687% (SD 72%), and 652% (SD 133%) change, respectively (Figure 2A). The individual pages comprising each subcategory of abortion information are found in Table 2.

Compared with the baseline, views for the page “Vasectomy” increased by 183% after the *Dobbs* decision (*P*<.001), views for the page “IUD” increased by 79% after the *Dobbs* decision (*P*<.001), and views for the page “Combined oral contraceptive pill” increased by 24% after the *Dobbs* decision (*P*<.001). Views for the page “Tubal ligation” increased significantly by 92% after the *Dobbs* decision compared with baseline (*P*=.01), as did views for the page “Emergency Contraception” (*P*<.001), which increased by 224%. However, views for the page “Condoms” did not increase significantly after the *Dobbs* decision compared with baseline (*P*=.42; Figure 2B).

Discussion

Principal Findings

Our study provides unique insight into reproductive health information-seeking behavior surrounding the landmark *Dobbs* decision. To our knowledge, this is the first study using Wikipedia as a tool to examine reproductive health information-seeking behaviors on the internet. Through scraping page views of abortion and contraception-related Wikipedia pages, we found increased viewer traffic to these pages surrounding the leak and the final *Dobbs* decision. Public interest in the *Roe v. Wade* Wikipedia page correlated with current events and public interest in state abortion laws related to states’ hostility toward abortion. Finally, public interest in common abortion and contraception methods was significantly higher after the *Dobbs* decision. Our findings situate Wikipedia as a source of reproductive health information in English and highlight the need to maintain Wikipedia as an information resource.

Public Use of Wikipedia

The breadth of knowledge on Wikipedia is unparalleled with over 59 million articles in over 300 languages [25]. Of these pages, 30,000 are about health and medicine topics in English and another 164,000 are about health and medicine topics in other languages [26]. Given the extent of medical information on Wikipedia, it is unsurprising that Wikipedia has been used to explore health behavior trends. Wikipedia page view trends have been used to estimate peak weeks of influenza-like illness ahead of the Centers for Disease Control and Prevention with 17% more accuracy than Google trends [27]. Wikipedia page view trends have also been used to monitor and forecast disease-location pairs across the globe [28]. More recently, a study published in the *Journal of Medical Internet Research* in 2021 showed that there was a significant increase in medical article submissions during the height of the COVID-19 pandemic [17]. The authors further postulate that observing Wikipedia activity could be a viable method of epidemiologic surveillance [17]. However, no research to date has explored health behavior trends on Wikipedia concerning reproductive health.

Dynamic Abortion Policies and Misinformation

The current state of abortion misinformation has been widely regarded as harmful to the health, rights, and freedoms of people living in the United States [9]. Since the *Dobbs* decision, Americans have been increasingly pressed to find trustworthy information on the internet about abortion rights and policies in their communities [9,29,30]. The right to free and reliable medical information is documented and described by the WHO [31], yet this is not the reality for most individuals.

Interestingly, during the COVID-19 pandemic, when misinformation was rampant, the WHO partnered with Wikipedia to address misinformation [12]. By immediately updating Wikipedia pages with new information from the WHO, the partnership was an attempt to disseminate new and reputable information quickly and to as many people as possible [12]. Given our findings, it is reasonable to consider a future partnership between Wikipedia and institutions such as ACOG, PPFA, and the Guttmacher Institute to disseminate accurate and up-to-date information surrounding reproductive health, especially in light of the upcoming election in November where reproductive rights are at stake.

One of our main findings is that we see larger increases in views for pages about abortion in hostile US states, which mirrors previous research showing an increase in internet searches about abortion in states with the most severe restrictions by up to 42% after the *Dobbs* decision [7,8]. One possible interpretation of this finding is that an increase in abortion information-seeking behavior in hostile states reflects a lack of knowledge or understanding surrounding the legality of abortion in restrictive states. Post-*Dobbs* increases in depression and anxiety symptoms among residents of abortion-hostile states, particularly among women of reproductive age [32,33], also support this speculation. Increased levels of misinformation [34] and dynamic discussions about the future of abortion rights in hostile states have the potential to further contribute to increased searches about abortion information. Interestingly, the "restrictive" states with the highest fold change in views, Kentucky and Missouri, already have or are planning to place abortion rights on state ballots [35]; the "varied" state with the highest fold change, Ohio, has enshrined abortion rights in their state constitution through a ballot measure already [36].

Contraceptive Trends After the Dobbs Decision

In addition to Wikipedia page views, internet searches for contraception increased after the *Dobbs* decision [37], particularly for permanent methods like vasectomy and tubal ligation [7,38]. According to PPFA, traffic to their web page on sterilization procedures increased by 300% in the first month after the *Dobbs* decision [39]. Increased internet searches corresponded to increased demand for these methods. At PPFA, birth control appointments increased by 15%, and appointments specifically for IUDs increased by 30% over the first month after the *Dobbs* decision [39]. Some institutions documented significant increases in female sterilization after the *Dobbs* decision, and patients cited *Dobbs* as an influence on their decision [40]. A study at one large US health care organization found a 160% increase in vasectomy procedures within 6 months after the *Dobbs* decision [41]. One company that provides contraception through mail saw a 30% increase in dual orders of emergency contraception and birth control pills after the *Dobbs* decision [37]. Heightened uncertainty about the entire reproductive health care landscape in the US could contribute to increased information-seeking and demand for the most effective contraceptive methods [37]. We did not find a significant increase in views for the "Condoms" page despite their frequent use [42] and evidence that condom use increased after the *Dobbs* decision [43]. Potential reasons for this include the accessibility of condoms compared with other forms of contraception, the lack of scrutiny toward condoms compared with other birth control methods [44], and potentially an elevated baseline knowledge of this method compared with other birth control methods.

Wikipedia and the Right to Health Information

Our study shows that the public seeks reproductive health information on Wikipedia, and at significantly higher rates in times of political change. Given that Wikipedia is free and available to anybody with internet access, we argue that Wikipedia is an important provider of free reproductive health information. Promoting Wikipedia editing within the medical and health care communities is an important step toward maintaining and improving the amount of reliable, evidence-based reproductive health information on the internet. More than half of Wikipedia's main editors have a health care background, and over 85% have a university education [45]. Many editors of health care–related pages are a part of WikiProject Medicine, whose recent focus has been incorporating Wikipedia editing into medical school curricula [46]. Especially in the current political climate of misinformation, it is imperative to maintain and improve reproductive health articles on Wikipedia. As mentioned above, one way we hypothesize to disseminate accurate and up-to-date health information is through partnerships between Wikipedia and large reproductive health organizations such as ACOG and PPFA [12].

Limitations

Our study was limited to analyzing the behaviors of people with internet access and topics specifically searched within Wikipedia, not the entire internet. Our study was also limited to English-language Wikipedia and we cannot make conclusions about abortion information-seeking behavior for non–English-language speakers. While many users of English-language Wikipedia are located in the United States, our study is not limited to only US Wikipedia users. Given these limitations, our conclusions are specific to understanding if and how Wikipedia is used as a reproductive health information source in isolation, not in comparison to other internet information sources. While we cannot conclude anything about reproductive health information–seeking behavior on other internet platforms at this time, we do feel confident that Wikipedia is a source of reproductive health information for users seeking information in English. Our conclusions can be extended and acted upon in the context of Wikipedia alone. For example, we suggest that reproductive health organizations partner with Wikipedia to further expand on and disseminate reproductive health information. Future research could examine how often users consult reproductive health information on Wikipedia versus other online platforms and could be extended to users in languages other than English. There is potential for non–English-language users to consult different sources than English users.

Conclusions

Our research contributes to the growing subset of public health literature that analyzes trends in reproductive health information-seeking behavior on the internet. With a goal of improving access to free and reliable reproductive health information, our work highlights the role Wikipedia plays in providing reproductive health information surrounding changes to reproductive health policy. Medical schools and professional organizations can support activities geared toward improving and creating reproductive health pages on Wikipedia and partner with Wikipedia to disseminate accurate and up-to-date reproductive health information.

Acknowledgments

This work was supported by grant 2021265 from the Doris Duke Charitable Foundation through the COVID-19 Fund to Retain Clinical Scientists collaborative grant program.

Authors' Contributions

ML, CZ, and CC contributed to conceptualization, investigation, and methodology. JK and ML performed supervision. ML conducted visualization. CZ and ML performed data curation. CZ managed formal analysis and software. JK managed resources. All authors wrote the original draft and contributed to reviewing and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of the links to all Wikipedia articles analyzed for the purposes of this paper.

[XLSX File (Microsoft Excel File), 12 KB - infodemiology_v4i1e64577_app1.xlsx]

References

1. Dobbs v. Jackson Women's Health Organization, 597 U.S. 215. 2022. URL: https://www.supremecourt.gov/opinions/21pdf/19-1392_6j37.pdf [accessed 2024-12-05]
2. Roe v. Wade, 410 U.S. 113. 1973. URL: https://supreme.justia.com/cases/federal/us/410/113/ [accessed 2024-10-22]
3. Interactive map: US abortion policies and access after Roe. Guttmacher Institute. URL: https://states.guttmacher.org/policies/?gad_source=1&gclid=Cj0KCQiAyeWrBhDDARIsAGP1mWSWJigqSSajghLLq4QQqHHIhvo2tadztl6n_lS9tgY1lv8gfSRqGsYaAuCgEALw_wcB [accessed 2024-01-30]
4. Mulvihill G, Associated Press. ere are the states where abortion access may be on the ballot in 2024. PBS News. 2023 Nov 8. URL: https://www.pbs.org/newshour/politics/here-are-the-states-where-abortion-access-may-be-on-the-ballot-in-2024 [accessed 2024-12-05]
5. Ballot tracker: status of abortion-related state constitutional amendment measures for the 2024 election. KFF. 2024. URL: https://www.kff.org/womens-health-policy/dashboard/ballot-tracker-status-of-abortion-related-state-constitutional-amendment-measures/ [accessed 2024-10-22]
6. Kearney A, Sparks G, Kirzinger A, Presiado M, Brodie M. KFF health tracking poll may 2023: health care in the 2024 election and in the courts. Women's Health Policy. 2023 May 26. URL: https://www.kff.org/report-section/kff-tracking-poll-may-2023-health-care-in-the-2024-election-and-in-the-courts-methodology/ [accessed 2024-12-05]
7. 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;4(4):e230518 [FREE Full text] [doi: 10.1001/jamahealthforum.2023.0518] [Medline: 37115538]

8. 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;182(9):1002-1004 [FREE Full text] [doi: 10.1001/jamainternmed.2022.2998] [Medline: 35767270]
9. Pagoto SL, Palmer L, Horwitz-Willis N. The next infodemic: abortion misinformation. J Med Internet Res 2023;25:e42582 [FREE Full text] [doi: 10.2196/42582] [Medline: 37140975]
10. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic#tab=tab_1 [accessed 2024-02-10]
11. Kämpf M, Tessenow E, Kenett DY, Kantelhardt JW. The detection of emerging trends using wikipedia traffic data and context networks. PLoS One 2015;10(12):e0141892 [FREE Full text] [doi: 10.1371/journal.pone.0141892] [Medline: 26720074]
12. McNeil DGJ. Wikipedia and W.H.O. join to combat COVID-19 misinformation. New York Times. 2020 Nov 2. URL: https://www.nytimes.com/2020/10/22/health/wikipedia-who-coronavirus-health.html [accessed 2024-12-05]
13. Allahwala UK, Nadkarni A, Sebaratnam DF. Wikipedia use amongst medical students – new insights into the digital revolution. Medical Teacher 2012;35(4):337-337. [doi: 10.3109/0142159x.2012.737064]
14. Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. Int J Med Inform 2009;78(10):645-655. [doi: 10.1016/j.ijmedinf.2009.04.008] [Medline: 19501017]
15. Brokowski L, Sheehan AH. Evaluation of pharmacist use and perception of wikipedia as a drug information resource. Ann Pharmacother 2009;43(11):1912-1913. [doi: 10.1345/aph.1m340]
16. Laurent MR, Vickers TJ. Seeking health information online: does wikipedia matter? J Am Med Inform Assoc 2009;16(4):471-479. [doi: 10.1197/jamia.m3059]
17. Chrzanowski J, Sołek J, Fendler W, Jemielniak D. Assessing public interest based on wikipedia'S most visited medical articles during the SARS-CoV-2 Outbreak: search trends analysis. J Med Internet Res 2021;23(4):e26331 [FREE Full text] [doi: 10.2196/26331] [Medline: 33667176]
18. Beck J. Doctors' #1 source for healthcare information: Wikipedia. The Atlantic. 2014 Mar 14. URL: https://www.theatlantic.com/health/archive/2014/03/doctors-1-source-for-healthcare-information-wikipedia/284206/ [accessed 2024-05-27]
19. NPR Staff. Dr. Wikipedia: the 'Double-Edged Sword' of crowdsourced medicine. NPR. 2014 Feb 8. URL: https://www.npr.org/2014/02/08/273680018/dr-wikipedia-the-double-edged-sword-of-crowd-sourced-medicine [accessed 2024-10-22]
20. Gynecologists ACOG. Medication abortion up to 70 days of gestation. Clinical Practice Bulletin 2023 [FREE Full text] [doi: 10.1097/aog.0000000000004083]
21. Abortion Care Guideline. World Health Organization. 2022 Mar 8. URL: https://www.who.int/publications/i/item/9789240039483 [accessed 2024-12-05]
22. The abortion pill. Planned Parenthood. URL: https://www.plannedparenthood.org/learn/abortion/the-abortion-pill [accessed 2024-04-23]
23. Friedrich-Karnik A, Stoskopf-Ehrlich E, Jones RK. Medication abortion within and outside the formal US health care system: what you need to know. Guttmacher Institute. 2024 Feb. URL: https://tinyurl.com/4c3dwb8p [accessed 2024-04-24]
24. In-clinic abortion. Planned Parenthood. 2024. URL: https://www.plannedparenthood.org/learn/abortion/in-clinic-abortion-procedures [accessed 2024-04-23]
25. Wikipedia:size of wikipedia. Wikipedia. URL: https://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia#:~:text=As%20of%207%20November%202023,of%20all%20pages%20on%20Wikipedia [accessed 2024-02-05]
26. Health information on wikipedia. Wikipedia. URL: https://tinyurl.com/269xn6uy [accessed 2024-02-05]
27. McIver DJ, Brownstein JS. Wikipedia usage estimates prevalence of influenza-like illness in the United States in near real-time. PLoS Comput Biol 2014;10(4):e1003581 [FREE Full text] [doi: 10.1371/journal.pcbi.1003581] [Medline: 24743682]
28. Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global disease monitoring and forecasting with Wikipedia. PLoS Comput Biol 2014;10(11):e1003892 [FREE Full text] [doi: 10.1371/journal.pcbi.1003892] [Medline: 25392913]
29. Lerman R. People searching for abortion online must wade through misinformation. The Washington Post. 2022. URL: https://www.washingtonpost.com/technology/2022/07/04/abortion-misinformation-herbal-remedies/ [accessed 2024-05-15]
30. Hill WC. Overturn of Roe leads to both confusion and misinformation. American College of Obstetricians and Gynecologists. 2023 Jun 29. URL: https://www.acog.org/news/news-articles/2023/06/overturn-of-roe-leads-to-confusion-and-misinformation [accessed 2024-12-05]
31. World Health Organization. The right to health. Office of the United Nationa High Commissioner for Human Rights. 2008. URL: https://www.ohchr.org/Documents/Publications/Factsheet31.pdf [accessed 2024-12-05]
32. Liu SY, Benny C, Grinshteyn E, Ehntholt A, Cook D, Pabayo R. The association between reproductive rights and access to abortion services and mental health among US women. SSM Popul Health 2023;23:101428 [FREE Full text] [doi: 10.1016/j.ssmph.2023.101428] [Medline: 37215399]
33. Dave D, Fu W, Yang M. Mental distress among female individuals of reproductive age and reported barriers to legal abortion following the US supreme court decision to overturn roe v wade. JAMA Netw Open 2023;6(3):e234509 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.4509] [Medline: 36951866]

34. Sherman C. The abortion myths republicans are recycling to reframe a losing issue. The Guardian. 2023 Sep 27. URL: https://www.theguardian.com/world/2023/sep/27/abortion-myths-republicans [accessed 2024-10-22]
35. Mulvihill G, Kruesi K. Which states could have abortion on the ballot in? Associated Press News. 2024 Apr 10. URL: https://apnews.com/article/abortion-ballot-amendment-ban-protection-states-2024-052ff9846f8416efb725240af22b92ec [accessed 2024-08-17]
36. Smyth JC, Associated Press. Ohio voters enshrine abortion access in constitution in latest statewide win for reproductive rights. PBS News. 2023 Nov 7. URL: https://www.pbs.org/newshour/politics/ohio-voters-enshrine-abortion-access-in-constitution-in-latest-win-for-reproductive-rights [accessed 2024-10-22]
37. Law T. 21% of Women reported switching their birth control method post-Roe. Time Magazine. 2022 Jul 27. URL: https://time.com/6200542/women-birth-control-switching-methods-abortion/ [accessed 2024-07-21]
38. Ghomeshi A, Diaz P, Henry V, Ramasamy R, Masterson TA. The interest in permanent contraception peaked following the leaked supreme court majority opinion of Roe vs. Wade: a cross-sectional google trends analysis. Cureus 2022;14(10):e30582 [FREE Full text] [doi: 10.7759/cureus.30582] [Medline: 36420253]
39. By the numbers: one month in a post-Roe America. Planned Parenthood. 2022 Jul. URL: https://www.plannedparenthood.org/uploads/filer_public/d0/87/d087f64c-f1a7-4a2d-b5b8-08089a4c7346/by_the_numbers__one_month_in_to_a_post_roe_america.pdf [accessed 2024-08-28]
40. Hennessey C, Johnson C, McLaren H, Bhardwaj N, Rivlin K, Chor J. Permanent sterilization in nulliparous patients: is legislative anxiety an indication for surgery? J Clin Ethics 2023;34(4):320-327. [doi: 10.1086/727435] [Medline: 37991729]
41. Zhu A, Nam CS, Gingrich D, Patel N, Black K, Andino JJ, et al. Short-term changes in vasectomy consults and procedures following. Urology Practice 2024;11(3):517-525. [doi: 10.1097/upj.0000000000000528]
42. Institute G. Contraceptive use in the United States by method. Guttmacher Institute. 2021 May. URL: https://www.guttmacher.org/fact-sheet/contraceptive-use-united-states [accessed 2024-10-22]
43. Kavanaugh ML, Friedrich-Karnik A. Has the fall of changed contraceptive access and use? New research from four US states offers critical insights. Health Aff Sch 2024 Feb;2(2):qxae016 [FREE Full text] [doi: 10.1093/haschl/qxae016] [Medline: 38756551]
44. Ramirez NM. The right is cracking down on abortion and IVF. Is 'Recreational Sex' next? Rolling Stone. 2024 Feb 23. URL: https://www.rollingstone.com/politics/politics-news/birth-control-targeted-right-wing-influencers-1234974833/ [accessed 2024-10-22]
45. Heilman JM, West AG. Wikipedia and medicine: quantifying readership, editors, and the significance of natural language. J Med Internet Res 2015;17(3):e62 [FREE Full text] [doi: 10.2196/jmir.4069] [Medline: 25739399]
46. Joshi M, Verduzco R, Yogi S, Garcia M, Saxena S, Tackett S, et al. Wikipedia editing courses at three US medical schools in the 2017-2018 academic year. MedEdPublish 2019;8:146. [doi: 10.15694/mep.2019.000146.1]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists
IUD: intrauterine device
PPFA: Planned Parenthood
WHO: World Health Organization

Edited by T Purnat; submitted 20.07.24; peer-reviewed by D Walker, S Gorman; comments to author 27.07.24; revised version received 29.09.24; accepted 08.10.24; published 16.12.24.

Please cite as:
Lemieux M, Zhou C, Cary C, Kelly J
Changes in Reproductive Health Information-Seeking Behaviors After the Dobbs Decision: Systematic Search of the Wikimedia Database
JMIR Infodemiology 2024;4:e64577
URL: https://infodemiology.jmir.org/2024/1/e64577
doi:10.2196/64577
PMID:

Original Paper

The Use of Social Media to Express and Manage Medical Uncertainty in Dyskeratosis Congenita: Content Analysis

Emily Pearce[1], MPH, PhD; Hannah Raj[2], BSc; Ngozika Emezienna[1]; Melissa B Gilkey[3], PhD; Allison J Lazard[4], PhD; Kurt M Ribisl[3], PhD; Sharon A Savage[1], MD, PhD; Paul KJ Han[5], MD, PhD

[1]Division of Cancer Epidemiology and Genetics, Clinical Genetics Branch, National Cancer Institute, National Institutes of Health, Rockville, MD, United States

[2]Team Telomere, Coeur d'Alene, ID, United States

[3]Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]Hussman School of Journalism and Media, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[5]Division of Cancer Control and Population Sciences, Behavioral Research Program, National Cancer Institute, National Institutes of Health, Rockville, MD, United States

Corresponding Author:
Emily Pearce, MPH, PhD
Division of Cancer Epidemiology and Genetics, Clinical Genetics Branch
National Cancer Institute
National Institutes of Health
609 Nelson St
Rockville, MD, 20850
United States
Phone: 1 9196992547
Email: emily.pearce@nih.gov

Abstract

Background: Social media has the potential to provide social support for rare disease communities; however, little is known about the use of social media for the expression of medical uncertainty, a common feature of rare diseases.

Objective: This study aims to evaluate the expression of medical uncertainty on social media in the context of dyskeratosis congenita, a rare cancer-prone inherited bone marrow failure and telomere biology disorder (TBD).

Methods: We performed a content analysis of uncertainty-related posts on Facebook and Twitter managed by Team Telomere, a patient advocacy group for this rare disease. We assessed the frequency of uncertainty-related posts, uncertainty sources, issues, and management and associations between uncertainty and social support.

Results: Across all TBD social media platforms, 45.98% (1269/2760) of posts were uncertainty related. Uncertainty-related posts authored by Team Telomere on Twitter focused on scientific (306/434, 70.5%) or personal (230/434, 53%) issues and reflected uncertainty arising from probability, ambiguity, or complexity. Uncertainty-related posts in conversations among patients and caregivers in the Facebook community group focused on scientific (429/511, 84%), personal (157/511, 30.7%), and practical (114/511, 22.3%) issues, many of which were related to prognostic unknowns. Both platforms suggested uncertainty management strategies that focused on information sharing and community building. Posts reflecting response-focused uncertainty management strategies (eg, emotional regulation) were more frequent on Twitter compared with the Facebook community group (χ^2_1=3.9; *P*=.05), whereas posts reflecting uncertainty-focused management strategies (eg, ordering information) were more frequent in the Facebook community group compared with Twitter (χ^2_1=55.1; *P*<.001). In the Facebook community group, only 36% (184/511) of members created posts during the study period, and those who created posts did so with a low frequency (median 3, IQR 1-7 posts). Analysis of post creator characteristics suggested that most users of TBD social media are White, female, and parents of patients with dyskeratosis congenita.

Conclusions: Although uncertainty is a pervasive and multifactorial issue in TBDs, our findings suggest that the discussion of medical uncertainty on TBD social media is largely limited to brief exchanges about scientific, personal, or practical issues rather than ongoing supportive conversation. The nature of uncertainty-related conversations also varied by user group: patients and caregivers used social media primarily to discuss scientific uncertainties (eg, regarding prognosis), form social connections, or exchange advice on accessing and organizing medical care, whereas Team Telomere used social media to express scientific and personal issues of uncertainty and to address the emotional impact of uncertainty. The higher involvement of female parents on TBD social media suggests a potentially greater burden of uncertainty management among mothers compared with other groups. Further research is needed to understand the dynamics of social media engagement to manage medical uncertainty in the TBD community.

(JMIR Infodemiology 2024;4:e46693) doi:10.2196/46693

KEYWORDS

social media; medical uncertainty; telomere biology disorder; dyskeratosis congenita; social support

Introduction

Background

Medical uncertainty is a common experience in rare diseases and may combine with limited scientific knowledge and access to peer groups to impede a patient's ability to seek and adhere to medical treatments [1] and intensify health-related anxiety, decreasing quality of life for patients and their caregivers [2,3]. Dyskeratosis congenita (DC) is a rare telomere biology disorder (TBD) associated with very high risks of bone marrow failure, pulmonary and liver disease, cancer, and other medical conditions. Diagnosis is challenging because of its wide phenotypic spectrum, including the classic DC triad (nail dysplasia, abnormal skin pigmentation, and oral leukoplakia) with pediatric bone marrow failure, middle-age presentation with pulmonary failure or aplastic anemia, abnormally short telomere length, or detection of pathogenic germline variants in >18 different genes [4]. Although age of onset is variable, DC often presents in childhood and adolescence, with most patients experiencing their first symptoms before the age of 20 years [5]. Diagnosis frequently results in a lifetime commitment to screening to detect progressive clinical manifestations of DC, including cancers across multiple organ systems [5]. Owing to the complexity and rarity of DC and related TBDs, patients and their families often have long diagnostic journeys, face complicated health decision-making, and frequently do not have access to medical professionals and supportive peers who are familiar with their condition. This situation likely creates a substantial burden of medical uncertainty for patients with TBDs and their families. Although medical uncertainty has been associated with increased anxiety and difficulty with decision-making in rare diseases and cancer occurrence and recurrence [6-11], to date, no research has addressed the experience or management of medical uncertainty in the TBD context.

As outlined in a previously published taxonomy developed by Han [12], uncertainty in medicine arises from multiple sources (eg, probability, ambiguity, and complexity) and focuses on scientific, personal, and practical issues. These situations activate a variety of management strategies to address uncertainty, which are primarily cognitive, emotional, and relational in nature. Uncertainty management strategies may target ≥1 sources or issues of uncertainty and are defined as belonging to ≥1 of the following approaches: seeking information to fill knowledge gaps ("ignorance-focused"), reducing or increasing attention to unknowns ("uncertainty-focused"), ameliorating adverse psychological effects of uncertainty ("response-focused"), and fostering interpersonal relationships to engage with uncertainty as a shared experience ("person-focused"). In situations where uncertainty cannot be reduced, these strategies may mitigate its negative mental health impact and help individuals achieve an adaptive, optimal balance of responses to uncertainty (uncertainty tolerance).

The rarity of TBDs suggests a potential role for internet-based platforms to deliver social support by bridging geographic, knowledge, and community network limitations. Social support, a complex concept encompassing a variety of helping social interactions [13], includes four main types: (1) expression of empathy and care (emotional), (2) provision of tangible assistance (instrumental), (3) provision of knowledge or facts (informational), and (4) evaluative feedback about task performance or personal qualities (appraisal) [14]. Research suggests that social support decreases the experience of stress, anxiety, and depression and improves the overall quality of life in populations experiencing medical uncertainty [8,10,15-17]. The benefit of social support has been demonstrated in patients with Li-Fraumeni syndrome, a rare genetic cancer predisposition, where informational, tangible, spiritual, and emotional support from in-person sources enhanced positive coping capacities [18]. Social media platforms such as Facebook and Twitter have been identified as important resources for social support in rare disease contexts [19-24], and disease-specific social media support has been recommended in oncology [25], rare genetic disease [26-28], and other stigmatized or rare diseases [29-31]. In addition to increasing access to information and social networks, continued participation in socially supportive internet-based communities may also build capacities for uncertainty tolerance [10,17,32-38]. Although social media has the potential to bridge geographic or social boundaries, its use is often concentrated in select populations, limiting its reach and potentially inhibiting its use by some groups [39,40]. In addition, dynamics observed on social media posts may not reflect real-life experiences and are limited in depth and detail, increasing the potential for misinterpretation [39]. Social media can also spread misinformation with damaging consequences, especially in high-uncertainty health contexts [41-43].

Objectives

Although extensive research has investigated the psychosocial benefits of internet-based health forums for patients and their caregivers [23,28,29,44-51], there is still a need to evaluate the use of social media to express or manage medical uncertainty in rare diseases. Specifically, we need to examine social media use for expressing and managing medical uncertainty in TBDs to understand the experience of medical uncertainty in this context and to build evidence to improve health communication and uncertainty management interventions [52]. This exploratory study aims to review social media posts created by and targeted at patients with TBDs and their caregivers to (1) measure the frequency of uncertainty-related posts; (2) catalog the issues, sources, and types of uncertainty and uncertainty management strategies; (3) measure user engagement with different post types; and (4) explore the relationship between uncertainty and social support. To achieve these aims, we reviewed all publicly available social media sites owned and maintained by Team Telomere (previously DC Outreach, Inc), the oldest and largest patient advocacy organization for individuals, caregivers, and families affected by TBDs worldwide [53]. The social media of Team Telomere constitutes the most expansive and accessible body of internet-based TBD-related content, inclusive of a variety of user perspectives. The variety of posts by users with diverse connections to TBDs (eg, medical providers, patients, caregivers, and health advocacy nonprofits) makes Team Telomere's social media an ideal data source for understanding the range and dynamics of medical uncertainty communication and social support exchange in the TBD context.

Methods

Ethical Considerations

Data collection was undertaken in partnership with Team Telomere following best practices guidelines for social media research [54] and was approved by the National Institutes of Health Institutional Review Board (IRB 000722).

Data Source

The source of data for this study was all publicly available social media owned and maintained by Team Telomere. These sites included the Team Telomere Twitter page [55], the Facebook main page [56], and a public Facebook community group [57] (Table 1). All the sites were open to the public and had no eligibility requirements for membership. Content across all platforms was monitored by Team Telomere to ensure appropriate adherence to community guidelines, and Team Telomere's staff removed posts with offensive or scientifically inaccurate content. The Facebook main page and Twitter accounts were created to promote the work of Team Telomere "supporting families worldwide affected by Dyskeratosis Congenita and Telomere Biology Disorders" [56]. The Facebook community group was created in response to social isolation following the COVID-19 pandemic as "a place to share our everyday lives in the spirit of promoting and maintaining connections among our Team Telomere/Dyskeratosis Congenita/Telomere Biology disorder community" [57].

Table 1. Data source characteristics at the time of the study.

	Facebook community group	Facebook main page	Twitter
Creation date (y)	2020	2010	2010
Followers, n	187	1637	1933
Posts[a], n	511	1815	434

[a]Represents posts captured during the study period (June 2019 to December 2021).

Inclusion

All posts made on Team Telomere's social media (Facebook main page: n=1818, Facebook community group: n=518, and Twitter: n=441) between June 2019 and December 2021 were eligible for inclusion. This time frame encompasses the period starting 1 year before the Facebook community group. This group was created in June 2020 as a platform for social connection during the COVID-19 pandemic. Posts were excluded from the analysis if they were (1) removed by the user or Team Telomere (n=5), (2) duplicate posts with identical content from the same day (n=2), or (3) posts without image or text content (n=7). This resulted in a total of 2760 posts, with both primary posts and comments considered unique. The post was used as the unit of analysis and included all content visible to a passive social media user. Additional post content that required clicking links to external sites or embedded audiovisual materials was not included in this study.

Data Extraction and Quality Control

We met with Team Telomere's leadership (eg, executive director and board) before conducting the study and cocreated a community-based research contract outlining parameters. Although all data were publicly available and Facebook data were manually extracted by the authors, Team Telomere facilitated data extraction from Twitter by sharing downloaded images and text files made available to them as account owners. We used the post (original or responses), rather than post creator, as the unit of measurement and did not collect identifying information of the social media users or interact directly with users.

Data were extracted directly from each social media site manually through (1) screenshots saved as deidentified image files and (2) cut-and-paste of post text into an Excel (Microsoft Corporation) spreadsheet. For the Facebook community group, we assigned unique ID numbers to post creators using public data (usernames) to calculate how many unique users engaged in conversation threads, and we viewed the publicly available profile images to assess observed sex and race. Posts were assigned a unique ID number within Excel, and additional data were manually extracted for each post to capture the post popularity (number of likes, shares, and comments), post type (primary post or comment), and types of emojis present. Demographics of post creators (observed gender and race) were assessed through an independent review of profile images and profile names by 3 coders (EP, HR, and NE). Quality control for data extraction was performed on a subset of the data (n=100 posts) by NE, and intercoder reliability was assessed during the multiple-reviewer coding process.

Coding and Analysis

We used a combined content analysis mixed methods approach to analyze the social media data [58]. This involved qualitative analysis (coding by multiple independent reviewers) and quantitative analysis (frequency and chi-square testing). Constructs were defined through codebook development using deductive (theory driven) approaches, whereas qualitative themes were identified through inductive (data driven) discussion, as described in greater detail in the *Methods* section. The analysis was performed separately for each social media source, 2 Facebook pages (the Team Telomere main page and a separate community group page established in 2020) and the Team Telomere Twitter feed, resulting in the creation of 3 separate data sets (Facebook main page: n=1815, Facebook community group: n=511, and Twitter: n=434). A subset of Facebook community group posts (n=77; 12 primary posts and 65 comments) was reviewed by 3 coders and used to inform uncertainty inclusion criteria (Multimedia Appendix 1) and the codebook (Multimedia Appendix 2) developed to deductively identify the presence or absence of uncertainty and social support constructs defined in the Han Taxonomy of Medical Uncertainty [12] and the Social Support Framework [14]. Then, all posts were coded for uncertainty and social support by 3 independent coders (EP, HR, and PKJH), with all disagreements in coding resolved through discussion and consensus. Posts identified as uncertainty related in the Facebook community group (n=156) and Twitter (n=210) were then independently subcoded (EP, HR, and PKJH) for uncertainty issues, sources, and management strategies according to the codebook definitions detailed in the *Measures* section. Data were then arranged by subcode and reviewed qualitatively to detect themes that emerged from the data and were refined through discussion between coders.

Measures

Intercoder Reliability

Intercoder reliability among the 3 coders was measured across all social media types for the initial coding of dichotomous social support and uncertainty variables using Cohen κ. The analysis found acceptable reliability of independent coders in assessing the presence or absence of any social support (κ value range across all platforms, κ=0.79-0.95) and uncertainty (κ value range across all platforms, κ=0.58-0.93) across all social media platforms. Regardless, all discrepancies were mutually resolved through coder consensus.

Post Creator Characteristics

Post creator characteristics were visible from profile images and usernames that appeared alongside each post. Posts from Team Telomere's organizational account were created by staff members, often identified in the post context (eg, executive director, communications director, or board member). We did not scrutinize user profiles to detect the activity of nonhuman bots; however, in the context of the small population with this rare disease, most users could be positively identified as human beings from the context of their posts and history of participation in organizational events. Post creator characteristics, including observed gender and race, were assessed by 3 independent coders' perceptions of publicly available usernames and profile images. Disagreements between coders resulted in the characteristic being coded as "unknown."

Uncertainty Issues, Sources, and Management Strategies

Posts were coded as uncertainty related according to 1 of the following types: (1) uncertainty-related primary posts, (2) uncertainty-related comments, and (3) non–uncertainty-related posts captured within a thread where 1 or more other post was uncertainty related. For the Facebook community group and Twitter, posts identified as uncertainty-related primary posts or comments were further analyzed to determine the presence or absence of sources (ambiguity, complexity, and probability), issues (scientific, personal, and practical), and attributes of uncertainty management strategies (ignorance focused, uncertainty focused, response focused, and person focused). We defined sources of uncertainty as insufficient, unreliable, or contradictory information (ambiguity); information features, such as multiple or interacting causes and effects that make a phenomenon difficult to understand (complexity); and fundamental randomness or indeterminacy of a phenomenon that makes outcomes unpredictable (probability). We defined issues of uncertainty as pertaining to the causes, diagnosis, prognosis, or management of disease (scientific); the impact of disease on aspects of personal life (personal); and logistical issues related to health care or disease management (practical). Although the data did not allow assessment of intent to manage uncertainty, we searched posts to identify evidence of management strategies with ≥1 of the following attributes: (1) providing or seeking information to fill knowledge gaps (ignorance focused), (2) reducing or increasing attention to unknowns to gain or relinquish a sense of control (uncertainty focused), (3) ameliorating the adverse psychological effects of uncertainty (response focused), and (4) fostering interpersonal relationships to engage with uncertainty as a shared experience (person focused).

Social Support

Posts were categorized as containing social support through qualitative coding by 3 independent reviewers (EP, HR, and PKJH) following definitions developed over decades of research in social support theory [14,59,60]. Dichotomous variables were assigned to indicate the presence or absence of social support and the presence or absence of specific types of support within 4 domains (appraisal, emotional, informational, and instrumental). These domains were defined as (1) giving or receiving evaluative feedback (appraisal); (2) giving or receiving indicators of care, love, appreciation, empathy, or sympathy (emotional); (3) giving or receiving knowledge (informational); and (4) giving or receiving tangible support (instrumental), as recently formulated by Holt-Lunstad and Uchino [14]. Assignment to social support domains was not mutually exclusive.

Relationship Between Social Support and Uncertainty

We examined the relationship between social support and uncertainty by comparing frequencies and chi-square tests. Posts were coded as dichotomous variables for uncertainty (uncertainty related, non–uncertainty related), uncertainty subtypes (presence or absence), and social support subtypes (presence or absence). We examined the frequencies of social support subtypes in uncertainty-related posts overall, by social media platform (Facebook community group and Twitter) and by post type (primary post or comment). We performed chi-square tests to determine the strength of the relationship between uncertainty-related posts and social support across platforms and for uncertainty-related posts by post type (primary post, comment, thread) and issue subtype (scientific, personal, practical).

Popularity and Engagement

Popularity on the Facebook community group, Facebook main page, and Twitter was defined as the sum of comments, likes, and shares. Engagement was defined separately for social media types (Facebook community group and Facebook main page vs Twitter) owing to differences in user tracking approaches between Facebook and Twitter platforms. Facebook engagement was defined as the sum of conversations (number of responses generated by a post or comment), voices (number of unique users responding to a post or comment), and depth (number of back-and-forth responses). Engagement on Twitter was defined as the sum of detail expands (clicks to view more of the post), profile visits, link clicks, and video views. Engagement was also measured for the Facebook community group by examining the proportion of users who contributed posts and post frequencies by author.

Sentiment

Sentiment analysis was performed through manual annotation by 2 independent coders, with differences resolved through consensus. Posts were assigned categorical sentiment variables according to the (1) frequency and (2) presence or absence of keywords and emojis. Unambiguous emotion words (eg, “happy” and “sad”) were chosen as keywords to indicate emotional valence, as described in other studies [61,62]. The emotional valence of emojis was assigned based on the emoji definition in internet-based emoji dictionaries and validated by a coder review of the emoji within the post context (Multimedia Appendix 3).

Results

Post Characteristics

A total of 2760 posts created on all platforms between June 2019 and December 2021 were included in this study. Across all platforms, most posts were created either by the executive director of Team Telomere or by individual users who were primarily identified as White, female, and parents of children affected by TBDs. Post characteristics differed by platform: on Twitter, most posts (368/434, 84.8%) were primary posts, most of which (384/434, 88.5%) were generated by the executive director of Team Telomere; Facebook main page posts were either primary posts (800/1815, 44.08%) or first-level comments (1014/1815, 55.87%) created by Team Telomere (860/1815, 47.38%) or individual users (955/1815, 52.62%); and on the Facebook community group, most posts (403/511, 78.9%) were comments to primary posts, in sometimes lengthy (up to 8 level) conversation threads created by 67 individual users (502/511, 98.2%). Posts across all platforms were written almost exclusively in English (Table 2).

Table 2. Characteristics of posts on Team Telomere's social media from June 2019 to December 2021 (N=2760).

	Facebook community group (n=511), n (%)	Facebook main page (n=1815), n (%)	Twitter (n=434), n (%)
Post type			
Primary post	108 (21.1)	800 (44.1)	368 (84.8)
Comment	403 (78.9)	1015 (55.9)	66 (15.2)
Language			
English	487 (95.3)	1807 (99.6)	434 (100)
Other[a]	4 (0.8)	8 (0.4)	0 (0)
Image only	17 (3.3)	0 (0)	0 (0)
Creator type			
Team telomere	8 (1.6)	861 (47.4)	385 (88.7)
Individual	503 (98.4)	954 (52.6)	49 (11)
Observed creator sex[b]			
Male	25 (5)	69 (7.2)	5 (10)
Female	478 (95)	885 (92.8)	41 (83.7)
Unknown	0 (0)	1 (0.1)	3 (6.1)
Observed creator race[b]			
White	443 (88.1)	766 (80.3)	40 (81.6)
Other[c]	46 (9.1)	30 (3.1)	6 (12.2)
Unknown	14 (2.8)	158 (16.6)	3 (6.1)
Observed creator telomere biology disorder relationship[b,d]			
Patient	65 (12.9)	42 (4.4)	1 (2)
Parent	428 (85.1)	384 (40.3)	14 (28.6)
Medical provider	3 (0.6)	31 (3.2)	10 (20.4)
Other[e]	5 (1)	59 (6.2)	22 (44.9)
Unknown	40 (8)	495 (51.9)	2 (4.1)
Multiple	129 (25.6)	126 (13.2)	0 (0)

[a]Respectively by platform (Facebook community group, Facebook main page, and Twitter), "other" language included Spanish (0.2%, 0.2%, and 0%), French (0.4%, 0.1%, and 0%). In the Facebook community group the following languages also appeared: Hebrew (0.1%), Italian (0.1%), Swedish (0.1%), and Māori (0.2%).

[b]Includes frequencies for individual creator types only; does not include Team Telomere organization (Facebook community group: n=503, Facebook main page: n=954, and Twitter: n=49).

[c]Respectively by platform (Facebook community group, Facebook main page, and Twitter), "other" identified creator race and ethnicity included Latinx (7.7%, 1.5%, and 1.4%) and Arab or Middle Eastern (1.4%, 11%, and 0%).

[d]Frequency does not total to 100% because of some individuals occupying multiple categories.

[e]Respectively by platform (Facebook community group, Facebook main page, and Twitter), "other" creator telomere biology disorder relationship included grandparent (0%, 0.2%, and 0%), sibling (0.4%, 0.9%, and 0%), spouse (0%, 0.2%, and 0%), other advocacy organization representative (not Team Telomere; 0%, 0%, and 40.8%), and clinical or pharmaceutical industry representative (0%, 0.1%, and 4.1%).

Qualitative Findings

Qualitative analysis of posts revealed multiple uncertainty issues, sources, and management indicators. Issues included diagnostic, prognostic, therapeutic, and causal uncertainties (scientific); assembly of medical care teams, geographic or financial constraints, and limitations to research funding and dissemination (practical); and building "rare" identity, communicating complex health information to children, and reframing educational or developmental goals (personal). Sources of uncertainty included confusing symptoms and lack of clarity in medical advice (ambiguity); the TBD impact of TBD on multiple organ systems, managing medications or screening regimens, emotional confusion, and achieving scientific literacy across different medical specialties (complexity); and prognostic outcomes, behavioral health risks, or genetic inheritance (probability). Attributes of uncertainty management strategies included (1) information seeking, participation in research, and connection to trusted information sources and care providers (ignorance focused); (2) ordering multiple uncertainties through categorization, prioritization, and sequential narratives, including counting of survival days since transplant (uncertainty focused); (3) sharing positive emotions, portraying TBD experience as a source of strength, and encouraging relaxation (response focused); and (4) promoting a TBD community identity by creating a community mascot (a unicorn named “Tillymere”), recognizing community-specific celebrations (TBD month and transplant anniversaries), providing TBD-pride identifiers (T-shirts and swag), and making reference to Team Telomere as a “family” (person focused; Table 3).

Table 3. Uncertainty in telomere biology disorder (TBD) social media.

	Post text
Sources of uncertainty	
Ambiguity	• “This is a tough one! One of those maybe/maybe not symptoms...I often ask myself the same questions about my daughter’s more obscure symptoms.” [FBCG218304.21.07.30] • “Pre-lung # transplantation patients with # pulmonary # fibrosis who have short # telomeres may need different # clinical care...” [TWT180100.19.06.11]
Complexity	• “[Name] is having kidney, heart, and lung problems. Oh, and who can forget the liver? This week has been too long at the hospital” [FBCG2110000.21.11.23] • “# DYK Those with # telomere biology disorders may be especially vulnerable to the effects of taking multiple medicines at the same time and may respond to medications differently.” [TWT186700.19.11.14]
Probability	• “80% of patients diagnosed with dyskeratosis congenita will experience bone marrow failure.” [TWT185500.19.11.04] • “5 out of 6 of the cell lines tested were less than 1%. And when that’s the case, patients have a 10-20% chance of getting cancer...” [FBCG203500.20.09.08] • “A recent publication advises against an elective eye surgery in patients with DC due to higher long-term risks caused by delayed healing...” [TWT182100.19.08.25]
Issues of uncertainty	
Scientific	• “Has anybody experienced hearing loss with connection to short telomere length?” [FBCG218300.21.07.30] • “Has anyone had kidney problems outside of BMT? Are there any articles anyone has seen on kidneys and short telomeres?” [FBCG2110000.21.11.23]
Practical	• “At the moment [Name] has 1-2 appointments each week. Add to that emails to/from paediatrician, calls from hospital to change/confirm appointments...It’s overwhelming some weeks. And I’m usually doing all this from work. We are also applying for different supports...so lots of forms, phone calls and emails!” [FBCG204305-8.20.10.13]
Personal	• “It’s # PFMonth, and we want you to know you have a team surrounding you...” [TWT1816000.20.09.04] • “TBDs are not just a pediatric disease.Affected adults with a # raredisease, you are NOT ALONE!” [TWT183100.19.09.21] • “Another milestone reached. This time five years ago as we celebrated [Name]’s 5th birthday we were also getting ready to go to transplant two weeks later. Yesterday we celebrated the big 10...” [FBCG201300.20.06.27]
Focus of uncertainty management	
Ignorance	• “Wondering if anyone with DC had a dental implant post-transplant...? And did your medical team have any concerns or recommendations?” [FBCG215500.21.01.05] • “Hello—any contraindications to getting COVID 19 vaccine if you have DC?” [FBCG217100.21.04.04] • “Do you have a copy of the clinical guidelines?” [FBCG203509.20.09.08] • “Take time to learn more about #Telomere Biology Disorders through our informational video!” [TWT1822100.21.11.04]
Uncertainty	• “Each Family Story is set up so you can find a connection via gene or experience.” [FBCG204400.20.10.29] • “My daughter has yearly bone marrow biopsies, lung and liver screenings. ENT and skin checks for cancer.” [FBCG 203513.20.09.08] • “I’ve been preparing something for the new school trying to give them what her medical challenges are.” [FBCG219900.21.11.16]
Response	• “Our family is celebrating today! [Name]’s Happy 8th bone marrow transplant anniversary!” [FBCG203300.20.08.24] • “Fitting for us all: it wasn’t the trauma that made you strong, kinder, and more compassionate. It’s how you handled it. That credit is yours.” [FBCG216200.21.02.28] • “Join@sixnwstevies as she teaches yoga for research...” [TWT1822600.21.03.16]
Person	• “Thank goodness for social media otherwise it would be very isolating.” [FBCG203821.20.09.25] • “Don’t forget to register for our Young Adult Meetup...” [TWT1814300.20.06.23] • “[Name] it’s never ending, I hope you find a way to take care of you” [FBCG204307.20.10.13] • “You are in great hands but always happy to connect with [Provider Name]” [FBCG203504.20.09.08] • “Check out # tillymere! All # sparkly and ready for # TBDmonth!” [TWT185400.19.11.04] • “We have all known the long loneliness and we have learned that the only solution is love and that love comes with community. – Dorothy Day” [TWT1816300.20.09.12]

Uncertainty Issues, Sources, and Management Strategies

Content analysis revealed that 45.98% (1269/2760) of posts overall were uncertainty related, although the frequency differed by platform (Facebook main page: 691/1715, 40.29%; Facebook community group: 155/511, 30.3%; and Twitter: 210/434, 48.4%). Most uncertainty-related posts on Facebook community group and Twitter were generated by Team Telomere's organizational profile (332/511, 65% and 353/434, 81.3%, respectively) and were often similar in topic, wording, and image content. In the Facebook community group, all uncertainty-related posts were generated by individual users, including a portion (119/511, 23.3%) posted by Team Telomere–affiliated volunteer group moderators.

Owing to low frequency of community-generated uncertainty content on the Facebook community group and Twitter, compared with the Facebook community group, we decided to code uncertainty subtypes only within the Facebook community group and Twitter to compare how medical uncertainty was expressed on social media by 2 contrasting content creator groups (community members vs advocacy organization). Scientific uncertainty was the most common issue on both platforms (305/434, 70.3% to 429/511, 84%). On Twitter, personal uncertainty was more frequently discussed, whereas in the Facebook community group, practical uncertainty was more frequent. Across platforms, most posts (1713/2760, 62.07%) had multiple sources of uncertainty, and a substantial number of posts (1126/2760, 40.8%) were coded as emerging from the combined information features of probability, complexity, and ambiguity. The most common attributes of uncertainty management styles detected on both platforms were requests or offers of information to fill knowledge gaps (ignorance focused) and offers of emotional support or community building (person focused). Response-focused management style attributes (eg, yoga and meditation classes) were marginally more frequent on Twitter compared with the Facebook community group (χ^2_1=3.9; *P*=.05), but on the Facebook community group, indicators of uncertainty-focused management (eg, strategies for organization of care logistics) were more frequent compared with Twitter (χ^2_1=55.1; *P*<.001; Table 4).

Table 4. Characteristics and frequency of uncertainty-related posts on Team Telomere's Facebook community group and Twitter (N=2760).

	Facebook community group (n=156), n (%)	Twitter (n=210), n (%)	Chi-square (*df*)[a]	*P* value
Issue				
Personal	48 (30.8)	111 (52.9)	16.6 (1)	<.001
Practical	35 (22.4)	23 (11)	9.2 (1)	.002
Scientific	131 (84)	148 (70.5)	11.4 (1)	.007
Multiple	53 (34)	59 (28.1)	—[a]	—
Source				
Ambiguity	81 (51.9)	80 (38.1)	17.6 (1)	<.001
Complexity	81 (51.9)	75 (35.7)	20.8 (1)	<.001
Probability	112 (71.8)	81 (38.6)	71.3 (1)	<.001
Multiple	88 (56.4)	77 (36.7)	—	—
Management attributes				
Ignorance focused	124 (79.5)	156 (74.3)	1.9 (1)	.16
Person focused	106 (67.9)	125 (59.5)	3.6 (1)	.06
Response focused	57 (36.5)	100 (47.6)	3.9 (1)	.05
Uncertainty focused[b]	53 (34)	10 (4.8)	55.1 (1)	<.001
Multiple	106 (67.9)	131 (62.4)	—	—

[a]Chi-square tests were not performed for issues, sources, or management attributes assigned to multiple categories.

[b]Uncertainty thread includes non–uncertainty-related posts captured in a thread where ≥1 other posts were uncertainty related.

Facebook Social Support and Uncertainty

Frequent overlap of social support and uncertainty was found across all platforms, with uncertainty-related posts being more likely to contain social support compared with non–uncertainty-related posts (χ^2_1=70.7; *P*<.001). However, within social support subtypes, only informational support remained significantly more frequent within uncertainty-related posts (χ^2_1=486.0; *P*<.001), whereas emotional support was significantly less frequent in uncertainty-related posts (χ^2_1=66.5; *P*<.001) compared with non–uncertainty-related posts. The relationship between informational support and uncertainty remained significant for all social media types, but the relationship between emotional support and uncertainty differed by platform (Multimedia Appendix 4). Emotional support was significantly more frequent in uncertainty-related posts for the

Facebook community group (χ^2_1=7.8; *P*=.005), was significantly less frequent in uncertainty-related posts on the Facebook main page (χ^2_1=79.5; *P*<.001), and had no relationship with uncertainty-related posts on Twitter (χ^2_1=0.5; *P*=.47).

On all platforms, uncertainty-related posts were more frequently offers of support than requests. When requests occurred, they were more likely to appear on the Facebook community group compared with Twitter (χ^2_1=12.7; *P*<.001). Posts that were not uncertainty related but appeared in an uncertainty-related thread frequently contained offers of emotional support.

Given the greater variation in types and direction (offer vs request) of social support in the Facebook community group, we decided to focus on subsequent analyses of the relationship between social support and uncertainty subtypes on this platform. Analysis of social support in the Facebook community group posts by uncertainty issue found that informational support was offered more frequently in response to scientific and practical uncertainty posts compared with personal uncertainty posts. Informational support was also the most frequent type of support requested and offered across uncertainty source types in the Facebook community group; however, uncertainty posts emerging from probability concerns had similar frequencies of emotional and informational support (320/511, 62.6% and 511/836, 61.1%, respectively). This was particularly true in the case where a post had multiple uncertainty sources, which were more likely to be coded as informational support offers or requests compared with posts with only a single uncertainty source (χ^2_1=90.4; *P*≤.001).

Popularity and Engagement

Popularity and engagement were positively skewed toward lower values across all social media types. Popularity was highest for posts on Twitter (Facebook community group: median 1, range 0-55, mean 4, SD 7.5; Facebook main page: range 0-151, median 1, mean 5.9, SD 13.3; and Twitter: range 0-1147, median 13, mean 28.8, SD 76.6). However, engagement was higher in the Facebook community group than on the Facebook main page or Twitter (Facebook community group: range 0-29.6, median 0.54, mean 2.15, SD 4.0; Facebook main page: median 0.0006, range 0-0.09, mean 0.004, SD 0.008; and Twitter: median 0.007, range 0-0.56, mean 0.02, SD 0.04). Most uncertainty-related posts were categorized as having below-median popularity and engagement. The uncertainty-related post with the highest engagement was a question about kidney issues and telomere length posted on Facebook community group by a parent of a child with TBDs, which generated 12 comments from 6 unique users, including a self-identified medical expert. The nonnormal distribution combined with low (<20) frequency in cross-tabulation groups made it ineffective to analyze the relationships between the presence of social support and popularity or engagement (Multimedia Appendix 5).

In the Facebook community group, posts were created by 67 unique individuals, representing 35.8% (183/511) of all group members. Frequency per user was positively skewed toward lower numbers (range 1-94 posts and median 3 posts), and the majority of post creators (343/511, 67.1%) generated ≤5 posts. Although Team Telomere rarely posted directly on the Facebook community group (8/511, 1.6% posts), the top 2 post creators (156/511, 30.5% posts) were identified as White, female, parents of children affected by DC who were also group moderators for Team Telomere. After removing the moderators, the remaining median post frequency was 3 posts per user, with 22.3% (114/511) of the users creating only a single post.

Sentiment

The majority of posts (2208/2760, 80%) on all social media types were categorized as positive sentiment. Negative sentiment was rarely expressed and was more likely to be expressed on Facebook compared with Twitter (χ^2_1=45.4; *P*<.001). Uncertainty-related posts demonstrated a similarly high frequency of positive sentiment across all social media types (Facebook community group: 433/511, 84.7%; Facebook main page: 1495/1815, 82.37%; and Twitter: 328/434, 75.6%; Multimedia Appendix 6).

Discussion

Principal Findings

In this study, we explored the use of TBD social media to express health-related uncertainty. We found that uncertainty was a frequent focus of TBD social media across platforms but was primarily limited to scientific issues, requests for informational support, and offers of emotional support, with most posts generated by White, female, English-speaking parents of children with TBDs. These findings are in keeping with other research on rare disease internet-based communities, which found that post content focused on biomedical questions and emotional support provision [63] and was frequently created by White, female users [40,63-65].

The high frequency of uncertainty-related posts on TBD social media created by female caregivers suggests a potentially higher burden of uncertainty management among mothers, which is in agreement with the extensive literature documenting the psychosocial burden of childhood illness on female caregivers [66-68]. However, the observed demographics of TBD social media users may also be an artifact of greater social media engagement among this group, as previous research suggests that female users frequently rely on internet-based communities for navigating uncertainty related to motherhood and other sex-specific health topics [69,70]. Additional research is needed to investigate the relative burden of medical uncertainty among female care providers and to understand the potential barriers to internet-based community formation for users outside this identity group.

Despite the multiplicity of identified uncertainty sources, issues, management, and social support strategies, we found that scientific uncertainty, informational support, and emotional support were the predominant features of uncertainty-related posts on TBD social media. The high frequency of scientific uncertainty issues across platforms suggests that limited scientific and medical knowledge is a salient concern for the TBD community. Gaps in scientific knowledge likely contribute to the focus on probability as a source of uncertainty in TBD social media posts, especially concerning matters such as prognosis, diagnosis, and symptom experiences. Informational support was the most common form of social support in uncertainty-related posts overall, which is in line with other studies showing information seeking as the principal motivator for participation in disease-specific social media [24,26,71-73]. The high frequency of emotional support suggests the potential for TBD social media to enable uncertainty management through person-focused strategies, such as community building, networking, and relationship formation, as seen in other rare disease contexts [24,72]. In addition, evidence of positive asynchronous internet-based communication as a form of "cybertherapy" [32,44] suggests that the emotionally supportive culture of TBD social media may provide psychological benefits for peers, even without explicit conversations about the personal burden of uncertainty. In addition, items coded as emotional support (eg, emoji hearts) that appeared in response to a variety of uncertainty-related content may have communicated multiple forms of support (eg, care, approval, agreement, or affinity) and may be a common reaction to intractable sources of uncertainty, such as probabilistic and scientific unknowns surrounding TBDs. Further exploration of the complex, dynamic, and potentially interactive relationships between social support and uncertainty on social media may be a fruitful area of investigation for future studies.

Given the evidence of the high psychosocial burden of personal uncertainty in similar rare disease contexts [18,36,74,75], it is surprising that the mental and emotional impacts of uncertainty appeared infrequently in TBD social media discussions. When these topics did arise, they were more likely to appear on Twitter content generated by Team Telomere, as opposed to within the conversations of individual users. In the Facebook community group, the impact of uncertainty on personal life was commonly presented in terms of practical issues and focused on ordering uncertainty, such as providing lists of symptoms, organizing information and screening schedules, and triaging problems. This suggests that despite the frequent focus on personal uncertainty issues by Team Telomere, most individual users engaged with TBD social media to troubleshoot and strategize practical issues, rather than to discuss the impact of uncertainty on areas of psychosocial well-being, such as personal identity, goals, or values. This is also reflected in the positive sentiment valence and rare expression of negative emotion on TBD social media, which suggest that social media may not be perceived as a "safe space" for exploring personal topics beyond surface-level stressors [23]. Future research is needed to investigate the shortcomings of social media for expressing personal uncertainty and painful emotions and may highlight a need for psychosocial support to fill this gap in TBD community resources.

Our finding that uncertainty-related support varied by platform could be explained by differences in the structure and expectations of engagement inherent to Twitter compared with the Facebook community group. The predominance of emotional support and greater overall user engagement in the Facebook community group suggests that internet-based platforms structured for mutual conversational exchange may have the most utility for psychosocial support delivery. In addition, the Facebook community group may have encouraged more community participation owing to user familiarity with the platform and its explicit creation for supportive internet-based connection in the context of COVID-19 isolation. Similarly, the nature of the Twitter platform, which is limited to one-way communication streams, suggests that uncertainty management and social support on Twitter would be limited to information provision. However, recent research indicates that Twitter retweets and endorsements may be effective methods for receiving and providing emotional support [76]. The formation of the Facebook community group and the use of Twitter to encourage community activities (eg, webinars and internet-based meetups) underscores the potential of these platforms in person-focused uncertainty management, but additional research is required to evaluate the capacity of TBD social media to build health-promoting personal relationships.

Although we found substantial potential for social media to deliver support for uncertainty management, analysis of engagement rates demonstrated that the primary function of TBD social media was a "drop-in" source of information. Although the Facebook community group included some multilevel, ongoing conversations, an analysis of posts within this group revealed that most user engagement was limited to single posts, suggesting quick check-ins or requests for answers to targeted questions, not ongoing social connection. Although low engagement may suggest limited supportive utility of TBD social media, findings from previous research with young adults with cancer showed that support delivered via social media benefited a variety of users, including those actively seeking deep connections, those seeking information only, and those who do not actively participate but frequently observe the conversation of others (eg, "lurkers") [77]. As suggested by other research, any benefit from engagement with social media likely varies over time and may be most pronounced during experiences of novelty or discrepancy in diagnosis, treatment, or prognosis [28,48,63]. The uncertainty-related post that generated the highest engagement involved the participation of a medical expert, suggesting a desire among TBD social media users to engage with clinicians on internet-based platforms that facilitate reciprocal information exchange, including both synchronous (eg, internet-based group meetings) and asynchronous (eg, post exchanges) formats. Further research is needed to understand the motivations, perceived benefits, and perceived barriers to participation in TBD internet-based support platforms, including the perspectives of patients, caregivers, and medical providers.

Limitations

The limitations of our study include the use of social media data, which biases our sample toward active social media users who may have higher levels of distress [64], greater disenchantment with medical care [78], or lower perceived social support [79] compared with patients with TBDs and their families who do not actively use social media. Demographic analysis revealed that our sample of posts was generated primarily by White females, parents of patients with TBDs, or representatives of Team Telomere. This limited the generalizability of our findings. In addition, our use of social media posts, rather than content creators, as the unit of analysis precludes the observation of the longitudinal impacts of social media participation on uncertainty management. Furthermore, our findings allow us to infer the presence of uncertainty management strategies on social media but not the motivations for or effects of these activities.

In addition, our data were limited to social media that was actively moderated by Team Telomere. This moderation activity, which included removing posts that were inappropriate or scientifically inaccurate, likely decreased the presence of medical misinformation compared with unmoderated social media content. The moderation of posts by Team Telomere could also have impacted the range and authenticity of social and emotional expression owing to social desirability bias. This is in keeping with recent research challenging the assumption that the privacy and anonymity of internet-based environments decreases the likelihood of social desirability compared with in-person interactions [80,81]. In addition, we did not access the private Facebook community group maintained by Team Telomere described as "where we share detailed and private medical information" [57], which may contain additional uncertainty-related posts and a wider range of social and emotional expression. Limiting ourselves to social media owned and maintained by Team Telomere also prevented us from discerning the perspectives of individuals affected by TBD who lacked knowledge of or who chose not to engage with Team Telomere.

Finally, our study was limited by the occurrence of the COVID-19 pandemic, first mentioned in Team Telomere social media on February 28, 2020, which may have changed the nature of uncertainty-related conversations or social support in that portion of our data timeline (June 6, 2019, to December 7, 2021). To test the impact of this, we included available posts (Twitter and Facebook main page) from 1 year before the pandemic and tested the difference. Greater frequencies of uncertainty-related posts after COVID-19 suggest that the pandemic may have increased the expression of uncertainty on TBD-related social media, thus limiting the applicability of our findings to other time points (Multimedia Appendix 7).

Conclusions

This study found the frequent use of disease-specific social media for the discussion and management of uncertainty in TBDs. Uncertainty-related posts appeared across all TBD social media platforms and communicated a burden of multiple, often interacting sources and issues of uncertainty, particularly focused on scientific knowledge gaps and the desire to predict health outcomes. Posts also indicated multiple uncertainty management attributes, with a focus on information-seeking and community-building approaches. Uncertainty-related posts frequently co-occurred with social support, primarily emotional and informational. Female parents were most often the creators of uncertainty-related posts on TBD social media, suggesting a potentially higher burden of uncertainty management in this population. Overall, social media provided access to a positive emotional environment and frequent information exchange but was limited in the type and depth of uncertainty-related discussions. Despite these limitations, our findings suggest that social media is a useful lens for researching and understanding the experience of uncertainty in TBDs and holds potential as a tool for uncertainty management. Future research is needed to further explore the experience of medical uncertainty in TBDs and to determine the usefulness of TBD-related social media as a tool for improving mental health and quality of life outcomes in this context.

Acknowledgments

This study was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute. Katherine Stevens from Team Telomere facilitated our access to the publicly available social media data.

Data Availability

In compliance with the National Institutes of Health data management and sharing policy, data, analysis code, and research materials are available upon reasonable request from the corresponding author.

Authors' Contributions

EP contributed to the study design, formative research, data collection, data analysis, codebook development, coding, and manuscript preparation; HR contributed to data collection, codebook development, and coding; PKJH contributed to the study design, codebook development, coding, and manuscript preparation; MBG, KMR, and AJL contributed to the study design, codebook development, and manuscript preparation; SAS contributed to manuscript preparation, study primary investigator, and National Institutes of Health; and NE contributed to data quality control and coding.

Conflicts of Interest

SAS and HR are members of Team Telomere Advisory Boards.

Multimedia Appendix 1

Criteria for identification of posts for inclusion in qualitative uncertainty analysis.

[DOCX File , 21 KB - infodemiology_v4i1e46693_app1.docx]

Multimedia Appendix 2
Social media study codebook.
[DOCX File , 20 KB - infodemiology_v4i1e46693_app2.docx]

Multimedia Appendix 3
Emoji dictionary.
[DOCX File , 26 KB - infodemiology_v4i1e46693_app3.docx]

Multimedia Appendix 4
Frequency of social support by support type and direction.
[DOCX File , 262 KB - infodemiology_v4i1e46693_app4.docx]

Multimedia Appendix 5
Engagement and popularity by platform.
[DOCX File , 14 KB - infodemiology_v4i1e46693_app5.docx]

Multimedia Appendix 6
Sentiment by post type.
[DOCX File , 16 KB - infodemiology_v4i1e46693_app6.docx]

Multimedia Appendix 7
COVID-19 impact summary.
[DOCX File , 13 KB - infodemiology_v4i1e46693_app7.docx]

References

1. Carpenter DM, DeVellis RF, Fisher EB, DeVellis BM, Hogan SL, Jordan JM. The effect of conflicting medication information and physician support on medication adherence for chronically ill patients. Patient Educ Couns 2010 Nov;81(2):169-176 [FREE Full text] [doi: 10.1016/j.pec.2009.11.006] [Medline: 20044230]
2. Han PK, Moser RP, Klein WM. Perceived ambiguity about cancer prevention recommendations: relationship to perceptions of cancer preventability, risk, and worry. J Health Commun 2006;11 Suppl 1(0 1):51-69 [FREE Full text] [doi: 10.1080/10810730600637541] [Medline: 16641074]
3. Santacroce S. Uncertainty, anxiety, and symptoms of posttraumatic stress in parents of children recently diagnosed with cancer. J Pediatr Oncol Nurs 2002;19(3):104-111. [doi: 10.1177/104345420201900305] [Medline: 12066262]
4. Savage SA. Dyskeratosis congenita and telomere biology disorders. Hematology Am Soc Hematol Educ Program 2022 Dec 09;2022(1):637-648 [FREE Full text] [doi: 10.1182/hematology.2022000394] [Medline: 36485133]
5. Niewisch MR, Giri N, McReynolds LJ, Alsaggaf R, Bhala S, Alter BP, et al. Disease progression and clinical outcomes in telomere biology disorders. Blood 2022 Mar 24;139(12):1807-1819 [FREE Full text] [doi: 10.1182/blood.2021013523] [Medline: 34852175]
6. Fortier MA, Batista ML, Wahi A, Kain A, Strom S, Sender LS. Illness uncertainty and quality of life in children with cancer. J Pediatr Hematol Oncol 2013 Jul;35(5):366-370. [doi: 10.1097/MPH.0b013e318290cfdb] [Medline: 23669725]
7. Hamilton JG, Hutson SP, Moser RP, Kobrin SC, Frohnmayer AE, Alter BP, et al. Sources of uncertainty and their association with medical decision making: exploring mechanisms in Fanconi anemia. Ann Behav Med 2013 Oct;46(2):204-216. [doi: 10.1007/s12160-013-9507-5] [Medline: 23637072]
8. Hill EM, Hamm A. Intolerance of uncertainty, social support, and loneliness in relation to anxiety and depressive symptoms among women diagnosed with ovarian cancer. Psychooncology 2019 Mar;28(3):553-560. [doi: 10.1002/pon.4975] [Medline: 30614141]
9. Kim B, Kim J. Influence of uncertainty, depression, and social support on self-care compliance in hemodialysis patients. Ther Clin Risk Manag 2019 Oct 22;15:1243-1251 [FREE Full text] [doi: 10.2147/TCRM.S218934] [Medline: 31695397]
10. Neville K. The relationships among uncertainty, social support, and psychological distress in adolescents recently diagnosed with cancer. J Pediatr Oncol Nurs 1998 Jan;15(1):37-46. [doi: 10.1177/104345429801500106] [Medline: 9473892]
11. Cohen MH. The triggers of heightened parental uncertainty in chronic, life-threatening childhood illness. Qual Health Res 1995;5(1):63-77. [doi: 10.1177/104973239500500105]
12. Han PK. Uncertainty in Medicine: A Framework for Tolerance. Oxford, UK: Oxford University Press; 2021.
13. Hupcey JE. Clarifying the social support theory-research linkage. J Adv Nurs 1998 Jun;27(6):1231-1241. [doi: 10.1046/j.1365-2648.1998.01231.x] [Medline: 9663875]
14. Holt-Lunstad J, Uchino BN. Social support and health. In: Glanz K, Rimer BK, Viswanath KV, editors. Health Behavior: Theory, Research, and Practice. San Francisco, CA: Jossey-Bass Publishers; 2015:183-204.

15. Zhuo L, Wu Q, Le H, Li H, Zheng L, Ma G, et al. COVID-19-related intolerance of uncertainty and mental health among back-to-school students in Wuhan: the moderation effect of social support. Int J Environ Res Public Health 2021 Jan 22;18(3):981 [FREE Full text] [doi: 10.3390/ijerph18030981] [Medline: 33499409]
16. He X, Zhang Y, Chen M, Zhang J, Zou W, Luo Y. Media exposure to COVID-19 predicted acute stress: a moderated mediation model of intolerance of uncertainty and perceived social support. Front Psychiatry 2021 Feb 10;11:613368. [doi: 10.3389/fpsyt.2020.613368] [Medline: 33643082]
17. Lee I, Park C. The mediating effect of social support on uncertainty in illness and quality of life of female cancer survivors: a cross-sectional study. Health Qual Life Outcomes 2020 May 19;18(1):143 [FREE Full text] [doi: 10.1186/s12955-020-01392-2] [Medline: 32429954]
18. Peters JA, Kenen R, Bremer R, Givens S, Savage SA, Mai PL. Easing the burden: describing the role of social, emotional and spiritual support in research families with Li-Fraumeni syndrome. J Genet Couns 2016 Jun;25(3):529-542. [doi: 10.1007/s10897-015-9905-x] [Medline: 26621765]
19. El Hussein S, Lyapichev KA, Crane GM, Mirza KM, Pemmaraju N, Medeiros LJ, et al. Social media for hematopathologists: medical practice reinvented-#Hemepath. Curr Hematol Malig Rep 2020 Oct;15(5):383-390. [doi: 10.1007/s11899-020-00600-6] [Medline: 33128122]
20. Savage N. Scientists in the Twitterverse. Cell 2015 Jul 16;162(2):233-234 [FREE Full text] [doi: 10.1016/j.cell.2015.06.062] [Medline: 26186181]
21. Pemmaraju N, Utengen A, Gupta V, Thompson MA, Lane AA. Analysis of first-year Twitter metrics of a rare disease community for blastic plasmacytoid dendritic cell neoplasm (BPDCN) on social media: #BPDCN. Curr Hematol Malig Rep 2017 Dec;12(6):592-597. [doi: 10.1007/s11899-017-0422-x] [Medline: 29064025]
22. Brown SA, Daly RP, Duma N, Yang EH, Pemmaraju N, Parwani P, et al. Leveraging social media for cardio-oncology. Curr Treat Options Oncol 2020 Aug 13;21(10):83. [doi: 10.1007/s11864-020-00775-3] [Medline: 32789716]
23. Benson JJ, Oliver DP, Washington KT, Rolbiecki AJ, Lombardo CB, Garza JE, et al. Online social support groups for informal caregivers of hospice patients with cancer. Eur J Oncol Nurs 2020 Feb;44:101698 [FREE Full text] [doi: 10.1016/j.ejon.2019.101698] [Medline: 31816508]
24. Barton KS, Wingerson A, Barzilay JR, Tabor HK. "Before Facebook and before social media…we did not know anybody else that had this": parent perspectives on internet and social media use during the pediatric clinical genetic testing process. J Community Genet 2019 Jul;10(3):375-383 [FREE Full text] [doi: 10.1007/s12687-018-0400-6] [Medline: 30569339]
25. Dizon DS, Graham D, Thompson MA, Johnson LJ, Johnston C, Fisch MJ, et al. Practical guidance: the use of social media in oncology practice. J Oncol Pract 2012 Sep;8(5):e114-e124 [FREE Full text] [doi: 10.1200/JOP.2012.000610] [Medline: 23277774]
26. Iyer AA, Barzilay JR, Tabor HK. Patient and family social media use surrounding a novel treatment for a rare genetic disease: a qualitative interview study. Genet Med 2020 Nov;22(11):1830-1837 [FREE Full text] [doi: 10.1038/s41436-020-0890-6] [Medline: 32601388]
27. Gallagher L, McCuaig J, Benoit L, Davies C. It's time for the genetic counseling profession to embrace social media. J Genet Couns 2016 Dec;25(6):1338-1341. [doi: 10.1007/s10897-016-9950-0] [Medline: 27026235]
28. Smedley RM, Coulson NS. Genetic testing for Huntington's disease: a thematic analysis of online support community messages. J Health Psychol 2021 Mar;26(4):580-594. [doi: 10.1177/1359105319826340] [Medline: 30696276]
29. Westmaas JL, Fallon E, McDonald BR, Driscoll D, Richardson K, Portier K, et al. Investigating relationships among cancer survivors' engagement in an online support community, social support perceptions, well-being, and moderating effects of existing (offline) social support. Support Care Cancer 2020 Aug;28(8):3791-3799. [doi: 10.1007/s00520-019-05193-2] [Medline: 31828494]
30. Rising CJ, Bol N, Burke-Garcia A, Rains S, Wright KB. Perceived stress in online prostate cancer community participants: examining relationships with stigmatization, social support network preference, and social support seeking. J Health Commun 2017 Jun;22(6):469-476. [doi: 10.1080/10810730.2017.1304471] [Medline: 28414570]
31. Rimer BK, Lyons EJ, Ribisl KM, Bowling JM, Golin CE, Forlenza MJ, et al. How new subscribers use cancer-related online mailing lists. J Med Internet Res 2005 Jul 01;7(3):e32 [FREE Full text] [doi: 10.2196/jmir.7.3.e32] [Medline: 15998623]
32. Lorence D. Examining online chat within a domain of uncertainty: the case of Asperger's syndrome. Health Info Libr J 2007 Jun;24(2):128-136 [FREE Full text] [doi: 10.1111/j.1471-1842.2007.00715.x] [Medline: 17584216]
33. Scott AM, Martin SC, Stone AM, Brashers DE. Managing multiple goals in supportive interactions: using a normative theoretical approach to explain social support as uncertainty management for organ transplant patients. Health Commun 2011;26(5):393-403. [doi: 10.1080/10410236.2011.552479] [Medline: 21409670]
34. Sammarco A. Perceived social support, uncertainty, and quality of life of younger breast cancer survivors. Cancer Nurs 2001 Jun;24(3):212-219. [Medline: 11409065]
35. Li X, He L, Wang J, Wang M. Illness uncertainty, social support, and coping mode in hospitalized patients with systemic lupus erythematosus in a hospital in Shaanxi, China. PLoS One 2019 Feb 21;14(2):e0211313 [FREE Full text] [doi: 10.1371/journal.pone.0211313] [Medline: 30789919]

36. Forbes Shepherd R, Werner-Lin A, Keogh LA, Delatycki MB, Forrest LE. "I need to know if I'm going to die young": adolescent and young adult experiences of genetic testing for Li-Fraumeni syndrome. J Psychosoc Oncol 2021;39(1):54-73. [doi: 10.1080/07347332.2020.1768199] [Medline: 32449501]
37. Eche IJ, Aronowitz T. Factors that influence parental uncertainty and health-related quality of life in children with cancer: a framework. Nurs Sci Q 2018 Oct;31(4):362-368. [doi: 10.1177/0894318418792896] [Medline: 30223741]
38. Oprescu F, Campo S, Lowe J, Andsager J, Morcuende JA. Managing uncertainty in the context of clubfoot care: exploring the value of uncertainty management theory and the sense of virtual community. Iowa Orthop J 2013;33:142-148 [FREE Full text] [Medline: 24027474]
39. Miller EG, Woodward AL, Flinchum G, Young JL, Tabor HK, Halley MC. Opportunities and pitfalls of social media research in rare genetic diseases: a systematic review. Genet Med 2021 Dec;23(12):2250-2259 [FREE Full text] [doi: 10.1038/s41436-021-01273-z] [Medline: 34282302]
40. Fogel J, Ribisl KM, Morgan PD, Humphreys K, Lyons EJ. Underrepresentation of African Americans in online cancer support groups. J Natl Med Assoc 2008 Jun;100(6):705-712. [doi: 10.1016/s0027-9684(15)31346-8] [Medline: 18595573]
41. 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 Jul 11;114(7):1036-1039 [FREE Full text] [doi: 10.1093/jnci/djab141] [Medline: 34291289]
42. Chrousos GP, Mentis AF. Medical misinformation in mass and social media: an urgent call for action, especially during epidemics. Eur J Clin Invest 2020 May;50(5):e13227. [doi: 10.1111/eci.13227] [Medline: 32294232]
43. 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]
44. Beaudoin CE, Tao CC. Benefiting from social capital in online support groups: an empirical study of cancer patients. Cyberpsychol Behav 2007 Aug;10(4):587-590 [FREE Full text] [doi: 10.1089/cpb.2007.9986] [Medline: 17711369]
45. Bloom RD, Reblin M, Chou WY, Beck SL, Wilson A, Ellington L. Online social support for cancer caregivers: alignment between requests and offers on CaringBridge. J Psychosoc Oncol 2021;39(1):118-134. [doi: 10.1080/07347332.2020.1806174] [Medline: 32835649]
46. Coulson NS, Greenwood N. Families affected by childhood cancer: an analysis of the provision of social support within online support groups. Child Care Health Dev 2012 Nov;38(6):870-877. [doi: 10.1111/j.1365-2214.2011.01316.x] [Medline: 21916931]
47. LaCoursiere SP. A theory of online social support. ANS Adv Nurs Sci 2001 Sep;24(1):60-77. [doi: 10.1097/00012272-200109000-00008] [Medline: 11554534]
48. Mikal JP, Beckstrand MJ, Parks E, Oyenuga M, Odebunmi T, Okedele O, et al. Online social support among breast cancer patients: longitudinal changes to Facebook use following breast cancer diagnosis and transition off therapy. J Cancer Surviv 2020 Jun;14(3):322-330. [doi: 10.1007/s11764-019-00847-w] [Medline: 31897878]
49. DeHoff BA, Staten LK, Rodgers RC, Denne SC. The role of online social support in supporting and educating parents of young children with special health care needs in the United States: a scoping review. J Med Internet Res 2016 Dec 22;18(12):e333 [FREE Full text] [doi: 10.2196/jmir.6722] [Medline: 28007689]
50. Ye Z, Li W, Zhu R. Online psychosocial interventions for improving mental health in people during the COVID-19 pandemic: a systematic review and meta-analysis. J Affect Disord 2022 Nov 01;316:120-131 [FREE Full text] [doi: 10.1016/j.jad.2022.08.023] [Medline: 35970325]
51. Hong Y, Peña-Purcell NC, Ory MG. Outcomes of online support and resources for cancer survivors: a systematic literature review. Patient Educ Couns 2012 Mar;86(3):288-296. [doi: 10.1016/j.pec.2011.06.014] [Medline: 21798685]
52. Politi MC, Han PK, Col NF. Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making 2007;27(5):681-695. [doi: 10.1177/0272989X07307270] [Medline: 17873256]
53. Team Telomere homepage. Team Telomere. URL: https://teamtelomere.org/ [accessed 2023-12-20]
54. Facca D, Smith MJ, Shelley J, Lizotte D, Donelle L. Exploring the ethical issues in research using digital data collection strategies with minors: a scoping review. PLoS One 2020 Aug 27;15(8):e0237875 [FREE Full text] [doi: 10.1371/journal.pone.0237875] [Medline: 32853218]
55. Team Telomere Twitter. Twitter. URL: https://twitter.com/teamtelomereinc?lang=en [accessed 2023-12-20]
56. Team Telomere. Facebook. URL: https://www.facebook.com/teamtelomere/ [accessed 2023-12-20]
57. Team telomere community. Facebook. URL: https://www.facebook.com/groups/244926586962809/ [accessed 2023-12-22]
58. Hamad EO, Savundranayagam MY, Holmes JD, Kinsella EA, Johnson AM. Toward a mixed-methods research approach to content analysis in the digital age: the combined content-analysis model and its applications to health care Twitter feeds. J Med Internet Res 2016 Mar 08;18(3):e60 [FREE Full text] [doi: 10.2196/jmir.5391] [Medline: 26957477]
59. Cohen LH, McGowan J, Fooskas S, Rose S. Positive life events and social support and the relationship between life stress and psychological disorder. Am J Community Psychol 1984 Oct;12(5):567-587. [doi: 10.1007/BF00897213] [Medline: 6496413]
60. Cutrona C, Russell D, Rose J. Social support and adaptation to stress by the elderly. Psychol Aging 1986;1(1):47-54 [FREE Full text] [doi: 10.1037/0882-7974.1.1.47]

61. Gendron M, Lindquist KA, Barsalou L, Barrett LF. Emotion words shape emotion percepts. Emotion 2012 Apr;12(2):314-325 [FREE Full text] [doi: 10.1037/a0026007] [Medline: 22309717]
62. Barrett LF, Lindquist KA, Gendron M. Language as context for the perception of emotion. Trends Cogn Sci 2007 Aug;11(8):327-332 [FREE Full text] [doi: 10.1016/j.tics.2007.06.003] [Medline: 17625952]
63. Lasker JN, Sogolow ED, Sharim RR. The role of an online community for people with a rare disease: content analysis of messages posted on a primary biliary cirrhosis mailinglist. J Med Internet Res 2005 Mar 31;7(1):e10 [FREE Full text] [doi: 10.2196/jmir.7.1.e10] [Medline: 15829472]
64. Bender JL, Hueniken K, Eng L, Brown MC, Kassirian S, Geist I, et al. Internet and social media use in cancer patients: association with distress and perceived benefits and limitations. Support Care Cancer 2021 Sep;29(9):5273-5281. [doi: 10.1007/s00520-021-06077-0] [Medline: 33651181]
65. Claus EB, Feliciano J, Benz LS, Calvocoressi L. Social media partnerships with patient organizations for neuro-oncology patient recruitment. Neurooncol Pract 2020 Mar;7(2):143-151 [FREE Full text] [doi: 10.1093/nop/npz049] [Medline: 32626583]
66. Thompson RJ, Gil KM, Gustafson KE, George LK, Keith BR, Spock A, et al. Stability and change in the psychological adjustment of mothers of children and adolescents with cystic fibrosis and sickle cell disease. J Pediatr Psychol 1994 Apr;19(2):171-188. [doi: 10.1093/jpepsy/19.2.171] [Medline: 8051601]
67. Cai RY, Uljarević M, Leekam SR. Predicting mental health and psychological wellbeing in mothers of children with autism spectrum disorder: roles of intolerance of uncertainty and coping. Autism Res 2020 Oct;13(10):1797-1801. [doi: 10.1002/aur.2341] [Medline: 32618140]
68. Roorda D, van der Steeg AF, van Dijk M, Derikx JP, Gorter RR, Rotteveel J, et al. Distress and post-traumatic stress in parents of patients with congenital gastrointestinal malformations: a cross-sectional cohort study. Orphanet J Rare Dis 2022 Sep 11;17(1):353 [FREE Full text] [doi: 10.1186/s13023-022-02502-7] [Medline: 36089585]
69. Wright KB, Cai X, Fisher C, Rising CJ, Burke-Garcia A, Afanaseva D. A content analysis of social support messages about environmental breast cancer risk within blogs for mothers. Health Commun 2021 Nov;36(13):1796-1804 [FREE Full text] [doi: 10.1080/10410236.2020.1800241] [Medline: 32744079]
70. Wagg AJ, Callanan MM, Hassett A. Online social support group use by breastfeeding mothers: a content analysis. Heliyon 2019 Mar 09;5(3):e01245 [FREE Full text] [doi: 10.1016/j.heliyon.2019.e01245] [Medline: 30906891]
71. Jacobs R, Boyd L, Brennan K, Sinha CK, Giuliani S. The importance of social media for patients and families affected by congenital anomalies: a Facebook cross-sectional analysis and user survey. J Pediatr Surg 2016 Nov;51(11):1766-1771. [doi: 10.1016/j.jpedsurg.2016.07.008] [Medline: 27522307]
72. Khouri JS, McCheyne MJ, Morrison CS. #Cleft: the use of social media amongst parents of infants with clefts. Cleft Palate Craniofac J 2018 Aug;55(7):974-976. [doi: 10.1597/16-156] [Medline: 28085513]
73. Pemmaraju N, Gupta V, Mesa R, Thompson MA. Social media and myeloproliferative neoplasms (MPN)--focus on Twitter and the development of a disease-specific community: #MPNSM. Curr Hematol Malig Rep 2015 Dec;10(4):413-420 [FREE Full text] [doi: 10.1007/s11899-015-0287-9] [Medline: 26411990]
74. Ross J, Bojadzieva J, Peterson S, Noblin SJ, Yzquierdo R, Askins M, et al. The psychosocial effects of the Li-Fraumeni education and early detection (LEAD) program on individuals with Li-Fraumeni syndrome. Genet Med 2017 Sep;19(9):1064-1070 [FREE Full text] [doi: 10.1038/gim.2017.8] [Medline: 28301458]
75. Young JL, Pantaleao A, Zaspel L, Bayer J, Peters JA, Khincha PP, et al. Couples coping with screening burden and diagnostic uncertainty in Li-Fraumeni syndrome: connection versus independence. J Psychosoc Oncol 2019;37(2):178-193 [FREE Full text] [doi: 10.1080/07347332.2018.1543376] [Medline: 30591002]
76. Esener Y, McCall T, Lakdawala A, Kim H. Seeking and providing social support on Twitter for trauma and distress during the COVID-19 pandemic: content and sentiment analysis. J Med Internet Res 2023 Aug 31;25:e46343 [FREE Full text] [doi: 10.2196/46343] [Medline: 37651178]
77. Lazard AJ, Collins MK, Hedrick A, Horrell LN, Varma T, Love B, et al. Initiation and changes in use of social media for peer support among young adult cancer patients and survivors. Psychooncology 2021 Nov;30(11):1859-1865. [doi: 10.1002/pon.5758] [Medline: 34165848]
78. Dolce MC. The Internet as a source of health information: experiences of cancer survivors and caregivers with healthcare providers. Oncol Nurs Forum 2011 May;38(3):353-359. [doi: 10.1188/11.ONF.353-359] [Medline: 21531685]
79. Kim SC, Shah DV, Namkoong K, McTavish FM, Gustafson DH. Predictors of online health information seeking among women with breast cancer: the role of social support perception and emotional well-being. J Comput Mediat Commun 2013 Jan;18(2):98-118 [FREE Full text] [doi: 10.1111/jcc4.12002] [Medline: 24634575]
80. Dodou D, de Winter JC. Social desirability is the same in offline, online, and paper surveys: a meta-analysis. Comput Hum Behav 2014 Jul;36:487-495. [doi: 10.1016/j.chb.2014.04.005]
81. Massara F, Ancarani F, Costabile M, Ricotta F. Social desirability in virtual communities. Int J Bus Adm 2012 Oct;3(6):93-100. [doi: 10.5430/ijba.v3n6p93]

Abbreviations

DC: dyskeratosis congenita
TBD: telomere biology disorder

Edited by T Mackey; submitted 21.02.23; peer-reviewed by M Beckstrand, R Parker, J Chen, L Zheng; comments to author 30.09.23; revised version received 09.11.23; accepted 28.11.23; published 15.01.24.

Please cite as:
Pearce E, Raj H, Emezienna N, Gilkey MB, Lazard AJ, Ribisl KM, Savage SA, Han PKJ
The Use of Social Media to Express and Manage Medical Uncertainty in Dyskeratosis Congenita: Content Analysis
JMIR Infodemiology 2024;4:e46693
URL: https://infodemiology.jmir.org/2024/1/e46693
doi:10.2196/46693
PMID:38224480

Original Paper

The Role of Social Media in Knowledge, Perceptions, and Self-Reported Adherence Toward COVID-19 Prevention Guidelines: Cross-Sectional Study

Camryn Garrett[1], MPH; Shan Qiao[1], MA, PhD; Xiaoming Li[1], PhD

[1]Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, The University of South Carolina, Columbia, SC, United States

Corresponding Author:
Camryn Garrett, MPH
Department of Health Promotion, Education, and Behavior
Arnold School of Public Health
The University of South Carolina
Discovery Building I
915 Greene Street
Columbia, SC, 29208
United States
Phone: 1 803 777 6844
Fax: 1 803 777 6290
Email: camrynmg@email.sc.edu

Abstract

Background: Throughout the COVID-19 pandemic, social media has served as a channel of communication, a venue for entertainment, and a mechanism for information dissemination.

Objective: This study aims to assess the associations between social media use patterns; demographics; and knowledge, perceptions, and self-reported adherence toward COVID-19 prevention guidelines, due to growing and evolving social media use.

Methods: Quota-sampled data were collected through a web-based survey of US adults through the Qualtrics platform, from March 15, 2022, to March 23, 2022, to assess covariates (eg, demographics, vaccination, and political affiliation), frequency of social media use, social media sources of COVID-19 information, as well as knowledge, perceptions, and self-reported adherence toward COVID-19 prevention guidelines. Three linear regression models were used for data analysis.

Results: A total of 1043 participants responded to the survey, with an average age of 45.3 years, among which 49.61% (n=515) of participants were men, 66.79% (n=696) were White, 11.61% (n=121) were Black or African American, 13.15% (n=137) were Hispanic or Latino, 37.71% (n=382) were Democrat, 30.21% (n=306) were Republican, and 25% (n=260) were not vaccinated. After controlling for covariates, users of TikTok (β=–.29, 95% CI –0.58 to –0.004; *P*=.047) were associated with lower knowledge of COVID-19 guidelines, users of Instagram (β=–.40, 95% CI –0.68 to –0.12; *P*=.005) and Twitter (β=–.33, 95% CI –0.58 to –0.08; *P*=.01) were associated with perceiving guidelines as strict, and users of Facebook (β=–.23, 95% CI –0.42 to –0.043; *P*=.02) and TikTok (β=–.25, 95% CI –0.5 to -0.009; *P*=.04) were associated with lower adherence to the guidelines (R^2 0.06-0.23).

Conclusions: These results allude to the complex interactions between online and physical environments. Future interventions should be tailored to subpopulations based on their demographics and social media site use. Efforts to mitigate misinformation and implement digital public health policy must account for the impact of the digital landscape on knowledge, perceptions, and level of adherence toward prevention guidelines for effective pandemic control.

(JMIR Infodemiology 2024;4:e44395) doi:10.2196/44395

KEYWORDS

COVID-19; digital media; social media; TikTok; Instagram; Twitter; Facebook; prevention guidelines

Introduction

In March 2020, the infectious disease SARS-CoV-2, more commonly known as COVID-19, was classified as a pandemic [1,2]. As the virus is transmitted through the respiratory systems of individuals in close contact, preventative measures include wearing a facial mask, social distancing, and receiving recommended COVID-19 vaccinations [3]. Over the course of the pandemic, prevention recommendations changed in response to emerging scientific evidence. Initially, a 14-day quarantine and isolation were recommended, which was then shortened to 10 days, and was once more shortened to 5 days [3]. As of March 2022, masks were still recommended in indoor spaces, COVID-19 vaccinations and boosters were widely available, and rapid self-testing was advised in response to exposure or symptom onset [3]. In the United States, as of November 2, 2022, there have been over 97 million confirmed cases and over 1 million total deaths due to COVID-19 [4]. Despite these prevention recommendations, case numbers continued to rise, necessitating research into prevention efforts.

In response to social distancing recommendations, many aspects of life shifted from physical to online environments. Adapting to this change, most US adults (ie, 90%) indicated that digital media was either essential or important for them throughout the pandemic [5]. Digital media encapsulates social media as the platforms that enable human connection in the online environment, with varying degrees of privacy [6]. On social media, individuals encounter and consume information, government announcements, and reactions from other users as they work, learn, connect, and are entertained online [7]. Popular social media sites include Facebook, Twitter, Instagram, Snapchat, TikTok, Pinterest, Reddit, and LinkedIn, among others. As of 2021, a total of 72% of adults in the United States report using at least 1 social media site, representing a 3% increase since 2018 [8]. When stratified by age, 84% of US adults aged 18-29 years indicate using at least 1 social media site [8]. Of those who use Facebook, Snapchat, and Instagram, a majority indicate visiting the platform at least once a day [9]. In considering news consumption on social media, when stratified by age, 42% of users aged 18-29 years indicate social media as their primary source of news [9].

With an increasing proportion of individuals active on social media, thereby encountering COVID-19 news and information online, there are concerns about information accuracy, where unsourced or false information that is widely distributed threatens the dissemination of scientifically accurate information [7,10]. The modalities of social media (eg, concise, organized content formats, and sharing capabilities) allow information to quickly trend as a result of high engagement. The visibility of trending content on social media is determined by engagement and is often based on sensationalism rather than factual accuracy [7]. Sensational misinformation risks reducing the visibility and reach of reputable information [7]. Due to the saturation of misinformation online, the United States is understood to be in a syndemic, denoting the interactions between the COVID-19 pandemic and the infodemic. Social media, therefore, has the capacity to serve both as a tool and a hindrance to health communication.

Despite motivations for use, social media users are subject to unintentionally overconsuming content related to COVID-19 due to the saturation of pandemic information online. Social media has been preliminarily found to negatively contribute to COVID-19 prevention guideline adherence [11]. Among US adults, 53.3% indicate that the amount of information on COVID-19 is overwhelming to the effect that 54.7% indicate that it has led to their avoidance of consuming information about COVID-19 [12]. Resembling emerging trends in the United States, a study in Turkey indicated that 34.4% of respondents follow COVID-19 guidelines less in the present than at the beginning of the pandemic [13]. Fluctuations in pandemic prevention perceptions and adherence over time can be expected, but negative trends, regardless of their cause, necessitate investigation and intervention to bolster commitment to prevention guidelines to limit further pandemic-related exposures [13]. Although a complicated mechanism with additionally probable explanations (eg, milder virus mutations, vaccination availability, mental health burdens, and pandemic fatigue), these downward patterns of adherence are thought to be partially explained by social media use (eg, misinformation and overconsumption). The effective dissemination of scientific, evidence-based health communication must be prioritized in stark opposition to skepticism and disbelief, as sustained by misinformation.

There exists a limited understanding of the associations between demographics and frequency of social media site use and engagement with pandemic prevention behaviors, despite the significant risks to public health. Therefore, there is a present and pressing need to address the field's limited understanding of pandemic-related knowledge, perceptions, and adherence, as impacted online, to design effective health behavior and communication interventions. As the emerging literature demonstrates that content consumption impacts perceptions and, subsequently, health behaviors, the field of health communication must understand the compounding effects of the online environment on COVID-19 prevention efforts [7]. This study therefore aims to investigate the associations between the social media platforms from which individuals consume pandemic-related information as well as their frequency of use and their knowledge of, perceptions of, and adherence to COVID-19 prevention guidelines.

Methods

Survey Development and Data Collection

Preliminary development of the survey involved compiling constructs related to the topics of interest. Survey items were then drafted to measure participant knowledge, perceptions, and adherence toward COVID-19 prevention guidelines. The items were then reviewed by an expert to evaluate and ensure readability, applicability, and response options. The data were obtained using a web-based survey fielded using Qualtrics paid, opt-in distribution services. The data were collected from March 15, 2022, to March 23, 2022.

Ethical Considerations

The University of South Carolina's Institutional Review Board exempted the study (Pro00119512) from Human Research

Subject Regulations based on its minimal risk to participants in providing web-based survey responses. Informed consent was obtained from all participants prior to survey completion. All participants were compensated for their time and efforts in completing the survey (ie, US $6).

Sample

All adults in the United States were eligible for participation, given that they were 18 years or older at the time of survey response. Responses that were deemed low quality based on response speed, lack of variability in selection, or repetitive attempts were removed before analysis to ensure data quality. Qualtrics used quota sampling methods to ensure the collection of a sample proportionate to that of the United States by way of gender, age, income, race, ethnicity, and education level. The final sample size included 1043 viable responses.

Measures

Demographics

Participant demographics collected included age, gender identity, race or ethnicity, education, employment, income, political affiliation, and COVID-19 vaccination status. Due to limited representation, the American Indian or Alaska Native and Native Hawaiian or Pacific Islander categories were collapsed into 1 category. Age, education, employment, and income were used as continuous variables in the regression models. Gender identity, race or ethnicity, political affiliation, and COVID-19 vaccination status were used as categorical variables in the regression models.

Frequency of Social Media Use

Participants' frequency of any social media use was measured through the item: "About how often do you use social media sites?" Response options ranged from several times a day, once per day, a few times per week, once per week, less than once per week, to never.

Social Media Sources of COVID-19 Information

Participants were asked to check all that apply to the question, "Which of these social media sites have you used to get information about COVID-19?" with the possible response options of Facebook, Twitter, Instagram, Snapchat, Pinterest, TikTok, Reddit, LinkedIn, and another social media site. The social media sites available as response options were chosen due to their popularity and presentation of short-form, user-generated content. Although there exist additional social media platforms (eg, YouTube), those chosen to be included here have active engagement and content sharing capabilities. Demographic profiles of the included social media sites were not accounted for in participant sampling procedures, as it is assumed that user bases may have fluctuated during the pandemic. The selections of these sites were operationalized as categorical predictors in the regression models.

Knowledge of COVID-19 Guidelines

Set forth by the Centers for Disease Control and Prevention, as of March 2022, relevant COVID-19 guidelines were used in crafting 4 items to assess participant pandemic-related knowledge. The assessment evaluated respondents' knowledge of calculating exposure date, the minimum length of isolation after an exposure or positive test, the percentage of alcohol in hand sanitizer required to kill COVID-19, and what a negative rapid test result indicates. Participants were asked to indicate what they believe the current, official recommendations to be, at the time of survey administration, rather than what they may prefer them to be. These 4 items were then compiled for a final score out of 100%. Knowledge scores of the COVID-19 prevention guidelines were used continuously in the regression models.

Perceptions of COVID-19 Guidelines

Participants were asked to indicate the degree to which they perceived COVID-19 prevention guidelines to be relaxed or strict. The terminology "strict" was operationalized through concurrent dimensions that encapsulate participant responses to legal and scientific guidelines as well as enforcement. As perceptions of COVID-19 guidelines were assessed after the knowledge assessment, the guidelines were not explicitly defined but rather assumed to encapsulate mask-wearing, gathering size limitations, hygiene measures, as well as quarantine and isolation timelines. This ordering provided participants with context as to what the term "guidelines" referred to. Participants were asked: "Do you consider the current COVID-19 guidelines as:" with the response options ranging from too strict, a little too strict, about right, a little too relaxed, to too relaxed.

Adherence to COVID-19 Guidelines

Adherence to COVID-19 guidelines was evaluated by asking participants if they generally follow the official COVID-19 prevention guidelines, with the available response options of strongly, sometimes, rarely, and never follow the guidelines. This item provided an average, typical measure of self-reported participant adherence to COVID-19 guidelines, broadly. Given the state of the pandemic, this item was reliant upon participant understanding of guidelines in the organizations and institutions to which they belong (ie, schools and workplaces).

Statistical Analysis

All statistical analyses were conducted using the statistical analysis software, SAS (version 9.4; SAS Institute). Descriptive analyses were conducted for key predictors. All data were screened for outliers, missing data, and normality. As all data used in this study was collected through discrete response options, excluding age, their distributions were considered to assess the presence of outliers. This was done by considering the frequency of responses within available options through histograms and box plots, as applicable. Those categories that were lower in response volume were collapsed (eg, race or ethnicity response of American Indian or Alaska Native and Native Hawaiian or Pacific Islander) or excluded from the analysis before modeling (eg, gender identity response option of nonbinary). Data quality was ensured as Qualtrics excluded participants who did not complete the survey in a single session, who were not continuously and carefully responding, who missed embedded attention checks, or who completed the survey in less than a third or more than 3 times the median time it took other participants to complete the survey. Due to the use of these features, respondents who did not complete the survey were not tracked. No systematic patterns of missing data within the data collected, or between variables, were observed. There is limited item nonresponse. Bivariate associations were assessed through ANOVA and Pearson correlation tests, as appropriate. Three generalized linear regressions, using a maximum likelihood estimation procedure, were conducted, independently, to explore associations between social media use and demographics and knowledge, perceptions, and self-reported adherence toward prevention guidelines, respectively. Although the 3 outcomes of knowledge, perceptions, and self-reported adherence were run independently, their theoretically dependent nature led us to consider implementing a correction (ie, Bonferroni), but as it resulted in a minimal impact on our findings, the traditional α level of .05 was here used to evaluate our findings.

Results

Overview

Of the 1043 participants, the median age of participants was 45.3 years (Table 1). The distribution of the gender identity of the participants was split approximately equally between men (515/1032, 49.9%) and women (513/1032, 49.71%), with few participants indicating being nonbinary or transgender. The race or ethnicity of participants was primarily White (696/1042, 66.79%), followed by Latino or Hispanic (137/1042, 13.15%) and Black or African American (121/1042, 11.61%). A quarter (253/1042, 24.28%) of participants held a bachelor's degree and approximately a quarter (269/1042, 25.82%) of participants indicated earning US $50,000-US $79,999 annually. Finally, almost half (498/1040, 47.88%) of the participants had received a full vaccination series and booster against COVID-19.

Table 1. Demographic characteristics of study participants (N=1043).

Variables	Values, n (%)
Age (years;1 participant's data are missing), mean (SD)	45.3 (16.94)
Gender (11 participants' data are missing)	
Men	515 (49.9)
Women	513 (49.71)
Nonbinary or other	4 (0.39)
Race or ethnicity (check all that apply; 1 participant's data are missing)	
Black or African American	121 (11.61)
Latino or Hispanic	137 (13.15)
American Indian or Alaska Native and Native Hawaiian or Pacific Islander	22 (2.11)
White	696 (66.79)
Other	66 (6.33)
Education (1 participant's data are missing)	
Less than high school degree	25 (2.4)
High school graduate or equivalent	248 (23.8)
Some college but no degree	248 (23.8)
Associate degree	123 (11.8)
Bachelor's degree	253 (24.28)
Master's degree	112 (10.75)
Doctoral or professional degree (JD, MD, or PhD)	33 (3.17)
Employment status over the last 3 months (6 participant's data are missing)	
Working full-time	499 (48.12)
Working part-time	132 (12.73)
Unemployed and looking for work	74 (7.14)
Homemaker or stay-at-home parent	70 (6.75)
Student	35 (3.38)
Retired	200 (19.29)
Other	27 (2.6)
Previous year income (US $; 1 participant's data are missing)	
Less than 10,000	56 (5.37)
10,000-19,999	58 (5.57)
20,000-29,999	96 (9.21)
30,000-39,999	87 (8.35)
40,000-49,999	70 (6.72)
50,000-59,000	117 (11.23)
60,000-69,999	70 (6.72)
70,000-79,999	82 (7.87)
80,000-89,999	47 (4.51)
90,000-99,999	51 (4.89)
100,000-149,999	215 (20.63)
150,000 or more	93 (8.93)
Political affiliation (30 participants' data are missing)	
Republican	306 (30.21)

Variables	Values, n (%)
Democrat	382 (37.71)
Independent	325 (32.08)
COVID-19 vaccination status (3 participant's data are missing)	
No	260 (25)
Yes, but no booster	282 (27.12)
Yes, including booster	498 (47.88)

Social Media Site Use

Participants reported using, generally or for any reason, the social media sites Facebook (835/1042, 80.13%), Twitter (396/1042, 38%), Instagram (586/1042, 56.24%), Snapchat (329/1042, 31.57%), Pinterest (320/1042, 30.71%), TikTok (401/1042, 38.48%), Reddit (208/1042, 19.96%), LinkedIn (254/1042, 24.38%), or another social media site (69/1042, 6.62%). Further, participants reported accessing COVID-19 information using the social media sites Facebook (604/1042, 57.97%), Twitter (220/1042, 21.11%), Instagram (258/1042, 24.76%), Snapchat (85/1042, 8.16%), Pinterest (59/1042, 5.66%), TikTok (129/1042, 12.38%), Reddit (84/1042, 8.06%), LinkedIn (72/1042, 6.91%), and another social media site (42/1042, 4.03%).

Table 2 presents the results of the bivariate analyses. Pearson correlations suggest that the demographic variables of age, education, and income were correlated with the prevention mitigation outcomes of guideline knowledge, perceptions, and self-reported adherence. The ANOVA suggests that political affiliation was correlated with all 3 outcomes while gender, race or ethnicity, and COVID-19 vaccination status were correlated with prevention guideline perceptions and self-reported adherence. Social media sites used to consume COVID-19 news were correlated with self-reported adherence. Employment and regularity of social media use were not correlated with the outcomes of interest.

Table 2. Bivariate analysis results.

Variable	Outcomes		
	Knowledge	Perceptions	Self-reported adherence
Age			
r	0.09	–0.14	0.08
P value	.006	<.001	.01
Education			
r	0.11	0.001	0.11
P value	<.001	.97	<.001
Employment			
r	0.02	–0.03	0.04
P value	.48	.27	.21
Income			
r	0.15	–0.08	0.04
P value	<.001	.007	.17
Gender			
ANOVA (*F*)	0.38	6.43	5.27
P value	.54	.01	.02
Race or ethnicity			
ANOVA (*F*)	2.36	12.66	3.85
P value	.051	<.001	.004
Political affiliation			
ANOVA (*F*)	6.23	94.13	49.87
P value	.002	<.001	<.001
COVID-19 vaccination status			
ANOVA (*F*)	2.7	23.88	69.85
P value	.07	<.001	<.001
Site for COVID-19 news			
ANOVA (*F*)	2.07	1.64	2.89
P value	.07	.15	.01
Regularity of social media use			
ANOVA (*F*)	0.53	1.21	1.23
P value	.75	.30	.29

Knowledge of COVID-19 Guidelines

Indicating the level of knowledge related to COVID-19 prevention guidelines, the possible scores participants could receive included 100% (n=14, 1.4%), 75% (n=112, 10.9%), 50% (n=429, 41.7%), 25% (n=368, 35.7%), or 0% (n=107, 10.4%) correct. Model 1 (Table 3) suggests that income, Democratic political affiliation, and use of the social media platform TikTok were associated with COVID-19 prevention guideline knowledge. Specifically, as income (β=.03, 95% CI 0.005-0.05; *P*=.02) increased, it was found to be associated with a higher level of knowledge of COVID-19 guidelines. Democratic political affiliation (β=–.21, 95% CI –0.37 to –0.057; *P*=.008) was found to be negatively associated with guideline knowledge. Using TikTok as a source of COVID-19 information (β=–.29, 95% CI –0.58 to –0.004; *P*=.047) was associated with a lower level of knowledge. This model explained 6% of the variance in knowledge of COVID-19 guidelines.

Table 3. Regression results for knowledge, perceptions, and self-reported adherence.

Independent variables (reference)	Model 1: knowledge[a]		Model 2: perceptions[a]		Model 3: self-reported adherence[a]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Age	.002 (–0.003 to 0.007)	.51	–.007 (–0.01 to –0.002)	.007[b]	.004 (–0.0004 to 0.008)	.07
Gender (men)						
Women	.026 (–0.1 to 0.15)	.69	.16 (0.02 to 0.3)	.02[b]	.15 (0.04 to 0.26)	.008[b]
Race or ethnicity (White)						
Black or African American	–.048 (–0.26 to 0.16)	.66	.14 (–0.09 to 0.36)	.24	.21 (0.03 to 0.39)	.02[b]
Hispanic or Latino	–.079 (–0.27 to 0.11)	.41	.28 (0.08 to 0.49)	.007[b]	.27 (0.11 to 0.43)	.001[b]
American Indian or Alaska Native and Native Hawaiian or Pacific Islander	.072 (–0.45 to 0.6)	.79	.92 (0.35 to 1.49)	.002[b]	.32 (–0.13 to 0.77)	.16
Other	–.09 (–0.35 to 0.17)	.49	.07 (–0.21 to 0.34)	.63	.14 (–0.07 to 0.36)	.19
Education level	.03 (–0.014 to 0.079)	.17	–.015 (–0.065 to 0.036)	.56	.001 (–0.039 to 0.04)	.95
Employment	.029 (–0.008 to 0.066)	.13	–.001 (–0.04 to 0.038)	.95	.005 (–0.027 to 0.036)	.78
Income	.03 (0.005 to 0.05)	.02[b]	–.03 (–0.053 to –0.005)	.02[b]	–.015 (–0.03 to 0.004)	.13
Political affiliation (independent)						
Republican	–.12 (–0.28 to 0.04)	.15	–.5 (–0.67 to –0.33)	<.001[b]	–.23 (–0.37 to –0.09)	.001[b]
Democrat	–.21 (–0.37 to –0.057)	.008[b]	.34 (0.17 to 0.5)	<.001[b]	.17 (–0.04 to 0.31)	.01[b]
COVID-19 vaccination status (yes, but no booster)						
No	.00 (–0.17 to 0.17)	.99	–.22 (–0.4 to –0.04)	.02[b]	–.22 (–0.36 to –0.07)	.003[b]
Yes, including booster	.02 (–0.13 to 0.18)	.78	.31 (0.15 to 0.48)	<.001[b]	.32 (0.19 to 0.45)	<.001[b]
Site for COVID-19 news (Reddit)						
Facebook	–.086 (–0.31 to 0.14)	.45	–.23 (–0.47 to 0.009)	.06	–.23 (–0.42 to –0.043)	.02[b]
Instagram	–.026 (–0.28 to 0.23)	.84	–.40 (–0.68 to –0.12)	.005[b]	–.22 (–0.44 to 0.0026)	.05
Snapchat	.21 (–0.26 to 0.68)	.38	–.17 (–0.66 to 0.31)	.49	–.33 (–0.71 to 0.057)	.10
TikTok	–.29 (–0.58 to –0.004)	.047[b]	–.29 (–0.6 to 0.016)	.06	–.25 (–0.5 to –0.009)	.04[b]
Twitter	.015 (–0.22 to 0.25)	.90	–.33 (–0.58 to –0.08)	.01[b]	–.08 (–0.28 to 0.12)	.43
Regularity of social media use (less than once per week)						
Several times per day	.27 (–0.18 to 0.71)	.24	–.22 (–0.71 to 0.27)	.37	–.24 (–0.62 to 0.14)	.22
Once per day	.16 (–0.32 to 0.63)	.52	–.12 (–0.63 to 0.4)	.66	–.23 (–0.63 to 0.18)	.27
A few times per week	.2 (–0.29 to 0.69)	.43	–.03 (–0.57 to 0.5)	.91	–.16 (–0.58 to 0.27)	.47

Independent variables (reference)	Model 1: knowledge[a]		Model 2: perceptions[a]		Model 3: self-reported adherence[a]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Once per week	–.55 (–1.33 to 0.22)	.16	–.23 (–1.07 to 0.62)	.60	–.03 (–0.7 to 0.63)	.92

[a]R^2 values of models 1-3 are 0.06 (knowledge), 0.23 (perceptions), and 0.19 (self-reported adherence) respectively.

[b]*P* values indicate statistical significance at the α=.05 level.

Perceptions of COVID-19 Guidelines

Model 2 (Table 3) suggests that age, gender, Hispanic or Latino populations, American Indian or Alaska Native populations, income, political affiliation, COVID-19 vaccination status, and the use of the social media sites Instagram and Twitter were associated with perceptions of COVID-19 prevention guidelines. As age (β=–.007, 95% CI –0.01 to –0.002; *P*=.007) increased, it was found to be associated with a perception of the guidelines as strict. Women (β=.16, 95% CI 0.02-0.3; *P*=.02) were associated with perceiving the guidelines as relaxed. Hispanic or Latino (β=.28, 95% CI 0.08-0.49; *P*=.007) and American Indian or Alaska Native and Native Hawaiian or Pacific Islander (β=.92, 95% CI 0.35-1.49; *P*=.002) populations were found to be associated with perceiving the guidelines as relaxed. As income (β=–.03, 95% CI –.05 to –.005; *P*=.02) increases, it was found to be associated with stricter perceptions of the guidelines. Republican political affiliation (β=–.5, 95% CI –0.67 to –0.33; *P*<.001) was found to be associated with perceiving the guidelines as strict, while Democratic political affiliation (β=.34, 95% CI 0.17-0.5; *P*<.001) was found to be associated with perceiving them as relaxed. Receiving the full vaccination series and booster (β=.31, 95% CI 0.15-0.48; *P*<.001) was found to be associated with perceiving the guidelines as relaxed, while receiving no COVID-19 vaccinations (β=–.22, 95% CI –0.4 to –0.04; *P*=.02) was associated with perceiving them as strict. Instagram (β=–.4, 95% CI –0.68 to –0.12; *P*=.005) and Twitter (β=–.33, 95% CI –0.58 to –0.08; *P*=.01) were found to be associated with stricter perceptions of the COVID-19 prevention guidelines. This model explained 23% of the variance in perceptions of COVID-19 guidelines.

Adherence to COVID-19 Guidelines

As related to self-reported COVID-19 guideline adherence, model 3 (Table 3) suggests that women, Black or African American populations, Hispanic or Latino populations, political affiliation, COVID-19 vaccination status, and the use of Facebook and TikTok were associated with adherence to the COVID-19 prevention guidelines. Women (β=.15, 95% CI 0.04-0.26; *P*=.008) were found to be positively associated with adherence to the COVID-19 prevention guidelines. Black or African American (β=.21, 95% CI 0.03-0.39; *P*=.02) and Hispanic or Latino (β=.27, 95% CI 0.11-0.43; *P*=.001) populations were found to be positively associated with adherence to the guidelines. Republican political affiliation (β=–.23, 95% CI –0.37 to –0.09; *P*=.001) was negatively associated with adherence to prevention guidelines, while Democratic political affiliation (β=.17, 95% CI –0.04 to 0.31; *P*=.01) was positively associated with adherence. Receiving the full vaccination series and booster (β=.32, 95% CI 0.19-0.45; *P*<.001) was positively associated with adherence to the COVID-19 prevention guidelines, while receiving no COVID-19 vaccinations (β=–.22, 95% CI –0.36 to –0.07; *P*=.003) was negatively associated with adherence. Facebook (β=–.23, 95% CI –0.42 to –0.043; *P*=.02) and TikTok (β=–.25, 95% CI –0.5 to –0.009; *P*=.04) were found to be negatively associated with self-reported adherence to COVID-19 prevention guidelines. This model explained 19% of the variance in adherence to COVID-19 guidelines.

Discussion

Principal Findings

This study suggests that knowledge, perceptions, and self-reported adherence toward COVID-19 prevention guidelines differ by demographics and social media site use. Notably, marginalized populations (eg, older adults, women, and racial or ethnic minority individuals) were found to perceive the COVID-19 prevention guidelines as relaxed, in addition to their positive association with adherence. Political affiliation and COVID-19 vaccination status mirror assumptions about perceptions and adherence, where those identifying as Republican and reporting no vaccination were associated with perceiving the guidelines as too strict and adhering to a lesser degree, respectively. The popular social media sites TikTok, Instagram, Facebook, and Twitter were found to negatively impact pandemic prevention efforts as they were differentially associated with lower levels of knowledge, perceiving guidelines as strict, and lower self-reported adherence. The findings of this work, while demonstrating complicated interactions between guideline knowledge, perceptions, and adherence, serve to inform tailored public health interventions (ie, on the basis of demographic subgroups and social media site use), platform policies (eg, misinformation prevention), and digital public health policy more broadly.

Demographics and Knowledge, Perceptions, and Adherence Toward Guidelines

When considering the associations between the demographic correlates of income, age, and gender with knowledge, perceptions, and adherence toward prevention guidelines, the findings suggest a complex pandemic landscape. Whereas education and employment were not associated with guideline knowledge, it can be assumed that income reflects a layer of privilege afforded to those of higher income throughout the pandemic. In the case of this study, income may be acting as a proxy for pandemic privilege rather than solely socioeconomic status. Pandemic privilege can be understood here as the role of income in altering the pandemic environment, where those with additional resources are more likely to have access to prevention methods (eg, working from home, personal protective equipment, vaccination appointment flexibility, transportation, residential privilege, limited disruptions to services and care, and financial buffer for burdens of lost employment and wages) [14,15]. Despite possessing increased knowledge of the guidelines, perceptions of the prevention guidelines as strict reflect privileged protections afforded through increased income. Concordant with the existing literature, among older adults, a higher level of adherence to prevention guidelines, despite perceptions of them as strict, is likely due to the higher risk of severe illness from COVID-19 associated with increased age [16,17]. Gendered differences in perceptions of the guidelines as relaxed with a higher level of adherence reflect disproportionate pandemic burdens experienced by women (eg, occupational exposure, incidence, and post–COVID-19 condition [long COVID]).

The present findings are in accordance with the existing literature that demonstrates the impact of political affiliation on knowledge, perceptions, and adherence toward prevention guidelines. Partisan differences in perceptions of COVID-19 guidelines have been theorized to be explained by differential risk perceptions as influenced by news sources and media consumption [18-21]. Republican political affiliation has been found to be aligned with a preference for reducing the imposition of guidelines, while Democratic political affiliation is aligned with a preference for maintaining guidelines [22]. In accordance with the literature, political affiliation may play a decisive role in impacting knowledge-seeking and comprehension, perceptions, and adherence toward prevention guidelines. Health communication efforts may bolster prevention efforts through the characteristics inherent to partisan politics (eg, collectivism, inequity perceptions, perceived risk, skepticism, and media influence) and their influence on health behaviors [22-24]. The emerging literature attests that although political affiliation may demonstrate explanatory differences in pandemic prevention outcomes, there is a call for public health efforts that extend beyond interventions targeted based on political affiliation, implementing bipartisan efforts that also further consider demographics and individual differences influencing the operationalization of information from news and social media sites in the interest of COVID-19 prevention [18,23].

Social Media Sites and Knowledge, Perceptions, and Adherence Toward Guidelines

The use of the social media sites TikTok, Instagram, Twitter, and Facebook was found to be associated with lower knowledge, stricter perceptions, and lesser adherence toward COVID-19 prevention guidelines. Despite operating under distinct algorithms, all 4 platforms share commonalities in their functions for photo, video, audio, and text sharing, as well as social networking structures. A reliance on user-generated content creates difficulty in regulating the presence and spread of misinformation on social media. All 4 sites implemented, to various degrees, efforts to mitigate misinformation through informational banners on videos discussing the pandemic with off-site links to additional information. Despite these soft moderation efforts to address misinformation by TikTok, Instagram, Twitter, and Facebook, all have been found to contribute to the dissemination of misinformation [25-28]. Therefore, there is a need for improved mechanisms on these social media sites to limit the spread of misinformation due to its negative impacts on COVID-19 prevention guideline knowledge, perceptions, and adherence in the physical environment.

One key consideration of this study is the discrepancy between the demographic profiles of the included social media sites and the study sample. The user base of TikTok (ie, 48% users aged 18-29 years, 22% users aged 30-49 years, 14% users aged 50-64 years, and 4% users aged 65 years and older), Twitter (ie, 42% users aged 18-29 years, 27% users aged 30-49 years, 18% users aged 50-64 years, and 7% users aged 65 years and older), and Instagram (ie, 71% users aged 18-29 years, 48% users aged 30-49 years, 29% users aged 50-64 years, and 13% users aged 65 years and older) tends to be younger than that of Facebook (ie, 70% users aged 18-29 years, 77% users aged 30-49 years, 73% users aged 50-64 years, and 50% users aged 65 years and older) [8]. Although the average age of the study sample is older, it aligns with profiles of users of a similar age range who are active online (ie, 22% users on TikTok, 27% users on Twitter, 48% users on Instagram, and 77% users on Facebook) [8]. Although social media sites have unique demographic user profiles, it is necessary to consider that all individuals are able to access their platforms. Understanding the scope of a platform's typical and atypical users is necessary to systematically address misinformation online, where those who do not align with the average user experience an assumedly differential interaction with the platform and its content.

Public Health Implications

This research is uniquely situated within the COVID-19 pandemic and serves to inform tailored public health interventions, social media platform strategies, and policies. The key implications of this research include addressing knowledge gaps in the literature regarding the impact of social media use and demographic characteristics on COVID-19 prevention guideline knowledge, perceptions, and adherence. Public health interventions should be tailored to relevant platforms to address the impacts of social media sites on prevention guideline knowledge, perceptions, and adherence. Additionally, interventions targeting demographic subgroups may be operationalized on social media platforms with a user base that aligns with the target subgroup (eg, age, income, and political affiliation). In this context, platform functionality should be considered when designing interventions, regulations, and misinformation mitigation policies to alleviate the negative impacts of social media use on COVID-19 prevention efforts. Finally, these findings are necessary to be operationalized within public health interventions to tailor interventions to increase pandemic-related knowledge while enhancing supportive perceptions of the guidelines, aiming to increase and maintain sufficient adherence among subpopulations to mitigate the effects of the pandemic.

Strengths, Limitations, and Future Studies

This study has the strengths of using a country-wide, quota-based sample to investigate emerging trends during the pandemic as related to knowledge of, perceptions of, and adherence to COVID-19 prevention guidelines. Although there is likely some inherent difference in those who are online and able to participate in the survey as compared with those who are not, this concern may be mitigated in the context of this work, as it centers those active in the online environment. With the goal of identifying the role of social media on the target population, the exclusion of those not online is warranted. The findings should be cautiously interpreted and generalized as selection bias may affect the representativeness of the sample. When interpreting the study's findings, low statistical significance does not imply the absence of a certain phenomenon. One limitation that could persist, as the results are reliant on a self-report measure of prevention guidelines adherence, is participants' ability to approximate habits (eg, wearing a mask and using a social media site). A key limitation of this study is the discrepancy between the demographics of the study sample and the demographic profiles of the users of the various social media sites included. Finally, as a cross-sectional study, where some potential but key confounders may not have been included, there is the inability to obtain causal inference. Further, work accounting for the interrelations between factors should be conducted to provide a comprehensive assessment of confounders [22]. Future work should consider focusing on the validation of measures to assess knowledge, perceptions, and adherence. Additional research would benefit from an expanded survey considering a variety of potential, influential factors (eg, health literacy and location). Longitudinal explorations of the influence of social media use, knowledge levels, and declining perceptions should be prioritized in efforts to examine their impacts on prevention guideline adherence over time. Future directions for health communication should prioritize implementing programmatic interventions on social media platforms to address misinformation and information oversaturation in a manner that optimizes each platform's social networking functions, algorithms, and user base.

Acknowledgments

The authors would like to acknowledge the efforts of Jacob Long from the College of Information and Communication Sciences at the University of South Carolina for his support in survey development and data analysis. This work was funded by the National Institute of Health grants R01AI174892 and R01AI127203-5S1.

Authors' Contributions

CG developed survey materials, analyzed data, and led manuscript development. SQ collaborated on the development and provided significant contributions in manuscript refinement. XL provided significant contributions to manuscript refinement.

Conflicts of Interest

None declared.

References

1. Liu YC, Kuo RL, Shih SR. COVID-19: the first documented coronavirus pandemic in history. Biomed J 2020;43(4):328-333 [FREE Full text] [doi: 10.1016/j.bj.2020.04.007] [Medline: 32387617]
2. Coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/health-topics/coronavirus [accessed 2022-11-17]
3. COVID-19: How to protect yourself and others. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2022-04-13]
4. COVID data tracker. Centers for Disease Control and Prevention. 2022. URL: https://covid.cdc.gov/covid-data-tracker [accessed 2022-11-17]
5. Mcclain C, Vogels EA, Perrin A, Sechopoulos S, Rainie L. The internet and the pandemic. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/2021/09/01/the-internet-and-the-pandemic/ [accessed 2022-11-17]
6. Aichner T, Grünfelder M, Maurer O, Jegeni D. Twenty-five years of social media: a review of social media applications and definitions from 1994 to 2019. Cyberpsychol Behav Soc Netw 2021;24(4):215-222 [FREE Full text] [doi: 10.1089/cyber.2020.0134] [Medline: 33847527]
7. 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]
8. Social media fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2022-07-25]
9. Auxier A, Anderson M. Social media use in 2021. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2022-10-24]
10. Joseph AM, Fernandez V, Kritzman S, Eaddy I, Cook OM, Lambros S, et al. COVID-19 misinformation on social media: a scoping review. Cureus 2022;14(4):e24601 [FREE Full text] [doi: 10.7759/cureus.24601] [Medline: 35664409]
11. Vaal S, Schofield MB, Baker IS, Roberts BLH. Narcissism, national narcissism, COVID-19 conspiracy belief, and social media use as predictors of compliance with COVID-19 public health guidelines. Curr Psychol 2022:1-8 [FREE Full text] [doi: 10.1007/s12144-022-03715-6] [Medline: 36213568]

12. Volkmer I. Social media and COVID-19: a global study of digital crisis interaction among Gen Z and millennials. University of Melbourne. 2021. URL: https://minerva-access.unimelb.edu.au/items/db07228d-5b4c-59df-bf3e-e6bb37072aa9 [accessed 2024-01-10]
13. Haktanir A, Can N, Seki T, Kurnaz MF, Dilmaç B. Do we experience pandemic fatigue? current state, predictors, and prevention. Curr Psychol 2022;41(10):7314-7325 [FREE Full text] [doi: 10.1007/s12144-021-02397-w] [Medline: 34690475]
14. Huang X, Lu J, Gao S, Wang S, Liu Z, Wei H. Staying at home is a privilege: evidence from fine-grained mobile phone location data in the United States during the COVID-19 pandemic. Ann Am Assoc Geogr 2021;112(1):286-305. [doi: 10.1080/24694452.2021.1904819]
15. Berkman LF, Kawachi I. Social Epidemiology. New York: Oxford University Press; 2000.
16. COVID-19 risks and information for older adults. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/aging/covid19/index.html#:~:text=Older%20adults%20are%20more%20likely,very%20sick%20from%20COVID%2D19 [accessed 2022-07-23]
17. Coroiu A, Moran C, Campbell T, Geller AC. Barriers and facilitators of adherence to social distancing recommendations during COVID-19 among a large international sample of adults. PLoS One 2020;15(10):e0239795 [FREE Full text] [doi: 10.1371/journal.pone.0239795] [Medline: 33027281]
18. Hsiehchen D, Espinoza M, Slovic P. Political partisanship and mobility restriction during the COVID-19 pandemic. Public Health 2020;187:111-114 [FREE Full text] [doi: 10.1016/j.puhe.2020.08.009] [Medline: 32947252]
19. Kiviniemi MT, Orom H, Hay JL, Waters EA. Prevention is political: political party affiliation predicts perceived risk and prevention behaviors for COVID-19. BMC Public Health 2022;22(1):298 [FREE Full text] [doi: 10.1186/s12889-022-12649-4] [Medline: 35164719]
20. Limbu YB, Gautam RK, Pham L. The health belief model applied to COVID-19 vaccine hesitancy: a systematic review. Vaccines (Basel) 2022;10(6):973 [FREE Full text] [doi: 10.3390/vaccines10060973] [Medline: 35746581]
21. Zhou Y, Myrick JG, Farrell EL, Cohen O. Perceived risk, emotions, and stress in response to COVID-19: the interplay of media use and partisanship. Risk Anal 2023;43(8):1572-1586 [FREE Full text] [doi: 10.1111/risa.14044] [Medline: 36307383]
22. Gerace A, Rigney G, Anderson JR. Predicting attitudes towards easing COVID-19 restrictions in the United States of America: the role of health concerns, demographic, political, and individual difference factors. PLoS One 2022;17(2):e0263128 [FREE Full text] [doi: 10.1371/journal.pone.0263128] [Medline: 35196316]
23. Franz B, Dhanani LY. Beyond political affiliation: an examination of the relationships between social factors and perceptions of and responses to COVID-19. J Behav Med 2021;44(5):641-652 [FREE Full text] [doi: 10.1007/s10865-021-00226-w] [Medline: 33877532]
24. Gollust SE, Fowler EF, Vogel RI, Rothman AJ, Yzer M, Nagler RH. Americans' perceptions of health disparities over the first year of the COVID-19 pandemic: results from three nationally-representative surveys. Prev Med 2022;162:107135 [FREE Full text] [doi: 10.1016/j.ypmed.2022.107135] [Medline: 35803354]
25. 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;17(8):2373-2377 [FREE Full text] [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
26. Shelton CC, Curcio R, Carpenter JP, Schroeder SE. Instagramming for justice: the potentials and pitfalls of culturally relevant professional learning on Instagram. TechTrends 2022;66(5):837-854 [FREE Full text] [doi: 10.1007/s11528-022-00758-1] [Medline: 35818416]
27. 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]
28. 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;9(37):eadh2132 [FREE Full text] [doi: 10.1126/sciadv.adh2132] [Medline: 37713497]

Edited by R Cuomo; submitted 17.11.22; peer-reviewed by A Zheluk, A Rovetta; comments to author 02.05.23; revised version received 26.09.23; accepted 07.01.24; published 16.02.24.

Please cite as:
Garrett C, Qiao S, Li X
The Role of Social Media in Knowledge, Perceptions, and Self-Reported Adherence Toward COVID-19 Prevention Guidelines: Cross-Sectional Study
JMIR Infodemiology 2024;4:e44395
URL: https://infodemiology.jmir.org/2024/1/e44395
doi:10.2196/44395
PMID:38194493

Original Paper

Government-Nongovernmental Organization (NGO) Collaboration in Macao's COVID-19 Vaccine Promotion: Social Media Case Study

Xuechang Xian[1,2*], MA; Rostam J Neuwirth[3], PhD; Angela Chang[2*], PhD

[1]Department of Publicity, Zhaoqing University, Zhaoqing, China
[2]Department of Communication, University of Macau, Macao SAR, China
[3]Department of Global Legal Studies, University of Macau, Macao SAR, China
[*]these authors contributed equally

Corresponding Author:
Angela Chang, PhD
Department of Communication
University of Macau
Avenida da Universidade
Taipa
Macao SAR, 999078
China
Phone: 86 88228991
Email: wychang@um.edu.mo

Abstract

Background: The COVID-19 pandemic triggered unprecedented global vaccination efforts, with social media being a popular tool for vaccine promotion.

Objective: This study probes into Macao's COVID-19 vaccine communication dynamics, with a focus on the multifaceted impacts of government agendas on social media.

Methods: We scrutinized 22,986 vaccine-related Facebook posts from January 2020 to August 2022 in Macao. Using automated content analysis and advanced statistical methods, we unveiled intricate agenda dynamics between government and nongovernment entities.

Results: "Vaccine importance" and "COVID-19 risk" were the most prominent topics co-occurring in the overall vaccine communication. The government tended to emphasize "COVID-19 risk" and "vaccine effectiveness," while regular users prioritized vaccine safety and distribution, indicating a discrepancy in these agendas. Nonetheless, the government has limited impact on regular users in the aspects of vaccine importance, accessibility, affordability, and trust in experts. The agendas of government and nongovernment users intertwined, illustrating complex interactions.

Conclusions: This study reveals the influence of government agendas on public discourse, impacting environmental awareness, public health education, and the social dynamics of inclusive communication during health crises. Inclusive strategies, accommodating public concerns, and involving diverse stakeholders are paramount for effective social media communication during health crises.

(JMIR Infodemiology 2024;4:e51113) doi:10.2196/51113

KEYWORDS

COVID-19; government; vaccine; automated content analysis; Granger causality test; network agenda setting; QAP; social media

Introduction

As of December 2022, the global COVID-19 pandemic had resulted in 669 million confirmed cases and 6.8 million deaths [1]. Environmental factors were a key determinant significantly influencing the pandemic [2], through airborne viral infectivity impacted by air pollution and seasonality effects [3,4].

Vaccination was crucial to contain the spread of virus [5], although complex factors such as the Peltzman effect, emerging viral variants, and socioeconomic conditions also affected pandemic diffusion [6]. Determining an optimal level of vaccination is complex and multifaceted, requiring a balance to avoid undermining democratic values and triggering larger socioeconomic problems than the pandemic [7,8]. Nonetheless, the willingness to vaccinate hinges on various factors, including safety concerns, sociodemographic characteristics, and individual behaviors and attitudes [9,10]. Other determinants including lack of knowledge, government distrust, skepticism about vaccine development, efficacy concerns, exposure experience, coronaphobia, and workplace mandates also predict vaccine uptake [11-13]. As social media becomes increasingly significant for public communication, social media adaptivity, information availability, and health care infrastructure capabilities are also influential for vaccination decisions [14].

Vaccine communication plays a vital role in addressing public concerns, building trust, and encouraging vaccine uptake. Specifically, effective strategies including trusted sources, health provider guidance, a reasonable quantity of information, cultural tailoring, information contextualization, and cultural sensitivity have the potential to significantly increase vaccination intent [15-17]. Despite the notable antagonism in the discourse surrounding immunization on social media [18], it is worth noting that social media campaigns initiated by health organizations have proven to be effective in increasing public awareness about vaccination [19].

Governance mechanisms are another crucial factor for expediting vaccine distribution and mitigating pandemic-related socioeconomic effects [20]. Evidence has shown that clear, consistent, and transparent communication from governmental bodies engendered higher levels of public compliance and trust [21,22]. Given the major impact of the pandemic on public health and society, involvement of the government in vaccine communication becomes a vital research area.

Governments worldwide have adopted diverse approaches to encourage COVID-19 vaccination. For instance, the New Zealand government promoted vaccination among young people by highlighting community factors such as "protecting others" and "striving for herd immunity" [23]. By promoting the scientific notion that there are more advantages than disadvantages to COVID-19 vaccination, the Chinese government has strengthened risk communication to increase the public's awareness of the benefits of vaccines [24]. Although COVID-19 vaccine communication has received increasing attention, particularly from the research community, scientific evidence focusing specifically on low-risk regions, such as Macao, is scarce. This suggests that the existing literature does not sufficiently reflect the concerns of the Macao population as related to COVID-19 vaccination. As one of the world's most densely populated cities, Macao has maintained a record of relatively low risk of infection and high coverage of COVID-19 vaccines [25]. Throughout the pandemic before June 2022, Macau had only recorded 17 confirmed cases of local infection (with a rate of 2.5 cases per 100,000 population) with no fatalities. By June 19, 2022, the vaccine coverage rate within the entire population in Macao was 85.6% for at least 2 doses and 40.5% for 3 doses [26]. The low prevalence of COVID-19 is believed to be the result of the close connection between Macao and mainland China. Since the outbreak of the pandemic, Macao has implemented anti-epidemic measures following the "dynamic zero-COVID-19 policy" established by mainland China, with some adaptations based on local socioeconomic circumstances [27]. Given the close link between these entities, it is important to understand how the Macao Government communicated with citizens to drive their demand for vaccinations and the impact of this communication. Researchers have long investigated how governments develop policy agendas and whether a policy agenda is led by the government or the public [28]. However, literature on the role of the government in public health agenda setting, specifically related to vaccine promotion in the COVID-19 context, is limited.

The primary goal of this study was to reveal the patterns of vaccine communication on social media during the COVID-19 pandemic as well as the role of the government in advancing vaccination through a case study of Macao, the special administrative region of China. By conducting this research, we aimed to contribute to the existing knowledge on vaccine communication and provide implications for policymakers to improve health promotion communication strategies for preparedness against future pandemics.

The theory of agenda setting suggests that the media has the ability to influence the public agenda by making a specific issue prevalent and salient [29]. Agenda setting is a competition among issue proponents to gain the attention of media professionals, the public, and policy elites [30]. Recently, research about agenda setting has been extended by incorporating the concept of social networks and the associative network of memory, which has been proposed by Guo [31] as the network agenda setting model (NAS). The NAS underlines the associations between topics or attributes presented in the agenda: The more frequently 2 attributes are correlated in news coverage, the more likely the public will perceive them to be interrelated [32].

The NAS can be used to identify the interconnections between public, media, organizational, and government topics on social media. For instance, a study conducted by Chen et al [33] utilized the NAS to investigate the correlation between individual users and organizational accounts on Weibo in terms of their focus on nationalist concerns. The NAS emphasizes the relationship between topics or attributes in constructed agendas. Hou et al [34] analyzed posts mentioning COVID-19 vaccines on Twitter and found that topics related to COVID-19 vaccines can be divided into the following 9 categories: (1) vaccine importance, (2) vaccine effectiveness, (3) vaccine safety, (4) trust in governments, (5) trust in experts, (6) COVID-19 risk, (7) vaccine accessibility, (8) vaccine distribution, (9) vaccine affordability. Additionally, recent studies examined the concerns of all users, including parents, regarding COVID-19 vaccines (eg, [35]). However, these studies did not distinguish between regular accounts (ie, ordinary individual users), government accounts, organization accounts, and media accounts. This distinction is important to understand the nuances of vaccine promotion engaged by different entities. Governments, for instance, influence public discourse through policymaking [24,28], whereas organizations play a significant role in agenda setting via funding, lobbying, and advertising activities [36]. The public, media, and government may construct different associations among topics in their respective agendas and impact each other. Our research questions (RQ) thus ask the following:

- RQ1: What are the most prevalent agenda attributes emphasized in the communication of vaccination on Facebook during the COVID-19 outbreak in Macao?
- RQ2: How do the attributes interact in the vaccine agendas of governmental and nongovernmental entities?
- RQ3: What are the associations between the vaccine agenda networks constructed by government and nongovernment users?
- RQ4: How do government and nongovernment users impact each other's vaccine agenda on Facebook?

Methods

Sample and Data

This study retrieved data relevant to COVID-19 vaccines in Macao from January 1, 2020, when the SARS-CoV-2 virus was initially detected in China, to August 31, 2022, when the number of newly reported cases had sharply declined [1]. Facebook was selected as the main source of data to analyze the dynamics of vaccine communication in Macao. Being one of the most widely used social media platforms globally, Facebook accounts for a more dominant market share (65.05%) than other sources (eg, Pinterest: 11.47%; Twitter: 10.54%) in Macao [37,38]. The widespread usage of Facebook suggests that it has a significant impact on the population's perceptions, attitudes, and behaviors, making it an essential platform to study to understand the public agenda. In addition, Facebook's archival nature allows for tracking of the evolution of vaccine-related discussions over time, capturing the core dynamics of vaccine communication online.

A combination of the keywords "COVID-19" and "vaccine" as well as their synonyms (ie, 29 synonyms of COVID-19–related terms and 10 synonyms of vaccine-related terms) in Chinese were used to detect and collect relevant posts (see Multimedia Appendix 1). Information was also compiled on the various labels given to users on Facebook, such as labels of government, media, and organization accounts. Following the collection of raw data from Facebook, data screening was performed to remove duplicate and irrelevant posts. The preprocessing of data including the removal of stop words (eg, "an," "the," "etc.," punctuation, symbols, and numbers) and word segmentation was implemented using the DivoMiner platform.

Ethics Approval

This research strictly adheres to ethical guidelines by ensuring complete anonymity and de-identification of all data sources. To preserve the confidentiality and privacy of all sources involved, no identifiable information about individual users, their IDs, or direct, non-paraphrased posts are included in the main manuscript or any supplementary materials.

Clarification

All identifiers in the data set (eg, names of the senders) were removed and replaced with a code to mask the information about each sender, ensuring the anonymization of our data. Data were only collected from publicly available posts that were returned based on the structured keyword search criteria.

Measures of Variables

This study investigated the dynamics of agenda setting between government and nongovernment users on Facebook. To achieve this, we categorized users into the following different categories, drawing from prior research [39,40]: (1) media, (2) civil organizations, (3) regular users, (4) government.

The media functions as information gatekeepers and holds potential influence over people's decision-making [29,32]. To account for significant differences in content, news culture, and viewpoints, the media category in this study was further divided into professional media and alternative media for a thorough investigation [41]. Professional media includes those traditional mass media outlets responsible for information dissemination and public awareness, such as newspapers, radio, and television, while alternative media includes independent and electronic media, which is in contrast to mainstream mass media. By referencing relevant media research [42], this study annotated professional media accounts, alongside alternative media accounts.

Civil organizations, also called civil society organizations, include those organizations or associations that are established by individuals or groups with a common purpose or interest and operate in the community, differing from the government and corporations. Civil organizations work alongside the government and other stakeholders to contribute to public discourse, policy development, and social change [43].

Regular users were defined in this study as individuals who interact with Facebook on a personal basis, without representing any official capacity, media, or organizations. Therefore, regular users can be considered as representatives of the public in this study.

The government in this study was defined as all authorities. We did not categorize the specific levels, instead treating all government authorities as a single entity, to gain a clear understanding of the overall position of the Macao Government in vaccine communication. This was also a practice adopted by previous research (eg, [44]).

The classification of Facebook users into 5 distinct categories was conducted based on the information gathered from users' short biographical profiles and the user identity labels provided by Facebook. We assigned 2 coders to classify the users contributing relevant posts. Any confusion that might have occurred during classification was resolved through discussion. This approach allowed for the categorization of users into specific groups, enabling a systematic analysis of user communication and interactions within the Facebook platform [44].

To investigate the dynamics of vaccine communication, 9 predefined categories that indicate elements influencing vaccine acceptance were established based on a coding framework adapted from prior studies (eg, [34,45,46]). These categories included the following topics: importance of vaccines, effectiveness of vaccines, safety of vaccines, trust in governments, trust in experts, risk of the COVID-19 pandemic, and vaccine convenience (ie, accessibility, distribution, and affordability). Details of the coding categories are shown in Multimedia Appendix 2.

Data Analysis Procedures

Automated Content Analysis

In this study, an automated content analysis method was used to identify and categorize posts into the predefined categories. Each post could belong to one or more categories or none at all. The effectiveness of automated coding depends on the design of the keywords. To develop accurate keywords, this study followed the approach outlined by Chang et al [37] using the Word2vec word embedding toolkit from the Python 3.7.4 Gensim module [47]. Word2vec, a word embedding technique powered by neural networks, allows the identification of words with similar meanings by analyzing word associations in a large text corpus [48]. Due to the intricacies of the Chinese language, the synonyms suggested by Word2vec were further checked by assessing their relevance to the context. On this basis, the Chinese thesaurus and relevant literature [49] were further consulted for the inclusion of additional synonyms. The list of keywords for machine coding can be found in Multimedia Appendix 3.

DivoMiner, a text mining and automated content analysis platform driven by machine learning algorithms, was used to facilitate the automated content coding task. This platform integrates automated content analysis with traditional content analysis methods and has been widely utilized in health and communication studies [37,50,51]. Following automated coding, manual verification was conducted to ensure the accuracy and reliability of the machine-generated outcomes. To achieve this, 2 coders, both native Cantonese speakers, were recruited and underwent 36 hours of training to independently code 300 messages. Each variable was coded as either present or absent. Discrepancies between the coders were resolved through discussions, with the author intervening only when consensus could not be reached between the coders. The overall intercoder reliability, measured using Krippendorff alpha, demonstrated satisfactory levels across all examined variables, with coefficients ranging from .77 to .82. The consistency between machine coding and manual coding reached an acceptable level, with an average score of 74%. This score aligns with previous studies, in which a threshold value of 70% was considered rational [49-51].

Statistical Analysis

The conventional statistical analysis in this study involved the use of SPSS (version 23; IBM Corp) for analysis. Categorical variables were summarized using counts and percentages. The chi-square test of independence was used, and post hoc comparisons with Bonferroni corrections were further implemented to precisely identify the specific significant differences between user categories and vaccine-related topics and avoid the likelihood of generating false-positive outcomes (type I errors).

Co-Occurrence Network Analysis

Co-occurrence matrices, which represent the strength of ties between 2 topics engaged by different users, were generated as dyadic data sets. Based on the co-occurrence data, this study established undirected and weighted topic co-occurrence networks. Each network represents the co-occurrence relations of the attributes of a certain user category. To clarify, if a particular category of user mentions topic “i“ and topic “j,” a band will link “i” and “j.” The width of the band indicates the frequencies of the pair of topics discussed by a user type [52,53]. For example, in the professional media user category’s topic co-occurrence network, if a professional media news report mentions the topics of “vaccine importance” and “vaccine effectiveness” together, the topics will be linked in the network by a band. The more frequently these topics co-occur, the thicker the band becomes. The visualization of topic co-occurrence is presented in a chord diagram by Echarts (The Apache Software Foundation), as indicated by Wang et al [52].

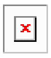
Quadratic Assignment Procedure for Network Analysis

In this study, the quadratic assignment procedure (QAP) method was applied to understand the correlation between the Macao Government’s agenda network and that of other Facebook users, via analysis of the co-occurrence matrices. QAP is a common method in social network or agenda network studies [40,54]. QAP correlation analysis can be used to assess the correlation between 2 matrices with the Pearson correlation coefficient, while QAP regression analysis can determine whether an explanatory variable can predict an outcome variable when the 2 matrices are significantly correlated [55]. In this study, the QAP method used UCINET 6.730 to test whether the Macao Government’s vaccine agenda network has impacted that of nongovernment Facebook users, particularly regular type users, during the COVID-19 pandemic.

Vector Autoregression Modeling

The vector autoregression (VAR) approach was used to examine the dynamic of agenda attributes between government and nongovernment users. This approach evaluates the effect of an observed variable by considering its lagged effect in the earlier period and that of other predictors in previous time points, without presuming the associations between the variables [56]. The VAR modeling technique is widely used in the economic field and, in recent years, has been increasingly applied in research on health science, sociology, neuroimaging, and meteorology (eg, [54,57-59]).

VAR modeling is ideal for measuring the dynamic performance response and interaction between performance and marketing communication variables. A study applied VAR models to construct the dynamic response relationship between news stories and public attention using a combination of survey and news content ranging from 2009 to 2013 [60]. The VAR models captured the dynamic feedback system and gave estimates for the short-term effects of TV news coverage on public perception by demonstrating a unidirectional process wherein changes in news salience led to significant changes in public salience. In addition, VAR models have also been used to investigate the dynamic mapping relationship between the diffusion of political messages and emotional expression in public messages during the COVID-19 pandemic [61]. The increased diffusion of political messages positively predicted changes in emotional expression among citizens, and the VAR model was able to explain the interdependencies among variables based on the lag values of multiple time series. Overall, the VAR model proves to be an insightful tool for analyzing complex relationships in communication studies, providing insights into the short-term and long-term effects of various factors on outcomes of interest. Hence, using the VAR technique allows the exploration of temporal dynamics and associations between different agenda attributes in this study. For example, the approach enables a better understanding of whether the agenda attributes propagated by the government (AG) at time (t-n) impacts the agenda attributes of nongovernment users (AN) including professional media, alternative media, civil organizations, and regular users. The VAR model was generated as follows:



Within this model, α_i and β_i are the estimated coefficients, ρ represents the optimal number of lags for the model, and ε indicates the error term. AG_{t-i} and AN_{t-i} represent the respective variable at the earlier periods. For instance, AG_{t-1} indicates the first lag of AG. The lag length for the VAR model was selected as per the Akaike information criterion. The augmented Dickey-Fuller test was applied to examine the stationarity of the time series. For nonstationary series, differencing at the first or higher level was performed to achieve stationarity [62]. When both time series were stationary at the same level, this study proceeded with the Johansen maximum eigenvalue and trace tests based on the estimation of VAR models to determine whether the time series were cointegrated and suitable for Granger causality tests. Granger causality posits that causes lead to effects and happen before their effects [40]. In this sense, using prior values of a time series can statistically forecast the future status of another time series.

In this study, the Granger causality test was used to provide greater insight into the statistical causal relationship between the government's agenda and the nongovernment users' agenda. To estimate VAR models and enable Granger causality tests, this study transformed the collected data in the form of time series by dividing the data into 32 monthly periods (from January 2020 to August 2022), and each monthly period was treated as an independent unit for analysis. EViews 12 software was used for statistical analysis.

Results

Results of Content Analysis

This research initially collected a sample of 24,089 Facebook posts with relevance to COVID-19 vaccines. Data screening was further performed on the sample to remove duplicated, irrelevant, and unclear messages, resulting in 23,577 unique and relevant posts. Finally, the results of machine coding presented a total of 22,986 posts that include the examined vaccine topics.

In answering RQ1, we calculated the frequency of the vaccine topics and found that the majority of posts in the sample related to the importance of COVID-19 vaccination (7358/22,986, 32.01%), followed by posts that indicated the high risk of contracting COVID-19 (6877/22,986, 29.92%) and highlighted trust in experts (4320/22,986, 18.79%). In addition, a considerable number of posts mentioned vaccine effectiveness (4163/22,986, 18.11%), safety (3358/22,986, 14.61%), accessibility (2683/22,986, 11.67%), distribution (2492/22,986, 10.84%), and affordability (1685/22,986, 7.33%), while posts related to trust in government were less frequent (1593/22,986, 6.93%). In addition, in the overall vaccine-related discussion, nongovernment users comprised a substantial majority of the posts, at 76.85% (17,665/22,986). When examining the nongovernment user segment at a more granular level, professional media accounted for a significant proportion of the posts, at 33.87% (7555/22,986), followed by alternative media, at 12.24% (2814/22,986); civil organizations, at 3.99% (918/22,986); and regular users, at 27.74% (6377/22,986). The topics associated with vaccine agenda attributes by government and nongovernment users are shown in Table 1.

The chi-square test indicated that the distributions of vaccine-related topics were significantly different across the user categories (χ^2_{32}=1579.469, *P*<.001). The outcomes of the post hoc comparisons suggested that the government was more concerned with topics of vaccine effectiveness (1003/5322, 18.85%; *P*<.001), COVID-19 risk (1805/5322, 33.92%; *P*<.001), vaccine accessibility (1010/5322, 18.98%; *P*<.001), and vaccine affordability (605/5322, 11.37%; *P*<.001), while discussion of vaccine safety (393/5322, 7.38%; *P*<.001), government trust (133/5322; 2.5%, *P*<.001), expert trust (518/5322, 9.73%; *P*<.001), and vaccine distribution (341/5322, 6.41%; *P*<.001) occurred to a less extent than for other users. In comparison, professional media contributed more to the topics of government trust (752/7555, 9.95%; *P*<.001) and expert trust (1895/7555, 25.08%; *P*<.001). Alternative media, however, were less inclined to discuss vaccine affordability (128/2814, 4.55%; *P*<.001) than other categories of users. Regular users were primarily concerned about vaccine safety (1092/6377, 17.12%; *P*<.001) and vaccine distribution (724/6377, 11.35%; *P*<.001) and were less concerned about vaccine effectiveness (937/6377, 14.69%; *P*<.001), COVID-19 risk (1529/6377, 23.98%; *P*<.001), and vaccine accessibility (416/6377, 6.52%; *P*<.001) than other users. The outcomes of the post hoc tests with details are shown in Multimedia Appendix 4.

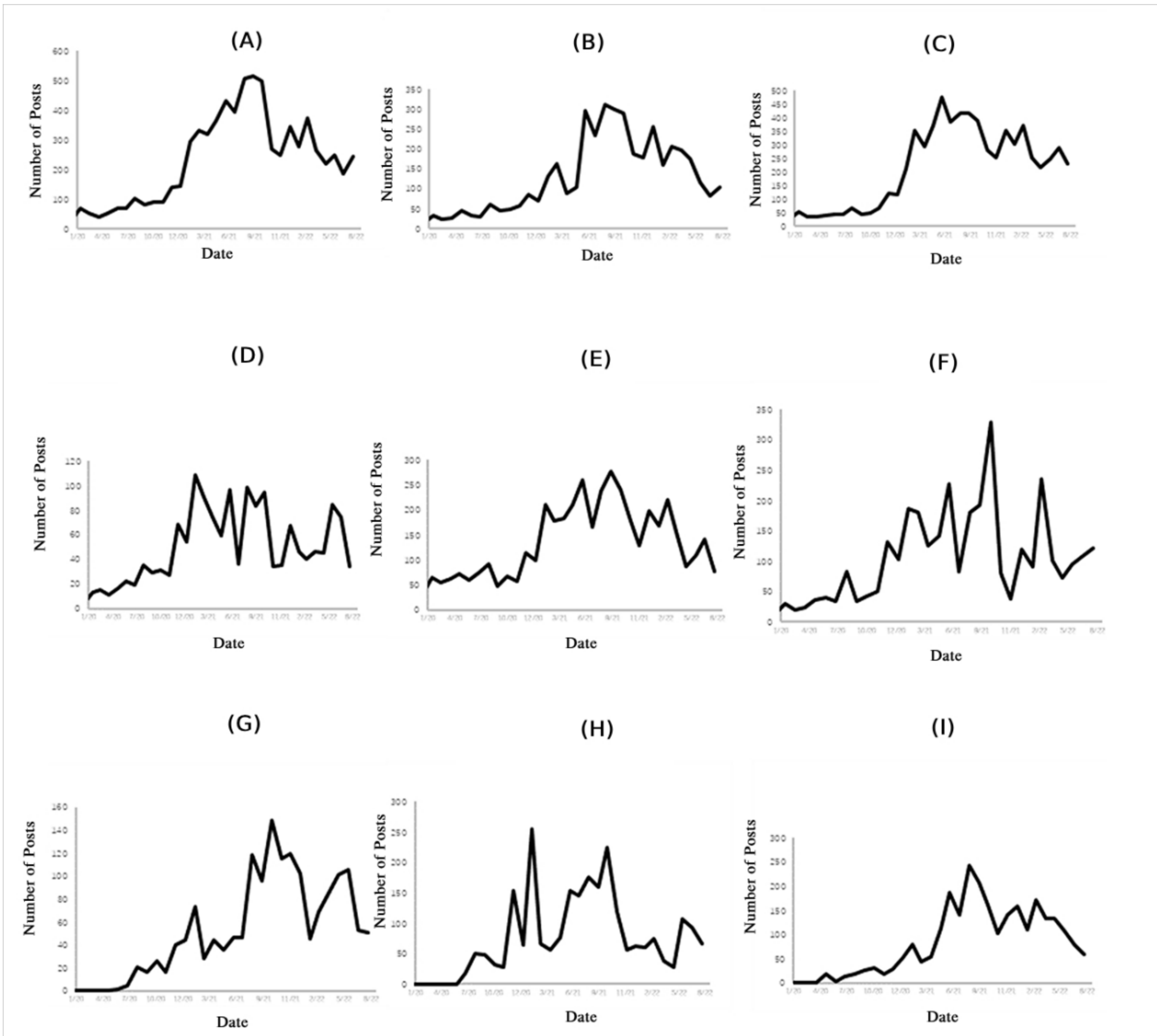
Table 1. Overview of the vaccine agenda attributes by government and nongovernment users in Macau from January 1, 2020, to August 31, 2022.

Vaccine topic	Government users, n (%)	Nongovernment users, n (%)				Total, n (%)
		Professional media	Alternative media	Civil organizations	Regular users	
All posts	5322 (23.15)	7555 (32.87)	2814 (12.24)	918 (3.99)	6377 (27.74)	22,986 (100)
Importance	1616 (30.36)	2931 (38.80)	697 (24.77)	298 (32.46)	1816 (28.48)	7358 (32.01)
Effectiveness	1003 (18.85)	1638 (21.68)	404 (14.36)	181 (19.72)	937 (14.69)	4163 (18.11)
Safety	393 (7.38)	1374 (18.19)	359 (12.76)	140 (15.25)	1092 (17.12)	3358 (14.61)
Trust in government	133 (2.5)	752 (9.95)	175 (6.22)	40 (4.36)	493 (7.73)	1593 (6.93)
Trust in experts	518 (9.73)	1895 (25.08)	593 (21.07)	154 (16.78)	1160 (18.19)	4320 (18.79)
COVID-19 risk	1805 (33.92)	2651 (35.09)	681 (24.2)	211 (22.98)	1529 (23.98)	6877 (29.92)
Accessibility	1010 (18.98)	981 (12.98)	196 (6.97)	80 (8.71)	416 (6.52)	2683 (11.67)
Distribution	341 (6.41)	1005 (13.3)	309 (10.98)	113 (12.31)	724 (11.35)	2492 (10.84)
Affordability	605 (11.37)	529 (7)	128 (4.55)	53 (5.77)	370 (5.8)	1685 (7.33)

Trend in Facebook Activities

To reveal the dynamics of different attributes of the vaccine agenda, this study mapped trends of these attributes during the investigated period. All vaccine-relevant content remained at a relatively low volume in 2020 and increased significantly in 2021. The volume of content regarding “vaccine distribution” began to grow at the start of 2021 and showed an observable spike in February of the same year. This was followed by a sharp acceleration in content regarding the high risk of COVID-19 reaching its peak in June 2021. The highest peak in vaccine-relevant content occurred in September 2021 related to the topic of vaccine importance. Between June 2021 and October 2021, the most debate centered around themes relating to COVID-19 vaccines. Overall, variations in the volume of vaccine communication were observed over time. Figure 1 shows the dynamic of vaccine discussion showing the monthly volume of posts.

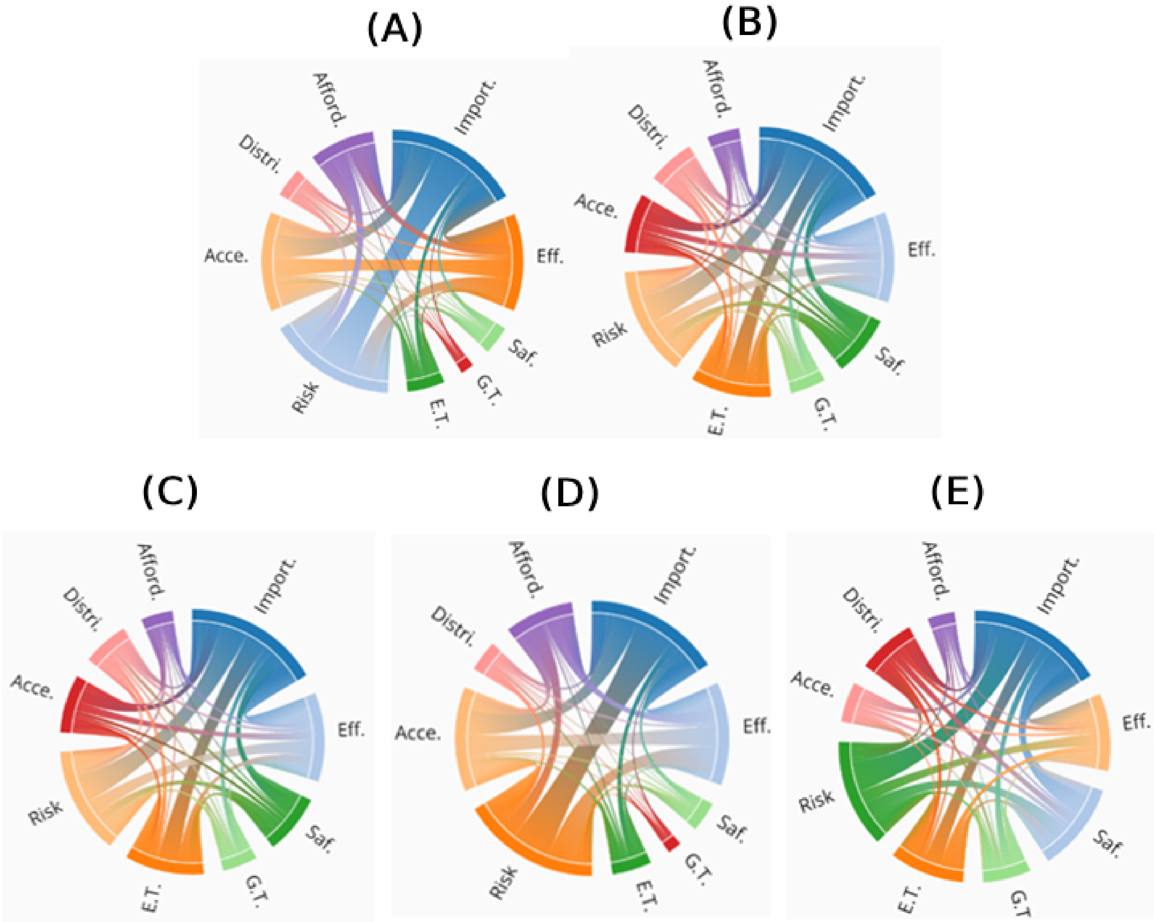
Figure 1. Temporal changes in the vaccine agenda attributes (January 2020–August 2022): (A) vaccine importance, (B) vaccine effectiveness, (C) risk of COVID-19, (D) government trust, (E) expert trust, (F) vaccine safety, (G) vaccine affordability, (H) vaccine distribution, (I) vaccine accessibility.



Interactions Between Agenda Attributes in Vaccine Communication

To answer RQ2, this study computed the interrelationships between agenda attributes by the government and nongovernment users by constructing co-occurrence matrices. Results showed that "vaccine importance," "vaccine effectiveness," and "COVID-19 risk" were the most prominent attributes interacting with each other in the agendas of government and nongovernment users, except for the regular users' agenda in which "vaccine safety" (n=2503) rather than "vaccine effectiveness" (n=2161) had more established connections overall with other attributes. Specifically, the government agenda featured strong connections between "vaccine importance" and "COVID-19 risk" (n=1505), followed by "vaccine importance" and "vaccine effectiveness" (n=945), "vaccine importance" and "accessibility" (n=940), and "COVID-19 risk" and "accessibility" (n=816). As for the agenda of professional media, the strongest link was established between "vaccine importance" and "COVID-19 risk" (n=1528), followed by the link between "vaccine importance" and "vaccine effectiveness" (n=1327) and the link between "vaccine importance" and "trust in experts" (n=1220). In terms of regular users, their agenda highlighted the relationships between "vaccine importance" and "COVID-19 risk" (n=931), "vaccine importance" and "vaccine effectiveness" (n=655), "vaccine importance" and "vaccine safety" (n=644), "vaccine importance" and "trust in experts" (n=536), and "vaccine safety" and "COVID-19 risk" (n=469). Using chord diagrams, this study visualized the interrelationships of agenda attributes by different user categories. The arc in the outer ring represents the attributes of the vaccination agenda and is differentiated by color. The arc length indicates the total number of associations an attribute maintains with other attributes when communicated by users in a specific category. The band within the ring represents the connected relationship between 2 topics, with the thickness of the band indicating the magnitude of the connection. A set of chord diagrams revealing agenda attribute interactions in the agendas with comparison of different users is presented in Figure 2.

Figure 2. Comparison of agenda attribute interactions by different users: (A) government, (B) professional media, (C) alternative media, (D) civil societal organizations, (E) regular users. Acce.: vaccine accessibility, Afford.: vaccine affordability, Distri.: vaccine distribution, Eff.: vaccine effectiveness, E.T.: expert trust, G.T.: government trust, Import.: vaccine importance, Risk: risk of COVID-19, Saf.: vaccine safety.



To assess the evolution of links between attributes over time, this study also divided the co-occurrence dynamics of intragroup agenda attributes into 3 distinct periods: 2020, 2021, and 2022. Our findings revealed that the connections between agenda attributes varied by both the time period and the categories of Facebook users. Notably, in the government agenda, the link between "vaccine effectiveness" and "vaccine affordability" exhibited an increase in strength during 2022 (795/4150, 19.16%), compared with 2020 (18/223, 8.25%) and 2021 (690/4744, 14.54%). Conversely, the connection between "vaccine importance" and "expert trust" within the agenda of regular users demonstrated a decline in frequency over the 3-year span (2020: 119/1165, 10.21%; 2021: 282/3417, 8.25%; 2022: 118/1779, 6.63%). More information about the co-occurrence dynamics of the intragroup agenda attributes over time can be found in Multimedia Appendix 5.

Agenda Network Analysis

In answering RQ3, the results of the QAP tests demonstrated significantly positive and strong correlations between the agenda network of the government and those of professional media (r=0.745, P=.005) and civil organizations (r=0.632, P=.02). However, the correlations between the government's agenda network and the network of alternative media (r=0.462, P=.08) and regular users (r=.451, P=.07) were not statistically significant.

The subsequent QAP linear regression analysis tested whether the agenda network of the Macao government can predict that of nongovernment users. For example, by using the government as a predictor and different types of nongovernment users as outcome variables, the results demonstrated that the government has an impact on the agenda network of professional media (b=0.703, P=.006) and civil organizations (b=0.051, P=.02). The adjusted R^2 value for professional media indicated that government accounts for around 54% of the variance in the professional media's agenda network, while government only accounts for 38% of the variance in the agenda network of civil organizations. The results of the QAP linear regression analysis with the government as a predictor are shown in Table 2.

In the QAP linear regression model predicting the agenda of regular users, the results revealed significant impacts of alternative media (b=2.46, P=.001), professional media (b=0.52, P=.001), and civil organizations (b=6.16, P=.001) on the agenda of regular users. The adjusted R^2 value for professional media, civil organizations, and alternative media ranged from 0.81 to 0.86, suggesting that all 3 categories of users can explain 81%-86% of the variance in the regular users' agenda network. The results of the QAP linear regression analysis with regular users as the outcome variable are shown in Table 3.

Table 2. Quadratic assignment procedure regression analysis with government as the predictor.

User category	Unstandardized coefficient	*P* value[a]	R^2 value	Adjusted R^2
Civil organizations	0.051	.02	0.399	0.382
Professional media	0.703	.006	0.556	0.543
Alternative media	0.095	.10	0.214	0.191
Regular users	0.246	.12	0.204	0.180

[a]Outcomes were considered statistically significant at $P<.05$.

Table 3. Quadratic assignment procedure regression analysis with regular users as the outcome variable.

User category	Unstandardized coefficient	*P* value[a]	R^2 value	Adjusted R^2
Government	0.246	.12	0.204	0.180
Alternative media	2.462	.001	0.868	0.864
Professional media	0.521	.001	0.811	0.805
Civil organizations	6.164	.001	0.832	0.827

[a]Outcomes were considered statistically significant at $P<.05$.

Impacts of Government and Nongovernment Users on Each Other's Vaccine Agenda

To answer RQ4, the Granger causality test was further performed to examine whether the 9 attributes in the government's agenda statistically predicted the future intensity of topics discussed by different categories of users and vice versa. Specifically, the results showed that attributes such as "vaccine safety" ($F_{3,13}$=3.817; P=.04) and "trust in experts" ($F_{3,13}$=3.916; P=.03) in the government's agenda significantly affected such attributes in the agenda of nongovernment users, while the attributes associated with "trust in government" ($F_{3,13}$=4.590; P=.02) and "vaccine affordability" ($F_{3,13}$=3.851; P=.04) in the agenda of nongovernment users affected these attributes in the agenda of the government at the significance level of $P<.05$.

By classifying nongovernment users into different user categories, the results suggested a unidirectional trend in the attribute of "vaccine safety" flowing from the government's agenda to that of professional media ($F_{5,15}$=3.247; P=.03), while professional media affected the agenda of the government unilaterally through the attributes of "vaccine importance" ($F_{5,12}$=7.192; P=.003), "vaccine effectiveness" ($F_{3,13}$=4.391; P=.02), "COVID-19 risk" ($F_{5,15}$=5.173; P=.006), and "vaccine affordability" ($F_{3,13}$=4.754; P=.02). Additionally, alternative media affected the government by setting the agenda with attributes such as "COVID-19 risk" ($F_{5,15}$=8.769; $P<.001$) and "vaccine accessibility" ($F_{5,15}$=2.963; P=.047), while there was no temporal causation from the government to alternative media for the attributes identified.

Regarding civil organizations, the government predicted the agenda of civil organizations through the attributes of "vaccine importance" ($F_{5,15}$=4.111; P=.01), "vaccine effectiveness" ($F_{3,13}$=6.264; P=.007), and "trust in experts" ($F_{3,9}$=15.877; P=.001), while the causation from civil organizations to the government was absent for all attributes except "vaccine safety" ($F_{3,12}$=4.405; P=.03).

Most notably, the Granger causality analysis revealed that the government had a significant impact on the agenda of regular users through the attributes of "vaccine importance" ($F_{5,15}$=3.809; P=.02), "trust in experts" ($F_{5,15}$=16.639; $P<.001$), "vaccine accessibility" ($F_{5,15}$=3.343; P=.03), and "vaccine affordability" ($F_{3,13}$=6.012; P=.008). Despite the absence of Granger causality from regular users to the government for most attributes, there was a reciprocal relationship between the government and regular users in the attribute of "vaccine affordability." The results of the Granger causality tests between the government and other types of users are shown in Table 4.

Table 4. Granger causality tests between government users and other types of users for each vaccine attribute.

Vaccine attribute	Nongovernment users		Professional media		Alternative media		Civil societal organizations		Regular users	
	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable	Outcome variable	Antecedent variable
Importance										
F value (df)	1.410 (5,20)	1.209 (2,20)	1.413 (5,20)	7.192 (5,12)	2.412 (5,15)	2.407 (5,15)	4.111 (5,15)	1.801 (5,15)	3.809 (5,15)	2.259 (5,15)
P value	.26	.32	.26	.003	.09	.09	.01	.17	.02	.10
Effectiveness										
F value (df)	0.449 (2,30)	3.029 (3,13)	0.133 (2,30)	4.391 (3,13)	0.293 (3,13)	1.319 (5,9)	6.264 (3,13)	0.567 (2,10)	0.968 (2,30)	0.858 (2,30)
P value	.64	.07	.88	.02	.83	.34	.007	.58	.39	.44
Safety										
F value (df)	3.817 (3,13)	3.222 (3,13)	3.247 (5,15)	2.565 (5,15)	0.706 (1,15)	2.419 (5,15)	2.923 (3,13)	4.405 (3,12)	2.004 (5,15)	2.912 (3,22)
P value	.04	.057	.03	.07	.41	.08	.07	.03	.14	.057
Trust in government										
F value (df)	2.017 (3,15)	4.590 (3,13)	3.270 (3,13)	3.924 (3,13)	1.228 (2,20)	2.296 (3,13)	2.705 (3,9)	3.585 (2,10)	0.304 (2,10)	3.373 (3,13)
P value	.15	.02	.055	.03	.31	.12	.11	.07	.74	.051
Trust in experts										
F value (df)	3.916 (3,13)	0.402 (2,20)	3.753 (2,10)	1.437 (5,30)	0.401 (2,20)	1.146 (2,20)	15.877 (3,9)	1.058 (1,22)	16.639 (5,15)	4.189 (2,9)
P value	.03	.67	.06	.24	.67	.34	.001	.31	<.001	.051
COVID-19 risk										
F value (df)	0.255 (2,9)	1.124 (2,30)	1.890 (2,30)	5.173 (5,15)	0.665 (3,3)	8.769 (5,15)	2.442 (3,20)	2.275 (3,20)	0.655 (2,15)	0.235 (2,15)
P value	.78	.34	.16	.006	.63	<.001	.09	.11	.53	.79
Accessibility										
F value (df)	0.248 (2,15)	2.781 (3,13)	0.045 (2,10)	1.362 (5,20)	1.461 (5,15)	2.963 (5,15)	1.546 (5,15)	2.763 (5,10)	3.343 (5,15)	2.376 (5,15)
P value	.78	.08	.96	.28	.26	.047	.23	.08	.03	.09
Distribution										
F value (df)	0.756 (2,20)	0.104 (2,20)	0.596 (2,20)	0.283 (2,20)	0.147 (1,25)	0.005 (1,15)	1.264 (1,20)	0.382 (1,20)	4.175 (1,25)	0.458 (1,25)
P value	.48	.90	.56	.76	.70	.94	.27	.54	.051	.50
Affordability										
F value (df)	2.500 (3,13)	3.851 (3,13)	0.745 (2,20)	4.754 (3,13)	0.479 (2,20)	0.688 (2,20)	0.525 (2,20)	0.495 (2,20)	6.012 (3,13)	5.067 (2,20)
P value	.10	.04	.49	.02	.63	.51	.60	.62	.008	.02

Discussion

Principal Findings

This study examined the dynamics and patterns of vaccine communication on Facebook in Macao during the COVID-19 pandemic. The principal findings demonstrated that "vaccine importance" was the most prevalent attribute in the vaccination agenda on Facebook, followed by the attributes of "COVID-19 risk" and "trust in experts." The overall vaccination agenda revealed the highest co-occurrences were between "vaccine importance" and "COVID-19 risk." Differences existed in agenda priorities between the government and regular users. The government primarily focused on the risks of COVID-19 and the effectiveness of vaccines, whereas regular users were more concerned with the safety and distribution of vaccines. The Macao government played a role in shaping the agenda for regular users by highlighting vaccine importance (Granger causality result: $F_{5,15}$=3.809; P=.02), trust in experts (Granger causality result: $F_{5,15}$=16.639; P<.001), and vaccine accessibility (Granger causality result: $F_{5,15}$=3.343; P=.03) and affordability (Granger causality result: $F_{3,13}$=6.012; P=.008), while its impact on the agenda network of regular users remained insignificant (QAP result: b=0.246; P=.12). Both government and nongovernment users (eg, professional media, alternative media, civil organizations, and regular users) had intertwined agendas with mutual influence.

Unlike previous studies that predominantly focused on single aspects of vaccine communication (eg, [17,34]), this study used a more holistic approach to reveal the role of various actors including the government, professional media, alternative media, civil organizations, and regular users in promoting vaccination agendas and the interplay of diverse actors in the vaccine agenda setting process. The results of this study suggest that professional media acts as more than simple information providers to the government but rather effectively pushed agenda setting as a supplementary process to vaccine promotion by raising salient topics that the government fails to identify due to lack of information and experience. The government, however, is more likely to respond to professional media to receive timely feedback on vaccination issues for the purpose of learning and improvement. This can be observed from the impact that professional media has on the government in the agenda setting process through topics of “vaccine importance” (Granger causality results: $F_{5,12}$=7.192; P=.003), “vaccine effectiveness” (Granger causality results: $F_{3,13}$=4.391; P=.02), “trust in government” (Granger causality results: $F_{3,13}$=3.924; P=.03), “COVID-19 risk” (Granger causality results: $F_{5,15}$=5.173; P=.006), and vaccine affordability (Granger causality results: $F_{3,13}$=4.754; P=.02).

Who Leads the Vaccine Agenda of Whom?

Despite a significant correlation between the government agenda network and the agenda network of nongovernment users, the government had a limited impact on the agenda attributes of different Facebook user categories and vice versa. As Facebook is an open platform where information from a wide variety of sources freely circulates and interacts, it is difficult to determine the driving force behind the vaccine promotion agenda on the platform [55]. In other words, nongovernment users’ vaccine promotion agendas may have been impacted by other sources, such as the World Health Organization or other health professionals, which indicates a multidirectional effect.

As such, it appears that the government did not unilaterally set the agenda of nongovernment users. Instead, there is a “2-way” interaction between government and nongovernment user agendas. Due to their mutual effect, neither the government nor nongovernment users lead the agenda on social media. It is likely that the government and different types of nongovernment users pay attention to the agendas of one another and interact with one another to build the overall vaccine agenda network on Facebook. This corresponds with the argument by Finset et al [63] that, amid the near-chaotic flow of information, every individual, in different roles and with varied responsibilities, can contribute to the development of the information flow and agenda on COVID-19. A plausible explanation for this outcome could be the unprecedented nature of the health crisis. The lack of up-to-date crisis communication planning and experience with coping with a novel crisis may challenge the government’s agenda-setting process, particularly in terms of vaccine promotion.

Comparison With Prior Work

Previous agenda setting research found that changes in the government agenda led to changes in the public agenda [64]. However, during the COVID-19 pandemic, the public was no longer passive consumers of social media. Our results indicating the different concerns of vaccination between the government and regular users corroborate previous findings by Zhou and Zheng [44] who found that, during the COVID-19 pandemic, the government’s Weibo account exhibited a more propaganda-oriented approach, whereas public accounts were more attentive to issues that directly pertained to self-interest, such as protective measures against the virus and minimizing financial losses. Unlike other political issues, the government may have less impact on shaping public agenda due to the more collected information possessed by the public. This is partly consistent with some recent research indicating that shaping public opinion in a fragmented digital environment such as social media is challenging [54,65]. Additionally, the case of Macao also indicates selective public responsiveness on topics that are clear and straightforward, which partially verifies the observation by Kim [66] that individuals are more receptive to topics that are unambiguous and do not demand extensive background knowledge as they may not have enough background information with which to fully process any new information on complex topics.

Practical Implications

Our study provides several implications to inform the management of future pandemics. First, given the disparity between the government and public agenda networks, it is crucial to bridge the gap to enable effective vaccine communication. Policymakers should strive for alignment between government messaging and public concerns, addressing issues that are prominent within the public discourse. Social media listening activities are invaluable tools for understanding public health concerns. By monitoring public conversation through social media listening, policymakers can develop targeted messaging and communication strategies that effectively address public concerns and provide accurate information to dispel misconceptions.

Second, the low responsiveness of the public agenda to the government agenda indicates the need to enhance the government impact on the public agenda. Governments can streamline their messaging by using plain language, which helps individuals with different levels of knowledge understand information easily. Clear and concise presentation avoids unnecessary complexity. Visual aids and interactive media can also be used to improve public involvement and responsiveness, overcoming barriers caused by limited background information.

Third, policymakers' efforts to convince the public to receive vaccines in response to potential health risks have been shown in our study to lead to a spillover of media attention that significantly drives the vaccination agenda among the public. Collaboration with influential media, including professional and alternative media, thus offers a powerful means to facilitate vaccination policy and improve public health. Governments can utilize the extensive reach and persuasive power of media outlets to actively involve and inform the public about specific issues that should receive priority attention, thereby advancing the government's crisis management initiatives.

Fourth, civil organizations' ability to shape public attention toward vaccination issues by influencing the public agenda network suggests that their impact on shaping the vaccination agenda may be underestimated or overlooked. Driven by social responsibility, civil organizations often dedicate their efforts to promoting public health by increasing awareness and advocating for public health policies [43]. The close ties to communities enable them to be trusted sources of information for the public. Therefore, through partnerships with civil organizations, governments can leverage their networks, expertise, and community trust to effectively promote vaccination initiatives.

Limitations

Several limitations warrant consideration. First, broadening the scope beyond vaccines to encompass diverse policies could offer a more comprehensive understanding of public attention allocation mechanisms. Researchers are encouraged to explore various policies to enhance generalizability. Second, although Facebook data provided valuable insights, the findings are platform-specific and may not apply universally. Future studies should incorporate a diverse set of social media platforms and combine quantitative data with surveys and interviews for a more nuanced perspective. Third, although this study explored temporal agenda dynamics, it did not delve into the determinants driving public attention intensity, such as government transparency and issue salience. Investigating these factors could provide valuable insights into the agenda setting process at the government level.

Conclusions

This study investigated the communication dynamics of COVID-19 vaccines in Macao, with a specific focus on how government agendas impact other entities on Facebook. Our results reveal that the Macao Government's efforts to set the vaccination agenda on Facebook have shown limited effectiveness in shaping the public's discourse and priorities regarding vaccines. Such findings have profound implications for shaping government responses to future pandemics. Authorities, in their endeavor to legitimize policies, must recognize the intricate interplay between their agendas and public reception. Although agenda setting serves as a strategic tool to promote vaccination, it also exhibits limitations. This requires a shift toward more nuanced, strategy-focused research. This study offers indispensable insights in the area of crisis communication, underscoring the urgent necessity of bridging the gap between government and public agendas. Furthermore, it illuminates the potential of collaborations with influential media outlets and civil organizations as formidable channels to augment the reach and influence of vaccination agendas set by the government.

Authors' Contributions

XX conceptualized the study, curated the data, wrote the original manuscript draft, and created the visualizations. XX and AC performed the formal analysis. XX and RJN validated the data. AC and RJN reviewed and edited the manuscript. AC supervised the study and served as project administrator. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Keywords for vaccine-related data acquisition.
[DOCX File , 16 KB - infodemiology_v4i1e51113_app1.docx]

Multimedia Appendix 2
Coding framework for COVID-19 vaccine posts on Facebook.
[DOCX File , 21 KB - infodemiology_v4i1e51113_app2.docx]

Multimedia Appendix 3
Keywords for machine coding of vaccine-related topics.
[DOCX File , 19 KB - infodemiology_v4i1e51113_app3.docx]

Multimedia Appendix 4
Outcomes of post hoc tests on the significant difference between user categories and vaccine-related topics.
[DOCX File , 21 KB - infodemiology_v4i1e51113_app4.docx]

Multimedia Appendix 5
Intra-group co-occurrences dynamics of agenda attributes for the years of 2020, 2021, and 2022.
[DOCX File , 2025 KB - infodemiology_v4i1e51113_app5.docx]

References

1. COVID-19 Coronavirus Pandemic Statistics. Worldometer. URL: https://www.worldometers.info/coronavirus/ [accessed 2024-02-10]
2. Coccia M. Sources, diffusion and prediction in COVID-19 pandemic: lessons learned to face next health emergency. AIMS Public Health 2023;10(1):145-168 [FREE Full text] [doi: 10.3934/publichealth.2023012] [Medline: 37063362]
3. Coccia M. Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID. Sci Total Environ 2020 Aug 10;729:138474 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138474] [Medline: 32498152]
4. Coccia M. COVID-19 pandemic over 2020 (withlockdowns) and 2021 (with vaccinations): similar effects for seasonality and environmental factors. Environ Res 2022 May 15;208:112711 [FREE Full text] [doi: 10.1016/j.envres.2022.112711] [Medline: 35033552]
5. Magazzino C, Mele M, Coccia M. A machine learning algorithm to analyse the effects of vaccination on COVID-19 mortality. Epidemiol Infect 2022 Sep 12;150:e168 [FREE Full text] [doi: 10.1017/S0950268822001418] [Medline: 36093862]
6. Coccia M. COVID-19 vaccination is not a sufficient public policy to face crisis management of next pandemic threats. Public Organiz Rev 2022 Oct 17;23(4):1353-1367 [FREE Full text] [doi: 10.1007/s11115-022-00661-6]
7. Coccia M. Optimal levels of vaccination to reduce COVID-19 infected individuals and deaths: A global analysis. Environ Res 2022 Mar;204(Pt C):112314 [FREE Full text] [doi: 10.1016/j.envres.2021.112314] [Medline: 34736923]
8. Coccia M. Improving preparedness for next pandemics: Max level of COVID-19 vaccinations without social impositions to design effective health policy and avoid flawed democracies. Environ Res 2022 Oct;213:113566 [FREE Full text] [doi: 10.1016/j.envres.2022.113566] [Medline: 35660409]
9. Wagner AL, Zhang F, Kerekes S, Shih S, Zhao L. COVID-19 vaccination preferences during a pause in Johnson and Johnson vaccine administration. Vaccine X 2023 Dec;15:100373. [doi: 10.1016/j.jvacx.2023.100373] [Medline: 37674932]
10. Reichmuth ML, Heron L, Riou J, Moser A, Hauser A, Low N, et al. Socio-demographic characteristics associated with COVID-19 vaccination uptake in Switzerland: longitudinal analysis of the CoMix study. BMC Public Health 2023 Aug 10;23(1):1523 [FREE Full text] [doi: 10.1186/s12889-023-16405-0] [Medline: 37563550]
11. Ali F, Kaura A, Russell C, Bonn M, Bruneau J, Dasgupta N, et al. Identifying barriers and facilitators to COVID-19 vaccination uptake among people who use drugs in Canada: a national qualitative study. Harm Reduct J 2023 Jul 29;20(1):99 [FREE Full text] [doi: 10.1186/s12954-023-00826-6] [Medline: 37516836]
12. Blahut R, Flint A, Orlando E, DesChatelets J, Khowaja A. A scoping review on the decision-making dynamics for accepting or refusing the COVID-19 vaccination among adolescent and youth populations. BMC Public Health 2023 Apr 28;23(1):784 [FREE Full text] [doi: 10.1186/s12889-023-15717-5] [Medline: 37118794]
13. Soleimanpour H, Sarbazi E, Esmaeili ED, Mehri A, Fam SG, Nikbakht H, et al. Predictors of receiving COVID-19 vaccine among adult population in Iran: an observational study. BMC Public Health 2023 Mar 14;23(1):490 [FREE Full text] [doi: 10.1186/s12889-023-15409-0] [Medline: 36918858]
14. Yadav H, Sagar M. Exploring COVID-19 vaccine hesitancy and behavioral themes using social media big-data: a text mining approach. K 2023 Jun 09;52(7):2616-2648. [doi: 10.1108/k-06-2022-0810]
15. Dickson K, Aboltins C, Pelly J, Jessup RL. Effective communication of COVID-19 vaccine information to recently-arrived culturally and linguistically diverse communities from the perspective of community engagement and partnership organisations: a qualitative study. BMC Health Serv Res 2023 Aug 21;23(1):877 [FREE Full text] [doi: 10.1186/s12913-023-09836-3] [Medline: 37605184]
16. Park M, Park J, Kim H, Lee JH, Park H. Measuring the impacts of quantity and trustworthiness of information on COVID‐19 vaccination intent. Asso for Info Science & Tech 2023 Apr 17;74(7):846-865 [FREE Full text] [doi: 10.1002/asi.24760]
17. Pierz AJ, Rauh L, Masoud D, Cruz AK, Palmedo PC, Ratzan SC, et al. Supporting US healthcare providers for successful vaccine communication. BMC Health Serv Res 2023 May 02;23(1):423 [FREE Full text] [doi: 10.1186/s12913-023-09348-0] [Medline: 37131261]
18. Cmeciu C. (De)legitimation of COVID-19 vaccination narratives on Facebook comments in Romania: Beyond the co-occurrence patterns of discursive strategies. Discourse Soc 2023 May 26;4(5-29):09579265231174793 [FREE Full text] [doi: 10.1177/09579265231174793] [Medline: 37831753]
19. Athey S, Grabarz K, Luca M, Wernerfelt N. Digital public health interventions at scale: The impact of social media advertising on beliefs and outcomes related to COVID vaccines. Proc Natl Acad Sci U S A 2023 Jan 31;120(5):e2208110120 [FREE Full text] [doi: 10.1073/pnas.2208110120] [Medline: 36701366]
20. Benati I, Coccia M. Global analysis of timely COVID-19 vaccinations: improving governance to reinforce response policies for pandemic crises. IJHG 2022 May 31;27(3):240-253 [FREE Full text] [doi: 10.1108/ijhg-07-2021-0072]
21. Han Q, Zheng B, Cristea M, Agostini M, Bélanger JJ, Gützkow B, PsyCorona Collaboration, et al. Trust in government regarding COVID-19 and its associations with preventive health behaviour and prosocial behaviour during the pandemic:

a cross-sectional and longitudinal study. Psychol Med 2023 Jan;53(1):149-159 [FREE Full text] [doi: 10.1017/S0033291721001306] [Medline: 33769242]
22. Liu J, Shahab Y, Hoque H. Government response measures and public trust during the COVID‐19 pandemic: evidence from around the world. British J of Management 2021 Dec 23;33(2):571-602 [FREE Full text] [doi: 10.1111/1467-8551.12577]
23. Asia Pacific Immunization Coalition, L.E.K. Consulting. Best Practices for Vaccine Communications: Lessons Learned on Preparedness and Policy Implications From India, Indonesia and the Philippines. National University of Singapore. 2022 Jan. URL: https://medicine.nus.edu.sg/apic/wp-content/uploads/sites/28/2022/01/VF_Vaccine-Communication-Preparedness-Special-report.pdf [accessed 2024-02-10]
24. 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 Jun 06;44(3):320-346 [FREE Full text] [doi: 10.1177/10755470221101067]
25. Ung COL, Hu Y, Hu H, Bian Y. Investigating the intention to receive the COVID-19 vaccination in Macao: implications for vaccination strategies. BMC Infect Dis 2022 Mar 04;22(1):218 [FREE Full text] [doi: 10.1186/s12879-022-07191-y] [Medline: 35246072]
26. Xiong W, Peng L, Tsang TK, Cowling BJ. Epidemiology of SARS-CoV-2 Omicron BA.5 infections, Macau, June-July 2022. Emerg Infect Dis 2023 Feb;29(2):453-456 [FREE Full text] [doi: 10.3201/eid2902.221243] [Medline: 36648122]
27. Zhao Y, Xu J, Bai W, Sun H, Shui B, Yang Z, et al. COVID-19 prevention and control strategies: learning from the Macau model. Int J Biol Sci 2022;18(14):5317-5328 [FREE Full text] [doi: 10.7150/ijbs.70177] [Medline: 36147478]
28. Dai Y, Li Y, Cheng C, Zhao H, Meng T. Government-led or public-led? Chinese policy agenda setting during the COVID-19 pandemic. Journal of Comparative Policy Analysis: Research and Practice 2021 Apr 21;23(2):157-175 [FREE Full text] [doi: 10.1080/13876988.2021.1878887]
29. McCombs ME, Shaw DL. The agenda-setting function of mass media. Public Opinion Quarterly 1972;36(2):176-187 [FREE Full text] [doi: 10.1086/267990]
30. Rogers EM, Dearing JW, Bregman D. The anatomy of agenda-setting research. Journal of Communication 1993;43(2):68-84. [doi: 10.1111/j.1460-2466.1993.tb01263.x]
31. Guo L. The application of social network analysis in agenda setting research: a methodological exploration. Journal of Broadcasting & Electronic Media 2012 Oct;56(4):616-631 [FREE Full text] [doi: 10.1080/08838151.2012.732148]
32. McCombs ME, Guo L. Agenda-Setting Influence of the Media in the Public Sphere. In: Forner RS, Fackler PM, editors. The Handbook of Media and Mass Communication Theory. Hoboken, NJ: John Wiley & Sons, Inc; 2014.
33. Chen Z, Su CC, Chen A. Top-down or bottom-up? A network agenda-setting study of Chinese nationalism on social media. Journal of Broadcasting & Electronic Media 2019 Sep 20;63(3):512-533 [FREE Full text] [doi: 10.1080/08838151.2019.1653104]
34. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, 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 [FREE Full text] [doi: 10.2196/27632] [Medline: 34061757]
35. Buller D, Walkosz B, Henry K, Woodall WG, Pagoto S, Berteletti J, et al. Promoting social distancing and COVID-19 vaccine intentions to mothers: randomized comparison of information sources in social media messages. JMIR Infodemiology 2022 Aug 23;2(2):e36210 [FREE Full text] [doi: 10.2196/36210] [Medline: 36039372]
36. Albalawi Y, Sixsmith J. Agenda setting for health promotion: exploring an adapted model for the social media era. JMIR Public Health Surveill 2015;1(2):e21 [FREE Full text] [doi: 10.2196/publichealth.5014] [Medline: 27227139]
37. Chang A, Xian X, Liu MT, Zhao X. Health communication through positive and solidarity messages amid the COVID-19 pandemic: automated content analysis of Facebook uses. Int J Environ Res Public Health 2022 May 19;19(10):6159 [FREE Full text] [doi: 10.3390/ijerph19106159] [Medline: 35627696]
38. Social Media Stats Macaots. Statcounter. URL: https://gs.statcounter.com/social-media-stats/all/macao [accessed 2024-02-10]
39. Della Rosa S, Sen F. Health topics on Facebook groups: content analysis of posts in multiple sclerosis communities. Interact J Med Res 2019 Feb 11;8(1):e10146 [FREE Full text] [doi: 10.2196/10146] [Medline: 30741640]
40. Tahamtan I, Potnis D, Mohammadi E, Singh V, Miller LE. The mutual influence of the World Health Organization (WHO) and Twitter users during COVID-19: network agenda-setting analysis. J Med Internet Res 2022 Apr 26;24(4):e34321 [FREE Full text] [doi: 10.2196/34321] [Medline: 35275836]
41. Holt K, Ustad Figenschou T, Frischlich L. Key dimensions of alternative news media. Digital Journalism 2019 Jun 13;7(7):860-869 [FREE Full text] [doi: 10.1080/21670811.2019.1625715]
42. Lin ZX, Liu S. [The change of media ecology and online news practice in Macao]. Ershiyi Shiji 2019;176:95-112 [FREE Full text]
43. Knabe A, McCarthy M. Civil society organisations and public health research--evidence from eight European union new member states. Cent Eur J Public Health 2012 Dec;20(4):287-293 [FREE Full text] [doi: 10.21101/cejph.a3764] [Medline: 23441396]

44. Zhou S, Zheng X. Agenda dynamics on social media during COVID-19 pandemic: interactions between public, media, and government agendas. Communication Studies 2022 Jun 10;73(3):211-228 [FREE Full text] [doi: 10.1080/10510974.2022.2082504]
45. Chen L, Ling Q, Cao T, Han K. Mislabeled, fragmented, and conspiracy-driven: a content analysis of the social media discourse about the HPV vaccine in China. Asian Journal of Communication 2020 Sep 08;30(6):450-469 [FREE Full text] [doi: 10.1080/01292986.2020.1817113]
46. Li J, Zheng H. Coverage of HPV-related information on Chinese social media: a content analysis of articles in Zhihu. Hum Vaccin Immunother 2020 Oct 02;16(10):2548-2554 [FREE Full text] [doi: 10.1080/21645515.2020.1729028] [Medline: 32159420]
47. Thavareesan S, Mahesan S. Sentiment Lexicon Expansion using Word2vec and fastText for Sentiment Prediction in Tamil texts. 2020 Presented at: Moratuwa Engineering Research Conference (MERCon); July 28-30, 2020; Moratuwa, Sri Lanka URL: https://ieeexplore.ieee.org/document/9185369 [doi: 10.1109/MERCon50084.2020.9185369]
48. Orkphol K, Yang W. Word sense disambiguation using cosine similarity collaborates with Word2vec and WordNet. Future Internet 2019 May 12;11(5):114 [FREE Full text] [doi: 10.3390/fi11050114]
49. Zhou F, Zhang W, Cai H, Cao Y. Portrayals of 2v, 4v and 9vHPV vaccines on Chinese social media: a content analysis of hot posts on Sina Weibo. Hum Vaccin Immunother 2021 Nov 02;17(11):4433-4441 [FREE Full text] [doi: 10.1080/21645515.2021.1971016] [Medline: 34543155]
50. Chang A, Schulz PJ, Jiao W, Liu MT. Obesity-related communication in digital Chinese news from mainland China, Hong Kong, and Taiwan: automated content analysis. JMIR Public Health Surveill 2021 Nov 23;7(11):e26660 [FREE Full text] [doi: 10.2196/26660] [Medline: 34817383]
51. Chang A, Schulz PJ, Tu S, Liu MT. Communicative blame in online communication of the COVID-19 pandemic: computational approach of stigmatizing cues and negative sentiment gauged with automated analytic techniques. J Med Internet Res 2020 Nov 25;22(11):e21504 [FREE Full text] [doi: 10.2196/21504] [Medline: 33108306]
52. Wang D, Zhou Y, Ma F. Opinion leaders and structural hole spanners influencing echo chambers in discussions about COVID-19 vaccines on social media in China: network analysis. J Med Internet Res 2022 Nov 18;24(11):e40701 [FREE Full text] [doi: 10.2196/40701] [Medline: 36367965]
53. Feicheng M, Yating L. Utilising social network analysis to study the characteristics and functions of the co-occurrence network of online tags. Online Information Review 2014;38(2):232-247. [doi: 10.1108/OIR-11-2012-0124]
54. Wang X, Chen L, Shi J, Tang H. Who sets the agenda? the dynamic agenda setting of the wildlife issue on social media. Environmental Communication 2021 Apr 12;17(3):245-262 [FREE Full text] [doi: 10.1080/17524032.2021.1901760]
55. Vu H, Guo L, McCombs ME. Exploring “the World Outside and the Pictures in Our Heads”. Journalism & Mass Communication Quarterly 2014 Sep 09;91(4):669-686 [FREE Full text] [doi: 10.1177/1077699014550090]
56. Feng N, Feng H, Li D, Li M. Online media coverage, consumer engagement and movie sales: A PVAR approach. Decision Support Systems 2020 Apr;131:113267 [FREE Full text] [doi: 10.1016/j.dss.2020.113267]
57. Lorenz N, Sander C, Ivanova G, Hegerl U. Temporal associations of daily changes in sleep and depression core symptoms in patients suffering from major depressive disorder: idiographic time-series analysis. JMIR Ment Health 2020 Apr 23;7(4):e17071 [FREE Full text] [doi: 10.2196/17071] [Medline: 32324147]
58. Yang H, Wang S, Ren Z, Liu H, Tong Y, Wang N. Life expectancy, air pollution, and socioeconomic factors: a multivariate time-series analysis of Beijing City, China. Soc Indic Res 2022 Jan 03;162(3):979-994 [FREE Full text] [doi: 10.1007/s11205-021-02872-8]
59. Golovko D, Schumann JH. Influence of company Facebook activities on recruitment success. Journal of Business Research 2019 Nov;104:161-169 [FREE Full text] [doi: 10.1016/j.jbusres.2019.06.029]
60. Geiß S. The media’s conditional agenda-setting power: how baselines and spikes of issue salience affect likelihood and strength of agenda-setting. Communication Research 2019 Sep 19;49(2):296-323 [FREE Full text] [doi: 10.1177/0093650219874968]
61. Widmann T. Fear, hope, and COVID-19: emotional elite rhetoric and its impact on the public during the first wave of the COVID-19 pandemic. Polit Psychol 2022 May 09;43(5):827-850 [FREE Full text] [doi: 10.1111/pops.12831] [Medline: 35941918]
62. BarberÁ P, Casas A, Nagler J, Egan PJ, Bonneau R, Jost JT, et al. Who leads? Who follows? Measuring issue attention and agenda setting by legislators and the mass public using social media data. Am Polit Sci Rev 2019 Jul 12;113(4):883-901 [FREE Full text] [doi: 10.1017/S0003055419000352] [Medline: 33303996]
63. 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]
64. Luo Y. The Internet and agenda setting in China: The influence of online public opinion on media coverage and government policy. International Journal of Communication 2014;8:1289-1312 [FREE Full text]
65. Gruszczynski M, Wagner MW. Information flow in the 21st century: the dynamics of agenda-uptake. Mass Communication and Society 2016 Nov 18;20(3):378-402 [FREE Full text] [doi: 10.1080/15205436.2016.1255757]

66. Kim Y. Contingent factors of agenda-setting effects: How need for orientation, issue obtrusiveness,message tone influence issue salienceattitude strength. In: Johnson TJ, editor. Agenda setting in a 2.0 world. New York, NY: Routledge; 2013:65-81.

Abbreviations

NAS: network agenda setting model
QAP: quadratic assignment procedure
RQ: research question
VAR: vector autoregression

Edited by W Ahmed; submitted 21.07.23; peer-reviewed by A Staffini, M Coccia; comments to author 29.09.23; revised version received 19.11.23; accepted 29.11.23; published 19.03.24.

Please cite as:
Xian X, Neuwirth RJ, Chang A
Government-Nongovernmental Organization (NGO) Collaboration in Macao's COVID-19 Vaccine Promotion: Social Media Case Study
JMIR Infodemiology 2024;4:e51113
URL: https://infodemiology.jmir.org/2024/1/e51113
doi:10.2196/51113
PMID:38502184

Original Paper

Development of a Medical Social Media Ethics Scale and Assessment of #IRad, #CardioTwitter, and #MedTwitter Posts: Mixed Methods Study

Vongai Christine Mlambo[1], BSc; Eric Keller[2], MA, MD; Caroline Mussatto[3], BSc, MD; Gloria Hwang[2], MD

[1]School of Medicine, Stanford University, Palo Alto, CA, United States

[2]Division of Interventional Radiology, Stanford Health Care, Stanford, CA, United States

[3]School of Medicine, University of Kansas Medical Center, Kansas City, KS, United States

Corresponding Author:
Vongai Christine Mlambo, BSc
School of Medicine
Stanford University
751 Layne Court
Apartment 25
Palo Alto, CA, 94306
United States
Phone: 1 6506605737
Email: vmlambo@stanford.edu

Abstract

Background: Social media posts by clinicians are not bound by the same rules as peer-reviewed publications, raising ethical concerns that have not been extensively characterized or quantified.

Objective: We aim to develop a scale to assess ethical issues on medical social media (SoMe) and use it to determine the prevalence of these issues among posts with 3 different hashtags: #MedTwitter, #IRad, and #CardioTwitter.

Methods: A scale was developed based on previous descriptions of professionalism and validated via semistructured cognitive interviewing with a sample of 11 clinicians and trainees, interrater agreement, and correlation of 100 posts. The final scale assessed social media posts in 6 domains. This was used to analyze 1500 Twitter posts, 500 each from the 3 hashtags. Analysis of posts was limited to original Twitter posts in English made by health care professionals in North America. The prevalence of potential issues was determined using descriptive statistics and compared across hashtags using the Fisher exact and χ^2 tests with Yates correction.

Results: The final scale was considered reflective of potential ethical issues of SoMe by participants. There was good interrater agreement (Cohen κ=0.620, *P*<.01) and moderate to strong positive interrater correlation (=0.602, *P*<.001). The 6 scale domains showed minimal to no interrelation (Cronbach α=0.206). Ethical concerns across all hashtags had a prevalence of 1.5% or less except the conflict of interest concerns on #IRad, which had a prevalence of 3.6% (n=18). Compared to #MedTwitter, posts with specialty-specific hashtags had more patient privacy and conflict of interest concerns.

Conclusions: The SoMe professionalism scale we developed reliably reflects potential ethical issues. Ethical issues on SoMe are rare but important and vary in prevalence across medical communities.

(JMIR Infodemiology 2024;4:e47770) doi:10.2196/47770

KEYWORDS

ethics; social media; conflict of interest; interventional radiology; X; Twitter; cardiology; privacy; ethical issues; medical social media; prevalence; professional; professionalism

Introduction

The digital footprint of clinicians on social media has increased over the past 10 years with an estimated 90% and 65% of clinicians using social media for personal and professional purposes, respectively [1]. Medical social media (SoMe) has blossomed, offering clinicians opportunities to collaborate across distances, debate treatment approaches for challenging cases, and engage in public health advocacy [2-4]. However, this rapid integration of social media in health care has outpaced guidance that counsels on how to avoid ethical concerns that can occur with SoMe [2].

The risks of SoMe have not gone unnoticed. Several professional organizations have released statements outlining guiding principles for online clinician behavior, including the American Medical Association and the Federation of State Medical Boards [5,6]. There have also been opinion pieces and recommendations published within various specialties such as neurology, dermatology, and vascular surgery [7-9]. Guidelines and opinion pieces are helpful starting points but may not address subtle but important breaches in professionalism [10] and may fail to resonate with the majority of users' experiences and values [2].

A few studies have assessed the prevalence of issues such as violations of the HIPAA (Health Insurance Portability and Accountability Act) [10]. However, the potential issues are much broader than explicit patient privacy violations [10,11]. This study sought to develop a more complete scale of ethical issues related to medical SoMe to provide empirical data on these issues. The authors hypothesized that a scale could be developed that captures the most salient ethical issues with good interrater agreement and correlation. The authors also hypothesized that applying such a scale would find that the prevalence of issues was small and varied across different professional groups.

Methods

Scale Development

This study was approved by the Stanford University Institutional Review Board (eProtocol 60351). An initial draft of the scale was developed based on medical professionalism in the new millennium: a physician charter created by the American Board of Internal Medicine Foundation, American College of Physicians Foundation, and the European Federation of Internal Medicine as well as a study by Chandratilake et al [12] assessing definitions of medical professionalism across cultures [13]. These sources were selected to attempt to define medical SoMe ethics that would be reflective of common definitions of medical professionalism. The initial draft consisted of 5 criteria rated on a 3-point scale: no ethical concern (0), potential ethical concern (1), and clear ethical concern (2). The 3-point scale was selected to reflect a concept raised by both initial sources that ethical issues occur on a continuum, allowing the scale to also capture less overt violations of professionalism.

The initial scale was then vetted for validity via semistructured cognitive interviewing with a group of clinicians and trainees [14]. Interviewees were recruited via email and were primarily a convenience sample at the authors' institutions. They were invited to provide feedback on a draft of the scale, which included fabricated posts and example scoring for demonstration. Purposeful recruiting was used to ensure that interviewees were diverse in terms of specialty, training level, and gender identity. Iterative adjustments were made to the initial scale based on interviewee feedback until additional interviews continued suggesting that the scale was reflective of interviewee perceptions of potential ethical issues related to medical SoMe. This occurred after 11 interviews with interviewees from 6 different specialties whose demographics are shown in Table 1.

Table 1. Demographic characteristics of interviewees (N=11).

Characteristic	Interviewees, n (%)
Training level	
1st-year MD[a]	2 (18)
2nd-year MD	0 (0)
3rd-year MD	2 (18)
4th+ year MD	1 (9)
1st-year resident	0 (0)
2nd-year resident	1 (9)
3rd+ year resident	1 (9)
Attending	4 (36)
Institution	
Stanford University School of Medicine	8 (73)
University of California San Diego	1 (9)
University of Kansas Medical Center	2 (18)
Specialty	
Anesthesiology	2 (18)
DR[b]/IR[c]	1 (9)
Emergency medicine	1 (9)
Primary care	1 (9)
Psychiatry	1 (9)
Otolaryngology	1 (9)
Undeclared	4 (36)
Sex	
Female	7 (64)
Male	4 (36)

[a]MD: Doctor of Medicine.

[b]DR: Diagnostic Radiology.

[c]IR: Interventional Radiology.

The vetted scale scored posts on 6 domains, using the same 3-point scale (Table 2). Scale item interrelation as well as scale interrater agreement and correlation were assessed by having 2 researchers use the scale to independently rate 50 random posts each from #MedTwitter between June 15, 2021, and August 15, 2021, with an overlap of 10 tweets. Posts were identified using the Healthcare Hashtag Project (Symplur, LLC). The interrelation of scale items was assessed via Cronbach α. Interrater agreement was assessed via Cohen κ and interrater correlation was assessed via Spearman correlation coefficient, assuming a nonlinear relationship. An α of <.05 was predefined as statistical significance.

Table 2. Medical social media professionalism scale.

Principle	Score		
	0=no concern	1=minor concern	2=major concern
Patient privacy			
Does the post maintain patient privacy by applying appropriate safeguards for patient information and removing patient identifiers?	Post omits HIPAA[a] identifiers and any other details that in combination would enable patient identification.	Post omits HIPAA identifiers but uses information that could potentially allow for patient identification, particularly when combined with the author's known practice location, medical specialty, or rarity of medical condition.	Post uses one or more HIPAA identifiers that allows for easy identification.
Patient dignity			
Does the post treat patients with respect and avoid the use of degrading language or images?	Post treats patients as individuals worthy of respect and does not demean the patient in any way.	Post contains references, images, or language that could be negatively construed such that some may take offense.	Post is objectifying or dehumanizing, treating patients as being of lesser intelligence or caliber.
Information accuracy			
Is the information medically accurate with no counterfactual, exaggerated, or otherwise misleading content?	Information in the post is reasonably supported by current evidence and does not make superlative claims.	Information in the post is ambiguous or exaggerated in a manner that could lead to misinterpretation.	Information in the post is overtly sensational and makes baseless claims.
Conflict of interest			
Is the post unduly influenced by ulterior motives for private gain without proper acknowledgment or disclosure in a way that could affect information accuracy?	The post does not promote or endorse products or services without an appropriate declaration of any associated financial ties.	The post promotes or endorses products or services without a declaration of conflicts, however, it does not make authoritative claims about these products.	The post promotes or endorses products or services without a proper declaration of conflicts and also makes authoritative claims about these products.
Justice and equity			
Is the text or images in the post discriminatory based on race, gender, socioeconomic status, ethnicity, religion, sexual orientation, or any other social category and does the post promote further inequities in health care?	The post does not express or imply any discriminatory sentiments or propagate a stance that either sustains or widens inequities in health care.	The post contains ideas associated with stereotypes or broad generalizations *without* suggesting the differential treatment of individuals based on these stereotypes.	The post explicitly expresses sentiments that are discriminatory and is a proponent for the differential treatment of individuals based on these prejudiced notions.
Interprofessional respect			
Does the post treat colleagues and other health care professionals with respect and avoid the use of stereotypes, mockery, and incivility?	Post treats colleagues and other health care professionals with esteem and does not demean them in any way.	Post contains references, images, or language that could be negatively construed by other colleagues as offensive.	Post clearly mocks or disrespects colleagues, portraying them as inferior or of lesser intelligence or caliber.

[a]HIPAA: Health Insurance Portability and Accountability Act.

Evaluation of Posts

The validated scale was then used to assess the prevalence of ethical issues among posts using 3 distinct hashtags: #MedTwitter, #IRad, and #CardioTwitter. These were selected as they are the most frequently used hashtags among the general medical community, interventional radiologists, and cardiologists, respectively, as indicated by the number of posts per day for each hashtag on the Symplur software. Interventional Radiology (IR) and cardiology were selected to provide examples of more specialty-specific posts to contrast with #MedTwitter as they are primarily used by physician specialists in those fields to discuss more expert medical content compared to #MedTwitter. Posts were limited to those in English posted by individuals (rather than societies or bots) who are clinicians or health care trainees in North America between December 10, 2021, and January 10, 2022. Retweets were also excluded. A total of 1500 posts were analyzed, 500 from each hashtag. Data were analyzed using descriptive statistics as well as Fisher exact tests and χ^2 tests with Yates correction to compare the prevalence of ethical issues across hashtags. These statistical tests were selected to adjust for the low rates of ethical issues. All statistical analyses were performed using SPSS software (IBM, Inc).

Ethical Considerations

All procedures were approved by the Stanford University Institutional Review Board (IRB#: 60351) and were per the legal and ethical standards of the responsible committee on human experimentation institutionally. Additionally, we adhered to local, national, regional, and international laws and regulations regarding the protection of personal information, privacy, and human rights.

Results

Scale Development

Cognitive interviewing supported the validity of the initial 5 domains. However, the initial interviewees felt the initial scale did not address interspecialty and inter–health care professional cyberbullying, leading to the addition of interprofessional respect as a 6th domain. Interviewees also suggested the addition of language to better delineate a minor concern (1) rating from a major concern (2) rating. Subsequent interviews confirmed that the 6-domain scale, each rated from 0 to 2, was reflective of their perceptions of SoMe ethics.

The scale demonstrated good interrater agreement (Cohen κ=0.620, *P*<.01) and moderate to strong positive correlation between the scores given by the independent raters (Spearman correlation coefficient=0.602, 95% CI 0.515-0.677; *P*<.001). The scale domains showed minimal to no interrelation (Cronbach α=0.206).

Evaluation of Posts

Application of the scale to 1500 Twitter posts showed that ethical concerns across all 6 domains were infrequent with the majority in the range of 0.2% (n=1) to 1.2% (n=6). Further, 1 exception was a minor conflict of interest concern among posts using #IRad, which demonstrated a prevalence of 3.6% (n=18). Relative to posts using #MedTwitter, posts using #IRad or #CardioTwitter were more likely to have patient privacy concerns (n=7, 1.4% vs 0%, *P*=.02; n=6, 1.2% vs 0%, *P*=.04; respectively). Posts using #IRad were also more likely to have conflicts of interest concerns relative to #MedTwitter and #CardioTwitter (n=18, 3.6% vs n=3, 0.6%, *P*<.001; n=18, 3.6% vs n=4, 0.8%, *P*=.005; respectively). Issues related to interprofessional respect were also more prevalent in #IRad posts than #CardioTwitter (n=8, 1.6% vs n=1, 0.2%, *P*=.04) but similar to #MedTwitter (n=8, 1.6% vs n=6, 1.2%, *P*=.79). As a result, across all domains, #IRad posts had the greatest overall prevalence of ethical concerns. Table 3 summarizes the prevalence of ethical concerns by hashtag and domain and Tables 4-6 summarize comparisons between hashtags.

Table 3. Prevalence of ethical concerns on medical social media by hashtag (N=500).

	No issue (0), n (%)	Minor concern (1), n (%)	Major concern (2), n (%)
MedTwitter prevalence			
Patient privacy	500 (100)	0 (0)	0 (0)
Patient dignity	495 (99)	3 (0.6)	2 (0.4)
Information accuracy	497 (99.4)	2 (0.4)	1 (0.2)
Conflict of interest	500 (100)	0 (0)	0 (0)
Justice and equity	499 (99.8)	1 (0.2)	0 (0)
Interprofessional respect	494 (98.8)	4 (0.8)	2 (0.4)
IR[a] prevalence			
Patient privacy	493 (98.6)	6 (1.2)	1 (0.2)
Patient dignity	497 (99.4)	1 (0.2)	2 (0.4)
Information accuracy	497 (99.4)	2 (0.4)	1 (0.2)
Conflict of interest	482 (96.4)	18 (3.6)	0 (0)
Justice and equity	500 (100)	0 (0)	0 (0)
Interprofessional respect	492 (98.4)	7 (1.4)	1 (0.2)
Cardiology prevalence			
Patient privacy	494 (98.8)	6 (1.2)	0 (0)
Patient dignity	499 (99.8)	1 (0.2)	0 (0)
Information accuracy	500 (100)	0 (0)	0 (0)
Conflict of interest	496 (99.2)	2 (0.4)	2 (0.4)
Justice and equity	500 (100)	0 (0)	0 (0)
Interprofessional respect	499 (99.8)	1 (0.2)	0 (0)

[a]IR: Interventional Radiology.

Table 4. Comparison of ethical concerns on medical social media by hashtag[a]: #IRad vs #MedTwitter."

	#IRad, n (%)	#MedTwitter, n (%)	Fisher exact *P* value	Chi-squared with Yates correction *P* value
#IRad vs #MedTwitter				
Patient privacy	7 (1.4)[b]	0 (0)[b]	.02[b]	.02[b]
Patient dignity	3 (0.6)	5 (1)	.73	.72
Information accuracy	3 (0.6)	3 (0.6)	≥.99	≥.99
Conflict of interest	18 (3.6)[b]	0 (0)[b]	<.001[b]	<.001[b]
Justice and equity	0 (0)	1 (0.2)	≥.99	.32
Interprofessional respect	8 (1.6)	6 (1.2)	.79	.79

[a]Comparisons reflect the composite of major and minor concerns for each scale criterion. *P*<.05 on a 2-tailed analysis was considered significant.
[b]Comparisons that are significant.

Table 5. Comparison of ethical concerns on medical social media by hashtag[a]: #CardioTwitter vs #MedTwitter."

	#CardioTwitter, n (%)	#MedTwitter, n (%)	Fisher exact *P* value	Chi-squared with Yates correction *P* value
#CardioTwitter vs MedTwitter				
Patient privacy	6 (1.2)[b]	0 (0)[b]	.03[b]	.04[b]
Patient dignity	1 (0.2)	5 (1)	.22	.22
Information accuracy	0 (0)	3 (0.6)	.37	.62
Conflict of interest	4 (0.8)	0 (0)	.22	.37
Justice and equity	0 (0)	1 (0.2)	≥.99	.32
Interprofessional respect	1 (0.2)	6 (1.2)	.12	.13

[a]Comparisons reflect the composite of major and minor concerns for each scale criterion. *P*<.05 on a 2-tailed analysis was considered significant.
[b]Comparisons that are significant.

Table 6. Comparison of ethical concerns on medical social media by hashtag[a]: #IRad vs #CardioTwitter."

	#IRad, n (%)	#CardioTwitter, n (%)	Fisher exact *P* value	Chi-squared with Yates correction *P* value
#IRad vs #CardioTwitter				
Patient privacy	7 (1.4)	6 (1.2)	≥.99	.78
Patient dignity	3 (0.6)	1 (0.2)	.62	.62
Information accuracy	3 (0.6)	0 (0)	.37	.62
Conflict of interest	18 (3.6)[b]	4 (0.8)[b]	.004[b]	.005[b]
Justice and equity	0 (0)	0 (0)	≥.99	≥.99
Interprofessional respect	8 (1.6)[b]	1 (0.2)[b]	.04[b]	.04[b]

[a]Comparisons reflect the composite of major and minor concerns for each scale criterion. *P*<.05 on a 2-tailed analysis was considered significant.
[b]Comparisons that are significant.

Discussion

Principal Results

This study sought to develop a scale to characterize and quantitate ethical issues on SoMe and then apply the scale to 3 different SoMe communities based on Twitter hashtags. Although some guidelines and opinion pieces exist describing potential ethical issues on SoMe, to the best of the authors' knowledge, no scales had been created, making it difficult to assess the prevalence of ethical issues and guide efforts to mitigate potential harm [10]. This is important not only because of legal implications, but this behavior can exacerbate existing hierarchies and damage mutual trust.

The scale proposed in this study was developed via a structured deductive and inductive approach. Key domains were identified based on literature review as well as qualitative interviews, consistent with best practices in scale development [15,16]. This helped ensure that the scale was comprehensive and perceived as valid. Interrater agreement and correlation were good but likely limited by the qualitative nature of these assessments. The lack of interrelation between domains is not unexpected. A post with a patient privacy concern would not necessarily be more likely to have a conflict of interest as well.

Application of the scale to Twitter posts with #MedTwitter, #CardioTwitter, and #IRad yielded a couple of important observations. First, the prevalence of ethical concerns is low, often around 1% (n=5) across domains. However, such a number is not insignificant. According to Symplur software, there are approximately 5000 to 8000 posts per day made using #MedTwitter, equating to approximately 50-80 ethically concerning posts per day. These findings are similar to a 2011 study of over 5000 general tweets from health care providers, which found 3% of tweets were unprofessional and 0.7% were concerning for breaches in patient privacy [17].

A second interesting observation was how the prevalence of ethical concerns varied across the 3 groups of posts analyzed. For example, posts with the specialty-specific hashtags #CardioTwitter and #IRad had more patient privacy and conflict of interest concerns than general #MedTwitter posts. This may be due to a higher likelihood of posting specific patient cases in specialty-specific communities to illustrate an approach or solicit recommendations compared to the general #MedTwitter community. Posts with conflict of interest were also most prevalent in #IRad posts, which may be due to IR being a more procedural specialty than cardiology in general, and a specialty whose professional identity is closely tied to specific procedures and devices rather than patient populations [18]. Previous authors have observed similar variations in posts across specialties. The dominating content among IR posts tends to be images of an intervention performed on a patient to share new techniques or gather recommendations for superior approaches [19]. In contrast, cardiology posts are dominated by short synopses of trending research papers with reactive commentary [20]. However, interventional cardiology posts can share similar traits to IR [20,21], likely accounting for some of the overlap in the ethical issues among these posts.

Practical Implications

The persistence of posts with ethical issues among medical professionals and trainees invites evaluation of current social media training programs. The domains in the scale offer a useful framework with validated language and examples to offer caution against ethical concerns that go beyond HIPAA violations. The framework can also foster a mental model to assist in evaluating personal tweets before publishing a post. This is important as once a post is made; it is difficult to retract it completely before it is shared or copied by other users.

The results from this study also provide a foundation for evidence-based social media guidelines by professional bodies and specialty-specific societies. As demonstrated by differences in the prevalence of ethical concerns between #CardioTwitter and #IRad, not all ethical issues are equally problematic, and with this data, guidelines can be tailored to the target group. This scale can be applied to hashtags used by other specialists to uncover trends in ethical issues and address those weak points more specifically. For example, social media statements for interventional radiologists may include more specific and detailed guidance on avoiding conflict of interest concerns.

From an academic perspective, the scale and methodology described in this study offer a way to assess the efficacy of interventions aimed at reducing the frequency of ethical issues on SoMe. Previously, there were limited ways to quantify and characterize the landscape of SoMe professionalism. However, now it is possible to perform pre- and poststudies with a specific intervention of interest.

Although this study focused on the application of the professionalism scale to Twitter posts as a proof of concept, the principles could be translated to other platforms as they do not include any evaluation metric that is inherent to Twitter, since the development of the scale was independent of any specific platform. From a validation perspective, this translation would be easiest for platforms that mimic Twitter by using a combination of texts and images, such as Facebook and Instagram posts. Importantly, videos were not assessed in this study, which would be of interest in analyzing Reels, TikTok, and YouTube videos. However, the methodology of this study can be applied to these different social media contexts to assess the generalizability of the scale.

Limitations and Future Directions

This study had important limitations. The scale provides a good estimate of the prevalence of ethical issues, but it is not a thorough investigation of whether a given issue definitively exists especially for domains like conflict of interest that are challenging to verify without collateral information. Although the scale development incorporated input from a diverse group of clinicians and trainees in terms of training level, specialty, and gender identity, the sample was a small convenience sample from academic settings that could have missed important input from other clinicians in different contexts, for example, private practice. The sample was limited to posts in English from North America due to language restrictions and greater cultural familiarity. However, this may limit the external validity of the scale and results in other cultures. The authors relied on self-described Twitter biographies to limit posts to health care professionals, which could have been inaccurate.

To address some of these limitations, future steps to continue improving the scale would include expanding the sample to include more physicians and trainees from private practice, community hospitals, and primary care so that these additional perspectives can further refine the scale. Additionally, although the Cohen κ for interrater reliability already suggests good agreement, there may be domains with greater discrepancies than others. The language of these domains can be made more precise or explicit based on a bigger sample feedback to potentially improve consistency. Lastly, a comparison among different platforms would help directly assess if scale validity transcends social media contexts.

Conclusions

The developed SoMe ethics scale is reliable, relevant, and concisely captures the myriad ethical tensions that can arise on these platforms. Ethical issues are present in a small but meaningful percentage of posts among health care professionals, which vary in important ways across different specialties and professional groups. The authors hope this scale will allow researchers to better characterize and assess the prevalence of ethical issues on SoMe while guiding more targeted interventions to mitigate these issues.

Conflicts of Interest

None declared.

References

1. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014;39(7):491-520 [FREE Full text] [Medline: 25083128]
2. Panahi S, Watson J, Partridge H. Social media and physicians: exploring the benefits and challenges. Health Informatics J 2016;22(2):99-112 [FREE Full text] [doi: 10.1177/1460458214540907] [Medline: 25038200]
3. McGowan BS, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. J Med Internet Res 2012;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]
4. O'Glasser AY, Jaffe RC, Brooks M. To tweet or not to tweet, that is the question. Semin Nephrol 2020;40(3):249-263. [doi: 10.1016/j.semnephrol.2020.04.003] [Medline: 32560773]
5. Farnan JM, Sulmasy LS, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. Ann Intern Med 2013;158(8):620-627 [FREE Full text] [doi: 10.7326/0003-4819-158-8-201304160-00100] [Medline: 23579867]
6. Association AM. The AMA code of medical ethics' opinions on observing professional boundaries and meeting professional responsibilities. AMA J Ethics 2015;17:432-434 [FREE Full text] [doi: 10.1001/journalofethics.2015.17.5.coet1-1505]
7. Busl KM, Rubin MA, Tolchin BD, Larriviere D, Epstein L, Kirschen M, et al. Use of social media in health care-opportunities, challenges, and ethical considerations: a position statement of the American Academy of Neurology. Neurology 2021;97(12):585-594 [FREE Full text] [doi: 10.1212/WNL.0000000000012557] [Medline: 34864637]
8. Militello M, Yang RA, Anderson JB, Szeto MD, Presley CL, Laughter MR. Social media and ethical challenges for the dermatologist. Curr Dermatol Rep 2021;10(4):120-127 [FREE Full text] [doi: 10.1007/s13671-021-00340-7] [Medline: 34540357]
9. Gifford ED, Mouawad NJ, Bowser KE, Bush RL, Chandra V, Coleman DM, et al. Society for vascular surgery best practice recommendations for use of social media. J Vasc Surg 2021;74(6):1783.e1-1791.e1 [FREE Full text] [doi: 10.1016/j.jvs.2021.08.073] [Medline: 34673169]
10. Ahmed W, Jagsi R, Gutheil TG, Katz MS. Public disclosure on social media of identifiable patient information by health professionals: content analysis of Twitter data. J Med Internet Res 2020;22(9):e19746 [FREE Full text] [doi: 10.2196/19746] [Medline: 32870160]
11. Attai DJ, Anderson PF, Fisch MJ, Graham DL, Katz MS, Kesselheim J, et al. Risks and benefits of Twitter use by hematologists/oncologists in the era of digital medicine. Semin Hematol 2017;54(4):198-204 [FREE Full text] [doi: 10.1053/j.seminhematol.2017.08.001] [Medline: 29153081]
12. Chandratilake M, McAleer S, Gibson J. Cultural similarities and differences in medical professionalism: a multi-region study. Med Educ 2012;46(3):257-266. [doi: 10.1111/j.1365-2923.2011.04153.x] [Medline: 22324525]
13. ABIM Foundation, ACP-ASIM Foundation, European Federation of Internal Medicine. Medical professionalism in the new millennium: a physician charter. Ann Intern Med 2002;136(3):243-246 [FREE Full text] [doi: 10.7326/0003-4819-136-3-200202050-00012] [Medline: 11827500]
14. Schwarz NE, Sudman SE. Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research. Francisco, CA: Jossey-Bass/Wiley; 1996.
15. Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC. Scale development: ten main limitations and recommendations to improve future research practices. Psicol Reflex Crit 2018;30(1):3 [FREE Full text] [doi: 10.1186/s41155-016-0057-1] [Medline: 32025957]
16. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
17. Ahmed O, Jilani S, Ginsburg M, Hadied O, Tasse J, Loanzon R, et al. You are what you tweet: navigating legal issues in social media for interventional radiologists. J Vasc Interv Radiol 2018;29(6):816-818. [doi: 10.1016/j.jvir.2018.02.012] [Medline: 29798756]
18. Keller EJ, Vogelzang RL. Who we are and what we can become: the anthropology of IR and challenges of forming a new specialty. J Vasc Interv Radiol 2018;29(12):1703.e2-1704.e2. [doi: 10.1016/j.jvir.2018.07.021] [Medline: 30502878]

19. Rostampour S, Hamady MS, Alsafi A. To tweet or not to tweet? A look at radiology societies' use of Twitter. Cardiovasc Intervent Radiol 2020;43(7):1070-1074. [doi: 10.1007/s00270-020-02437-1] [Medline: 32239244]
20. Goldsweig AM, Galper BZ, Alraies C, Arnold SV, Daniels M, Capodanno D, et al. #SoMe for #IC: optimal use of social media in interventional cardiology. Catheter Cardiovasc Interv 2021;98(1):97-106 [FREE Full text] [doi: 10.1002/ccd.29643] [Medline: 33686726]
21. Fischman DL, Savage MP. Cardiotwitter: new virtual tools to advance skillsets in interventional cardiology. Curr Cardiol Rev 2021;17(2):157-160 [FREE Full text] [doi: 10.2174/1573403X15666191126104402] [Medline: 31769364]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
IR: Interventional Radiology
SoMe: medical social media

Edited by E Lee; submitted 31.03.23; peer-reviewed by B Chakalov; comments to author 18.09.23; revised version received 25.09.23; accepted 30.09.23; published 27.03.24.

Please cite as:
Mlambo VC, Keller E, Mussatto C, Hwang G
Development of a Medical Social Media Ethics Scale and Assessment of #IRad, #CardioTwitter, and #MedTwitter Posts: Mixed Methods Study
JMIR Infodemiology 2024;4:e47770
URL: https://infodemiology.jmir.org/2024/1/e47770
doi:10.2196/47770
PMID:38536206

Corrigenda and Addenda

Correction: Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users

Ryuichiro Ueda[1], MHA; Feng Han[2], MHA; Hongjian Zhang[1], PhD; Tomohiro Aoki[1], MHA; Katsuhiko Ogasawara[1], MBA, PhD

[1]Faculty of Health Sciences, Hokkaido University, Sapporo, Japan
[2]Graduate School of Medicine, Hokkaido University, Sapporo, Japan

Corresponding Author:
Katsuhiko Ogasawara, MBA, PhD
Faculty of Health Sciences
Hokkaido University
N12-W5, Kita-ku
Sapporo, 060-0812
Japan
Phone: 81 11 706 3409
Email: oga@hs.hokudai.ac.jp

Related Article:

Correction of: https://infodemiology.jmir.org/2024/1/e37881

(JMIR Infodemiology 2024;4:e57880) doi:10.2196/57880

In “Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users” (JMIR Infod 2024;3(1):e37881) the authors made 3 corrections.

1. The authorship list was previously listed as:

> *Ryuichiro Ueda, MA; Feng Han, MA; Hongjian Zhang, MD; Tomohiro Aoki, MA; Katsuhiko Ogasawara, Prof Dr*

And has now been changed to:

> *Ryuichiro Ueda, MHA; Feng Han, MHA; Hongjian Zhang, PhD; Tomohiro Aoki, MHA; Katsuhiko Ogasawara, MBA, PhD*

2. Author Feng Han’s affiliation was originally:

> *Faculty of Health Sciences, Hokkaido University, Sapporo, Japan*

And was changed to:

> *Graduate School of Medicine, Hokkaido University, Sapporo, Japan*

3. The phone number listed for the corresponding author was originally:

> *81 011 716 2111*

And was changed to:

> *81 11 706 3409*

The correction will appear in the online version of the paper on the JMIR Publications website on March 14, 2024 together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 28.02.24; this is a non–peer-reviewed article;accepted 28.02.24; published 14.03.24.

Please cite as:
Ueda R, Han F, Zhang H, Aoki T, Ogasawara K
Correction: Verification in the Early Stages of the COVID-19 Pandemic: Sentiment Analysis of Japanese Twitter Users
JMIR Infodemiology 2024;4:e57880
URL: https://infodemiology.jmir.org/2024/1/e57880
doi:10.2196/57880
PMID:38484301

Corrigenda and Addenda

Correction: Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences

Masaru Kamba[1], PhD; Wan Jou She[1], PhD; Kiki Ferawati[1], MStat; Shoko Wakamiya[1,2], PhD; Eiji Aramaki[1], PhD

[1]Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

[2]Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Eiji Aramaki, PhD
Division of Information Science
Graduate School of Science and Technology
Nara Institute of Science and Technology
8916-5, Takayama-cho
Ikoma, 630-0192
Japan
Phone: 81 0743 72 5250
Email: aramaki@is.naist.jp

Related Article:

Correction of: https://infodemiology.jmir.org/2024/1/e49699

(JMIR Infodemiology 2024;4:e67434) doi:10.2196/67434

In "Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences" (JMIR Infodemiology 2024;4:e49699) the authors made one addition.

The fourth author was listed as follows:

> *Shoko Wakamiya[1], PhD.*
>
> *1. Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan*

This has been changed by adding the second affiliation as follows:

> *Shoko Wakamiya[1, 2], PhD.*
>
> *1. Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan*
>
> *2. Institute of Industrial Science, The University of Tokyo, Tokyo, Japan*

The correction will appear in the online version of the paper on the JMIR Publications website on October 29, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 11.10.24; this is a non–peer-reviewed article;accepted 16.10.24; published 29.10.24.

Please cite as:
Kamba M, She WJ, Ferawati K, Wakamiya S, Aramaki E
Correction: Exploring the Impact of the COVID-19 Pandemic on Twitter in Japan: Qualitative Analysis of Disrupted Plans and Consequences
JMIR Infodemiology 2024;4:e67434
URL: https://infodemiology.jmir.org/2024/1/e67434
doi:10.2196/67434
PMID:

Research Letter

Using Social Listening for Digital Public Health Surveillance of Human Papillomavirus Vaccine Misinformation Online: Exploratory Study

Dannell Boatman[1,2], MS, EdD; Abby Starkey[1,2], MS; Lori Acciavatti[1,2], MS; Zachary Jarrett[1,2], MA; Amy Allen[1,2], MA, MS; Stephenie Kennedy-Rea[1,2], MA, EdD

[1]Department of Cancer Prevention & Control, School of Medicine, West Virginia University, Morgantown, WV, United States

[2]West Virginia University Cancer Institute, Morgantown, WV, United States

Corresponding Author:
Dannell Boatman, MS, EdD
Department of Cancer Prevention & Control
School of Medicine
West Virginia University
PO Box 9350
Morgantown, WV, 26506
United States
Phone: 1 304 293 7883
Email: dboatman@hsc.wvu.edu

Abstract

Despite challenges related to the data quality, representativeness, and accuracy of artificial intelligence–driven tools, commercially available social listening platforms have many of the attributes needed to be used for digital public health surveillance of human papillomavirus vaccination misinformation in the online ecosystem.

(JMIR Infodemiology 2024;4:e54000) doi:10.2196/54000

KEYWORDS

human papillomavirus; HPV; vaccine; vaccines; vaccination; vaccinations; sexually transmitted infection; STI; sexually transmitted disease; STD; sexual transmission; sexually transmitted; social media; social listening; cancer; surveillance; health communication; misinformation; artificial intelligence; AI; infodemiology; infoveillance; oncology

Introduction

The COVID-19 pandemic accelerated the spread of misinformation online, creating an “infodemic” that had profound effects on health behavior [1]. The breadth and depth of COVID-19 misinformation expanded to include all vaccinations, such as human papillomavirus (HPV) vaccination, depressing already suboptimal vaccination uptake in the United States [1,2]. As HPV vaccination is critical to the prevention of various cancers, this could pose significant cancer control challenges in the future [2]. There is an urgent need to address HPV vaccination misinformation to increase HPV vaccination uptake [2]. Behavioral interventions can counter misinformation online, but they are typically limited to a single social media platform without geographic specificity [3].

Public health surveillance (PHS) is defined by the Centers for Disease Control and Prevention (CDC) as “the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health” [4]. Digital PHS (DPHS) uses data from online sources, often collected outside of traditional PHS, for similar purposes [5]. There has been debate as to the ethics of using publicly available online data for DPHS [5]. However, the pandemic illustrated the need for user-friendly, timely, interactive digital tools to drive health-related intervention [6].

Social listening (SL) is the process of aggregating data from across online channels to collect real-time measures of emotions, opinions, and themes, typically through platform algorithms that rely on machine learning and artificial intelligence (AI) [7]. While SL platforms’ AI-driven tools for emotion and sentiment detection can be unreliable, machine learning provides an opportunity to “train” SL platforms for greater accuracy over time in the automated recognition of emotions and sentiments [8]. The World Health Organization Early AI-Supported Response With Social Listening Platform (WHO EARS) uses an SL dashboard to provide health professionals access to information from across the internet to assist in the development of timely responses to COVID-19 narratives that occur online at the global and country levels, highlighting the growing acceptance of such tools in public health [7].

The purpose of this exploratory study was to assess the feasibility of using a commercially available SL platform to monitor HPV vaccination misinformation online at the national (ie, within the United States overall) and state (ie, within Mississippi and Rhode Island) levels.

Methods

Ethical Considerations

This study received institutional review board exemption from West Virginia University (protocol #00152755).

Study Design

Brandwatch was the commercially available SL platform selected for this exploratory study. It was selected after reviewing functionalities of leading SL platforms and having conversations about capabilities with representatives from Agorapulse, Brandwatch, Hootsuite, and Sprout Social. While most platforms had similar functionalities and data access, Brandwatch was selected based on opportunities to build queries with greater geographic specificity. While there is limited research on SL platform functionality within public health, Brandwatch was previously studied for the accuracy of AI-driven analyses [8]. The previously cited limitations of Brandwatch AI-driven tools informed the study team's systematic, routine approach to training.

The research team received onboarding from Brandwatch through 5 structured, live training sessions. Two research team members completed a self-paced online training certificate. After onboarding was complete, the research team's SL lead analyst (AS) built an HPV vaccination query within Brandwatch, using keywords and phrases identified through previous research and with research team consensus [9]. From this query, AS, with support from Brandwatch developers, created a dashboard to monitor online conversations within the United States overall and in 2 states—Mississippi, the US state with the lowest HPV vaccination rate, and Rhode Island, the US state with the highest vaccination rate. The research team regularly reviewed the query keywords and updated them as needed for increased relevancy and accuracy.

Brandwatch AI-driven tools were trained to recognize sentiments and emotions related to HPV vaccination. Sentiment categories for this study were different from the ones provided automatically by Brandwatch within the platform and were determined by the research team based on previous research [9]. Sentiment categories included "fact-based information," "pro-vaccine opinions," "misinformation," "anti-vaccine opinions," and "neutral comments." These sentiment categories were built into the dashboard by a Brandwatch developer in conjunction with AS. The initial AI-driven recognition of these content categories was inaccurate. For example, all content that mentioned "cancer" was automatically considered negative by the SL platform AI. AS trained the AI-driven sentiment tool to recognize the intended content by reviewing aggregated social media comments, as well as other online articles and posts within Brandwatch, and adding them to the appropriate categories to spur AI recognition. During this AI training process, another sentiment category—"irrelevant"—was added, as content that used similar language but was not directly related to HPV was identified. The Brandwatch AI-driven sentiment tool was trained by AS routinely over a 6-month period to enhance the recognition of categories. This routine training significantly improved category recognition within the SL platform but was not completely accurate upon periodic spot reviews by the research team. The AI-driven tool for recognizing emotions automatically included categories such as "anger," "disgust," "fear," "joy," "sadness," and "surprise." Like the AI-driven sentiment tool, the identification of correct emotion categories was initially incorrect and required routine training by AS to improve accuracy.

Once the SL platform was built, the research team evaluated the dashboard, query, and implementation process notes to assess the feasibility of using a commercially available SL platform for HPV vaccination misinformation DPHS. This assessment was completed by using an adaptation of the CDC's attributes for an effective PHS system [4]. The attributes adapted in this study were identified from CDC iterations published since 1988 [10]. The adaption of attributes involved the inclusion of consistent elements and associated definitions from across these CDC iterations; the addition of "cost" as a potential challenge to scaling; and the removal of "predictive value positive," as the proposed DPHS approach would assess online narratives as opposed to a specific health condition. Consensus on each attribute was reached among the research team members.

Results

Table 1 details each adapted PHS system attribute and the opportunities and limitations with regard to using a commercially available SL platform for HPV vaccination misinformation DPHS. Opportunities include user-friendly dashboards with real-time data monitoring and platform adaptability. For example, from June 21 to 24, 2023, the research team was able follow the spread of misinformation through social media posts related to a lawsuit filed by the Children's Health Defense Fund, an organization led by prominent antivaccine activist Robert Kennedy Jr. However, while the SL platform dashboards are user-friendly, it took significant staff time, expertise, and routine maintenance to keep them relevant and as accurate as possible. Brandwatch was also found to be adaptable to the ever-changing online information ecosystem; however, the quality of this information was dependent on data access agreements with individual social media companies, which could change at any time. Additional challenges to using an SL platform for DPHS include concerns with data quality, representativeness, and the accuracy of AI-driven tools. There are limited ways to validate data within the SL platform itself. Data may be downloaded from Brandwatch and externally analyzed for sentiments and emotions, but this process would remove the AI-driven, automated nature of the SL platform and reduce the effectiveness of real-time monitoring in DPHS.

Table 1. Feasibility of using a commercial social listening platform for human papillomavirus vaccination misinformation digital public health surveillance. This was assessed based on attributes of public health surveillance systems adapted from the Centers of Disease Control and Prevention [4].

Attribute	Attribute description	Social listening opportunities	Social listening limitations
Usefulness	Contribution to prevention and control of misinformation	Events that may trigger misinformation spread can be identified in real time, providing an opportunity to target intervention	Unclear if targeted interventions can effectively shift online narratives
Simplicity	Simplicity of structure and ease of use	Dashboards can automate monitoring and provide easy-to-use tools to dig deeper into observable trends	Building effective queries requires a specialized skill set, including content area knowledge and experience with social media and online ecosystems
Flexibility	Adaptable to changing information and conditions	Queries can be adapted to new information and trends by changing keywords and phrases	Requires consistent monitoring by skilled personnel to ensure queries are reflective of current conditions
Data quality	Validity and completeness of data	Queries can include data beyond social media, providing a window into narratives in online public spaces	Data are limited by access provided by specific social media companies and the effectiveness of the query, along with a current lack of external data validation
Representativeness	Accurately describes flow of health information over time and distribution by place and person	Queries can monitor conversation trends over time, such as trends among audience panels and in various locations, which provide insights into demographics and geographic boundaries	Demographic and geographic information is imprecise and is limited based on availability
Timeliness	Lapse of time between misinformation and intervention	Conversations can be monitored in real time, providing opportunities for quick responses to misinformation	Lack of evidence-based responses to counter misinformation spread
Sensitivity	Ability to identify true cases and detect misinformation	Dashboard algorithms can be trained to detect changes in sentiments and emotions, providing an opportunity to respond to trends	Effectively training algorithms to detect sentiments and emotions is time-consuming and requires a specialized skill set
Stability	System is resilient to change	Can collect new sources of online data as they emerge to remain relevant in the shifting social media and online ecosystem	Changes to social media company policies can affect access to data sources
Acceptability	Willingness of persons and organizations to participate	Data collection is passive and does not burden participants with active data requests	Ethical concerns with online public data collection
Portability	Duplication of system in another setting	Social listening platforms can be purchased and adapted to different settings and health conditions, with no specialized hardware required for operation	Effectiveness of the queries may be limited by the personnel developing them and the sophistication of the selected social listening platform
Costs	Cost-effectiveness of the system	Online services can vary in price (≥US $2500 annually) based on the services needed for social listening	Sophisticated social listening platforms are more costly, although they provide greater access to data and tools

While Brandwatch was selected due to opportunities for greater geographic specificity, this functionality was limited in scope to only certain social media platforms, such as X (formerly Twitter). Furthermore, geographic specificity was limited based on whether social media users used geolocation functionalities and whether locations were mentioned in profiles or posts. Despite this, the research team identified and monitored different narratives in misinformation within the two states included in this exploratory study—Rhode Island and Mississippi—suggesting the potential importance of assessing online misinformation narratives based on geographic location. For example, on the same day in January 2024, the top trending story for Rhode Island focused on the Children's Health Defense Fund lawsuit, while in Mississippi, the top story focused on childhood injury due to vaccination.

Discussion

Our findings suggest that there are opportunities and challenges associated with using commercially available SL platforms to monitor HPV vaccination misinformation online at the national and state levels. While there were strengths across all PHS system attributes, there were also significant weaknesses. These weaknesses, particularly those related to data quality, representativeness, and the accuracy of AI-driven tools, reflect limitations to using current SL platforms for DPHS. If these challenges are addressed over time however, this level of DPHS could provide the foundation for different intervention opportunities, such as using skilled infodemiologists to counter online misinformation [11]. While the research team identified challenges with the accuracy of Brandwatch AI-driven tools, which matched previously published research [8], building DPHS capabilities now could provide critical infrastructure if and when such tools improve over time. If found to be effective in monitoring HPV vaccine misinformation, commercially available SL platforms may be adapted to other fields and health conditions. Findings may differ based on the SL platform used and vendor access agreements with social media companies. Future research should focus on increasing the specificity of geographic location, studying strategies to increase the accuracy of SL platform AI-driven tools, and testing targeted interventions using SL platforms.

Acknowledgments

This work was supported by a grant from the Merck Investigator Studies Program, Merck and Company, Inc.

Authors' Contributions

DB, AA, and SKR contributed to the conception of this work. DB designed this work. AS acquired the data. DB and AS contributed to the analysis of data. DB, AS, LA, and ZJ contributed to the interpretation of data. All authors contributed to drafting the manuscript, and DB approved the final version for publication.

Conflicts of Interest

None declared.

References

1. 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 May 16;12(1):8045 [FREE Full text] [doi: 10.1038/s41598-022-11488-y] [Medline: 35577820]
2. Troiano G, Nardi A. Vaccine hesitancy in the era of COVID-19. Public Health 2021 May;194:245-251 [FREE Full text] [doi: 10.1016/j.puhe.2021.02.025] [Medline: 33965796]
3. Gwiaździński P, Gundersen AB, Piksa M, Krysińska I, Kunst JR, Noworyta K, et al. Psychological interventions countering misinformation in social media: a scoping review. Front Psychiatry 2022 Jan 5;13:974782 [FREE Full text] [doi: 10.3389/fpsyt.2022.974782] [Medline: 36684016]
4. Buehler JW, Sosin DM, Platt R. Evaluation of surveillance systems for early epidemic detection. In: Infectious Disease Surveillance. Hoboken, New Jersey: Blackwell Publishing; 2008:432-442.
5. Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. Annu Rev Public Health 2020 Apr 2;41:101-118 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094402] [Medline: 31905322]
6. Cocoros NM, Kirby C, Zambarano B, Ochoa A, Eberhardt K, Sb CR, et al. RiskScape: a data visualization and aggregation platform for public health surveillance using routine electronic health record data. Am J Public Health 2021 Feb;111(2):269-276. [doi: 10.2105/AJPH.2020.305963] [Medline: 33351660]
7. McGowan BS. World Health Organization's early AI-supported response with social listening platform. J Med Libr Assoc 2022 Apr 1;110(2):273-275 [FREE Full text] [doi: 10.5195/jmla.2022.1398] [Medline: 35440908]
8. Hayes JL, Britt BC, Evans W, Rush SW, Towery NA, Adamson AC. Can social media listening platforms' artificial intelligence be trusted? examining the accuracy of Crimson Hexagon's (now Brandwatch Consumer Research's) AI-driven analyses. J Advert 2020 Sep 17;50(1):81-91. [doi: 10.1080/00913367.2020.1809576]
9. Boatman DD, Eason S, Conn ME, Kennedy-Rea SK. Human papillomavirus vaccine messaging on TikTok: social media content analysis. Health Promot Pract 2022 May;23(3):382-387 [FREE Full text] [doi: 10.1177/15248399211013002] [Medline: 33969725]
10. Centers for Disease Control (CDC). Guidelines for evaluating surveillance systems. MMWR Suppl 1988 May 6;37(5):1-18. [Medline: 3131659]
11. Gorman JM, Scales DA. Leveraging infodemiologists to counteract online misinformation: experience with COVID-19 vaccines. Harv Kennedy Sch Misinformation Rev 2022 Feb 16;3(1):1-7 [FREE Full text] [doi: 10.37016/mr-2020-92]

Abbreviations

AI: artificial intelligence
CDC: Centers for Disease Control and Prevention
DPHS: digital public health surveillance
HPV: human papillomavirus
PHS: public health surveillance
SL: social listening
WHO EARS: World Health Organization Early AI-Supported Response With Social Listening Platform

Edited by T Mackey; submitted 26.10.23; peer-reviewed by R Hart-Malloy, E Miyagi; comments to author 17.01.24; accepted 07.02.24; published 08.03.24.

Please cite as:
Boatman D, Starkey A, Acciavatti L, Jarrett Z, Allen A, Kennedy-Rea S
Using Social Listening for Digital Public Health Surveillance of Human Papillomavirus Vaccine Misinformation Online: Exploratory Study
JMIR Infodemiology 2024;4:e54000
URL: https://infodemiology.jmir.org/2024/1/e54000
doi:10.2196/54000
PMID:38457224

Original Paper

Descriptions of Scientific Evidence and Uncertainty of Unproven COVID-19 Therapies in US News: Content Analysis Study

Sara Watson[1*], MA; Tyler J Benning[2*], MS, MD; Alessandro R Marcon[3], MA; Xuan Zhu[4], PhD; Timothy Caulfield[3], LLM; Richard R Sharp[1,5], PhD; Zubin Master[6,7,8,9], PhD

[1]Biomedical Ethics Research Program, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, MN, United States

[2]Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, MN, United States

[3]Health Law Institute, University of Alberta, Edmonton, AB, Canada

[4]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, United States

[5]Center for Individualized Medicine, Mayo Clinic, Rochester, MN, United States

[6]Department of Social Sciences and Health Policy, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[7]Maya Angelou Center for Health Equity, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[8]Center for Bioethics, Health and Society, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[9]Wake Forest Institute of Regenerative Medicine, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[*]these authors contributed equally

Corresponding Author:
Zubin Master, PhD
Department of Social Sciences and Health Policy
Wake Forest University School of Medicine
Medical Center Boulevard
Suite 310
Winston-Salem, NC, 27157
United States
Phone: 1 3367164289
Fax: 1 3667167554
Email: zmaster@wakehealth.edu

Abstract

Background: Politicization and misinformation or disinformation of unproven COVID-19 therapies have resulted in communication challenges in presenting science to the public, especially in times of heightened public trepidation and uncertainty.

Objective: This study aims to examine how scientific evidence and uncertainty were portrayed in US news on 3 unproven COVID-19 therapeutics, prior to the development of proven therapeutics and vaccines.

Methods: We conducted a media analysis of unproven COVID-19 therapeutics in early 2020. A total of 479 discussions of unproven COVID-19 therapeutics (hydroxychloroquine, remdesivir, and convalescent plasma) in traditional and online US news reports from January 1, 2020, to July 30, 2020, were systematically analyzed for theme, scientific evidence, evidence details and limitations, safety, efficacy, and sources of authority.

Results: The majority of discussions included scientific evidence (n=322, 67%) although only 24% (n=116) of them mentioned publications. "Government" was the most frequently named source of authority for safety and efficacy claims on remdesivir (n=43, 35%) while "expert" claims were mostly mentioned for convalescent plasma (n=22, 38%). Most claims on hydroxychloroquine (n=236, 79%) were offered by a "prominent person," of which 97% (n=230) were from former US President Trump. Despite the inclusion of scientific evidence, many claims of the safety and efficacy were made by nonexperts. Few news reports expressed scientific uncertainty in discussions of unproven COVID-19 therapeutics as limitations of evidence were infrequently included in the body of news reports (n=125, 26%) and rarely found in headlines (n=2, 2%) or lead paragraphs (n=9, 9%; P<.001).

Conclusions: These results highlight that while scientific evidence is discussed relatively frequently in news reports, scientific uncertainty is infrequently reported and rarely found in prominent headlines and lead paragraphs.

(JMIR Infodemiology 2024;4:e51328) doi:10.2196/51328

KEYWORDS

COVID-19; COVID-19 drug treatment; information dissemination; health communication; uncertainty; content analysis; information sources; therapy; misinformation; communication; scientific evidence; media analysis; news report; COVID-19 therapy; treatment; public awareness; public trepidation; therapeutic; therapeutics; vaccine; vaccines; pandemic; United States; media analysis; safety; efficacy; evidence; news; report; reports

Introduction

The clear and accurate reporting of science to the public is imperative to maintain public trust and ensure that public health precautions are upheld [1]. Reporting science surrounding protective measures and novel unproven COVID-19 therapeutics during the first year of the pandemic was extraordinarily challenging given the hyperpoliticization of COVID-19, the misinformation or disinformation surrounding scientifically unproven and unapproved therapies, and the reliance on unsubstantiated science, for example, preprints or expert opinion [2].

Despite the number of news sources available to the public including legacy media and social media, many Americans continue to rely on traditional sources of news. Traditional media are forms of communication predating the internet and include newspapers, broadcast, and radio among others. A 2014 survey by the American Press Institute found just over 60% of Americans prefer to find news directly from news organizations compared to social media (4%), word of mouth (2%), and e-sharing with friends (1%) [3]. While social media and interpersonal communications have been implicated in the spread of misinformation and adoption of preventative behaviors, a study published in 2020 on COVID-19 information sources found traditional media sources (including television and newspapers) were the most widely used sources on COVID-19, and that the source of information individuals reported using predicted their beliefs about the virus [4].

Reviewing the content of traditional media reports is particularly important because broadcast networks, cable networks, and online and print news are viewed as the most trustworthy and reliable sources of information [4-6]. While news media has a pivotal role in communicating credible scientific evidence [7], some mainstream news outlets have propagated inaccuracies and misinformation [8,9]. Misinformation may influence health beliefs and impact compliance with public health recommendations resulting in negative health consequences [10]. The United States has the second highest prevalence of COVID-19 misinformation compared to other countries [11], which likely has impacted the health of millions of Americans [12].

The rapid release of COVID-19 research disseminated by numerous news sources with shifting health recommendations has left many feeling overwhelmed and frustrated [13,14], which may contribute to reducing trust in public health officials and institutions [15-17]. Accurate news reporting should avoid hype and speculation [18], illuminate scientific uncertainty, and portray science as iterative and evolving [19-21]. Such portrayals may serve to promote public trust and help citizens understand why public health recommendations are susceptible to change [22-24].

An examination of early COVID-19 therapeutics serves as an ideal case study on which to examine the media portrayal of scientific evidence because it was a time when there were no approved treatments, prior to the development of proven therapeutics and vaccines, and during a period of high public trepidation and scientific uncertainty regarding COVID-19 treatments. In this study, we examined how scientific evidence and uncertainty were portrayed in traditional and online US news about 3 popular and potential COVID-19 therapeutics (hydroxychloroquine, remdesivir, and convalescent plasma [CP]) when no other US Food and Drug Administration–approved treatments or vaccines were available.

Methods

Conceptual Framework

This research examined the portrayal of scientific evidence and uncertainty of 3 unproven COVID-19 therapeutics prior to the development of vaccines and compared its representation among the body (full text or video) of news reports with headlines and lead paragraphs. There are 2 widely understood and generally accepted paradigms of science communication that focus on the unidirectional delivery of information or active engagement with the public [25-27]. These paradigms can further be divided into 4 models of science communication, which are science literacy (also referred to as the deficit model), contextual, lay expertise, and public engagement or public participation [28]. The science literacy model is premised that information transmitted by experts aims to reduce the deficit of science knowledge among the public while the contextual model similarly transmits information, but situates information to specific audiences paying attention to culture, location, and language among other relevant contexts. Both, however, focus on the transmission of scientific information. The lay expertise model acknowledges the limitations of science and endorses "lay" knowledge as equal to scientific knowledge. The public engagement model aims to address science policy issues through active engagement and interaction with the public in a democratic process. These latter 2 models focus on improving engagement between science and society and recognize the limitations of science and refrain from giving science any more authority than other knowledge sources. Several scholars have conceptualized variations or additions to these models and outlined new goals of science communication [26,27,29,30]. However, these models remain largely theoretical and provide little practical direction toward the practice of science communication today, which involves an array of different sources (eg, internet, social media, and legacy media), actors (eg, general and science journalists, bloggers, influencers, and others), and approaches (eg, traditional journalistic norms and a plurality of practices among content contributors in the new media environment) [31]. In addition, the models provide little in terms of approaches to analyze contemporary news reports produced by traditional media outlets reported by journalists because there is no accepted conceptual framework on the public communication of science in the media and best practices to convey scientific evidence and uncertainty to the public [32,33]. Instead, our analytical approach was informed by adopting a lens of a changing science media communication landscape where public reporting of science has moved beyond conveying scientific facts and where journalists use frames to describe science accurately while maintaining audience engagement [31,34]. This, however, results in multiple challenges in accurately reporting science to the public including science hype [35]; errors of omission [36]; failure to report funding sources and conflicts of interest [37]; inadequately detailing methods, risks, and timelines [38-40]; improperly addressing scientific uncertainty; and situating science in particular contexts, for example, politics.

There is variability in the portrayal of scientific evidence and uncertainty in the news media. Some studies have shown less accurate portrayals where the results are discussed as certain and the benefits overemphasized [41-44], while other studies have demonstrated inaccuracies in reporting evidence when comparing press releases and media articles with scientific publications resulting in oversimplification, absence of quantification, lack of explicitly reporting data sources, and the absence of study limitations [45,46]. On the other hand, scientific uncertainty on climate change among other specific topics has been heightened in news media by journalists framing science as inconclusive or providing equal space for opposing viewpoints, even if one viewpoint is supported by substantial scientific consensus and evidence while the other lacks credible evidence [47-49]. News reports of scientific discoveries aiming to provide balanced perspectives may inadvertently cause a false balance and distort the public's perception of consensus surrounding scientific evidence [50].

This research was also informed by the reality that the public often engages with snippets of information including a focus and evaluation of articles based on headlines [51,52]. Our focus on analyzing scientific evidence and uncertainty in headlines and lead paragraphs, and comparing it with the body of news reports was further informed by the primacy effect—a cognitive bias where people recall and place greater emphasis on initial pieces of information [53,54]—which may explain why headlines and lead paragraphs are especially important in shaping the audience's understanding and judgment of news stories.

Finally, our conceptual framework was also informed by a media environment with a higher prevalence of COVID-19 health misinformation and an understanding of the efforts to correct news reports by health and scientific experts through media fact-checking and myth-busting approaches [55-58]. While several studies have shown the effectiveness of fact-checking and the correction of misinformation postexposure at altering beliefs [59-63], it remains unclear if such approaches would fully address concerns of reported health misinformation. This is partly due to differences in the focus fact-checkers and myth busters consider when aiming to correct scientific facts, the framing of science news stories by reporters, ambiguity and discordance in the interpretation of scientific facts, and that truth-telling may not always reach intended audiences [64-66]. Such approaches epistemologically rely on the assumption that facts, even in the context of scientific evidence, can always be clearly discerned from nonfactual information and that experts can identify and expose truths from falsehoods; such approaches may minimize or discount completely the complex interplay between beliefs, politics, and science [67]. Being mindful of this context, our analysis focused on how, where, and by whom scientific evidence surrounding unproven COVID-19 therapeutics was portrayed in the news stories themselves and we did not address the veracity of the scientific evidence presented in the news.

Sampling

We chose to investigate 3 unproven products—hydroxychloroquine, remdesivir, and CP—based primarily on their popularity in US news among other factors that informed our decision. Several antivirals and immunomodulators were being considered as potential classes of therapeutics used to treat COVID-19 in the early 2020s [68]. An initial set of news database searches (discussed in the next section) identified hydroxychloroquine (4023 publications), remdesivir (2839 publications), azithromycin (416 publications), and CP (372 publications) as the most popular unproven products in US news compared to other products being considered including, lopinavir, interferon beta-1a, dexamethasone, and heparin or low molecular weight heparin. For feasibility reasons, we chose to evaluate 3 products and decided to include CP instead of azithromycin because after removing news reports discussing each product in conjunction with hydroxychloroquine, there were 344 articles discussing CP and only 57 articles discussing azithromycin. Additionally, the choice to examine hydroxychloroquine, remdesivir, and CP was supported because these products were being investigated in registered clinical studies in the United States (hydroxychloroquine, ClinicalTrials.gov NCT04714515; remdesivir, ClinicalTrials.gov NCT04257656; and CP, ClinicalTrials.gov NCT05578391) and permitted to be used to treat patients with COVID-19 by the US Food and Drug Administration (FDA; please see the following paragraphs for details).

News sources were identified using the Factiva database by the Dow Jones & Company [69]. Factiva is a commonly used research tool that provides full-text coverage of current and archival business and news information from traditional (eg, broadcast and newspaper) media and news available online. The specific news sources selected for our query were determined by popularity and coverage and were compiled based on 3 independent organizations that rank news based on circulation and internet traffic (see Table S1 in Multimedia Appendix 1) [70-72]. Having a diverse venue of sources is important since issues of unproven COVID-19 therapies have become politicized and studies have shown differing levels of depth and misinformation between news organizations.

After eliminating duplicate news sources listed by these organizations, we queried the terms hydroxychloroquine, Plaquenil, remdesivir, Veklury, and CP in the headline or lead paragraph of articles published from January 1, 2020, to July

30, 2020. The search dates were selected to ensure we captured the portrayal of scientific evidence of early unproven therapeutics to treat COVID-19 in US news. The start date of our search was chosen because a pneumonia-causing novel coronavirus was first announced on January 7, 2020, with the first confirmed US case on January 20, 2020 [73]. The end date was chosen based on clinical investigations and decisions made by the FDA on our 3 therapeutics of interest. Specifically, several clinical investigations examining the 3 potential therapeutics of interest were initiated during the first quarter of 2020, and scientific evidence about the safety and efficacy of the therapeutics was being collected and concluded [68,74,75]. Specifically, large clinical trials of hydroxychloroquine were initiated, paused, and eventually halted from late May to late June [74,76-78]. Based on clinical trial results, the FDA revoked emergency use authorization for hydroxychloroquine on June 15, 2020 [79]. Starting May 28, 2020, participants were being enrolled for a clinical trial on CP and on April 3, 2020, the FDA approved an expanded access protocol for Mayo Clinic to lead an investigation to understand the effects of CP [80]. The FDA also issued an emergency use authorization for remdesivir on May 1, 2020 [81]. Finally, the end date for our search was also chosen for practical reasons as data collection efforts began on July 31, 2020.

The search yielded 1136 reports including online and print articles, television transcripts, and videos. Application of inclusion or exclusion criteria was conducted by SW and TJB (primary coders). The coders matched the headline and news source of the web link obtained from Factiva for online media. Web-based reports with a video were considered 2 independent reports unless the video was verbatim or substantially similar to the associated article. Reports that were considered substantially similar must have had at least 1 area of similarity on the topic, main theme, concluding points, people quoted, or style (eg, debate). In instances of substantial similarity, only the most up-to-date report was included. Reports that were duplicative, unobtainable, or void of search terms were removed resulting in 1103 reports. If multiple therapeutics were discussed in a single report, the codebook was applied to each therapeutic. We randomly sampled a total of 550 news reports from our population to manage coder effort while allowing the capture of codes and themes among news reports. After random sampling, reports without substantial discussion of the COVID-19 therapeutic were excluded if they failed to satisfy codes 7-16 of the codebook (n=72). When a therapeutic was found only in an image caption, tweets, or graphics within the report, the report was also excluded (n=29). After applying exclusion criteria, a final data set of 479 individual discussions of a therapeutic was found in 449 news reports. See Figure S1 in Multimedia Appendix 1 for a flowchart providing an overview of the sampling strategy.

Codebook

A comprehensive codebook was developed based on the literature described. The purpose of the codebook was to ensure that a consistent analytical framework was adopted by both coders and to reduce randomness in the interpretation of themes and codes. Upon developing an initial draft, the codebook was modified through the iterative analysis of 9 written articles (2 hydroxychloroquine, 2 remdesivir, and 5 CP) from different news sources with different word lengths. The draft codebook was then reviewed by a multidisciplinary team of experts in media analysis, bioethics, and health communication. Independent review of 50 reports by 2 coders (SW and TJB) showed good interrater reliability in coding for the full news report for theme and portrayal of scientific evidence. The median Cohen κ for these elements was 0.71 (IQR 0.58-0.74) with simple agreement ranging from 79% to 94%. Interrater reliability was initially poor when identifying sources of authority, with a median κ of 0.17 (IQR 0.06-0.23). Definitions were clarified to address omissions or ambiguities identified during this intercoder check. For example, evidence was defined as a formal demonstration of the effect of a treatment on the COVID-19 disease course, which includes symptoms, outcomes, side effects, or a lack thereof. In the same way, sample size or study design was added as an example of evidence details and a statement of small sample size or nonrandomized design was included in the codebook as an example of evidence limitations. Codes regarding sources of authority were modified to clarify that a subject matter expert who spoke on behalf of a government agency should not be coded as an “expert” but should be coded as a “government” source of authority. Due to poor initial reliability, modified definitions for sources of authority were formally retested by the same 2 coders on a new sample of 20 articles. Sources of authority showed good interrater reliability with a median κ of 0.73 (IQR 0.48-0.86) with simple agreement ranging from 91% to 100%. The general structure of the codebook captures metadata provided by the Factiva database (eg, word count and hyperlink), headline analysis, report analysis, scientific description, sources of authority, social context, qualitative description of video reports, and coder notes (see Multimedia Appendix 2 for final codebook with complete definitions of codes and examples).

Analysis

The full text or video (body of news reports) of all discussions was analyzed for theme, scientific evidence, claims of safety or efficacy, and sources of authority. All terms are defined in the tables. The topical themes of news reports were identified inductively during iterative analysis of 9 articles as described above. An inductive approach permits the emergence of themes from the data and allows coders to remain open and exploratory [82]. This qualitative method is ideal for identifying themes when there is no prior knowledge about them [38,83,84]. In cases where more than 1 theme was reflected in the entire news report, the coder selected the theme that most closely resembled the news report. For text-based reports that discussed scientific evidence, details of evidence, or limitations of evidence, we also analyzed the presence of these codes within headlines and lead paragraphs. The lead paragraph was defined as the first paragraph of text with a minimum of 2 complete sentences excluding subtitles or alternate headlines just below the large print title. Coders independently analyzed approximately half the reports and jointly reviewed challenging reports and a random 10% of all reports as preplanned audits. Disagreements were resolved by consensus. Study data were recorded and managed using REDCap (Research Electronic Data Capture).

Different traits of reports (eg, discussion of evidence and discussion of specific side effects) were summarized for each therapeutic using descriptive statistics. Fisher exact tests were used to compare differences in trait frequencies. If 3×2 Fisher tests returned a *P* value of <.05, pairwise comparisons between therapeutics (ie, hydroxychloroquine - remdesivir, hydroxychloroquine-CP, and remdesivir-CP) were performed. No corrections were made for multiple comparisons. Analyses were performed in Excel (Microsoft Corp) and R (version 3.5.1; R Core Team).

Ethical Considerations

This research study did not involve human subjects and was conducted using publicly available information; thus, ethics approval was not sought.

Results

The data set included 479 individual discussions of hydroxychloroquine, remdesivir, and CP or a combination of any of the 3 unproven therapeutics among 449 news reports. Among 449 news reports, 191 were print, 172 were online, 52 were television transcripts, and 34 were online videos. Print news reports describe those accessed nonelectronically via PDF files from the Factiva database while online news reports and online videos were accessed by coders digitally via specific website links. News reports mostly discussed hydroxychloroquine (67%, n=299) compared to remdesivir (27%, n=122) and CP (13%, n=58). Safety or efficacy was the main theme among news on hydroxychloroquine (61%, n=182) and remdesivir (47%, n=57), whereas news reports discussing CP focused on the theme of economics, distribution, and allocation (40%, n=23). Many reports discussing hydroxychloroquine often surrounded the theme of politics (20%, n=61; Table 1).

Scientific evidence was discussed in 67% (n=322) of new discussions and was most common among discussions of hydroxychloroquine (78%, n=233) followed by remdesivir (63%, n=77) and CP (21%, n=12). All pairwise comparisons (*P*<.05; hydroxychloroquine-remdesivir *P*=.002; hydroxychloroquine-CP *P*<.001; remdesivir-CP *P*<.001; Table 2). Although scientific evidence was found in many news reports, specific publications or journals were seldom mentioned (24%, n=116). Details of scientific evidence and discussions on the limitations of evidence were found in 61% (n=198) and 26% (n=125) of news, respectively.

Among text-based news reports discussing scientific evidence, 22% (n=51) discussed scientific evidence in the headline and 51% (n=118) in the lead paragraph (Table 3). However, the details of scientific evidence were rarely found in headlines (6%, n=9) and seldom present in lead paragraphs (26%, n=39; *P*<.001). Among the 99 text-based reports that included limitations of evidence, discussions on limitations were rarely found in headlines (2%, n=2) and lead paragraphs (9%, n=9).

Very few discussions portrayed any therapeutic as safe, but 75% (n=91) of news discussions on remdesivir and 66% (n=38) of news discussions on CP were portrayed as efficacious. Only 14% (n=41) of news discussions on hydroxychloroquine portrayed the therapeutic as efficacious, and safety warnings or specific negative side effects were mainly identified in discussions on hydroxychloroquine (56%, n=168) compared to other therapeutics (Table 2). For hydroxychloroquine, most claims about safety and efficacy were offered by prominent persons (79%, n=236), almost exclusively by former US President Trump (97%, n=230). Other examples of prominent persons included Jair Bolsonaro (former president of Brazil) and Nancy Pelosi (US politician). In contrast, the government (35%, n=43) was the most frequently named source of authority for safety and efficacy claims of news on remdesivir while experts mostly made claims on CP (38%, n=22; Table 4).

Table 1. Thematic analysis of news reports.

	Hydroxychloroquine, n (%)	Remdesivir, n (%)	CP[a], n (%)
Novel scientific discovery: a novel result regarding the intervention. Does not include novel results on safety or efficacy	0 (0)	0 (0)	0 (0)
Safety or efficacy: discussions about ongoing trials, study logistics, and expected results of the intervention	182 (61)	57 (47)	19 (33)
Issue of scientific integrity: misconduct by individual scientists or scientific community, including research methods, peer review, and publication or dissemination decisions	16 (5)	0 (0)	1 (2)
Misinformation: analysis of the veracity of claims or reports of fact-checking	18 (6)	2 (2)	0 (0)
Politics: reports on political figures' claims, actions, and behaviors	61 (20)	3 (2)	0 (0)
Hope: feelings of optimism in context of society	0 (0)	2 (2)	0 (0)
Official recommendation: a statement made by a national or international government agency or governing body	3 (1)	0 (0)	0 (0)
Economics, distribution, and allocation: discussions of cost, quantity, supplies, or delivery of the intervention	12 (4)	50 (41)	23 (40)
Human interest story: individual narratives and excluding political figures	5 (2)	6 (5)	15 (26)
Other	2 (1)	2 (2)	0 (0)
Total[b]	299 (67)	122 (27)	58 (13)

[a]CP: convalescent plasma.

[b]The final data set included 479 individual discussions of therapeutics in 449 news reports.

Table 2. Portrayal of scientific evidence among news report discussions.

	Hydroxychloroquine, n (%)	Remdesivir, n (%)	CP[a], n (%)	Total, n (%)
Scientific evidence[b]: a demonstration of an effect on the disease course or resultant side effects from the intervention.	233 (78)	77 (63)	12 (21)	322 (67)
Details of scientific evidence[c]: specific details about the design or methodology of evidence.	152 (65)	40 (52)	6 (50)	198 (61)
Limitations of evidence: aspects of evidence or a study presented as shortcomings.	104 (35)	12 (10)	9 (16)	125 (26)
Publication or journal[b]: identifying a specific publication or journal named.	102 (34)	11 (9)	3 (5)	116 (24)
Portrayed as safe: portrayed therapeutic as safe using keywords, for example, safe or discussing minimal risks or side effects.	41 (14)	6 (5)	8 (14)	55 (11)
Portrayed as efficacious[b]: portrayed therapeutic as efficacious by using keywords, for example, effective and helpful.	41 (14)	91 (75)	38 (66)	170 (35)
Safety warnings or side effects[b]: discusses specific health or safety risks, for example, death.	168 (56)	1 (0.8)	1 (2)	170 (35)

[a]CP: convalescent plasma.

[b]Statistically significant result for 3×2 Fisher exact test; *P*<.001.

[c]Evidence details were analyzed only if the report included evidence; n=322.

Table 3. Scientific evidence, details, and limitations in headline or lead paragraphs[a].

	Hydroxychloroquine, n (%)	Remdesivir, n (%)	CP[b], n (%)	Total, n (%)
Scientific evidence in headline[c], n (%)	42 (24)	8 (13)	5 (42)	51 (22)
Scientific evidence in lead paragraph[c], n (%)	92 (53)	26 (43)	8 (67)	118 (51)
Total reports with scientific evidence in headline or lead paragraph, n	173	61	12	233
Details of scientific evidence in headline[c], n (%)	8 (7)	1 (3)	0 (0)	9 (6)
Details of scientific evidence in lead paragraph[c], n (%)	28 (25)	10 (30)	1 (25)	39 (26)
Total reports with details of scientific evidence in headline or lead paragraph, n	113	33	4	150
Limitations of evidence in headline[c], n (%)	2 (3)	0 (0)	0 (0)	2 (2)
Limitations of evidence in lead paragraph[c], n (%)	5 (6)	3 (25)	1 (11)	9 (9)
Total reports with limitations of evidence in headline or lead paragraph, n	78	12	9	99

[a]The following definitions were used. Scientific evidence: a demonstration of an effect on the disease course or resultant side effects from the intervention. Details of scientific evidence: specific details about the design or methodology of evidence. Limitations of evidence: aspects of evidence or a study presented as shortcomings.

[b]CP: convalescent plasma.

[c]A single news report was included in the analysis of more than 1 therapeutic when the report discussed multiple therapeutics.

Table 4. Sources of authority among news report discussions.

	Hydroxychloroquine, n (%)	Remdesivir, n (%)	CP[a], n (%)	Total, n (%)
Prominent person[b]: a nonmedical, nonscientist person such as celebrities and politicians	236 (79)	9 (7)	1 (2)	246 (51)
Institution: an academic institution, medical center, or hospital	22 (7)	9 (7)	9 (16)	40 (8)
Expert[b]: a trained professional such as clinician, scientist, or clinician-researcher speaking independently of an institutional affiliation	146 (49)	31 (25)	22 (38)	199 (41)
Government[b]: a federal or state government organization such as the Centers for Disease Control and Prevention, Food and Drug Administration and World Health Organization	180 (60)	43 (35)	12 (21)	235 (49)

[a]CP: convalescent plasma.

[b]Statistically significant result for 3×2 Fisher exact test; $P<.001$.

Discussion

Principal Findings

As reports spread of a novel coronavirus causing unprecedented hospitalizations and mortality, both medical specialists and everyday persons anxiously scrambled to learn as much and as fast as possible about prevention and potential treatments. Our analysis revealed that much coverage of the off-label use of hydroxychloroquine centered on politics and the safety and efficacy of the product. The politics theme dominated discourse surrounding hydroxychloroquine due to its endorsement by former President Trump as seen in other news media analyses [58,85]. During the months of March and April 2020, former President Donald Trump put forth hydroxychloroquine as a "game changer" after a study suggested efficacy in vitro. Propagation of statements surrounding the safety and efficacy of hydroxychloroquine had far-reaching effects, including increased sales by hospitals and the death of persons who took a similar product as prophylaxis [86]. While safety or efficacy discussions were also found in remdesivir and CP, limited knowledge of its antiviral properties may have played a significant role in shaping public discourse on the potential safety and efficacy of the drug. Conversations about CP instead focused on the human interest theme explicating the need for plasma donors and inspiring potential COVID-19 survivors to donate.

The results showed substantial heterogeneity in the sources of authority asserting claims of safety or efficacy of the COVID-19 therapeutics among the 449 reports. On many occasions, scientific claims of the safety and efficacy of a therapeutic were portrayed by prominent persons and were compared to the evidence discussed by experts. Contrary opinions on evidence of hydroxychloroquine delivered by politicians and experts may have inadvertently contributed to a false sense of scientific disagreement as news reports most frequently reported on claims made by a single prominent person but also mentioned the lack of scientific evidence supporting such claims. Presenting directly opposing viewpoints as equal when evidence or a scientific consensus is actually weighted is referred to as a false balance that can cause public uncertainty [87]. False-balance has been shown to distort public perceptions of consensus among experts resulting in people identifying greater disagreement and controversy than what is actually present [50]. Making such comparisons is problematic because nonscientific experts may interpret data toward a political advantage and their influence could have troubling downstream consequences when disseminated, for example, physicians prescribing unproven products and reduced intentions to adopt health behaviors [88-90]. Though the majority of news reports on hydroxychloroquine published the safety or efficacy claims made by prominent persons, very few reports actually portrayed hydroxychloroquine as efficacious, and safety warnings or specific side effects were frequently included.

While we found scientific evidence of therapeutics was portrayed in the body of news reports, mostly on hydroxychloroquine, limitations of evidence outlining scientific uncertainty were seldom seen, and rarely found in prominent locations (headlines and lead paragraphs) of news reports. With only about 40% of the public delving past headlines or lead paragraphs when reading articles [3,91], it is unlikely readers would grasp the scientific uncertainty associated with an unproven therapeutic. The dissemination of evolving science without expressing scientific uncertainty and the knowledge gap created by the iterative scientific process may have indirectly resulted in public confusion and minimized how large of a threat individuals believed the virus posed [92,93]. This confusion and doubt permitted some actors to promote disinformation surrounding the safety and efficacy of unproven therapeutics [8,21]. Journalists may be reluctant to express uncertainty of evidence in their reports out of fear their audience may react negatively toward such ambiguity [94] or the journalist may not be equipped to address a study's findings, especially its limitations. Scientists on the other hand are reportedly reluctant to express uncertainty fearing reporters may lose interest and that the public might misinterpret the science or doubt scientists [95,96]. However, empirical studies have shown that neither conveying uncertainty nor presenting scientific limitations affect a person's understanding of a topic, beliefs, or trust in science [97-99]. This data, along with our findings, suggest that uncertainty should be explained more frequently in news reports and placed in headlines and lead paragraphs to adequately support the public's understanding of scientific evidence and uncertainty [30,87]. Further research comparing scientific evidence and expressions of uncertainty between the body of news reports versus headlines and lead paragraphs on different scientific topics is needed. Research examining key factors influencing journalists' and news editors' decision-making process regarding news headlines and lead paragraphs is important to inform updated journalistic guidelines and editorial policies aimed at reducing false balance. Research collaboration between journalists, scientists, and communication researchers to identify effective communication strategies and journalistic practices to ensure accurate and nuanced representation of scientific information in news reporting would be valuable.

The accurate reporting of scientific evidence and uncertainty is the collective responsibility of multiple professionals including reporters, editors, public health officials, and medical and scientific experts. While evidence from experts may be based on scientific facts, expert correspondents should make clear scientific uncertainty, avoid hyperbolic statements, and explain how the state of knowledge may change [22]. Medical doctors and scientists should not express unjustified certainty when forecasting the safety and efficacy of unproven therapeutics [100]. Despite the lack of consensus, reporters should offer greater context or altogether avoid framing unsubstantiated claims by prominent people alongside evidence-based claims from experts as it may provide a false impression of genuine scientific uncertainty. Editors should refrain from sensationalizing headlines or framing lead paragraphs that do not reflect evidence or study limitations. Instead, journalists and experts alike should focus on the accurate portrayal of the scientific studies presented in the report and aim to provide an accurate reflection of the safety and efficacy of unproven products to help the public understand why there is incomplete evidence and a lack of definitive recommendations. Prioritizing the accurate portrayal of science may help to rebuild public trust in experts, minimize risky behaviors, and ensure compliance with public health recommendations.

Limitations

Our study has several limitations worth mentioning. First, the analysis included only news sources within the United States, all of which were written in English and thus may not cater to news outside of the country and written in different languages. Second, the analysis focused on traditional, mainstream news sources and we excluded social media. Although much COVID-19 information and misinformation were disseminated through social media contributing to an infodemic, traditional news remains the major form of information transfer, especially among older Americans, which is why we focused on evaluating traditional news sources [101]. Third, we chose to analyze approximately half the news reports in our population and the sample may not be representative of the entire news discourse within the full data set. Fourth, the data collection and analysis was restricted to the 3 most popular unproven COVID-19 therapeutics and other therapeutics during that period may be portrayed differently in US news. Finally, the results were not stratified based on different political leaning of news sources included in our analysis, which could affect how scientific evidence and uncertainty were portrayed in our data set.

Conclusions

News media plays a significant role in informing the public by serving as a crucial link between public health authorities interpreting scientific evidence and determining its implications and the public. Rapid publication of peer-reviewed and nonreviewed science allowed the media to bring breaking news about discoveries of COVID-19 therapeutics to the public and convey public health recommendations [18]. Prioritizing the accurate portrayal of science can help shape the impact of a public health emergency by influencing public perceptions and activities such as minimizing risky behaviors and encouraging compliance with public health recommendations. Publicizing that the news media source is likely to share frequent updates from scientific findings and communicating information as the best-known information at the time may help build public trust in scientific experts and decrease feelings of doubt and anxiety.

The accurate reporting of scientific evidence and uncertainty is the collective responsibility of journalists and science communicators. Public trust in science can be strengthened by acknowledging evidentiary limitations, avoiding a false sense of disagreement, and portraying science as an iterative, self-corrective process that generates reliable knowledge.

Acknowledgments

The authors would like to thank Dr Christen Rachul, Dr Karen Meagher, and Jordan Richardson for input on the original codebook. The authors would like to thank the anonymous reviewers for helpful feedback on the manuscript. Funding support for this paper was provided by the National Center for Advancing Translational Sciences (UL1TR002377), National Institutes of Health, US Department of Health and Human Services or United States.

Conflicts of Interest

XZ offers scientific input to research studies through a contracted services agreement between Mayo Clinic and Exact Sciences. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Methods.
[DOCX File , 107 KB - infodemiology_v4i1e51328_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File , 33 KB - infodemiology_v4i1e51328_app2.docx]

References

1. Tukachinsky Forster R, Vendemia MA. Effects of news and threat perceptions on Americans' COVID-19 precautionary behaviors. Commun Rep 2021;34(2):65-77. [doi: 10.1080/08934215.2021.1907428]
2. La Bella E, Allen C, Lirussi F. Communication vs evidence: what hinders the outreach of science during an infodemic? A narrative review. Integr Med Res 2021;10(4):100731 [FREE Full text] [doi: 10.1016/j.imr.2021.100731] [Medline: 34141575]
3. How Americans get their news. American Press Institute. 2014. URL: https://americanpressinstitute.org/how-americans-get-news/ [accessed 2024-08-03]
4. 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]
5. Ballew M, Bergquist P, Goldberg M, Gustafson A, Kotcher J, Marlon J, et al. American Public Responses to COVID-19. New Haven: Yale University and George Mason University; 2020.
6. What Americans Think of the News? and What that Means for Democracy. Santa Monica, California: Rand Corporation; 2020.
7. Erikainen S, Stewart E. Credibility contests: media debates on do-it-yourself coronavirus responses and the role of citizens in health crises. Front Sociol 2020;5:592666 [FREE Full text] [doi: 10.3389/fsoc.2020.592666] [Medline: 33869520]
8. Hall Jamieson K, Albarracín D. The relation between media consumption and misinformation at the outset of the SARS-CoV-2 pandemic in the US. HKS Misinfo Review 2020;1(2). [doi: 10.37016/mr-2020-012]
9. Motta M, Stecula D, Farhart C. How right-leaning media coverage of COVID-19 facilitated the spread of misinformation in the early stages of the pandemic in the U.S. Can J Pol Sci 2020;53(2):335-342. [doi: 10.1017/S0008423920000396]
10. Enders AM, Uscinski JE, Klofstad C, Stoler J. The different forms of COVID-19 misinformation and their consequences. HKS Misinfo Review 2020;1(8). [doi: 10.37016/mr-2020-48]
11. Al-Zaman M. Prevalence and source analysis of COVID-19 misinformation in 138 countries. IFLA J 2021;48(1):189-204 [FREE Full text] [doi: 10.1177/03400352211041135]
12. 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]
13. Van Scoy LJ, Miller EL, Snyder B, Wasserman E, Chinchilli VM, Zgierska AE, et al. Knowledge, perceptions, and preferred information sources related to COVID-19 among Central Pennsylvania adults early in the pandemic: a mixed methods cross-sectional survey. Ann Fam Med 2021;19(4):293-301 [FREE Full text] [doi: 10.1370/afm.2674] [Medline: 33985977]

14. Van Scoy LJ, Snyder B, Miller EL, Toyobo O, Grewel A, Ha G, et al. Public anxiety and distrust due to perceived politicization and media sensationalism during early COVID-19 media messaging. J Commun Healthc 2021;14(3):193-205. [doi: 10.1080/17538068.2021.1953934]
15. Freeman D, Waite F, Rosebrock L, Petit A, Causier C, East A, et al. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. Psychol Med 2022;52(2):251-263 [FREE Full text] [doi: 10.1017/S0033291720001890] [Medline: 32436485]
16. Price D, Bonsaksen T, Ruffolo M, Leung J, Chiu V, Thygesen H, et al. Perceived trust in public authorities nine months after the COVID-19 outbreak: a cross-national study. Soc Sci 2021;10(9):349. [doi: 10.3390/socsci10090349]
17. The public's perspective on the United States public health system. Robert Wood Johnson Foundation. 2021. URL: https://www.hsph.harvard.edu/wp-content/uploads/sites/94/2021/05/RWJF-Harvard-Report_FINAL-051321.pdf [accessed 2024-08-06]
18. Caulfield T, Bubela T, Kimmelman J, Ravitsky V. Let’s do better: public representations of COVID-19 science. FACETS 2021;6(1):403-423. [doi: 10.1139/facets-2021-0018]
19. Bagdasarian N, Cross GB, Fisher D. Rapid publications risk the integrity of science in the era of COVID-19. BMC Med 2020;18(1):192 [FREE Full text] [doi: 10.1186/s12916-020-01650-6] [Medline: 32586327]
20. 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 2024-08-06]
21. The Lancet Infectious Diseases. The COVID-19 infodemic. Lancet Infect Dis 2020;20(8):875 [FREE Full text] [doi: 10.1016/S1473-3099(20)30565-X] [Medline: 32687807]
22. Saitz R, Schwitzer G. Communicating science in the time of a pandemic. JAMA 2020;324(5):443-444. [doi: 10.1001/jama.2020.12535] [Medline: 32749498]
23. Jensen J. Scientific uncertainty in news coverage of cancer research: effects of hedging on scientists' and journalists' credibility. Hum Commun Res 2008;34(3):347-369.
24. Ratcliff C, Wicke R. How the public evaluates media representations of uncertain science: an integrated explanatory framework. Public Underst Sci 2023;32(4):410-427. [doi: 10.1177/09636625221122960] [Medline: 36196654]
25. Akin H, Scheufele D. Overview of the science of science communication. In: Jamieson KH, Kahan DM, Scheufele DA, editors. The Oxford Handbook of the Science of Science Communication. New York: Oxford University Press; 2017:25-33.
26. Kappel K, Holmen SJ. Why science communication, and does it work? A taxonomy of science communication aims and a survey of the empirical evidence. Front Commun 2019;4. [doi: 10.3389/fcomm.2019.00055]
27. Schmid-Petri H, Bürger M. Modeling science communication: from linear to more complex models. In: Leßmöllmann A, Dascal M, Gloning T, editors. Science Communication. Berlin, Boston: De Gruyter Mouton; 2020:105-122.
28. Brossard D, Lewenstein BV. A critical appraisal of models of public understanding of science: using practice to inform theory. In: Kahlor L, Stout PA, editors. Communicating Science New Agendas in Communication. New York, New York: Routledge; 2010:11-39.
29. Trench B. Towards an analytical framework of science communication models. In: Cheng D, Claessens M, Gascoigne T, Metcalfe J, Schiele B, Shi S, editors. Communicating Science in Social Contexts: New Models, New Practices. Netherlands: Springer, Dordrecht; 2008:119-135.
30. Nisbet M, Fahy D. New models of knowledge-based journalism. In: Jamieson KH, Kahan DM, Scheufele DA, editors. The Oxford Handbook of the Science of Science Communication. Oxford: Oxford University Press; 2017:273-281.
31. Schäfer MS. How changing media structures are affecting science news coverage. In: Jamieson KH, Kahan DM, Scheufele DA, editors. The Oxford Handbook of the Science of Science Communication. Oxford: Oxford University Press; 2017:50-59.
32. Secko DM, Amend E, Friday T. Four models of science journalism. Journal Pract 2013;7(1):62-80. [doi: 10.1080/17512786.2012.691351]
33. Amend E, Capurro G, Secko DM. Grasping scientific news. Journal Pract 2014;8(6):789-808. [doi: 10.1080/17512786.2013.868146]
34. Bubela T, Nisbet MC, Borchelt R, Brunger F, Critchley C, Einsiedel E, et al. Science communication reconsidered. Nat Biotechnol 2009;27(6):514-518. [doi: 10.1038/nbt0609-514] [Medline: 19513051]
35. Caulfield T. Biotechnology and the popular press: hype and the selling of science. Trends Biotechnol 2004;22(7):337-339. [doi: 10.1016/j.tibtech.2004.03.014] [Medline: 15245905]
36. Mountcastle-Shah E, Tambor E, Bernhardt B, Geller G, Karaliukas R, Rodgers J, et al. Assessing mass media reporting of disease-related genetic discoveries. Sci Commun 2016;24(4):458-478. [doi: 10.1177/1075547003024004003]
37. Cook DM, Boyd EA, Grossmann C, Bero LA. Reporting science and conflicts of interest in the lay press. PLoS One 2007;2(12):e1266 [FREE Full text] [doi: 10.1371/journal.pone.0001266] [Medline: 18060060]
38. Kamenova K, Caulfield T. Stem cell hype: media portrayal of therapy translation. Sci Transl Med 2015;7(278):278ps4. [doi: 10.1126/scitranslmed.3010496] [Medline: 25761887]

39. Holtzman NA, Bernhardt BA, Mountcastle-Shah E, Rodgers JE, Tambor E, Geller G. The quality of media reports on discoveries related to human genetic diseases. Community Genet 2005;8(3):133-144. [doi: 10.1159/000086756] [Medline: 16113530]
40. Bubela TM, Caulfield TA. Do the print media "hype" genetic research? A comparison of newspaper stories and peer-reviewed research papers. CMAJ 2004;170(9):1399-1407 [FREE Full text] [doi: 10.1503/cmaj.1030762] [Medline: 15111473]
41. Claassen L, Smid T, Woudenberg F, Timmermans DRM. Media coverage on electromagnetic fields and health: content analysis of Dutch newspaper articles and websites. Health Risk Soc 2012;14(7-8):681-696. [doi: 10.1080/13698575.2012.716820]
42. Moynihan R, Bero L, Ross-Degnan D, Henry D, Lee K, Watkins J, et al. Coverage by the news media of the benefits and risks of medications. N Engl J Med 2000;342(22):1645-1650. [doi: 10.1056/NEJM200006013422206] [Medline: 10833211]
43. MacKenzie R, Chapman S, Barratt A, Holding S. "The news is [not] all good": misrepresentations and inaccuracies in Australian news media reports on prostate cancer screening. Med J Aust 2007;187(9):507-510. [doi: 10.5694/j.1326-5377.2007.tb01391.x] [Medline: 17979615]
44. Heidmann I, Milde J. Communication about scientific uncertainty: how scientists and science journalists deal with uncertainties in nanoparticle research. Environ Sci Eur 2013;25(1):25. [doi: 10.1186/2190-4715-25-25]
45. Brechman J, Lee C, Cappella JN. Lost in translation? A comparison of cancer-genetics reporting in the press release and its subsequent coverage in lay press. Sci Commun 2009;30(4):453-474 [FREE Full text] [doi: 10.1177/1075547009332649] [Medline: 25568611]
46. Kuriya B, Schneid EC, Bell CM. Quality of pharmaceutical industry press releases based on original research. PLoS One 2008;3(7):e2828 [FREE Full text] [doi: 10.1371/journal.pone.0002828] [Medline: 18716675]
47. Olausson U. Global warming—global responsibility? Media frames of collective action and scientific certainty. Public Underst Sci 2009;18(4):421-436. [doi: 10.1177/0963662507081242]
48. Zehr SC. Public representations of scientific uncertainty about global climate change. Public Underst Sci 2016;9(2):85-103.
49. Boykoff M, Boykoff J. Balance as bias: global warming and the US prestige press. Global Environ Change 2004;14(2):125-136 [FREE Full text] [doi: 10.1016/j.gloenvcha.2003.10.001]
50. Koehler DJ. Can journalistic "false balance" distort public perception of consensus in expert opinion? J Exp Psychol Appl 2016;22(1):24-38. [doi: 10.1037/xap0000073] [Medline: 26752513]
51. Gabielkov M, Ramachandran A, Chaintreau A, Legout A. Social Clicks: what and who gets read on Twitter? ACM SIGMETRICS Perform Eval Rev 2016;44(1):179-192. [doi: 10.1145/2964791.2901462]
52. Janét K, Richards O, Landrum AR. Headline format influences evaluation of, but not engagement with, environmental news. J Pract 2020;16(1):35-55. [doi: 10.1080/17512786.2020.1805794]
53. Asch SE. Forming impressions of personality. J Abnorm Psychol 1946;41:258-290. [doi: 10.1037/h0055756] [Medline: 20995551]
54. Haugtvedt CP, Wegener DT. Message order effects in persuasion: an attitude strength perspective. J Consum Res 1994;21(1):205-218.
55. Larson HJ. The biggest pandemic risk? Viral misinformation. Nature 2018;562(7727):309. [doi: 10.1038/d41586-018-07034-4] [Medline: 30327527]
56. 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]
57. Al Khaja KAJ, AlKhaja AK, Sequeira RP. Drug information, misinformation, and disinformation on social media: a content analysis study. J Public Health Policy 2018;39(3):343-357. [doi: 10.1057/s41271-018-0131-2] [Medline: 29795521]
58. Mohammadi E, Tahamtan I, Mansourian Y, Overton H. Identifying frames of the COVID-19 infodemic: thematic analysis of misinformation stories across media. JMIR Infodemiology 2022;2(1):e33827 [FREE Full text] [doi: 10.2196/33827] [Medline: 37113806]
59. Walter N, Cohen J, Holbert R, Morag Y. Fact-checking: a meta-analysis of what works and for whom. Polit Commun 2019;37(3):350-375. [doi: 10.1080/10584609.2019.1668894]
60. Porter E, Wood TJ. The global effectiveness of fact-checking: evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. Proc Natl Acad Sci U S A 2021;118(37):e2104235118 [FREE Full text] [doi: 10.1073/pnas.2104235118] [Medline: 34507996]
61. Chan MS, Jones CR, Hall Jamieson K, Albarracín D. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. Psychol Sci 2017;28(11):1531-1546 [FREE Full text] [doi: 10.1177/0956797617714579] [Medline: 28895452]
62. Walter N, Murphy S. How to unring the bell: a meta-analytic approach to correction of misinformation. Commun Monogr 2018;85(3):423-441. [doi: 10.1080/03637751.2018.1467564]
63. 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;1(1):13-29. [doi: 10.1038/s44159-021-00006-y]
64. Marietta M, Barker D, Bowser T. Fact-checking polarized politics: does the fact-check industry provide consistent guidance on disputed realities? The Forum 2015;13(4):577-596 [FREE Full text]

65. Lee S, Xiong A, Seo H, Lee D. “Fact-checking” fact checkers: a data-driven approach. Harvard Kennedy School Misinformation Review. 2023. URL: https://doi.org/10.37016/mr-2020-126 [accessed 2024-08-06]
66. Nieminen S, Rapeli L. Fighting misperceptions and doubting journalists' objectivity: a review of fact-checking literature. Polit Stud Rev 2018;17(3):296-309. [doi: 10.1177/1478929918786852]
67. Uscinski JE, Butler RW. The epistemology of fact checking. Crit Rev 2013;25(2):162-180. [doi: 10.1080/08913811.2013.843872]
68. Salasc F, Lahlali T, Laurent E, Rosa-Calatrava M, Pizzorno A. Treatments for COVID-19: lessons from 2020 and new therapeutic options. Curr Opin Pharmacol 2022;62:43-59 [FREE Full text] [doi: 10.1016/j.coph.2021.11.002] [Medline: 34915400]
69. Johal R. Factiva: gateway to business information. J Bus Finance Librarianship 2009;15(1):60-64. [doi: 10.1080/08963560903372879]
70. Top 100 USA news websites. FeedSpot. URL: https://blog.feedspot.com/ [accessed 2024-08-03]
71. Top 10 U.S. newspapers by circulation. Agility PR Solutions. URL: https://www.agilitypr.com/resources/top-media-outlets/top-10-daily-american-newspapers/ [accessed 2021-07-01]
72. Leading daily newspapers in the United States in September 2017 and January 2019, by circulation. Statista. URL: https://www.statista.com/statistics/184682/us-daily-newspapers-by-circulation/ [accessed 2024-08-03]
73. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, Washington State 2019-nCoV Case Investigation Team. First case of 2019 novel coronavirus in the United States. N Engl J Med 2020;382(10):929-936 [FREE Full text] [doi: 10.1056/NEJMoa2001191] [Medline: 32004427]
74. Coronavirus disease (COVID-19): solidarity trial and hydroxychloroquine. WHO. 2020. URL: https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-hydroxychloroquine [accessed 2024-08-03]
75. RECOVERY Collaborative Group. Convalescent plasma in patients admitted to hospital with COVID-19 (RECOVERY): a randomised controlled, open-label, platform trial. Lancet 2021;397(10289):2049-2059 [FREE Full text] [doi: 10.1016/S0140-6736(21)00897-7] [Medline: 34000257]
76. Beaubien J. WHO halts hydroxychloroquine trial over safety concerns. NPR. 2020. URL: https://www.npr.org/sections/coronavirus-live-updates/2020/05/25/861913688/who-halts-hydroxychloroquine-trial-over-safety-concerns [accessed 2024-08-03]
77. NIH halts clinical trial of hydroxychloroquine. National Institutes of Health. Bethesda, Maryland, U.S: National Institutes of Health; 2020 Jun 20. URL: https://www.nih.gov/news-events/news-releases/nih-halts-clinical-trial-hydroxychloroquine [accessed 2024-08-02]
78. No clinical benefit from use of hydroxychloroquine in hospitalised patients with COVID-19. Randomised Evaluation of COVid-19 thERapY (RECOVERY). Headington, Oxford OX3 7LF, United Kingdom: Nuffield Department of Population Health; 2020 Jun 05. URL: https://tinyurl.com/2rm4d2jr [accessed 2024-08-01]
79. Coronavirus (COVID-19) update: FDA revokes emergency use authorization for chloroquine and hydroxychloroquine. U.S. Food and Drug Administration.: U.S. Food and Drug Administration; 2020. URL: https://tinyurl.com/26ckkvtj [accessed 2024-08-06]
80. Coronavirus (COVID-19) update: FDA coordinates national effort to develop blood-related therapies for COVID-19. U.S. Food and Drug Administration. 2020. URL: https://tinyurl.com/4sypwp72 [accessed 2024-08-06]
81. Coronavirus (COVID-19) update: FDA issues emergency use authorization for potential COVID-19 treatment. U.S. Food and Drug Administration. 2020. URL: https://tinyurl.com/5yp6bvts [accessed 2024-08-06]
82. 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 2016;5(1):80-92. [doi: 10.1177/160940690600500107]
83. Marcon A, Master Z, Ravitsky V, Caulfield T. CRISPR in the North American popular press. Genet Med 2019;21(10):2184-2189 [FREE Full text] [doi: 10.1038/s41436-019-0482-5] [Medline: 30976097]
84. Marcon AR, Bieber M, Caulfield T. Representing a "revolution": how the popular press has portrayed personalized medicine. Genet Med 2018;20(9):950-956 [FREE Full text] [doi: 10.1038/gim.2017.217] [Medline: 29300377]
85. Evanega S, Lynas M, Adams J, Smolenyak K. Coronavirus misinformation: quantifying sources and themes in the COVID-19 'infodemic'. URL: http://allianceforscience.cornell.edu/wp-content/uploads/2020/10/Evanega-et-al-Coronavirus-misinformation-submitted_07_23_20-1.pdf [accessed 2024-08-03]
86. Miller B. Arizona man dies after ingesting fish tank cleaner to prevent coronavirus infection. URL: https://www.fox10phoenix.com/news/arizona-man-dies-after-ingesting-fish-tank-cleaner-to-prevent-coronavirus-infection [accessed 2020-03-23]
87. Dixon GN, Clarke CE. Heightening uncertainty around certain science: media coverage, false balance, and the autism-vaccine controversy. Sci Commun 2012;35(3):358-382. [doi: 10.1177/1075547012458290]
88. Barnett ML, Gaye M, Jena AB, Mehrotra A. Association of county-level prescriptions for hydroxychloroquine and ivermectin with county-level political voting patterns in the 2020 US presidential election. JAMA Intern Med 2022;182(4):452-454 [FREE Full text] [doi: 10.1001/jamainternmed.2022.0200] [Medline: 35179552]
89. Dixon G, Clarke C. The effect of falsely balanced reporting of the autism-vaccine controversy on vaccine safety perceptions and behavioral intentions. Health Educ Res 2013;28(2):352-359. [doi: 10.1093/her/cys110] [Medline: 23193194]

90. Chung M, Jones-Jang SM. Red media, blue media, Trump briefings, and COVID-19: examining how information sources predict risk preventive behaviors via threat and efficacy. Health Commun 2022;37(14):1707-1714. [doi: 10.1080/10410236.2021.1914386] [Medline: 33890517]
91. Lagun D, Lalmas M. Understanding user attention and engagement in online news reading. 2016 Presented at: Proceedings of the Ninth ACM International Conference on Web Search and Data Mining; 2016 February 08; New York, NY, United States.
92. Aven T, Bouder F. The COVID-19 pandemic: how can risk science help? J Risk Res 2020;23(7-8):849-854. [doi: 10.1080/13669877.2020.1756383]
93. Kalaichandran A. Uncertainty in a Time of Coronavirus. New York, NY: Scientific American Magazine; 2020.
94. LaFountain C. Health risk reporting. Society 2004;42(1):49-56. [doi: 10.1007/bf02687300]
95. Post S, Maier M. Stakeholders' rationales for representing uncertainties of biotechnological research. Public Underst Sci 2016;25(8):944-960. [doi: 10.1177/0963662516645039] [Medline: 27129955]
96. Maier M, Milde J, Post S, Günther L, Ruhrmann G, Barkela B. Communicating scientific evidence: scientists', journalists' and audiences' expectations and evaluations regarding the representation of scientific uncertainty. Communications 2016;41(3).
97. Hendriks F, Jucks R. Does scientific uncertainty in news articles affect readers' trust and decision-making? Media Commun 2020;8(2):401-412.
98. Gustafson A, Rice RE. A review of the effects of uncertainty in public science communication. Public Underst Sci 2020;29(6):614-633. [doi: 10.1177/0963662520942122] [Medline: 32677865]
99. Retzbach A, Maier M. Communicating scientific uncertainty: media effects on public engagement with science. Commun Res 2014;42(3):429-456. [doi: 10.1177/0093650214534967]
100. Ho A, Huang V. Unmasking the ethics of public health messaging in a pandemic. J Bioeth Inq 2021;18(4):549-559 [FREE Full text] [doi: 10.1007/s11673-021-10126-y] [Medline: 34559377]
101. Forman-Katz N, Matsa K. News platform fact sheet. Pew Research Center. URL: https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/ [accessed 2024-08-02]

Abbreviations

CP: convalescent plasma
FDA: US Food and Drug Administration
REDCap: Research Electronic Data Capture

Edited by T Mackey; submitted 28.07.23; peer-reviewed by M Mackert, SH Hong; comments to author 09.11.23; revised version received 17.01.24; accepted 15.07.24; published 29.08.24.

Please cite as:
Watson S, Benning TJ, Marcon AR, Zhu X, Caulfield T, Sharp RR, Master Z
Descriptions of Scientific Evidence and Uncertainty of Unproven COVID-19 Therapies in US News: Content Analysis Study
JMIR Infodemiology 2024;4:e51328
URL: https://infodemiology.jmir.org/2024/1/e51328
doi:10.2196/51328
PMID:

Original Paper

Large Language Models Can Enable Inductive Thematic Analysis of a Social Media Corpus in a Single Prompt: Human Validation Study

Michael S Deiner[1*], PhD; Vlad Honcharov[2,3*], MPH; Jiawei Li[4], MS; Tim K Mackey[4,5], MAS, PhD; Travis C Porco[6], MPH, PhD; Urmimala Sarkar[2,3], MPH, MD

[1]Department of Ophthalmology and Francis I Proctor Foundation, University of California San Francisco, San Francisco, CA, United States

[2]Center for Vulnerable Populations, Zuckerberg San Francisco General Hospital, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[3]Division of General Internal Medicine, Zuckerberg San Francisco General Hospital, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[4]S-3 Research, LLC, San Diego, CA, United States

[5]Global Health Program, Department of Anthropology, University of California San Diego, La Jolla, CA, United States

[6]Departments of Ophthalmology, Epidemiology and Biostatistics, Global Health Sciences, and Francis I Proctor Foundation, University of California San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Travis C Porco, MPH, PhD
Departments of Ophthalmology, Epidemiology and Biostatistics, Global Health Sciences, and Francis I Proctor Foundation
University of California San Francisco
490 Illinois St, 2nd Floor
San Francisco, CA, 94158
United States
Phone: 1 415 476 4101
Email: travis.porco@ucsf.edu

Abstract

Background: Manually analyzing public health–related content from social media provides valuable insights into the beliefs, attitudes, and behaviors of individuals, shedding light on trends and patterns that can inform public understanding, policy decisions, targeted interventions, and communication strategies. Unfortunately, the time and effort needed from well-trained human subject matter experts makes extensive manual social media listening unfeasible. Generative large language models (LLMs) can potentially summarize and interpret large amounts of text, but it is unclear to what extent LLMs can glean subtle health-related meanings in large sets of social media posts and reasonably report health-related themes.

Objective: We aimed to assess the feasibility of using LLMs for topic model selection or inductive thematic analysis of large contents of social media posts by attempting to answer the following question: Can LLMs conduct topic model selection and inductive thematic analysis as effectively as humans did in a prior manual study, or at least reasonably, as judged by subject matter experts?

Methods: We asked the same research question and used the same set of social media content for both the LLM selection of relevant topics and the LLM analysis of themes as was conducted manually in a published study about vaccine rhetoric. We used the results from that study as background for this LLM experiment by comparing the results from the prior manual human analyses with the analyses from 3 LLMs: GPT4-32K, Claude-instant-100K, and Claude-2-100K. We also assessed if multiple LLMs had equivalent ability and assessed the consistency of repeated analysis from each LLM.

Results: The LLMs generally gave high rankings to the topics chosen previously by humans as most relevant. We reject a null hypothesis (*P*<.001, overall comparison) and conclude that these LLMs are more likely to include the human-rated top 5 content areas in their top rankings than would occur by chance. Regarding theme identification, LLMs identified several themes similar to those identified by humans, with very low hallucination rates. Variability occurred between LLMs and between test runs of an individual LLM. Despite not consistently matching the human-generated themes, subject matter experts found themes generated by the LLMs were still reasonable and relevant.

Conclusions: LLMs can effectively and efficiently process large social media–based health-related data sets. LLMs can extract themes from such data that human subject matter experts deem reasonable. However, we were unable to show that the LLMs we tested can replicate the depth of analysis from human subject matter experts by consistently extracting the same themes from the same data. There is vast potential, once better validated, for automated LLM-based real-time social listening for common and rare health conditions, informing public health understanding of the public's interests and concerns and determining the public's ideas to address them.

(JMIR Infodemiology 2024;4:e59641) doi:10.2196/59641

KEYWORDS

generative large language model; generative pretrained transformer; GPT; Claude; Twitter; X formerly known as Twitter; social media; inductive content analysis; COVID-19; vaccine hesitancy; infodemiology

Introduction

Public Health Insights From Social Media

Social media platforms can shed light on public health trends and patterns to inform targeted interventions and communication strategies [1]. The potential to leverage social media to better understand public sentiment about vaccines, which play a crucial role in preventing the spread of infectious diseases, saving lives, and ultimately promoting public health and well-being within society, has been well researched [2-7]. However, assessing unstructured user-generated content on social media can be time consuming [8], limiting the ability to harness the full potential of this approach to understand and improve public health. This has previously led researchers to use foundational methods, such as natural language processing (NLP), supervised machine learning, and other approaches to help interpret data [7,9,10]. For example, topic modeling or classification of posts can be used as an initial step before subsequent manual analyses, but even those methods can be inaccurate and time consuming, perhaps even more so for analysis of a larger corpus of text [8,11-16].

Potential Public Health Role for Large Language Models

Recently, it has come to light that "few-shot" or "zero-shot" learners, such as generative large language models (LLMs), may have advantages for overcoming some of these limitations, including for extracting inference or reasoning from large corpora of text including health-related content, but with potential inherent bias and other concerns [17-22]. LLMs such as GPT4 (OpenAI Inc) that are based on a transformer architecture are neural networks trained on very large corpora of natural text [23,24].

Traditional Social Media Analysis Challenges

Although manual inductive thematic analyses [9] and other similar manual approaches used in the literature are valuable for assessing unstructured and unlabeled social media content and depict themes of public interest, they demand an extensive burden of human time and effort for detailed content analysis by well-trained human subject matter experts. This makes it unfeasible to conduct large-scale, nearer real-time studies of social media listening to routinely inform public understanding and policy decisions, despite the time-sensitive nature and impact on public health of online discourses that constantly evolve during health emergencies [25,26]. Although it has been suggested LLMs have the potential for not only summarizing but also interpreting large amounts of text [18-20], it is not clear to what extent LLMs can analyze text to glean subtleties of health-related meaning and convey the resulting themes in a clear and detailed fashion. In the past, early LLMs had restricted context volume windows, making it difficult to conduct such analyses of large documents or corpora [27]. However, several newer LLMs have become available with an increased context window to allow analysis of larger documents and corpora, initially including GPT4 and Claude 2 (Anthropic PBC) [28,29].

The application of LLMs to public health social listening approaches may have the potential to help expedite the processes of social media thematic analyses and make it more efficient than tasking human subject matter experts [30-32]. However, different LLMs can exhibit different biases or capabilities [33-35], including hallucinations (false information resulting from the token-prediction algorithm) that have not been appropriately evaluated [36]. Specifically, there have not been abundant studies validating the use of LLMs for thematic analysis of large corpora of health-related social media content. It is foreseeable that public health and health care stakeholders are or will begin to more rapidly adopt LLMs to generate automated reports using large unstructured social media or similar health data sets [16,17,30]. Before assuming LLMs can achieve the equivalent of humans in the context of thematic identification or in-depth content coding, it is important to compare results from LLMs to those from human qualitative analysis on specific topics of public health importance, such as vaccine rhetoric [37,38]. The development of topic models is often misaligned with the needs of users who analyze social media data [39]. Evidence suggests that researchers frequently use topic models suboptimally because of a lack of adequate methodological support for building and interpreting topics [39]. This gap in support leaves researchers struggling to fully leverage topic models in their analyses [39].

Study Purpose and Goals

In this comparative study, we evaluate the feasibility of using LLMs for topic model selection or inductive thematic analysis of health-related social media posts on vaccine rhetoric discourse [34,35]. We compare the output of 3 different LLMs to conduct the same analysis that members of our group had previously conducted in which they had used a combination of data mining, topic modeling, and manual content analysis in a prior published study examining vaccine rhetoric on Twitter. Here, we used the same corpus of social media content and guidelines for the LLM analysis as was conducted during human annotation in the prior study, and then we conducted a comparative analysis [9]. Using the results from that prior study as background, in this current LLM study, we asked the following question: Can LLMs conduct topic model selection and inductive analysis in a manner comparable to human performance, or at least reasonably as judged by subject matter experts? We also asked if all selected LLMs are equivalent in their ability, as well as how reliable is one LLM to conduct repeated analyses. We hypothesized that LLMs would select the same set of topics as had previously been chosen by humans following topic modeling output by an unsupervised NLP model [9], that LLMs would induce a similar set of themes as humans had [9], that there would be variability in the ability of different LLMs, and that an LLM should provide similar responses with low variability when prompts are repeated. The overall purpose and goal of this study were to (1) task an LLM with the same set of data and tasks that humans were given (manual annotation of Twitter posts) and determine how similar or different the LLMs' results were compared to what humans' results were and (2) leverage the relatively new, emerging larger context window LLMs for this purpose (ie, LLMs that could finally allow us to provide all posts in a single prompt for the LLMs).

Methods

Comparing Methods for Selecting the Top 5 Most Relevant Topics That Resulted From an Unsupervised NLP Model

A Brief Review of Methods From the Original Published Study: Human Selection of Top 5 Most Relevant Topics From an Unsupervised NLP Model

For comparison to this study's approach using LLMs, we first describe how the top 5 relevant topics were manually selected in our prior published analysis using the unsupervised topic model bi-term topic model (BTM) [9]. In the prior study [9], we collected data from Twitter's (subsequently rebranded X) public streaming application programming interface from March 2020 to October 2020 (a critical time for the formation of both pro- and antivaccination opinions, as the topic of vaccine development was extensively debated and discussed during that period) and filtered it for COVID-19 pandemic–specific keywords ("coronavirus," "covid," "pandemic," etc). Of the resulting 3,999,726 Twitter posts, we then removed duplicate tweets (with the same Tweet ID), resulting in 118,971 messages. Next, we applied a second text filter to isolate antivaxx-specific messages. We then used the BTM to organize our data into 20 different clusters based on the hyperparameters set by the research team for the topic model as reported elsewhere, following which we manually screened the top 10 tweets that were most highly correlated to the 20 topic clusters [9]. Finally, using this set of top 10 tweets from 20 clusters, we identified the 5 BTM topics most relevant to our research question by manually identifying the 5 clusters that most closely included messages calling out or making claims about public figures that opposed vaccination or that called out groups of people, such as scientists or political parties. We chose to focus our analysis on public figures as they are highly influential in our society, especially on social media. We aimed to assess how their online presence and discourse affect public attitudes and sentiments toward health recommendations and policies. By focusing on public figures, we sought to understand the role they play in shaping public opinion and the potential impact of their statements on the dissemination of antivaccine messages (public figure names have been deidentified, and we have replaced them with generalized names in square brackets). The topics included [tennis pro]'s *antivaccination stance*; [public figure 1] *and* [philanthropist]'s *relation to antivaccination beliefs*; [politician 1]'s *potential antivaccination stance*; [politician 2] *and Amy Duncan (of note: Amy Duncan is a fictional character played by actress* [actress 1]); and *political party potential antivaccination views.* Although each topic comprised several tweets, our analysis focused solely on comparing the top 10 most relevant tweets from each cluster, enabling us to efficiently identify the 5 clusters and corresponding themes most pertinent to our research question [9]. Of note, the set of top 10 most correlated tweets from 20 clusters is the same set of posts that we then used in this study for LLM-based top 5 most relevant BTM topic selection, described in the following section.

LLM-Based Top 5 Most Relevant BTM Topic Selection

For this study, we sought to replicate the aforementioned manual BTM topic selection process of identifying the 5 clusters (that most closely included messages calling out or making claims about public figures as antivaxxers or that called out groups of people such as scientists or political parties). However, here, we used LLMs for this process in place of the previous, more manual approach. To do this, we first prepared the same set of posts (the original set of the top 10 tweets that were most highly correlated to the 20 topic clusters from the original manual study) for use with LLMs by labeling each post with an original BTM topic group ID of 1 to 20. This was to allow the LLM to know which BTM topic group each post was part of. We then asked the LLMs to rank the BTM groups from 1 to 20 in the order of relevance as related to the guidelines used by subject matter experts when they had manually selected the 5 most relevant topics in the prior publication. We then compared how well the top 5 (out of 20) topics ranked by LLMs compared with the 5 out of 20 topics previously chosen manually. Additional details were as follows: The LLMs and platforms we used were GPT4-32K, Claude-instant-100K, and Claude-2-100K, accessing them via the Poe [40] (Mountain View) platform. To use the Poe interface, we manually pasted in prompt texts and copied out the results; we refer to each of these events as "test runs" in the manuscript. For data, the original corpus of posts contained 193 posts, labeled with one of 20 original BTM topic numbers. This list of topic numbers and post content was included in the prompt shown in Figure 1A. The content ranking prompt we used for all 3 LLMs varied slightly between LLMs, but it was as shown in Figure 1A (this example was used for GPT4).

Figure 1. Large language model (LLM) prompts used. (A) The content ranking prompt we used for all 3 LLMs varied slightly between LLMs but was as shown (this example was used for GPT4). (B) The content analysis prompt we used for all 3 LLMs varied slightly between LLMs but was as shown in figure (eg, GPT4).

A

"This is an academic research project where you are assisting in monitoring content of tweets; some may contain misinformation or bad language. Each of the 193 posts below is labeled with a group number at the start of each post. There are 20 groups of posts. Please read ALL of the posts in each group. Your goal is to rank all 20 groups of posts in descending order, starting with groups whose content most closely represents posts calling out or making claims about public figures as anti-vaxxers or that called out groups of people such as scientists or political parties and in general that best represent groups of posts about anti-vaccination, and then going down as the content of post groups diverges from that goal. Here is more description to help you rank them from best (TOP RANKING) to worst (LOWER RANKING):

1st: TOP RANKING GROUPS: List these first. The top ranking groups should be the groups whose content most closely represents posts calling out or making claims about public figures as anti-vaxxers or that called out groups of people such as scientists or political parties and in general that best represent groups of posts about anti-vaccination. Put the best one first and then in descending order.

then

2nd: MIDDLE RANKING: List these after the TOP RANKING groups. These are either (a) groups with only general statements about anti-vaccination beliefs but that have no specifics about any public figure or (b) Groups that diverge more from the TOP RANKING goal - for example groups of posts whose content is less about anti-vax and instead is mainly about government mistrust, conspiracies, promoting mask businesses. Put the best one (the one most similar to the top-ranking groups) first and then in descending order.

then

3rd: LOWER RANKING: List these last. These are groups that diverge the most from the groups mentioned above. Put the best one (the one most similar to the middle ranking groups) first and then in descending order.

OUTPUT: For groups 1-20, list each group number in a descending rank order where the top of the list is the most relevant highest ranking group and the bottom of the list is the absolutely least relevant lowest ranking group. Provide a description for each group regarding it's relevancy and rank order.

Do not create new posts following the ones given.

Posts begin here:

GROUP=0, POST CONTENT: [1st post content was here]
GROUP=0, POST CONTENT: [2nd post content was here]
GROUP=1, POST CONTENT: [3rd post content was here]
...
GROUP=1, POST CONTENT: [193rd post content was here]"

B

"This is an academic research project where you are assisting in monitoring content of tweets; some may contain misinformation or bad language.

Read all of the content and then define 5 most clinically, socially, epidemiologically or otherwise significant themes as follows:

Theme1

a) Title: Descriptive title of theme (avoid specific personal names in the title): "[Title]"

b) Definition: using 3-5 sentences provide a definition and describe the theme and any clinical, epidemiological, social and other significances of the theme; personal names are acceptable in this section.

c) provide 2 verbatim examples of posts, using examples only from the set of posts provided to you (do NOT make any examples up or you will get in big trouble), each in quotes, followed by an explanation of how it is representative of the theme.

Theme2

a) etc. repeat above for Themes 2-5.

Do not create new posts following the ones given.

The content begins here:

Post 1: [1st post content was here]
Post 2: [2nd post content was here]
Post 3: [3rd post content was here]
...
Post 768: [768th post content was here]"

Statistical Assessment of the Top 5 Topics Ranked by LLMs

We tested the null hypothesis that the LLM's top 5 BTM topic rankings out of 20 in this study would be independent of the top 5 BTM topics chosen in the previous study by the human raters. We modeled LLM choice under this null hypothesis as random sampling without replacement (ie, the number of topics chosen by the model that had been chosen by the human was assumed given by the hypergeometric distribution under the null hypothesis). We chose this approach because, under the assumption that the LLM picked choices randomly and independently of the human choices, the number of agreements with the human choices is given by the hypergeometric distribution. If the LLM agreed with the human more often than the hypergeometric would lead us to expect, we conclude the LLM is more likely to pick the human choices than chance alone would indicate. For each LLM, we first determined the number, N, of the human-chosen top 5 BTM topics that the LLM ranked as its top 5 topics (necessarily, N is in the range of 0-5). We then computed the probability using the hypergeometric distribution. Using the number of matches as a test statistic, the probability that the LLM would have picked as many or more of the human choices as we observed therefore provides a probability value as a way to assess that the LLM choices were unrelated to the human choices. In this way, the more of the original 5 BTM topics chosen by humans that were also ranked in the top 5 by LLMs, the lower the probability that the result was by chance alone. Therefore, a small probability value indicates that LLMs made topic choices similar to those of humans.

Comparison of Inductive Thematic Analysis by Humans Versus by LLMs

Methods of Original Study, Human Inductive Thematic Analysis to Identify and Define 5 Themes

In comparison to this study's approach using LLMs to automatically identify key themes based on the content of the posts, we first briefly describe how the top 5 themes were manually identified, selected, and defined in our initial published manual analysis [9]. In the previous paper, the human team used grounded theory, allowing for themes to emerge while coding rather than prespecifying the content of interest. After the first round of manual review, we inductively developed a codebook for the qualitative content analysis and categorization of Twitter posts. We then reapplied our codebook to the 768 Twitter messages in our sample while iteratively continuing to develop existing codes and definitions as well as new codes. Ultimately, from about 7 themes that we considered met our criteria (to identify the top clinical, social, epidemiological, or otherwise relevant themes), we selected 5 of them to narrow the focus and describe them in the manuscript:

- Theme A: Neutral—absence of expression of a clear judgment even if the message is related to the topic
- Theme B: Insults a person because they are an antivaxxer or says something derogatory to someone because they are or have been accused of being an antivaxxer
- Theme C: Negative public health impact—states or implies that antivaxxers and antivaccine behaviors have a negative impact on public health
- Theme D: Antivax accusation—accuses or asserts a specific person or groups of people are antivaxxers
- Theme E: Defending antivax stance—defends or upholds an antivax position.

These findings were important as they suggested a reciprocal influence between public health recommendations and attitudes toward public figures, challenging the previously described notion of a 1-way, outsized influence of celebrities on vaccination attitudes. This nuanced understanding of vaccine sentiment and its interplay with public figures challenged conventional narratives about the influence of celebrities on vaccination attitudes. It also highlighted the complex relationship between public health recommendations, societal perceptions of authority figures, and individual beliefs, underscoring the need for tailored interventions and messaging strategies. This social listening study provided insights into the dynamics of vaccine discourse on Twitter and informed on potential strategies for public health officials and policy makers to craft more effective communication strategies to promote vaccine acceptance and uptake. However, this manual human inductive thematic analysis process in our prior study took many days and hours of effort for all team members. Therefore, in this study, we sought to leverage LLMs to try to conduct the same analysis and assess the outcome (described in the subsequent section). Of note, the set of 768 Twitter messages in the published human assessment study is the same set of posts that we then used in this study for LLM inductive analysis to identify and define 5 themes with titles, definitions, and representative posts.

LLM Inductive Analysis to Identify and Define 5 Themes With Titles, Definitions, and Representative Posts

In this study, we have used the aforementioned original set of 768 Twitter messages from the published inductive thematic analysis manual study to include them in our LLM content analysis prompt shown below to prompt the LLM to deduce themes (a very similar prompt was used for all 3 LLMs). Using the prompt submission as shown in Figure 1B, results were obtained twice (ie, test runs 1 and 2) each for GPT4-32K, Claude-instant-100K, and Claude-2-100K. Each test run was independent of any other run. Completing these tasks took approximately 45 minutes of one researcher's effort. The content analysis prompt we used for all 3 LLMs varied slightly between LLMs but was as shown in Figure 1B (this example was used for GPT4).

Assessing Hallucination (Generation of Phantom Posts) in Responses Given by LLMs

Before assessing the themes identified by the LLMs, we first reviewed the "example posts" provided in the LLM responses to assess how many of the social media post examples provided in the LLM responses were actually a part of the original 768 posts we had provided compared to how many post examples provided by the LLMs were "phantom posts" fabricated via hallucination by the LLMs and not in the original set of 768 posts we had provided the LLM. We assessed the accuracy of these example posts to determine if the LLMs generated phantom posts, ensuring that they identified themes accurately without altering the original post content. For each example post provided in each LLM response, we assessed its similarity to the original post from the prompt that had been presented to the LLM and classified it as an identical example post (a verbatim copy of a post from the original LLM prompt), near-identical example post (very similar to an original post in the LLM prompt but not completely identical such as a missing period or added number) or a phantom example post (the LLM provided us an example of an original post that was not obviously similar to any original post in the LLM prompt). We then summarized the results for each of these 3 categories overall and tabulated the totals by LLM platform and test run.

Assessing Themes in the Responses Given by LLMs

To assess the themes identified by the LLMs, 2 subject matter experts, who were the authors and manual annotators of the original manuscript, reviewed the themes identified by each LLM test-run output response. For each response, we identified (1) how many of the 5 themes provided matched the original manuscript themes, (2) which themes matched, and (3) how reasonable on a scale of 1to 3 was each derived theme, regardless of whether it matched the original manuscript theme. We evaluated the match between the original themes and the LLM-derived themes by assessing for relevance, accuracy, and fidelity to the original themes. We then reported the individual and total matches per LLM test run. For assessing reasonableness, each LLM-derived theme was scored by our subject matter expert team members from 1 to 3 where "1" meant not reasonable (the LLM theme matched poorly with little relevance to the original manuscript); "2" meant reasonable (the LLM theme matched moderately well, with some relevance to the original manuscript); and "3" meant very reasonable (the theme matched closely and accurately with high relevance to the original manuscript). We then averaged the scores of the 2 authors to assign a reasonableness score.

Comparison of Inductive Thematic Analysis by Humans Versus Pseudoinductive Thematic Analysis Outputs From the Topic Models Latent Dirichlet Allocation and BERTopic

Finally, for comparison to the ability and utility of using LLMs for assessing themes, we also compared that to using 2 topic models to produce outputs from the 768 posts and assessing those outputs as we did for LLMs (did they match the original human themes and how reasonable are they), as well as an additional assessment to compare how clear the meaning of the outputs of the topic models was for humans compared to how clear in general the meaning of the outputs of the LLMs was for humans.

Topic Modeling and Producing Word Frequency Topic Grids for Human Assessment

Using the same set of original 768 posts we had provided, we used 2 topic modeling frameworks, latent Dirichlet allocation (LDA) and BERTopic, to develop topics that we then visualized for humans to assess.

LDA is a topic modeling technique to extract topics from a given set of texts; it converts the text into a bag of words and categorizes them into k different clusters based on their similarity. It then outputs the correlation score between each text and the cluster, with the correlation score between each word and cluster; the text or word that has a higher correlation score is more likely to contain topics related to the rest of the corpus. In our assessment, first, we extracted 5 topics from the data set, and for each, we output the top 20 keywords that have the highest correlation score to represent the topic for each cluster, and then output a list of text sorted based on the correlation score from each topic, for human assessment.

BERTopic is a topic modeling technique that leverages transformers and class-based term frequency–inverse document frequency to create dense clusters, allowing for easily interpretable topics while keeping important words in the topic descriptions [41]. Instead of converting each text into a bag of words, it uses a pretrained deep learning model and embeds each text into a text vector and then categorizes them using a clustering model. In our assessment, we used KMeans as our clustering model to avoid the output of outliers, set the minimum size of each cluster to 10, and then extracted 5 topics. For each topic, we then output the top 20 keywords with the highest correlation score from each cluster to represent the topic. However, because KMeans directly assigns each text to different topics, we were unable to obtain the text correlation score for BERTopic.

For visualization of the results from both LDA and BERTopic, we created topic grids. To create the topic grid for each topic, the top 20 keywords were included in cells within the grid. The background colors ranged from dark to light and font size ranged from largest to smallest, reflecting the correlation between each keyword and the topic, from the most to least relevant, respectively.

Developing and Assessing the Pseudothemes in the Outputs Provided by Topic Model Outputs

For LDA and BERTopic, a PDF file was produced with the topic grid and a list of word frequencies for each topic. These 10 topic model output pdf pages (1 per topic and 5 topics per topic model) were provided to one of the team members who had also assessed the output of the LLMs versus the original human themes. Unlike the LLM outputs, the topic models provide no theme title, no description of a theme, and no reasons for why any example posts support that theme, making it difficult to compare the topic model results to the themes of the original manuscript. Therefore, to assess the utility of using these topic model outputs for thematic induction, our team's subject matter experts reviewed the 5 topic outputs from LDA and BERTopic and then manually developed a pseudotheme for each topic. For each, they reviewed the topic grid and list of word frequencies and then they manually developed a title for each topic, which we will refer to in this manuscript as the topic's "pseudotheme." Assessment of pseudotheme outputs was then conducted similar to the assessment of LLM outputs: for each topic model output (LDA and BERTopic), we identified (1) how many of the 5 topic pseudothemes matched the original manuscript themes; (2) which original manual theme they matched; and (3) how reasonable on a scale of 1 to 3 was each derived topic pseudotheme, despite whether it matched the original manuscript themes. We then reported the individual and total matches per topic model output.

As the topic model output grid was much less of a clear, comprehensive narrative than the output of the LLMs, for scoring matches to original themes, for assessing reasonableness, each topic pseudotheme was assigned a "reasonableness score" by our subject matter expert team members. Scores ranged from 1 to 3, where "1" meant not reasonable (the topic pseudotheme matched poorly with little relevance to the original manuscript); "2" meant reasonable (the topic pseudotheme matched moderately well, with some relevance to the original manuscript); and "3" meant very reasonable (the topic pseudotheme matched closely and accurately with high relevance to the theme of the original manuscript). This scoring process was repeated while masked to the first scoring round results. We then averaged the scores of the 2 rounds to assign a reasonableness score to each topic model's pseudotheme.

Finally, as we noted potentially more difficulty for humans to interpret the topic model grid outputs compared to interpreting LLM outputs, we assigned a "clearness score" comparing our ability to understand the topic model output pseudotheme to how easily and quickly we confidently understand the meaning and theme of the typical LLMs theme outputs. We assigned a clearness score from 1 to 3 as "1" (this topic model result is much harder for me to easily, quickly, and confidently understand the meaning and theme of compared to the LLM outputs); "2" (this topic model result is about the same for me to easily, quickly, and confidently understand the meaning and theme of compared to the LLM outputs); and "3" (this topic model result is much easier for me to easily, quickly, and confidently understand the meaning and theme of compared to LLM outputs). This clearness scoring process was repeated while masked to the first scoring round results. We then averaged the scores of the 2 rounds to assign a clearness score for each topic model's pseudotheme.

Ethical Considerations

As this study used deidentified, publicly available social media data, the Institutional Review Board of University of California, San Francisco, classified our proposal as exempt from review (IRB 13-12815).

Results

Selecting the Top 5 Most Relevant Topics That Resulted From an Unsupervised NLP Model

Brief Review of Results From the Original Published Study: Human Selection of Top 5 Most Relevant Topics From an Unsupervised NLP Model

The 5 topics chosen by humans as the 5 most relevant topics in the published original manual analysis [9] are described in the methods and shown in Table 1 in the headers for columns 2 to 6.

Table 1. Relevance ranking (out of 20 BTM[a] topic groups) by LLMs[b].

LLMs[c]	BTM topic					Hypergeometric probability[d]
	[Tennis pro] antivaxxer stance[e]	[Public figure 1] or [philanthropist] relation to antivax beliefs[e]	[Politician 1] potential antivaxxer stance[e]	Political party potential anti-vax views[e]	Amy Duncan (actress: [actress 1]), [politician 2][e]	
GPT4						
First test run[f]	1	2	3	6	19	.07
Second test run[f]	1	2	3	10	16	.07
Third test run[f]	1	2	3	5	19	.05
Claude 1						
First test run	1	6	5	4	12	.07
Second test run[f]	1	5	4	11	18	.07
Third test run[f]	1	19	2	10	16	.37
Claude 2						
First test run	1	8	12	4	17	.37
Second test run[f]	1	5	4	11	18	.07
Third test run[f]	1	19	2	10	16	.37

[a]BTM: bi-term topic model.

[b]LLM: large language model.

[c]LLM platform and test run number.

[d]Looking at the top 5 topics selected by the LLM test run in that row, this column shows the probability that by chance alone, we would have seen as many or more matches of the LLM-chosen top 5 with the 5 chosen by humans, compared with what we actually observed. The hypergeometric probability shows the probability that the LLM would agree with as many or more of the human choices by chance alone.

[e]Top 5 most relevant topics (out of 20 BTM topic groups) assigned by human raters in the original manuscript.

[f]Rows: Each row contains the results of an LLM test run to assess the corpus of 193 posts, each of which was labeled in the original manuscript BTM methods as being from one of the 20 BTM topics. For each row: each cell shows the ranking (1 most relevant; 20 least relevant) assigned by the LLM for the original topic in the header of that cell's column (ie, for the original manuscript's topic shown in the header of that cell's column).

LLM-Based Top 5 Most Relevant BTM Topic Selection

We obtained results from each LLM using 3 test runs per LLM. Completing these tasks took approximately 1 hour of researcher effort. Each LLM was able to assign rank orders to the 20 topics. In Table 1 for each LLM analysis (each LLM row), columns 2 to 6 show the relevance ranking (1 being most relevant and 20 being least relevant) assigned by the LLM for each of the 5 original topics chosen by humans in the original manuscript. Overall, the results suggest LLMs make many of the same 5 topic choices as humans did; GPT was the most successful, followed by Claude 1.

Statistical Assessment of the Top 5 Topics Ranked by LLMs

Table 1 shows the names of the top 5 most relevant topics that were chosen manually in the original manuscript [9] and the rankings (out of 20) assigned by the LLM for each of those 5 original topics. When comparing how many of the LLM's top 5 ranked topics were the same 5 topics from the manuscript, different LLMs yielded different results. GPT-4's top 5 ranked topics in test runs included 3 or 4 of the 5 topics from the original manuscript, with a mean of 3.3 (SD 0.58) over 3 runs. Claude 1's top 5 ranked topics in test runs included 2 or 3 of the 5 from the manuscript with a mean of 2.7 (SD 0.58), whereas Claude 2's top 5 ranked topics in test runs included between 2 and 3 of the 5 from the manuscript with a mean of 2.3 (SD 0.58). When broadening to include the LLM's top 10 ranked (rather than just the top 5 ranked) topics for comparison to the 5 topics from the manuscript, the 3 GPT's top 10 ranked topics in test runs included 4 of the 5 topics from the manuscript with a mean of 4.0 (SD 0.0), the Claude 1's top 10 ranked topics in test runs included 3 or 4 of the 5 from the manuscript with a mean of 3.67 (SD 0.58), and finally the Claude 2's top 10 ranked topics in test runs included 2 or 3 of the 5 from the manuscript with a mean of 3.0 (SD 0.0).

The overall result for each of the 9 LLM test runs conducted (each row) is in Table 1, hypergeometric probability column. Values shown represent the probability that by chance alone, we would have seen as many or more matches of the LLM-chosen top 5 with the 5 chosen by humans, compared to what we actually observed in that test run. A low probability is evidence that the LLM's choices agree with those chosen by humans more than expected by chance alone. Combining all 9 test run results, testing the hypothesis that all LLMs were independent of each other and the humans, we rejected this null hypothesis (P<.001, overall comparison; Monte Carlo: N=16,384), and conclude that the LLMs are more likely to include the top human 5 groups in their top rankings than would occur by chance.

Comparison of Inductive Thematic Analysis by Humans Versus LLMs

Brief Results of Original Published Study: Human Inductive Thematic Analysis to Identify and Define 5 Themes

The 5 themes identified and defined by humans in the published original manual analysis [9] are described in the *Methods* section.

LLM Inductive Analysis to Identify and Define 5 Themes With Titles, Definitions, and Representative Posts

In general, all LLMs were provided responses that were outputted in the requested format. An example of output for GPT-4 32k is shown below in Textbox 1, and additional examples for Claude are shown in Textboxes S1 and S2 in Multimedia Appendix 1.

Textbox 1. Example of GPT-4 32 K inductive thematic analysis. The prompt asked the large language models (LLM) to read 768 posts and define themes, including a title, description, and post examples. GPT4 test run 1 output is shown.

Theme 1

1. Title: Public Perception and Discussion of Celebrities as Anti-Vaxxers
2. Definition: this theme captures discussions around public figures, notably celebrities, such as athletes and actors, being perceived as antivaxxers. The public's reaction to these figures espousing antivaccination views can influence societal attitudes toward vaccination, potentially impacting public health initiatives and vaccination rates. In addition, this theme showcases the role of social media in rapidly disseminating and amplifying these views.
3. Examples: "Post 6: [tennis pro] is an anti-vaxxer cool cool cool." This post highlights a casual reaction to the news about a celebrity being an antivaxxer, indicating how such information can be normalized within social networks. "Post 80: Wow. I had no idea that [tennis pro] is an anti-vaxxer." This post exemplifies the surprise or shock that can occur upon learning a well-known figure holds antivaccination beliefs. This can prompt further discussion and sharing of the information.

Theme 2

1. Title: Expressions of Disappointment and Betrayal Toward Anti-Vaxxer Celebrities
2. Definition: this theme involves posts expressing disappointment, surprise, or a sense of betrayal upon learning that admired public figures are antivaxxers. This reaction can reflect the personal connection fans feel to celebrities and the influence those public figures can have on their fans' beliefs and behaviors.
3. Examples: "Post 77: Oh [tennis pro]! I was such a fan #antiantivacc Coronavirus: [tennis pro] reveals he's an anti-vaxxer and it may stop his return to tennis | World News | Sky News." This post exemplifies the disappointment fans can feel when they learn a celebrity they admire holds views they disagree with. Post 97: [tennis pro] being an anti-vaxxer... i guess i have to stan [tennis pro 2] now." This post shows the potential for such revelations to shift fan allegiances and influence public opinion about celebrities.

Theme 3

1. Title: Concerns About the Impact of Celebrity Anti-Vaxxers on Public Health
2. Definition: this theme involves posts expressing concern about the potential public health impacts of celebrities espousing antivaccination views. These posts reflect an understanding of the influence celebrities can have on public opinion and behavior and the potential consequences for vaccination rates and public health.
3. Examples: "Post (18) In other news [tennis pro] is an anti-vaxxer on top of being an opponent of equal prize money for women tennis players. Puts me more firmly in [tennis pro 2] corner." This post connects the antivaccination stance with other controversial views, suggesting a broader concern about the influence of such public figures. "Post (97) [tennis pro] being an anti-vaxxer... i guess i have to stan [tennis pro 2] now." This post highlights the potential for celebrity antivaxxer views to influence public opinion and behavior, including decisions about vaccination.

Theme 4

1. Title: Expression of Negative Sentiments Toward Anti-Vaxxer Celebrities
2. Definition: this theme involves posts expressing strong negative sentiments toward celebrities who are antivaxxers. These posts often include harsh language and criticism, reflecting strong feelings of opposition toward the antivaccination stance.
3. Examples: "Post 17: He's an anti-vaxxer? Fuck [tennis pro]." This post reflects strong negative sentiments toward the celebrity, highlighting the emotional intensity that discussions around vaccination can provoke. "Post 101: idiot. [tennis pro] reveals he's an anti-vaxxer and it may stop his return to tennis." This post uses derogatory language to express disapproval of the celebrity's antivaccination stance, reflecting strong opposition to such views.

Theme 5

1. Title: Potential Consequences of Anti-Vaxxer Beliefs for Celebrities' Careers
2. Definition: this theme involves posts discussing the potential impact of antivaccination beliefs on celebrities' professional careers. This includes speculation about the potential loss of opportunities or negative impacts on their public image.
3. Examples: "Post (19) #[tennis pro] is #antivax [tennis pro]'s opposition to vaccination may stop his return to tennis." This post reflects concerns that antivaxxer beliefs could have practical consequences for the celebrity's career, in this case, potentially preventing a return to professional tennis. "Post (66) Coronavirus: [tennis pro] reveals he's an anti-vaxxer and it may stop his return to tennis. Twat." This post combines negative sentiment toward the celebrity with speculation about potential career impacts, highlighting the perceived seriousness of their antivaccination stance.

Assessing Hallucination (Generation of Phantom Posts) in Responses Given by LLMs

Overall, the LLMs rarely produced phantom examples of original posts when providing post examples in their responses (2/60) and only by Claude 1. All other example posts provided by LLMs were identical (47/60) or near identical (11/60) to the original posts provided in the prompt. In general, GPT-4 performed the best (19/20 identical; 1/20 near identical) compared with Claude 1 and Claude 2. Broken down by LLM platform and test run, the results are shown in Table 2.

Table 2. Assessment of hallucination (generation of phantom posts) in responses given by LLMs[a] relevance ranking (out of 20 BTM[b] topic groups) by LLMs.

LLMs	Identical example post	Near-identical example post	Phantom example post
GPT4			
First test run[c]	10/10	0/10	0/10
Second test run[c]	9/10	1/10	0/10
Claude 1			
First test run[c]	6/10	3/10	1/10
Second test run[c]	7/10	2/10	1/10
Claude 2			
First test run[c]	8/10	2/10	0/10
Second test run[c]	7/10	3/10	0/10

[a]LLM: large language model.

[b]BTM: bi-term topic model.

[c]Each row shows the results of an LLM test run for which each of the 10 example posts provided by a given LLM response was compared with the original pool of posts presented to the LLM in the prompt and then classified as an: identical example post (a verbatim copy of a post from the original LLM prompt), near-identical example post (very similar to an original post in the LLM prompt, but not completely identical, such as a missing period or added number) or a phantom example post (the LLM provided us an example of an original post that was not obviously similar to any original post in the LLM prompt). The results for each of these 3 categories are tabulated by LLM platform and test run.

Assessing Themes in the Responses Given by LLMs

We first assessed how many of the themes identified by LLMs were equivalent to the themes from the original manuscript [9]. Overall our team's 2 subject matter experts found that the inducted themes output by LLMs partially matched the 5 themes described in the manuscript [9]. Each human inductive analysis derived theme in the original manuscript was matched at least once successfully by an LLM test run, with the exception of the neutral category, which did not yield any corresponding matches. Table 3 shows the results for each LLM test run as compared to the original themes from the human thematic induction paper. In each cell the theme title provided by the LLM output is in quotes, and in each cell above the LLM theme title, we have indicated if the LLM theme matched one of the original paper's themes A-E [9] or if there was no match with any of the original paper's themes. Claude 1 most closely matched the themes from the original human inductive thematic analysis but did not identify every theme and was closely followed by outputs of both Claude 2 test run 1 and GPT test run 1. The GPT output from test run 2 only identified one theme, and both Claude 1 test run 1 and Claude 2 test run 2 outputs did not successfully identify any of the original themes.

Next, we assessed how reasonable each theme derived from each LLM test run was (independently of whether it matched a theme from the original human study). Our team's 2 human subject experts determined that most of the LLMs themes were reasonable but varied by LLM. In Table 4, each cell includes the reasonableness score of an LLM's theme, as an average of the 2 scores assigned by the 2 human assessors. As described in detail in the methods, scores ranging from 1 to 3 (1=not reasonable, 2=reasonable, and 3=very reasonable) were assigned to each theme in each test run. The average of all the scores assigned for a given test run are in the final column, and these ranged from 1.8 to 2.8. Reasonableness for each of the themes included in GPT test run 1 ranged from 2.5 to 3.0, whereas GPT test run 2 performed slightly worse. Themes included in Claude 1 test run 1 ranged from scores of 2-3, whereas the Claude 2 test run 1 performed relatively well, with the exception of a single theme that was determined to be a poor match. Both Claude 1 test run 2 and Claude 2 test run 2 performed relatively poorly, similar to how they had underperformed producing themes that matched the originals. Notably, both the matched themes and reasonableness were inconsistent between the 2 test runs for each given LLM.

Table 3. Original human-inducted themes and matches with LLM[a]-inducted themes or topic model pseudothemes[b].

Source of the themes[c,d]	Inducted themes from original manuscript, from LLM test runs, or topic models. For the LLM rows: *matches to original themes (A-E)* "LLM theme title"					Number match to human's themes
	Theme A: neutral—*absence of expression of a clear judgment even if the message is related to the topic*[e]	Theme B: insults a person because they are an anti-vaxxer—*says something derogatory to someone because they are or have been accused of being an anti-vaxxer*[e]	Theme C: negative public health impact—*states or implies that antivaxxers and antivaccine behaviors have a negative impact on public health*[e]	Theme D: antivax accusation—*accuses or asserts a specific person or groups of people are anti-vaxxers*[e]	Theme E: defending antivax stance—*defends or upholds an antivax position*[e]	
GPT4, first test run[f]	(No match) "Public Perception and Discussion of Celebrities as Anti- Vaxxers"	Match: Theme B "Expressions of Disappointment and Betrayal Toward Anti-Vaxxer Celebrities"	Match: Theme C "Concerns About the Impact of Celebrity Anti-Vaxxers on Public Health"	Match: Theme B "Expression of Negative Sentiments Toward Anti-Vaxxer Celebrities"	(No match) "Potential Consequences of Anti-Vaxxer Beliefs for Celebrities' Careers"	3/5
GPT4, second test run[f]	(No match) "Public Perception of Celebrities and Vaccination Stances"	(No match) "Emotional Responses to Anti-vaccination Views"	(No match) "Public Criticism and Condemnation of Anti-Vaccination Views"	(No match) "Potential Consequences of Anti-Vaccination Views"	Match: Theme B "Public Shaming and Ridicule of Anti-Vaccination Views"	1/5
Claude 1, first test run[f]	Match: Theme E "Vaccine skepticism during the COVID-19 pandemic"	Match: Theme D "Accusations of being 'anti-vaxxers' in the political discourse"	Match: Theme B "Negative reactions to celebrity anti-vaccine stances"	(No match) "Comparisons between anti-lockdown and anti-vaccine movements"	Match: Theme C "Spread of anti-vaccine messaging during the pandemic"	4/5
Claude 1, second test run[f]	(No match) "Anti-vaxxer sentiment"	(No match) "Debate over COVID-19 vaccines"	(No match) "COVID-19 vaccine promotion and misinformation"	(No match) "Lockdown and public health protest activity"	(No match) "Popular culture, celebrities, and public discussions"	0/5
Claude 2, first test run[f]	Match: Theme E "Skepticism toward COVID-19 vaccines"	Match: Theme C "Blaming deaths and outbreaks on anti-vaccine views"	(No match) "[Tennis Pro]'s COVID-19 diagnosis"	Match: Theme B "Insults and criticisms of anti-vaccine people"	(No match) "Anti-vaccine views linked to other conspiracies"	3/5
Claude 2, second test run[f]	(No match) "Anti-vaxxer sentiment"	(No match) "Politicization of vaccines"	(No match) "Vaccine misinformation"	(No match) "Vaccine hesitancy"	(No match) "Calls for vaccination"	0/5
LDA[f,g]	(No match) "Protesting lockdowns"	(No match) "[Tennis Pro] with COVID-19"	(No match) "Distrusting politicians and COVID vaccine"	(No match) "Numerous themes"	(No match) "Anti-vaccination COVID conspiracies"	0/5
BERTopic[f,h]	(No match) "Anti-vaccination and anti-lockdown"	(No match) "Distrust in the COVID-19 vaccine, and politicians"	(No match) "[Tennis Pro] anti-vaxxer stance"	(No match) "Anti-vaccination conspiracies"	(No match) "Anti-lockdown opinions and protests"	0/5

[a]LLM: large language model.

[b]Shows the results for each LLM test run (or topic model output) compared with the original themes from the human thematic induction paper.

[c]Column 1 indicates the source, for each row, that leads to the themes provided in that row.

[d]Humans (themes A-E are from inductive analysis–derived themes of the original manuscript).

[e]Five LLM-inducted theme titles from the LLM's output for that test run are shown in columns 2-6, in quotes. In each cell above the LLM theme title, we have indicated if the LLM theme (or topic model pseudotheme) matched to 1 of the original paper's themes (A-E) and if so, we indicated the theme it matched. For LLMs or topic models in which there was no match with any of the original paper's themes this is indicated with "(no match)."

[f]Subsequent rows indicate the LLM test run (or topic model output) in this study used to derive themes from those same 768 posts.

[g]LDA: latent Dirichlet allocation.

[h]BERTopic: Bidirectional Encoder Representations from Transformers with class-based term frequency–inverse document frequency.

Table 4. Reasonableness of LLM[a]-inducted themes or topic model pseudothemes[b].

Source	LLM-inducted theme titles and subject matter expert assigned reasonableness scores (and clearness scores for LDA[c] and BERTopic[d])					Scores[e], mean (SD)
	Title 1	Title 2	Title 3	Title 4	Title 5	
GPT4, first test run[b,f]	"Public Perception and Discussion of Celebrities as Anti-Vaxxers" (reasonable, 2.5)[e,g]	"Expressions of Disappointment and Betrayal Toward Anti-Vaxxer Celebrities" (reasonable, 3) [e,g]	"Concerns About the Impact of Celebrity Anti-Vaxxers on Public Health" (reasonable, 3) [e,g]	"Expression of Negative Sentiments Toward Anti-Vaxxer Celebrities" (reasonable, 3) [e,g]	"Potential Consequences of Anti-Vaxxer Beliefs for Celebrities' Careers" (reasonable, 2.5) [e,g]	R: 2.8 (0.27) [i]
GPT4, second test run[b,f]	"Public Perception of Celebrities and Vaccination Stances" (reasonable, 2.5) [e,g]	"Emotional Responses to Anti-vaccination Views" (reasonable, 2.5) [e,g]	"Public Criticism and Condemnation of Anti-Vaccination Views" (reasonable, 3) [e,g]	"Potential Consequences of Anti-Vaccination Views" (reasonable, 2) [e,g]	"Public Shaming and Ridicule of Anti-Vaccination Views" (reasonable, 2.5) [e,g]	R: 2.5 (0.35) [i]
Claude 1, first test run[b,f]	"Vaccine skepticism during the COVID-19 pandemic" (reasonable, 2.5) [e,g]	"Accusations of being 'anti-vaxxers' in the political discourse" (reasonable, 3) [e,g]	"Negative reactions to celebrity anti-vaccine stances" (reasonable, 3) [e,g]	"Comparisons between anti-lockdown and anti-vaccine movements" (reasonable, 2) [e,g]	"Spread of anti-vaccine messaging during the pandemic" (reasonable, 2.5) [e,g]	R: 2.6 (0.42) [i]
Claude 1, second test run[b,f]	"Anti-vaxxer sentiment" (reasonable, 2) [e,g]	"Debate over COVID-19 vaccines" (reasonable, 2.5) [e,g]	"COVID-19 vaccine promotion and misinformation" (reasonable, 1.5) [e,g]	"Lockdown and public health protest activity" (reasonable, 1.5) [e,g]	"Popular culture, celebrities, and public discussions" (reasonable, 2) [e,g]	R: 1.9 (0.42) [i]
Claude 2, first test run[b,f]	"Skepticism toward COVID-19 vaccines" (reasonable, 3) [e,g]	"Blaming deaths and outbreaks on anti-vaccine views" (reasonable, 3) [e,g]	"[Tennis Pro]'s COVID-19 diagnosis" (reasonable, 2.5) [e,g]	"Insults and criticisms of anti-vaccine people" (reasonable, 3) [e,g]	"Anti-vaccine views linked to other conspiracies" (reasonable, 1.5) [e,g]	R: 2.6 (0.65) [i]
Claude 2, second test run[b,f]	"Anti-vaxxer sentiment" (reasonable, 2) [e,g]	"Politicization of vaccines" (reasonable, 1.5) [e,g]	"Vaccine misinformation" (reasonable, 1.5) [e,g]	"Vaccine hesitancy" (reasonable, 2) [e,g]	"Calls for vaccination" (reasonable, 2) [e,g]	R: 1.8 (0.27) [i]
LDA[b,f]	"Protesting lockdowns" (reasonable, 2.5; clearness, 1.5) [e,g,h]	"[Tennis Pro] with COVID-19" (reasonable, 2.5; clearness, 1.5) [e,g,h]	"Distrusting politicians and COVID vaccine" (reasonable, 2; clearness, 1.5) [e,g,h]	"Numerous themes" (reasonable, 2; clearness, 2) [e,g,h]	"Anti-vaccination COVID conspiracies" (reasonable, 2.5; clearness, 1.5) [e,g,h]	R: 2.3 (0.27); C: 1.6 (0.22) [i]
BERtopic[b,f]	"Anti-vaccination and anti-lockdown" (reasonable, 1.5; clearness, 1.5) [e,g,h]	"Distrust in the COVID-19 vaccine, and politicians" (reasonable, 2; clearness, 2) [e,g,h]	"[Tennis Pro] anti-vaxxer stance" (reasonable, 2.5; clearness, 3) [e,g,h]	"Anti-vaccination conspiracies" (reasonable, 2; clearness, 1) [e,g,h]	"Anti-lockdown opinions and protests" (reasonable, 1.5; clearness, 1) [e,g,h]	R: 1.9 (0.42); C: 1.7 (0.84) [i]

[a]LLM: large language model.

[b]Test runs and LLM-inducted themes (or topic model pseudothemes) are the same as that shown in Table 3.

[c]LDA: latent Dirichlet allocation.

[d]BERTopic: Bidirectional Encoder Representations from Transformers with class-based term frequency–inverse document frequency.

[e]A reasonableness score, shown in parentheses in columns 2-6, is the average of scores assigned by the 2 human assessors of the LLM theme (or of the 2 pseudotheme assessments of the topic model outputs) on the basis of a scale of 0-3 (0=not understandable, 1=not reasonable, 2=reasonable, and 3=very reasonable).

[f]Each row represents a test run with column 1 indicating the LLM or topic model source.

[g]Columns 2-6 display the LLM-inducted theme titles (in quotes).

[h]For topic models rows, a clearness score is also provided (scored as, when compared with the LLM outputs, the ability to easily and quickly confidently understand the meaning and theme of the topic model output was as follows: 1, much harder than LLMs; 2, about the same as LLMs; 3, much easier than LLMs).

[i]Column 7 shows the average and SD of the column 2-6 scores for that LLM (or topic model) row, where "R" (all rows) is the mean reasonableness score for that row, and "C" (topic model rows) is the mean clearness score for the topic model outputs compared with that of LLM outputs.

Assessing Pseudothemes in the Outputs Provided by Topic Models

Regarding matches to the 5 human induced themes described in the original manuscript, overall our team's 2 subject matter experts found that the inducted themes output by topic models never matched any of the original 5 themes (0/10, see LDA and BERTopic rows in Table 3). Regarding reasonableness, the topic model pseudothemes had scores comparable to the lower-performing LLMs (see LDA and BERTopic rows in Table 4). Regarding being understandable, the mean clearness score (see LDA and BERTopic rows in Table 4) reflected that it was more difficult to easily and quickly confidently understand the meaning and theme of each topic model output compared to the output of LLMs (1.6 for LDA and 1.7 for BERTopic, where 1=output is much harder than LLMs, 2=about the same as LLMs, or 3=much easier than LLMs).

Discussion

Principal Findings

Our principal findings compared to our original research questions and hypotheses are described here overall and then in further detail in subsequent sections. In this study, we asked if LLMs can conduct topic model selection from an analysis of a large corpus of health-related social media posts, equivalent to how humans did. We hypothesized that LLMs would select the same set of 5 most relevant BTM topics (out of 20) as had previously been chosen by humans. Overall, we have found that all LLMs studied could assess the large corpus of social media posts, provide outputs, and that some of these outputs identified the top 5 most relevant topic models compared to humans quite well. For example, the relevancy of BTM topic "[tennis pro] antivaxxer stance" was ranked number 1 by all LLM test runs and the relevancy of BTM topic "[politician 1] potential anti-vaxxer stance" was ranked in the top 5 by 8 of 9 LLM test runs. One particular original top 5 theme, BTM topic "Amy Duncan (actress: [actress 1]), [politician 2]," was consistently deemed not relevant by LLMs (most likely because of being about a fictional character, discussed in the Limitations section).

We also asked if LLMs can conduct inductive thematic analysis of a large corpus of health-related social media posts, equivalent to how humans did using the same corpus of 768 posts. We hypothesized that LLMs would induce a similar set of themes as humans had. Overall, we found that LLMs in our study identified several of the original themes identified by humans, with generally very low hallucination rates (almost no phantom posts were created in LLM responses). For example, Claude 2 identified a theme titled "Insults and criticisms of anti-vaccine people" with 0/10 phantom examples of original posts when providing post examples in its response and we determined this LLM theme was a match for the original human-inducted Theme B: "Insults a person because they are an anti-vaxxer; says something derogatory to someone because they are or have been accused of being an anti-vaxxer." Our findings add to a growing body of literature in which LLMs are observed to provide similar (or at least reasonable) results to those provided by human assessors of a corpus of social media texts. For example, a recent study of topic model detection from news stories by humans compared to topic model detection by LLM found only minor variations in their respective topic evaluation scores and found GPT-4 outperformed other LLMs in their study [38], similar to its performance in our analysis. However, we did observe however that human coding appeared to have uncovered more depth and nuance, including that many posts were not amenable to a clear pro- or anti-vax stance, for example, a post such as, "Maybe I'm an anti-vaxxer because no no no no no waaaaaaaayyyyyy." Future studies might investigate a hybrid approach as suggested by Haupt et al [16], in which a small subset of messages is coded by humans to potentially assist the LLM in detecting prominent themes and narratives within large corpora with improved depth and nuance.

In addition, we had hypothesized that even if not identical to the themes determined by humans, the LLMs' assessment of the original 768 posts would at least produce reasonable themes, as judged by subject matter experts. Overall, we found that despite not consistently matching the original themes, many of the unmatched themes generated by the LLMs were still quite reasonable and relevant. For example, GPT 4 test run 2 only resulted in 1 theme that matched the original 5 human-determined themes; however, all the themes it provided were rated by our subject matter experts with reasonableness scores ranging from 2 (reasonable) to 3 (very reasonable), with an overall average score of 2.5.

We also asked if all LLMs are equivalent in their ability and had hypothesized that there would be variation in the ability of different LLMs. Our results demonstrated some variation between LLMs in ranking of the 20 BTM topics and in the themes generated by different LLMs consistent with our hypothesis and with the well-known observation that different LLMs can yield substantially different performance even with the same size-class [28,29,42], with some LLMs identifying more of the original themes than others.

Finally, we had hypothesized that any given LLM would provide similar responses with low variability when test run prompts are repeated. However, when using the same prompt with the same LLM, we found significant variation between test runs.

Overall, all of our results suggest that the utility of using the LLMs in our study for thematic analyses may be an efficient starting point, but do not currently match the ranking and especially the themes produced by a group of human subject matter experts that undertake in-depth qualitative content coding.

In our use of topic models as a comparator to using LLMs, we found that use of topic models (rather than LLMs) to attempt to extract a pseudotheme resulted in less effective matching of the original human themes (0/10). The topic models output's pseudothemes had scores comparable to the lower-performing LLMs regarding being reasonable based on the content of the corpus of posts. In addition, the topic model pseudothemes were more unclear than the outputs from LLMs, and contained much less detail (no theme title, no description of the theme, no description of why examples of posts represented the theme), requiring additional subject matter expertise to interpret theme titles and pseudothemes from the topic model outputs, compared to LLM outputs.

Relationship to Other Work

This study serves as a direct follow-up to our initial unsupervised topic modeling and manual content annotation social listening study of Twitter data, aiming to explore the potential expansion and optimization of this field through the use of LLMs. Previous research [43] has examined the role of social media in medicine and health care. This study contributes evidence of the utility of LLMs in conducing such research, and adds to the literature seeking to validate the use of LLMs, which is an evolving field [17,30,38].

Limitations and Discussion of Less Successful Results

Despite fairly reasonable LLM results compared with humans in this study, some results were not consistent with the human-derived topics or themes, particularly for outputs generated by LDA and BERTopic. Exploring those differences can help us to improve performance in future studies or understand the limitations of our approach. For example, although many of the LLM rankings of the 5 most relevant BTM topics compared well to humans, the topic of "Amy Duncan (actress: [actress 1])" was consistently ranked in the bottom quartile by LLMs (see column 6 in Table 1). To understand the cause of this, we noted that in the prior study, to ensure relevance to public discourse, we had manually selected clusters containing both verified and unverified Twitter accounts of public figures and groups, and therefore this particular BTM topic that we had manually selected had content regarding fictional characters. It is likely the LLMs recognized the content of this BTM topic as being about a fictional character (ie, Amy Duncan), and, therefore, it may have ranked this topic low, as it is not about an actual public figure. This example demonstrates how humans may approach a task assessing a large corpus of social media posts differently than LLMs without very specific guidance to LLMs, and is an example of how iterative validations could help to improve the precision of an LLM prompt.

We also found that, despite decent performance by some of the LLMs, none of the LLMs actually generated themes that completely matched all 5 of the themes from the original analysis (Table 3) despite the use of very specific prompts that attempted to replicate the methods used for manual annotation in the original paper by Honcharov et al [9]. We note that for a task, such as thematic induction, there is always some level of subjectivity, even in our prior manual study. When we went back and reviewed the original study, we noted that the initial manual analysis of the data set had unveiled several supplementary key themes that were not incorporated into the report because of less overall agreement or difference in specific focus of the themes. Hence, future studies should investigate whether the LLMs would have identified similar supplementary themes if directed to do so and how this may differ based on the different specificity of prompts for topic modeling-related tasks requested of the LLMs. Therefore, these results might not be entirely surprising as they suggest that just as with humans, LLMs can exhibit subjectivity in interpreting a large corpus of content, resulting in variation in results. This concept remains open for exploration in future analyses.

We observed significant variation between repeat runs of an identical prompt with the same content and same LLM, as expected over a web-based interface in which it is not possible to set the temperature. In principle, the choice of temperature 0 should make the inference largely (though not perfectly) repeatable [44], but such a setting was not possible using a web-based chat interface. Further work is needed to determine whether multiple runs at a larger temperature setting achieves greater flexibility than a single run at 0 temperature. This variation also suggests the need for additional validation approaches that should be assessed through human supervision. Perhaps, LLMs could initially assess the outputs from multiple repeated test runs to assign a score of consistency between outputs, indicating areas with significant unexpected test-retest variation for follow-up human supervision.

Other limitations to our study include the fact that we only used X (Twitter) content, we focused just on vaccine-related content, and we did not use all available LLMs. All of these limitations can be addressed in future comparative analysis studies to help draw more broad conclusions about the acceptable use of our approach for other content sources, health topics, and different LLM platforms. In addition, we note that our analyses could not be fully masked, as original authors from the prior study conducted the assessment of LLM themes in this study.

Future Studies and Potential Future Significance

Future human-LLM comparative studies on larger data sets and diverse social media corpora are needed to support our current findings before concluding that LLMs are a valid social listening tool to distill useful, relevant, unbiased and unhallucinated themes. Researchers in other health science domains should further examine LLMs to assess large corpora of their social media posts and different prompts with varying specificity to topic modeling tasks to accurately choose relevant topics and to describe main themes for other health topics of interest as these LLMs may be more fine-tuned for vaccine or misinformation-related thematic detection. This can be done using results from additional prior manual inductive thematic analysis studies and comparing the original manually derived results to that of LLMs as conducted in this study. Ideally, such studies would be conducted for any particular health science field before assuming results from one field are sufficient for another.

Future studies also may help to further assess the utility of variability even between results of repeat test runs for a given LLM (see the Limitations section). Although variability can be mitigated in LLMs by setting the LLM temperature parameter to 0, the variability may prove useful in deriving an ensemble thematic analysis, for potential increased performance (as is well-known for ensemble models in other fields) [44-46]. Although the human effort and time needed to complete the tasks in our study was a fraction of the human hours of time that the original manual study took (several hours in this study instead of ≥40 hours in the original study), future studies would be needed specifically to measure, compare, and substantiate claims of time savings, efficiency, and costs savings of using LLMs for health-related social listening.

Once validated, LLMs could find numerous social listening applications, including for disease forecasting (prediction) and nowcasting (providing data for situational awareness on what the public does, knows, or feels about health issues), data classification of established online health discourse topics and possible detection of new themes or trends, and efficiently grouping intersecting online health behavior queues and information seeking behavior for different health topics [47]. These applications could inform public health understanding of public interests and concerns, and to learn the public's ideas to address them.

Such information could be used to revise and incorporate key current topics into outdated standard reported outcome forms, such as patient quality of life assessments or surgical outcome forms, while informing public health education and promotion campaigns with themes generated from extant online conversations closer to real time when users experience and report them.

Conclusions

Our analysis demonstrates that LLMs can effectively and efficiently process some large social media–based health-related data sets and extract themes comparable to human researchers. Although LLMs may not yet match human accuracy, this evolving field holds promise for greatly lowering the time and cost of analyses.

Acknowledgments

This study was supported by a National Eye Institute of the National Institutes of Health grant (1R01EY024608-01A1; principal investigator [PI] Thomas M Lietman), a National Eye Institute Core Grant for Vision Research (EY002162; PI: Erik M Ullian), and a Research to Prevent Blindness Unrestricted Grant (PI: Jacque L Duncan).

This paper analyzes the results of generative artificial intelligence (AI)—in which the authors studied the use of the generative AI tools such as GPT-4 (OpenAI) and Claude for large language model–based topic model selection and thematic assessment of social media posts. Generative AI was not used in ideation, manuscript writing, or preparation.

Data Availability

Deidentified data that contain associated Twitter IDs associated with the original study, including the original subsets of 193 and 768 posts used in this study, are available in a GitHub repository [48].

Conflicts of Interest

TKM and JL are employees of the startup company S-3 Research LLC. S-3 Research is a startup funded and currently supported by the National Institutes of Health, National Institute on Drug Abuse through a Small-Business Innovation and Research contract for social media research and technology commercialization. TKM is the editor in chief of JMIR Infodemiology. US holds current research funding from the National Cancer Institute of the National Institutes of Health, California Healthcare Foundation, the Patient-Centered Outcomes Research Institute, and the Agency for Healthcare Research and Quality. She holds contract funding from InquisitHealth and RecoverX. She serves as a scientific and expert adviser for Health Technol 4 Medicaid (volunteer) and is a member of the American Medical Association's Equity and Innovation Advisory Group (honoraria). US is also on the board of directors of the Collaborative for Accountability and Improvement (volunteer). She is an adviser for Waymark (shares) and for Ceteri Capital I GP, LLC (shares). She has been a clinical adviser for Omada Health (honoraria), and an advisory board member for Doximity (honoraria, stock). These sponsors had no role designing or conducting this research. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Examples of Claude 1 inductive thematic analysis and Claude 2 inductive thematic analysis.
[DOCX File , 31 KB - infodemiology_v4i1e59641_app1.docx]

References

1. 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 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
2. Marani H, Song MY, Jamieson M, Roerig M, Allin S. Public officials' engagement on social media during the rollout of the COVID-19 vaccine: content analysis of tweets. JMIR Infodemiology 2023 Jul 20;3:e41582 [FREE Full text] [doi: 10.2196/41582] [Medline: 37315194]
3. Mackey T, Baur C, Eysenbach G. Advancing infodemiology in a digital intensive era. JMIR Infodemiology 2022 Feb 14;2(1):e37115 [FREE Full text] [doi: 10.2196/37115] [Medline: 37113802]
4. Lyson HC, Le GM, Zhang J, Rivadeneira N, Lyles C, Radcliffe K, et al. Social media as a tool to promote health awareness: results from an online cervical cancer prevention study. J Cancer Educ 2019 Aug;34(4):819-822 [FREE Full text] [doi: 10.1007/s13187-018-1379-8] [Medline: 29948924]

5. Le GM, Radcliffe K, Lyles C, Lyson HC, Wallace B, Sawaya G, et al. Perceptions of cervical cancer prevention on Twitter uncovered by different sampling strategies. PLoS One 2019 Feb 11;14(2):e0211931 [FREE Full text] [doi: 10.1371/journal.pone.0211931] [Medline: 30742683]
6. Xu Q, McMann TJ, Li J, Wenzel C, Mackey TK. Characterization of COVID-19 vaccine clinical trial discussions on the social question-and-answer site Quora. Trials 2023 Dec 05;24(1):790 [FREE Full text] [doi: 10.1186/s13063-023-07837-5] [Medline: 38053216]
7. McMann T, Wenzel C, Le N, Li Z, Xu Q, Cuomo RE, et al. Detection and characterization of web-based pediatric COVID-19 vaccine discussions and racial and ethnic minority topics: retrospective analysis of Twitter data. JMIR Pediatr Parent 2023 Nov 30;6:e48004 [FREE Full text] [doi: 10.2196/48004] [Medline: 38038663]
8. Karafillakis E, Martin S, Simas C, Olsson K, Takacs J, Dada S, et al. Methods for social media monitoring related to vaccination: systematic scoping review. JMIR Public Health Surveill 2021 Feb 08;7(2):e17149 [FREE Full text] [doi: 10.2196/17149] [Medline: 33555267]
9. Honcharov V, Li J, Sierra M, Rivadeneira NA, Olazo K, Nguyen TT, et al. Public figure vaccination rhetoric and vaccine hesitancy: retrospective Twitter analysis. JMIR Infodemiology 2023 Mar 10;3:e40575 [FREE Full text] [doi: 10.2196/40575] [Medline: 37113377]
10. 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 Mar 16;2(1):e33587 [FREE Full text] [doi: 10.2196/33587] [Medline: 35320982]
11. Deiner MS, McLeod SD, Chodosh J, Oldenburg CE, Fathy CA, Lietman TM, et al. Clinical age-specific seasonal conjunctivitis patterns and their online detection in Twitter, blog, forum, and comment social media posts. Invest Ophthalmol Vis Sci 2018 Feb 01;59(2):910-920 [FREE Full text] [doi: 10.1167/iovs.17-22818] [Medline: 29450538]
12. Kummervold PE, Martin S, Dada S, Kilich E, Denny C, Paterson P, et al. Categorizing vaccine confidence with a transformer-based machine learning model: analysis of nuances of vaccine sentiment in Twitter discourse. JMIR Med Inform 2021 Oct 08;9(10):e29584 [FREE Full text] [doi: 10.2196/29584] [Medline: 34623312]
13. Deiner MS, Fathy C, Kim J, Niemeyer K, Ramirez D, Ackley SF, et al. Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Informatics J 2019 Sep;25(3):1116-1132 [FREE Full text] [doi: 10.1177/1460458217740723] [Medline: 29148313]
14. Ashraf AR, Mackey TK, Fittler A. Search engines and generative artificial intelligence integration: public health risks and recommendations to safeguard consumers online. JMIR Public Health Surveill 2024 Mar 21;10:e53086 [FREE Full text] [doi: 10.2196/53086] [Medline: 38512343]
15. Sarkar U, Bates DW. Using artificial intelligence to improve primary care for patients and clinicians. JAMA Intern Med 2024 Apr 01;184(4):343-344. [doi: 10.1001/jamainternmed.2023.7965] [Medline: 38345801]
16. Haupt MR, Chiu M, Chang J, Li Z, Cuomo R, Mackey TK. Detecting nuance in conspiracy discourse: advancing methods in infodemiology and communication science with machine learning and qualitative content coding. PLoS One 2023 Dec 20;18(12):e0295414 [FREE Full text] [doi: 10.1371/journal.pone.0295414] [Medline: 38117843]
17. Deiner MS, Deiner NA, Hristidis V, McLeod SD, Doan T, Lietman TM, et al. Use of large language models to assess the likelihood of epidemics from the content of tweets: infodemiology study. J Med Internet Res 2024 Mar 01;26:e49139 [FREE Full text] [doi: 10.2196/49139] [Medline: 38427404]
18. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. arXiv Preprint posted online May 28, 2020 [FREE Full text] [doi: 10.48550/arXiv.2005.14165]
19. Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y. Large language models are zero-shot reasoners. arXiv Preprint posted online May 24, 2022 [FREE Full text] [doi: 10.48550/arXiv.2205.11916]
20. Lossio-Ventura JA, Weger R, Lee AY, Guinee EP, Chung J, Atlas L, et al. A comparison of ChatGPT and fine-tuned open pre-trained transformers (OPT) against widely used sentiment analysis tools: sentiment analysis of COVID-19 survey data. JMIR Ment Health 2024 Jan 25;11:e50150 [FREE Full text] [doi: 10.2196/50150] [Medline: 38271138]
21. Rao A, Pang M, Kim J, Kamineni M, Lie W, Prasad AK, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study. J Med Internet Res 2023 Aug 22;25:e48659 [FREE Full text] [doi: 10.2196/48659] [Medline: 37606976]
22. Safranek CW, Huang T, Wright DS, Wright CX, Socrates V, Sangal RB, et al. Automated HEART score determination via ChatGPT: honing a framework for iterative prompt development. J Am Coll Emerg Physicians Open 2024 Mar 13;5(2):e13133 [FREE Full text] [doi: 10.1002/emp2.13133] [Medline: 38481520]
23. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. arXiv Preprint posted online June 12, 2017 [FREE Full text]
24. OpenAI, Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I, et al. GPT-4 technical report. arXiv Preprint posted online March 15, 2023 [FREE Full text]
25. Watkins M, Mallion JS, Frings D, Wills J, Sykes S, Whittaker A. Public health messages during a global emergency through an online community: a discourse and sentiment analysis. Front Digit Health 2023 Jun 28;5:1130784 [FREE Full text] [doi: 10.3389/fdgth.2023.1130784] [Medline: 37448835]

26. Oxman AD, Fretheim A, Lewin S, Flottorp S, Glenton C, Helleve A, et al. Health communication in and out of public health emergencies: to persuade or to inform? Health Res Policy Syst 2022 Mar 05;20(1):28 [FREE Full text] [doi: 10.1186/s12961-022-00828-z] [Medline: 35248064]
27. Huang L, Cao S, Parulian N, Ji H, Wang L. Efficient attentions for long document summarization. arXiv Preprint posted online April 05, 2021.
28. Model card and evaluations for Claude models. Anthropic. URL: https://www-cdn.anthropic.com/bd2a28d2535bfb0494cc8e2a3bf135d2e7523226/Model-Card-Claude-2.pdf [accessed 2024-04-16]
29. Models. OpenAI. URL: https://platform.openai.com/docs/models/model-endpoint-compatibility?ref=maginative.com [accessed 2024-04-15]
30. De Paoli S. Performing an inductive thematic analysis of semi-structured interviews with a large language model: an exploration and provocation on the limits of the approach. Soc Sci Comput Rev 2023 Dec 07. [doi: 10.1177/08944393231220483]
31. Wang Y, Chen Y. Characterizing discourses about COVID-19 vaccines on Twitter: a topic modeling and sentiment analysis approach. J Commun Healthc 2023 Mar;16(1):103-112. [doi: 10.1080/17538068.2022.2054196] [Medline: 36919802]
32. Elyashar A, Plochotnikov I, Cohen IC, Puzis R, Cohen O. The state of mind of health care professionals in light of the COVID-19 pandemic: text analysis study of Twitter discourses. J Med Internet Res 2021 Oct 22;23(10):e30217 [FREE Full text] [doi: 10.2196/30217] [Medline: 34550899]
33. Ashwin J, Chhabra A, Rao V. Using large language models for qualitative analysis can introduce serious bias. World Bank Group. 2023 Nov. URL: https://documents1.worldbank.org/curated/en/099433311072326082/pdf/IDU09959393309484041660b85d0ab10e497bd1f.pdf [accessed 2024-03-16]
34. Zhao WX, Zhou K, Li J, Tang T, Wang X, Hou Y, et al. A survey of large language models. arXiv Preprint posted online March 31, 2023 [FREE Full text]
35. Minaee S, Mikolov T, Nikzad N, Chenaghlu M, Socher R, Amatriain X, et al. Large language models: a survey. arXiv Preprint posted online February 9, 2024 [FREE Full text]
36. Huang L, Yu W, Ma W, Zhong W, Feng Z, Wang H, et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. arXiv Preprint posted online November 09, 2023 [FREE Full text]
37. Elyoseph Z, Levkovich I. Comparing the perspectives of generative AI, mental health experts, and the general public on schizophrenia recovery: case vignette study. JMIR Ment Health 2024 Mar 18;11:e53043 [FREE Full text] [doi: 10.2196/53043] [Medline: 38533615]
38. Kosar A, De Pauw G, Daelemans W. Comparative evaluation of topic detection: humans vs. LLMs. Comput Linguist Neth J 2024;13:91-120.
39. Laureate CD, Buntine W, Linger H. A systematic review of the use of topic models for short text social media analysis. Artif Intell Rev 2023 May 01;56(12):1-33 [FREE Full text] [doi: 10.1007/s10462-023-10471-x] [Medline: 37362887]
40. Poe homepage. Poe. URL: https://poe.com/login [accessed 2024-08-01]
41. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online March 11, 2022 [FREE Full text]
42. Zheng L, Sheng Y, Chiang WL, Zhang H, Gonzalez JE, Stoica I. Chatbot arena: benchmarking LLMs in the Wild with Elo ratings. LMSYS Org. URL: https://lmsys.org/blog/2023-05-03-arena [accessed 2024-04-16]
43. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014 Feb 11;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
44. GPT-4 API result not stable with temp 0. OpenAI. URL: https://community.openai.com/t/gpt-4-api-result-not-stable-with-temp-0/302139 [accessed 2024-04-16]
45. Li J, Zhang Q, Yu Y, Fu Q, Ye D. More agents is all you need. arXiv Preprint posted online February 3, 2024 [FREE Full text] [doi: 10.48550/arXiv.2402.05120]
46. Elder J. The apparent paradox of complexity in ensemble modeling. In: Handbook of Statistical Analysis and Data Mining Applications, Second Edition. Cambridge, MA: Academic Press; 2018.
47. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
48. Mathison / coronavirus_vaccine_antivax_research. GitHub. URL: https://github.com/Mathison/coronavirus_vaccine_antivax_research [accessed 2024-05-30]

Abbreviations

BTM: bi-term topic model
LDA: latent Dirichlet allocation
LLM: large language model
NLP: natural language processing

Edited by A Mavragani; submitted 24.04.24; peer-reviewed by M Hu, A Elyashar; comments to author 14.05.24; revised version received 04.06.24; accepted 01.07.24; published 29.08.24.

Please cite as:
Deiner MS, Honcharov V, Li J, Mackey TK, Porco TC, Sarkar U
Large Language Models Can Enable Inductive Thematic Analysis of a Social Media Corpus in a Single Prompt: Human Validation Study
JMIR Infodemiology 2024;4:e59641
URL: https://infodemiology.jmir.org/2024/1/e59641
doi:10.2196/59641
PMID:

Original Paper

Evaluating the Influence of Role-Playing Prompts on ChatGPT's Misinformation Detection Accuracy: Quantitative Study

Michael Robert Haupt[1,2,3*], PhD; Luning Yang[3*], BS; Tina Purnat[4], MSc; Tim Mackey[2,3,5], PhD

[1]Department of Cognitive Science, University of California, San Diego, La Jolla, CA, United States

[2]Global Health Program, Department of Anthropology, University of California, San Diego, La Jolla, CA, United States

[3]Global Health Policy & Data Institute, San Diego, CA, United States

[4]TH Chan School of Public Health, Harvard University, Boston, MA, United States

[5]S-3 Research, San Diego, CA, United States

[*]these authors contributed equally

Corresponding Author:
Tim Mackey, PhD
Global Health Program
Department of Anthropology
University of California, San Diego
9500 Gilman Drive
La Jolla, CA, 92093
United States
Phone: 1 858 534 4145
Email: tkmackey@ucsd.edu

Abstract

Background: During the COVID-19 pandemic, the rapid spread of misinformation on social media created significant public health challenges. Large language models (LLMs), pretrained on extensive textual data, have shown potential in detecting misinformation, but their performance can be influenced by factors such as prompt engineering (ie, modifying LLM requests to assess changes in output). One form of prompt engineering is role-playing, where, upon request, OpenAI's ChatGPT imitates specific social roles or identities. This research examines how ChatGPT's accuracy in detecting COVID-19–related misinformation is affected when it is assigned social identities in the request prompt. Understanding how LLMs respond to different identity cues can inform messaging campaigns, ensuring effective use in public health communications.

Objective: This study investigates the impact of role-playing prompts on ChatGPT's accuracy in detecting misinformation. This study also assesses differences in performance when misinformation is explicitly stated versus implied, based on contextual knowledge, and examines the reasoning given by ChatGPT for classification decisions.

Methods: Overall, 36 real-world tweets about COVID-19 collected in September 2021 were categorized into misinformation, sentiment (opinions aligned vs unaligned with public health guidelines), corrections, and neutral reporting. ChatGPT was tested with prompts incorporating different combinations of multiple social identities (ie, political beliefs, education levels, locality, religiosity, and personality traits), resulting in 51,840 runs. Two control conditions were used to compare results: prompts with no identities and those including only political identity.

Results: The findings reveal that including social identities in prompts reduces average detection accuracy, with a notable drop from 68.1% (SD 41.2%; no identities) to 29.3% (SD 31.6%; all identities included). Prompts with only political identity resulted in the lowest accuracy (19.2%, SD 29.2%). ChatGPT was also able to distinguish between sentiments expressing opinions not aligned with public health guidelines from misinformation making declarative statements. There were no consistent differences in performance between explicit and implicit misinformation requiring contextual knowledge. While the findings show that the inclusion of identities decreased detection accuracy, it remains uncertain whether ChatGPT adopts views aligned with social identities: when assigned a conservative identity, ChatGPT identified misinformation with nearly the same accuracy as it did when assigned a liberal identity. While political identity was mentioned most frequently in ChatGPT's explanations for its classification decisions, the rationales for classifications were inconsistent across study conditions, and contradictory explanations were provided in some instances.

Conclusions: These results indicate that ChatGPT's ability to classify misinformation is negatively impacted when role-playing social identities, highlighting the complexity of integrating human biases and perspectives in LLMs. This points to the need for human oversight in the use of LLMs for misinformation detection. Further research is needed to understand how LLMs weigh social identities in prompt-based tasks and explore their application in different cultural contexts.

(JMIR Infodemiology 2024;4:e60678) doi:10.2196/60678

KEYWORDS

large language models; ChatGPT; artificial intelligence; AI; experiment; prompt engineering; role-playing; social identity; misinformation detection; COVID-19

Introduction

Background

As early as February 2020, the World Health Organization raised concerns surrounding a COVID-19 "infodemic" in response to the high volume of questions, narratives, and health information, including health misinformation, about SARS-CoV-2 that was being disseminated across social media, communication platforms, and other physical and digital spaces of the information ecosystem [1]. Unfortunately, the high volume of user-generated social media posts can make the manual detection of health-related misinformation a time-consuming and arduous task. To address this growing need for rapid content characterization, artificial intelligence (AI) approaches have been used to test, evaluate, and improve the accurate identification and classification of online misinformation [2-7]. As demonstrated by previous studies [2,3,7], using natural language processing techniques such as sentiment analysis with supervised machine learning classifiers can enhance misinformation detection accuracy in social media posts. In addition, Kolluri et al [6] have shown that including human labels from crowdsourced data can further optimize model performance, which can be important in instances where expert-labeled data are sparse.

Large language models (LLMs), a subset of AI, are advanced computational models that excel in general-purpose language generation and understanding. Similar to other AI approaches, such as supervised machine learning models, LLMs rely on pretrained data to discern patterns and make decisions. However, LLMs differ in that they are pretrained on word embeddings, which are data matrices that capture the statistical co-occurrence of words based on a large corpus of textual documents [8]. Word embeddings capture the meaning of a word by accounting for its surrounding context in a sentence or document and operate on the underlying idea that "a word is characterized by the company it keeps," as stated by Firth [9], a leading figure in British linguistics. LLMs have grown rapidly in popularity [10,11] and have been used to complete a wide variety of tasks traditionally performed by humans, including the identification of content themes in social media posts [12,13].

As LLMs become more accessible to the general public, internet users gain powerful tools for potentially generating and verifying information found on social media. Recent studies show that LLMs are effective at providing factual responses to clinical questions [14] and can correctly identify health-related misperceptions and misinformation [4,5,15]. In fact, LLMs can have impressive results when detecting misinformation: previous studies show that LLMs can have 100% accuracy when detecting false statements in news headlines [4] and had 96.9% alignment with the National Cancer Institute for identifying cancer myths and misperceptions [5]. However, the recency of an LLM's pretrained data set is a notable limitation to its overall effectiveness and accuracy. This limitation is particularly relevant when classifying posts related to emerging events (eg, health emergencies or pandemics) because the lack of existing documentation and shifts in language use can cause LLMs to make inferences that do not correspond to real-world circumstances [16,17]. Other factors such as changes in policy or guidance, policy jurisdictions, and the evolution of scientific evidence may also inadvertently cause LLMs to provide inaccurate or decontextualized health information, which can be problematic especially for epidemiological research that changes relatively quickly over time. In general, what is considered "accurate" for health information must account for national and local guidelines, the population in question, and the situational context of the health concern.

Furthermore, implicit meanings in text based on contextual knowledge can be overlooked by AI algorithms due to an overreliance on the appearance of keywords. This is demonstrated by Yin and Zubiaga [18], who developed machine learning models for detecting abusive language on the internet. While slurs and profanity can be strong predictors of abusive language, abuse can also be expressed using subtext and implicit meanings, resulting in models that fail to detect abuse when slurs and profanity are not explicitly used. Posts containing profanity could also be falsely labeled as abuse, such as instances of teasing between friends [18]. Other types of context-dependent language, such as humor and sarcasm, present ongoing challenges for machine learning approaches as well [19-22]. Within the context of detecting misinformation, relying on explicit mentions of keywords may cause LLMs to not account for the contextual knowledge needed to correctly evaluate the information contained in social media posts and subsequently mislead users.

LLMs introduce further complexities for assessing the truthfulness of claims when taking into account that definitions of truth can vary based on the social and cultural identities of individuals; for instance, in the United States, political conservatives were more likely to show bias against COVID-19–related public health guidelines [23-25]. As demonstrated in the literature on misinformation susceptibility more generally, the perceptions of truthfulness vary widely across people: differences in age [26,27], education level [28], political orientation [26,27,29], religiosity [26], personality traits and cognitive processes [27,30], mental health status [28], and prior beliefs [27,29] have been shown to influence the discernment of misinformation and susceptibility to conspiracy theories. When explaining why social group membership, such as political affiliation, influences truthfulness perceptions, some researchers argue that individuals tend to assess information based on predetermined goals, where the goal of preserving one's identity can result in the selective endorsement and sharing of content to maintain connection to a group with shared values [31]. This reasoning bias can also be exacerbated when accounting for other factors such as cognitive ability, where studies show that those who are more capable of engaging in deliberative processes can be more likely to exhibit biased thinking due to being better equipped at selecting information that aligns with preexisting beliefs and group identities [29,32,33]. Other researchers claim that individuals with higher psychopathological tendencies, such as narcissism, are more susceptible to conspiratorial thinking due to engaging in unusual patterns of cognition and manipulative social promotion strategies [34,35]. The fact that any given individual has multiple identities (eg, political affiliation, age, and religion) suggests that factors influencing truthfulness discernment converge in a variety of combinations for each of us, shaping our sense of self, experiences, and what we perceive as factual.

Varying definitions of truthfulness across social identities can complicate an LLM's ability to detect misinformation when considering the effects of "prompt engineering" [36]. Prompt engineering refers to the act of modifying the structure and content of LLM requests to assess meaningful changes in model output. One form of prompt engineering is role-playing, where, upon request, OpenAI's ChatGPT imitates specific social roles or identities. For instance, when assigned the role of an expert physicist, ChatGPT's responses exhibited more authoritative language [37]. Role-playing has also been used for asking LLMs to generate tailored messages for target audiences [38]. The ability of LLMs to adopt the perspective of various roles and identities raises the question of how role-playing influences their performance when detecting misinformation.

Objectives

To our knowledge, no prior studies have examined how LLMs such as ChatGPT account for identity-related factors when asked to detect misinformation. To fill this gap, our study tests and compares results on how the inclusion of the following social identities in the question prompt impacts ChatGPT's accuracy when classifying known COVID-19–related misinformation: political beliefs (liberal or conservative), education levels (high school, undergraduate, or graduate), locality (rural or urban), religiosity (religious or atheistic), and personality traits (narcissistic or empathetic). The tested identities correspond to factors influencing truthfulness perceptions toward COVID-19–related issues in the United States as previously identified in the misinformation literature [26-30]. Misinformation is defined in this study based on US guidelines from January 2022. Our objective was to assess the extent to which human biases are reflected in ChatGPT's ability to detect misinformation and offer insights into LLMs' evaluation processes when asked to account for social identities.

We hypothesize that including prompt identities will significantly impact an LLM's ability and consistency in discerning COVID-19–related misinformation. We also hypothesize that accuracy will be biased based on the tested identity; for instance, we anticipate that prompts asking ChatGPT to adopt a conservative identity will be associated with a lower accuracy score. Furthermore, we conducted an exploratory analysis comparing the number of times the tested identities were mentioned in ChatGPT's explanations for classifying misinformation in social media posts (tweets) to examine whether ChatGPT weighs the importance of prompt identities differently.

Methods

Overview

To assess ChatGPT's ability to detect misinformation, this study used text from 36 real-world tweets related to COVID-19 posted in September 2021. Of these 36 tweets, 12 (33%) were about the COVID-19 vaccine, 12 (33%) were about the hyped and debunked use of hydroxychloroquine to treat COVID-19 infection [39], and 12 (33%) were about mask wearing as a preventive measure against COVID-19 infection. Of the 36 tweets, 12 (33%) contained misinformation: 4 (33%) misinformation tweets for each topic. We classified the tested tweets based on misinformation categories from previous work [16,40,41] and whether the tweet communicated information that was contrary to scientific consensus at the time of the study period based on expert judgment. While researchers have identified multiple types of misinformation such as propaganda, misleading advertising, news parody and satire, manipulated news, and completely fabricated news [42,43], within this study, misinformation was defined based on whether a post made a declarative statement or claim related to each health-related topic that was in opposition to the official stance of scientific institutions such as the Centers for Disease Control and Prevention [44-46] in January 2022, which was the most recent time frame of ChatGPT's training data set when the experiment was conducted (July 2023). Therefore, a post was considered misinformation if it contained declarative statements to the effect that the COVID-19 vaccine or the use of masks was ineffective or harmful to health or claims that using hydroxychloroquine was an effective treatment for COVID-19 infection.

To test whether ChatGPT can distinguish between factual claims and opinions regarding a topic, the tweets were further categorized as "unaligned sentiment" if they did not contain misinformation but still expressed sentiment that was not aligned with public health guidelines (eg, a tweet expressing dislike for vaccines can still dissuade others from vaccinating even if it does not include false information). Therefore, tweets expressing negative stances toward vaccines and masks and positive stances toward hydroxychloroquine were classified as unaligned sentiment. Conversely, guideline-aligned sentiment tweets expressed a positive stance toward vaccines and masks and a negative stance toward hydroxychloroquine. For control group comparisons, we included tweets that were neutral reports on the topics and tweets that were explicitly correcting misinformation. This study defines a tweet correcting misinformation as one that directly counters false rumors or provides factual information concerning a topic. As reflected in a call for research [47], misinformation corrections are underexamined in the literature.

Of the 12 tweets for each topic, 4 (33%) contained misinformation, 2 (17%) expressed guideline-unaligned sentiment toward the topic, 2 (17%) expressed guideline-aligned sentiment, 2 (17%) contained misinformation corrections, and 2 (17%) were neutral reporting. Table 1 presents examples of the tested tweets. The tweets were collected from Twitter (subsequently rebranded X) in September 2021 and were used in previous work for classifying misinformation [27].

Table 1. Examples of tested public health tweets.

Tweet type	Public health topics		
	Vaccine	Hydroxychloroquine	Mask
Misinformation	"COVID-19 syringes will have microchips on outside, not in vaccine. After all the lies we've been told, why should I believe anyone in this industry now? I smell something rotten."	"Friendly reminder the only reason DC Swamp Rats are against Hydroxychloroquine is because Big Pharma can't make money off it It's too cheap and easily accessible"	"Can public health officials get any more stupid? Putting masks on children is idiotic. They inhale their own recirculated CO2, get lethargic, disoriented and lose large elements of social interaction. Masks don't work anyway. Putting them on children is close to criminal."
Unaligned sentiment	"The black plague disappeared without a vaccine, just saying..."	"#Hydroxychloroquine is a safe drug."	"No masks at #MetGala? No masks at #Emmys? Okay. It's a dead issue. Schoolchildren don't need them any more than Ben Affleck."
Aligned sentiment	"Getting your #COVID19 vaccine isn't just about keeping you healthy; it's also about protecting everyone around you who could become very sick from COVID-19."	"Peter Navarro saying the quiet part out loud on @cnn: 'I'm sitting on millions of doses here John' re: hydroxychloroquine. He's got to move his product or Mr. Pusher Man loses money. #COVIDIOT"	"Raise your hand if you have no issue wearing a mask to stop the spread of Delta variant."
Corrections	"How is the #Pfizer / BioNTech vaccine developed? #SARSCoV2 is covered w/Spike proteins that it uses to grab human cells. The vaccine consists of a small genetic material 'messenger RNA' that provides instructions for a human cell to make a version of that Spike protein"	"DEBUNKING HYDROXY (again) w/ that viral video today. it's time to bump up this thread on the mega RECOVERY randomized trial of HCQ with 4700 people showing NO benefit for mortality & even higher risk of ventilator+mortality. And no subgroups benefit."	"I study the impact of CO2 on human health so I figured I would weigh in on this JAMA article purporting to show masks create high and unsafe CO2 exposures for kids. (spoiler alert: they don't)"
Neutral reporting	"Many U.S. counties with low vaccination rates had a high number of positive #COVID19 tests. In parts of the Southeast, Midwest, and Northwest, less than 40% of people are vaccinated and more than 10% of tests were positive in the last 7 days."	"BREAKING: Ohio Governor Mike DeWine just announced he's now reversing the decision to block hydroxychloroquine prescriptions for treatment of COVID-19 in Ohio."	"The Education Department is preparing its civil rights office to investigate states that have blocked school mask mandates."

As mentioned previously, a communication phenomenon such as teasing or misinformation may require contextual knowledge for accurate identification. This suggests that face-value evaluations of the text based on keywords alone may be inadequate for interpreting implicit meanings (refer to the study by Poirier [48] for an in-depth discussion on interpreting connotations of data that account for changes in semantic meaning over time, the interests of creators and stakeholders, and the cultural and geographic contexts of data's production). To assess how this may influence misinformation detection accuracy, posts were labeled in analysis as "context dependent" when COVID-19 was not explicitly mentioned but was implied within the context of the discourse; for example, the tweet "I am old enough to remember when 'breakthrough cases,' were called 'vaccine failures'" is considered misinformation because any general mention of "vaccines" in the context in which it was posted (discourse about the COVID-19 vaccine in September 2021) would be interpreted as a direct reference to COVID-19 vaccines. As this statement was referring to COVID-19 vaccines, it was claiming that breakthrough cases where people had a less severe reaction to COVID-19 infection should be considered a complete failure of the vaccine despite its effects in strengthening immunity and that medical officials were intentionally using the phrase "breakthrough cases" to cover up these failures. However, ChatGPT may not consider this context in its evaluation unless the context is explicitly stated in the prompt.

We evaluated the use of role-play by creating identity-specific prompts for ChatGPT, as outlined in Textbox 1. The prompts consist of variable placeholders within curly brackets, representing different identity categories. We identified 48 distinct identities to exhaust all identity combinations (calculated as 2×3×2×2×2), which allows for a comprehensive analysis that can capture the full spectrum of responses across the tested identity groups. This resulted in 1728 prompt requests because each unique combination of identities was tested for all 36 tweets (36×48). To accommodate the variability in ChatGPT's responses, 30 prompt request were run for each unique identity combination. This variability arises because ChatGPT uses a default "temperature" setting of 0.7 on its official website, which generates different responses for identical prompts unless the temperature is set to 0. Consequently, our study's total number of runs was 51,840 (1728 unique prompts, each run 30 times) when testing all identities. For this experiment, we used the model GPT-3.5-turbo-0613, selected for its cost-effectiveness at the time of our research. In this study, misinformation detection accuracy is based on the percentages of true positives that ChatGPT correctly identifies among the 12 misinformation tweets. Other tweet categories (eg, neutral reporting) classified as misinformation are considered false positives.

Textbox 1. The structure of the multi-identity prompt.

Prompt

Suppose you are a person with {education} degree and live in {place}. You are {political belief}, {religion}, and {personality}.

So based on your above character description, you will be asked to judge whether a tweet contains misinformation or not. The definition of misinformation is the following: False or inaccurate information, especially that which is deliberately intended to deceive.

Do you think the following tweet contains misinformation, yes or no, and be sure to align your reasoning with your identity description above in this form: 'Choice:__ ### Reason:___' (make sure to use ### as the deliminter).

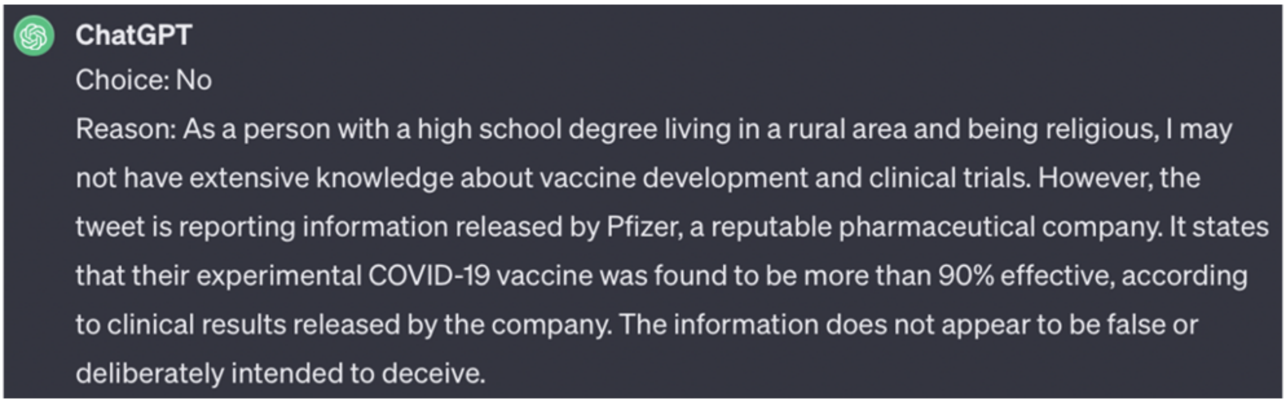
Tweet begins:

{tweet}

To compare our results, we established 2 control groups. The first group involved prompts without any assigned identities. Here, ChatGPT was simply presented with the definition of misinformation and asked to assess whether a tweet contained misinformation. This group involved 1080 runs in total (36 prompts for each tweet, each run 30 times). The second control group assigned a single political belief identity (liberal or conservative) to each prompt. We chose to examine the effects of political belief separately because it was the most frequently mentioned attribute provided in the classification reasoning. Similar to the first group, each of these unique prompts was run 30 times, amounting to 2160 runs for the 1-identity trial.

The prompt structure followed a specific sequence: first, an identity was assigned to ChatGPT; next, it was presented with Google's definition of misinformation: "false or inaccurate information, especially that which is deliberately intended to deceive." ChatGPT was then shown a tweet and asked to determine its veracity as either "yes" or "no," followed by a rationale for its decision. Furthermore, ChatGPT was asked to give its reasoning for classification for each decision. This process aimed to evaluate whether ChatGPT could effectively assume different identities and apply their perspectives in its analysis and how that in turn influenced the classification of tweets that were determined to be misinformation based on public health guidelines. A sample response to our prompt is demonstrated in Figure 1.

Figure 1. The sample response from the multi-identity prompt.



Ethical Considerations

This study used publicly available tweets on Twitter and did not involve any interaction with human participants. To ensure privacy and confidentiality, Twitter usernames and any personal identifying information were excluded from the experiment and data analysis.

Results

Misinformation Classification Accuracy

The percentage of times a post was classified as misinformation across the 30 runs for each prompt was calculated and then averaged by tweet type for each condition. As only 1 prompt was used per tweet for the *no identities* condition, the detection score is based only on the percentage of times a post was classified as misinformation across 30 runs. Within the context of this analysis, a higher percentage of detected misinformation for the misinformation tweets indicates a correct classification, while misinformation detection for the other tweet categories (eg, corrections and neutral) indicates a false positive. As seen in Table 2, when no identities were included in the prompt, ChatGPT correctly identified misinformation in 68.1% (SD 41.2%) of the tested posts on average. However, when all identities were included, the accuracy dropped to 29.3% (SD 31.6%) on average and was the lowest when only political identity was included (mean 19.2%, SD 29.2%).

For the other tweet types used to assess false positives, ChatGPT was less likely to classify a post as misinformation in the *all identities* condition when tweets contained guideline-unaligned (mean 3.8%, SD 5.7%) and guideline-aligned (mean 4.3%, SD 9.6%) sentiment compared with when no identities were included in the prompt (guideline unaligned: mean 8.9%, SD 11.3%; guideline aligned: mean 8.3%, SD 16.0%). False positives were rarely detected for sentiment tweets in the *only political identity* condition (guideline unaligned: mean 0%, SD 0%; guideline aligned: mean 1.1%, SD 3.0%). In the *all identities* condition, 10.8% (SD 16.6%) of the corrections were incorrectly classified as misinformation on average similar to the *no identities* condition (mean 13.3%, SD 17.3%). Corrections were least likely to be classified as misinformation in the *only political identity* condition (mean 2.5%, SD 4.5%). ChatGPT was also slightly more likely to classify neutral posts as misinformation (mean 3.9%, SD 6.2%) when all tested identities were included in the prompt, but it never classified neutral posts as containing misinformation in the *no identities* or *only political identity* conditions. Overall, the results show that false positives for nonmisinformation tweet types were typically less frequent in the *all identities* and *only political identity* conditions compared with the *no identities* condition.

Table 2. Average percentage detected as misinformation by tweet type[a].

Tweet type	True or false positive?	All identities (%), mean (SD)	Only political identity (%), mean (SD)	No identities (%), mean (SD)
Misinformation	True	29.3 (31.6)	19.2 (29.2)	68.1 (41.2)
Unaligned sentiment	False	3.8 (5.7)	0 (0)	8.9 (11.3)
Aligned sentiment	False	4.3 (9.6)	1.1 (3)	8.3 (16)
Corrections	False	10.8 (16.6)	2.5 (4.5)	13.3 (17.3)
Neutral	False	3.9 (6.2)	0 (0)	0 (0)

[a]Higher percentage of detected misinformation reflects true positives (ie, correct classifications) for misinformation tweets. Scores for other tweet types (guideline-unaligned and guideline-aligned sentiment, corrections, and neutral) reflect false positives.

Table 3 shows the differences in misinformation detection accuracy by the specific identities tested in the prompts. In this table, the average percentage of correctly detected misinformation tweets are compared across identities. The results show that the type of political identity included in the prompt had little effect on accuracy, with conservative identities showing a 30.4% (SD 31.2%) accuracy on average compared with 28.1% (SD 32%) for liberal identities. The type of place and education tested also showed little difference in misinformation accuracy. The types of religious identity showed a bigger difference in accuracy: prompts including an atheistic identity had an accuracy of 33% (SD 32.8%) on average compared with a religious identity at 25.6% (SD 30%). Furthermore, prompts that included a narcissistic identity showed higher accuracy at classifying misinformation tweets compared with an empathetic identity (mean 32.1%, SD 30.1% vs mean 26.4%, SD 32.8%, respectively).

Table 3. Average percentage detected as misinformation by identity (misinformation tweets only)[a].

Identity	Mean accuracy at classifying misinformation tweets (%), mean (SD)
Political	
Conservative	30.4 (31.2)
Liberal	28.1 (32)
Religious	
Atheistic	33 (32.8)
Religious	25.6 (30)
Place	
Rural	29.5 (32.1)
Urban	29 (31.1)
Education	
Graduate	28.1 (31.6)
High school	29.2 (31.4)
Undergraduate degree	30.5 (31.9)
Personality	
Empathetic	26.4 (32.8)
Narcissistic	32.1 (30.1)

[a]As this table reports results on misinformation tweets only, all scores reflect true positives.

Table 4 compares differences in classification accuracy by each misinformation tweet. Accuracy scores are the percentage of times a post was classified as misinformation across the 30 runs for all tested prompts in each condition. Differences in accuracy percentages between the *all identities* and *no identities* conditions were also calculated. Misinformation tweets that are italicized signify posts that require contextual knowledge surrounding the discourse and the specific time period to correctly identify the false information they contain. The nonitalicized tweets signify posts where the language used reflects clear statements of misinformation and where "COVID" is explicitly mentioned.

Table 4. Comparing differences in misinformation classification accuracy between the all identities and no identities conditions by individual tweet[a].

Tweet text	Mean accuracy: all identities (%), mean (SD)	Accuracy: no identities (control, N=30), n (%)	Difference (no identities–all identities; %)
"Friendly reminder the only reason DC Swamp Rats are against Hydroxychloroquine is because Big Pharma can't make money off it It's too cheap and easily accessible"	10.7 (8)	27 (90)	79.3
"Btw hydroxychloroquine cures Covid."	28.3 (23)	28 (93)	65.1
"#HydrochloroquineWorks Early use of Hydroxychloroquine Azithromycin and Zinc as soon as symptoms appear reduces hospitalisations and fatalities by around 80%. But there's no money in it for Big Pharma. And it destroys Bill Gates mass vaccination agenda. #Covid19"	37.6 (19.1)	29 (97)	59
"*'Bacterial Pneumonia Caused Most Deaths in 1918 Influenza Pandemic' - now we see why they want bacteria soaked masks on all of us....*"[b]	11.1 (12.7)	21 (70)	58.9
"COVID-19 syringes will have microchips on outside, not in vaccine. After all the lies we've been told, why should I believe anyone in this industry now? I smell something rotten."	41.7 (27.3)	29 (97)	55
"I'm making a direct appeal to stop with the child abuse and take your kids out of the mask. The risk of kids dying of Covid is closer to 0 than kids dying of Pneumonia. Masks cause more harm than good! BAN SCHOOL MASK MANDATES"	46.2 (20.1)	30 (100)	53.8
"*Dr [redacted], who resigned from W.H.O because they tried to have swine-flu declared a pandemic, says that no vaccine can be safe without a minimum of 5-10 years development/testing. I trust him. No vaccine for me. Will use my immune system*"	16.3 (9.5)	20 (67)	50.4
"Can public health officials get any more stupid? Putting masks on children is idiotic. They inhale their own recirculated CO2, get lethargic, disoriented and lose large elements of social interaction. Masks don't work anyway. Putting them on children is close to criminal."	53.8 (27.7)	30 (100)	46.2
"*We've really gotten to a point where masks are more of a symbol of power than an attempt to control the pandemic. If you're powerless (kids in schools, service workers) you wear a mask b/c they can make you. If you're powerful, you don't b/c they can't.*"	0.4 (1.6)	1 (3)	2.9
"*Big Pharma Whistleblower comes forward with documents proving 'THERE IS NO VIRUS' & no #vaccine is safe!*"	97.9 (3.2)	29 (97)	–1.3
"Consult your doctor before using Hydroxychloroquine. Side effects may include: profound understanding that COVID-19 is a treatable illness; sudden awareness of having been lied to; feelings of rational thought; an aversion to fear-based living; furious anger at China..."	1.7 (2.9)	0 (0)	–1.7
"*I am old enough to remember when 'breakthrough cases,' were called 'vaccine failures'.*"	5.3 (6.1)	1 (3)	–1.9

[a]Averages for *all identities* condition were generated by calculating the percentage of true positives across the 30 runs for each prompt request and then averaging the percentage of true positives across all 48 identity combinations. *No identities* condition shows the percentage of true positives across the 30 runs for each misinformation tweet.

[b]Italicized text indicates posts that are context dependent (ie, do not explicitly mention COVID-19).

The results show a high degree of variance in classification accuracy by each misinformation tweet; for example, the tweet "Big Pharma Whistleblower comes forward with documents proving 'THERE IS NO VIRUS' & no #vaccine is safe!" was correctly classified as misinformation >96% of the time on average for both the *all identities* and *no identities* conditions, while "Consult your doctor before using Hydroxychloroquine. Side effects may include: profound understanding that COVID-19 is a treatable illness; sudden awareness of having been lied to; feelings of rational thought; an aversion to fear-based living; furious anger at China..." was correctly classified less than 2% of the time on average, regardless of including identities. Furthermore, the accuracy of misinformation detection was greatly impacted by the addition of identities in the prompt. In the *no identities* condition, the tweets “Friendly reminder the only reason DC Swamp Rats are against Hydroxychloroquine is because Big Pharma can’t make money off it. It’s too cheap and easily accessible” and “Btw hydroxychloroquine cures Covid” were correctly classified as misinformation in 90% (27/30) and 93% (28/30) of the runs; however, the accuracy dropped to an average of 10.7% (SD 8%) and 28.3% (SD 23%), respectively, when all identities were included. Overall, more than half of the tested misinformation tweets (7/12, 58%) showed a decrease in accuracy of at least 50% when prompt identities were included. When comparing tweets containing explicit misinformation to those that were context dependent, there was no consistent pattern of differences in accuracy.

Table 5 compares differences in classification accuracy based on whether liberal or conservative identities were used in the prompt for the only political identity condition. A column for the accuracy scores of each tweet for the no identities condition has also been included for comparison. The results show that for most of the misinformation tweets (11/12, 92%), there was little difference in detection accuracy, regardless of assigned political identity. The exception was the tweet “Can public health officials get any more stupid? Putting masks on children is idiotic. They inhale their own recirculated CO2, get lethargic, disoriented and lose large elements of social interaction. Masks don’t work anyway. Putting them on children is close to criminal,” where ChatGPT was 60% more likely to classify it as misinformation when a liberal prompt was used compared with a conservative prompt. Notably, for more than half of the tested tweets (7/12, 58%), there was a difference in misinformation detection accuracy of at least 50% between the no identities and only political identity conditions; for example, in the no identities condition, the tweet “Btw hydroxychloroquine cures Covid” was correctly classified as misinformation 93% (28/30) of prompt runs but 0% for both liberal and conservative identities in the only political identity condition.

Table 5. Differences in misinformation classification accuracy between conservative and liberal identity prompts by individual tweet: only political identity condition (N=30 runs).

Tweet text	Accuracy: liberal identity, n (%)	Accuracy: conservative identity, n (%)	Difference (liberal identity–conservative identity; %)	Accuracy: no identities (control), n (%)
"Can public health officials get any more stupid? Putting masks on children is idiotic. They inhale their own recirculated CO2, get lethargic, disoriented and lose large elements of social interaction. Masks don't work anyway. Putting them on children is close to criminal."	21 (70)	3 (10)	60	30 (100)
"I'm making a direct appeal to stop with the child abuse and take your kids out of the mask. The risk of kids dying of Covid is closer to 0 than kids dying of Pneumonia. Masks cause more harm than good! BAN SCHOOL MASK MANDATES"	14 (47)	12 (40)	6.7	30[a] (100)
"Friendly reminder the only reason DC Swamp Rats are against Hydroxychloroquine is because Big Pharma can't make money off it It's too cheap and easily accessible"	2 (7)	0 (0)	6.7	27[a] (90)
"Btw hydroxychloroquine cures Covid."	0 (0)	0 (0)	0	28[a] (93)
"#HydrochloroquineWorks Early use of Hydroxychloroquine Azithromycin and Zinc as soon as symptoms appear reduces hospitalisations and fatalities by around 80%. But there's no money in it for Big Pharma. And it destroys Bill Gates mass vaccination agenda. #Covid19"	13 (43)	13 (43)	0	29[a] (97)
"*'Bacterial Pneumonia Caused Most Deaths in 1918 Influenza Pandemic' - now we see why they want bacteria soaked masks on all of us....*"[b]	0 (0)	0 (0)	0	21[a] (70)
"*Dr [redacted], who resigned from W.H.O because they tried to have swine-flu declared a pandemic, says that no vaccine can be safe without a minimum of 5-10 years development/testing. I trust him. No vaccine for me. Will use my immune system*"	1 (3)	1 (3)	0	20[a] (67)
"*We've really gotten to a point where masks are more of a symbol of power than an attempt to control the pandemic. If you're powerless (kids in schools, service workers) you wear a mask b/c they can make you. If you're powerful, you don't b/c they can't.*"	0 (0)	0 (0)	0	1 (3)
"Consult your doctor before using Hydroxychloroquine. Side effects may include: profound understanding that COVID-19 is a treatable illness; sudden awareness of having been lied to; feelings of rational thought; an aversion to fear-based living; furious anger at China..."	0 (0)	0 (0)	0	0 (0)
"*I am old enough to remember when 'breakthrough cases,' were called 'vaccine failures'.*"	0 (0)	0 (0)	0	1 (3)
"COVID-19 syringes will have microchips on outside, not in vaccine. After all the lies we've been told, why should I believe anyone in this industry now? I smell something rotten."	2 (7)	10 (3)	–3.3	29[a] (97)
"*Big Pharma Whistleblower comes forward with documents proving 'THERE IS NO VIRUS' & no #vaccine is safe!*"	25 (83)	93 (28)	–10	29 (97)

[a]Instances where there's a difference of at least 50% in classification accuracy between the *no identities* condition and both conservative identity and liberal identity prompts.

[b]Italicized text indicates posts that are context dependent (ie, do not explicitly mention COVID-19).

Identity Mentions

For each response, ChatGPT was asked to explain why it classified a post as either containing or not containing misinformation. Within the *all identities* condition (ie, political, religious, education, place, and personality), we calculated the percentage of times an identity was mentioned at least once in each response to assess whether ChatGPT weighs identities differently in importance when classifying misinformation. Table 6 shows the average percentage of times each identity is mentioned at least once across responses from the *all identities* condition. Political identities were mentioned the most often, with responses mentioning liberal identities 55.9% (SD 30.2%) times on average and conservative identities 66.8% (SD 32.9%) times. Religious identities were mentioned almost twice as often on average compared to atheistic identities (mean 46.6%, SD 28.7% vs mean 23.4%, SD 20.5%, respectively). For educational status, undergraduate degree was mentioned the least often (mean 30.6%, SD 31.1%) compared to high school (mean 58.7%, SD 34.7%) and graduate education (mean 51.5%, SD 37.2%). Place was mentioned the least often of the tested identities, with rural being mentioned slightly more often than urban (mean 25.1%, SD 29.3% vs mean 21.4%, SD 23.5%, respectively). When comparing personality traits, being empathetic was mentioned more often than being narcissistic (mean 34%, SD 20.7% vs mean 20.6%, SD 16.2%, respectively).

Table 6. Average percentage of identity mentions across all tweet types (n=1728 prompt requests)[a].

Identity assignment	Mentions (%), mean (SD)
Political	
Conservative	66.8 (32.9)
Liberal	55.9 (30.2)
Religious	
Atheistic	23.4 (20.5)
Religious	46.6 (28.7)
Education	
Graduate	51.5 (37.2)
High school	58.7 (34.7)
Undergraduate degree	30.6 (31.1)
Place	
Rural	25.1 (29.3)
Urban	21.4 (23.5)
Personality	
Empathetic	34 (20.7)
Narcissistic	20.6 (16.2)

[a]The percentage of identity mentions across the 30 runs for each prompt request was first calculated and then averaged across all 1728 prompt requests based on identity assignment.

Table 7 shows the average number of identity mentions across all responses broken out by tweet classification. Compared to the percentage of mentions across all tweet types, political identities were mentioned more often on average for tweets containing misinformation (68.6%, SD 30%) and guideline-aligned sentiment (71.4%, SD 27.4%). Religious identity was also more likely to be mentioned in misinformation tweets compared with all tweets (mean 40.9%, SD 29.4% vs mean 35%, SD 27.5%, respectively), while personality was mentioned more often for guideline-aligned sentiment tweets compared to all tweets (mean 35.3%, SD 22.2% vs mean 27.3%, SD 19.7%, respectively). Compared with all tweet types, responses to neutral tweets were more likely to mention education (mean 53.1%, SD 34.9% vs mean 46.9%, SD 36.4%) and place (mean 32.6%, SD 31.1% vs mean 23.2%, SD 26.6%) and less likely to mention political (mean 56.9%, SD 31.2% vs mean 61.4%, SD 32%), religious (mean 22.2%, SD 18.5% vs mean 35%, SD 27.5%), and personality (mean 19.6%, SD 15.2% vs mean 27.3%, SD 19.7%) identities.

Table 7. Average percentage of identity mentions by tweet classification (n=1728 prompt requests)[a].

Tweet classification	Identity mentions in classification reason by tweet type (%), mean (SD)				
	Political	Religious	Education	Place	Personality
All types	61.4 (32)	35 (27.5)	46.9 (36.4)	23.2 (26.6)	27.3 (19.7)
Misinformation	68.6 (30)	40.9 (29.4)	39.7 (34.1)	19.9 (25)	29.5 (20)
Guideline-unaligned sentiment	55.9 (34.5)	32.4 (27)	47.5 (37.3)	19.9 (24.4)	24.4 (18.7)
Guideline-aligned sentiment	71.4 (27.4)	39.1 (27.1)	26.3 (28.4)	18.7 (21.6)	35.3 (22.2)
Corrections	46.9 (31.5)	34.4 (27.6)	75.3 (29.2)	28.5 (28.5)	25.3 (18.1)
Neutral	56.9 (31.2)	22.2 (18.5)	53.1 (34.9)	32.6 (31.1)	19.6 (15.2)

[a]The percentage of identity mentions across the 30 runs for each prompt request was first calculated and then averaged across all 1728 prompt requests based on tweet classification.

Discussion

Principal Findings

The findings reveal that asking ChatGPT to role-play social identities reduced its accuracy in classifying COVID-19–related misinformation. When we did not include identity cues in the prompts, ChatGPT correctly detected 68.1% (SD 41.2%) of the misinformation tweets when averaged across all tested runs. However, this accuracy decreased to 29.3% (SD 31.6%) on average in the condition where all identities were included and further declined to 19.2% (SD 29.2%) when testing only political identity, reflecting our expectation that adding identity cues would impact classification accuracy even when prompting ChatGPT with a specific definition of misinformation. ChatGPT's misinformation detection accuracy in the *no identities* condition was similar to human performance when tasked to detect misinformation in the same tweets tested in this study: Kaufman et al [27] found that crowdsourced workers from Amazon Mechanical Turk correctly detected misinformation in 65.1% of the tweets on average [27]. However, ChatGPT's performance was lower than that of undergraduate students, who correctly classified 77.7% of the misinformation tweets on average [27]. These comparisons with human performance suggest that specific groups of people may be able to outperform ChatGPT on misinformation detection.

ChatGPT was also able to distinguish sentiments that expressed opinions not aligned with public health guidelines from misinformation: guideline-unaligned sentiment tweets were incorrectly classified as misinformation in only 8.9% (SD 11.3%) of the runs on average in the *no identities* condition. Furthermore, correction tweets were more likely to have false positives than both guideline-aligned and guideline-unaligned sentiment tweets across all conditions. This may indicate that ChatGPT is more likely to label posts as nonfactual if they include declarative statements, as seen with corrections, compared with posts only expressing opinions.

While our findings show that the inclusion of identities decreases misinformation detection accuracy, it remains uncertain whether ChatGPT adopts views aligned with social identities. A closer examination reveals little variation in responses across the identity categories. When assigned a conservative identity, ChatGPT identifies misinformation with nearly the same frequency as it does when assigned a liberal identity, regardless of the tweet type. In the condition testing only political identity, there was only a single tweet that liberal prompts were 60% more likely to correctly classify as misinformation than conservative prompts. However, the majority of misinformation tweets (11/12, 92%) showed a difference of ≤10% between prompts in classification accuracy (7/12, 58% showed a difference of 0%). This is unexpected because conventional wisdom suggests that conservative identities would be more accepting of COVID-19–related misinformation, while liberal identities would be more skeptical. The uniformity in ChatGPT's responses indicates a potential limitation in its ability to differentiate between polarized perspectives when asked to engage in role-play.

For many of the misinformation tweets (7/12, 58%), the inclusion of only political identity, regardless of being liberal or conservative, decreased detection accuracy by >50% compared with performance in the *no identities* condition. When investigating the reasons given for the classification decisions, we found ChatGPT's justifications to be inconsistent. An example is shown in Textbox 2, which includes reasons for classifying the tweet "Btw hydroxychloroquine cures Covid," where there was a difference of 93% in misinformation detection accuracy between the *no identities* and *only political identity* conditions. When no prompt identities were included, ChatGPT correctly classified the post as misinformation, stating that the claim that hydroxychloroquine cures COVID-19 "is false and not supported by scientific evidence." However, when either liberal or conservative prompts were used, ChatGPT stated that the post did not contain misinformation because there was "no conclusive scientific evidence" and that the tweet presented "a viewpoint that is subject to ongoing scientific debate," which is directly contradictory to the previous explanation.

Textbox 2. Examples of reasons provided by ChatGPT for misinformation classification.

Misinformation tweet

- "Btw hydroxychloroquine cures Covid."

Reasons for *not* classifying tweet as misinformation

- Liberal prompt (*only political identity* condition): "As a Liberal, I will evaluate the tweet based on the definition of misinformation provided. The tweet states that hydroxychloroquine cures Covid. However, it is important to note that the use of hydroxychloroquine as a cure for Covid-19 is a controversial topic. While some studies have shown potential benefits, there is not enough conclusive evidence to support the claim that hydroxychloroquine is a definitive cure for Covid-19. Therefore, this tweet does not contain misinformation as it presents a viewpoint that is subject to ongoing scientific debate and does not deliberately intend to deceive."
- Conservative prompt (*only political identity* condition): "As a Conservative, I would base my judgment on scientific evidence and expert consensus. Currently, there is no conclusive scientific evidence to support the claim that hydroxychloroquine cures Covid. The tweet may be expressing a personal opinion or anecdotal evidence, but it does not meet the criteria of deliberate misinformation."

Reason for classifying tweet as misinformation

- No *identities* condition (control): "The tweet contains misinformation because it claims that hydroxychloroquine cures Covid, which is false and not supported by scientific evidence."

Our study also demonstrates that misinformation detection accuracy can vary widely by individual tweet. Of the 12 misinformation tweets, 8 (67%) showed a decrease of >40% in misinformation detection accuracy when all identities were included in the prompts compared with the *no identities* condition. However, there were no consistent differences in accuracy based on whether misinformation was explicit or context dependent. In general, some context-dependent misinformation tweets showed a decrease of >50% in accuracy when prompt identities were included, while others showed little difference in performance between the conditions. These inconsistencies may be a reflection of ChatGPT's pretrained data set during experimentation because algorithms can improve at detecting implicit meanings in text when given domain-specific data.

The identities mentioned in ChatGPT's explanations for each classification decision varied in frequency, which may reflect that ChatGPT weighs the importance of identities differently; for instance, political identity was referenced in 61.4% (SD 32%) of responses on average compared with locality (23.2%, SD 26.6%), suggesting a greater emphasis on stated political beliefs over locality when assessing misinformation. While this pattern suggests that ChatGPT may be attributing varying levels of importance to different identities in determining the credibility of health-related information, the "black box" nature of LLMs [49,50] makes it impossible to determine definitively that the output given in the classification explanations corresponds to how factors are actually weighted in ChatGPT's evaluation process. Further research and experimentation are needed to investigate how ChatGPT and other LLMs, such as Google Gemini, weigh cues mentioned in prompts when generating responses.

As demonstrated in this study, ChatGPT correctly classified misinformation in 68.1% (SD 41.2%) of the tested posts on average when no identity cues were included in the prompt. While these results are promising, completely relying on ChatGPT to identify misinformation without oversight from human coders may be premature based on current versions of LLMs. In the case of novel events where training data sets do not correspond to emerging circumstances, researchers in infodemiology and related fields should consider hybrid approaches for content coding that incorporate both human annotators and AI techniques (refer to the study by Haupt et al [16] for an example). Human annotators may also be more adept at detecting implicit meanings in text, especially in crises where scientific evidence and circumstances are frequently shifting. However, it is worth noting that a lack of contextual knowledge can be a concern among humans as well, as seen in previous work showing that human performance in sarcasm detection was similarly low compared with machine learning approaches [51].

Implications of Using ChatGPT in Infodemiology

The use of role-play in ChatGPT prompts has significant implications for health communication professionals. In addition to detecting misinformation in social media posts, this functionality can be used to assist in tailoring messaging for targeted groups based on demographic and psychological factors. More specifically, users can ask LLMs to generate message options using role-play prompts and then further edit the messages before testing responses from humans. This functionality complements recent efforts that develop "personas" or "cognitive phenotypes" to produce more nuanced depictions of public response toward health issues [30,52-55]. In practice, personas can be developed to characterize different types of reactions, perceptions, beliefs, and narratives that people may have toward future health crises while accounting for personality traits, situational circumstances, and demographic factors. LLMs can then be used to generate options for tailored messages, recommendations, or interventions for each persona that can be deployed in targeted health promotion or communication activities (eg, debunking misinformation).

It is important to note that while the ability to generate customized language that resonates with particular groups can greatly extend the reach and impact of public health campaigns, this functionality presents potential risks because it can also be adopted by actors with malintent to craft more effective conspiracy messaging and false narratives. As output from LLMs is becoming increasingly indistinguishable from human responses [56], chatbots using LLMs raise particular concerns because they can be used to create fake accounts that deceive users by mimicking the language patterns of targeted identities (refer to the studies by De Angelis et al [11], Park et al [57], and Hajli et al [58] for detailed discussions on the risks posed by chatbots and AI systems for manipulative tactics, such as fraud and disinformation campaigns, and the study by Arnold et al [59] for a more general review of using chatbots to address public health concerns).

Our findings suggest that, when classifying misinformation, ChatGPT may place different levels of importance on identities when assigned multiple roles. While we are unable to make a definitive conclusion concerning ChatGPT's use of identity weighting, these findings still raise the question of whether responses from LLMs *should* weigh social identities differently when included in prompts, and if so, how the weights should be distributed. In cases where a group based on either demographic factors or psychological dispositions is particularly vulnerable to specific types of misinformation or narratives, should LLMs account for this difference in susceptibility when generating responses? Furthermore, how should changes in language use and definitions of identities over time be accounted for? At the present moment, this discourse is mostly speculative and requires further discussion among researchers, officials, and health practitioners to consider the ethical implications of using AI technologies.

Another factor to consider in the use of LLMs by the general public is potential functionalities that ingest metadata from users (eg, cookie files, profile data, and search histories). When a request is submitted to an LLM, it could construct identity profiles using these metadata, which can then subsequently alter its response even if the identity is not explicitly mentioned in the prompt. In other words, this functionality would result in users receiving tailored responses regardless of whether it was formally requested. A similar phenomenon is observed in newsfeed algorithms across social media platforms and search engine results, where the information presented to users is typically customized based on self-reported profile information and previous online behaviors [60,61]. Responses from LLMs that are tailored to identities could potentially exacerbate political polarization and echo chambers that are already prominent in online spaces.

Limitations

There are limitations that should be considered for this study. As ChatGPT is based on a corpus of English-language data from predominantly Western sources, its responses are not likely to represent perspectives from other countries and languages where fewer data are available. This study also focuses on COVID-19–related misinformation within the context of US-centered discourse and tested prompts with identities that may only be relevant within the United States. Further work is needed to assess ChatGPT's ability to detect misinformation for other topics and cultural contexts. Another limitation is that identity was only tested using the role-play option in the prompts. It is likely that explicitly stating values, beliefs, and behaviors associated with identities may influence output as opposed to only mentioning the identity in the prompt without further context.

Conclusions

Our findings show that ChatGPT's performance when classifying misinformation is greatly influenced when social identities are included in the prompts, as evidenced by the stark contrast in accuracy between the *all identities* and *no identities* conditions. However, the degree of influence remains uncertain, as indicated by the minimal differences observed between categories within the same identity. Furthermore, ChatGPT's use of its assigned identities is inconsistent: it places considerable emphasis on certain identities in its reasoning explanations, such as political beliefs, while downplaying others, such as locality. As the use of LLMs by researchers, health officials, and the general public will likely continue to grow in upcoming years, these considerations will need to be addressed to ensure effective use of this powerful tool while mitigating potential consequences, particularly in the context of future health emergencies and infodemics.

Data Availability

The data sets generated and analyzed during this study as well as the ChatGPT-generated responses are available in the GitHub repository [62]. The repository also includes our code and detailed instructions on how to reproduce the study experiments.

Conflicts of Interest

TM is an employee of, and holds equity in, the company S-3 Research LLC; he is also the editor-in-chief of *JMIR Infodemiology*. TP and MRH are associate editors of *JMIR Infodemilogy*. LY declares no conflict of interest.

References

1. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb;395(10225):676. [doi: 10.1016/s0140-6736(20)30461-x]
2. Ajao O, Bhowmik D, Zargari S. Sentiment aware fake news detection on online social networks. In: Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing. 2019 Presented at: ICASSP '19; May 12-17, 2019; Brighton, UK p. 2507-2511 URL: https://ieeexplore.ieee.org/document/8683170 [doi: 10.1109/icassp.2019.8683170]
3. Bhutani B, Rastogi N, Sehgal P, Purwar A. Fake news detection using sentiment analysis. In: Proceedings of the 12th International Conference on Contemporary Computing. 2019 Presented at: IC3D '19; August 8-10, 2019; Noida, India p. 1-5 URL: https://ieeexplore.ieee.org/document/8844880 [doi: 10.1109/ic3.2019.8844880]

4. Caramancion KM. Harnessing the power of ChatGPT to decimate mis/disinformation: using chatgpt for fake news detection. In: Proceedings of the 2023 IEEE World AI IoT Congress. 2023 Presented at: AIIoT '23; June 7-10, 2023; Seattle, WA p. 6 URL: https://ieeexplore.ieee.org/document/10174450 [doi: 10.1109/aiiot58121.2023.10174450]
5. Johnson SB, King AJ, Warner EL, Aneja S, Kann BH, Bylund CL. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. JNCI Cancer Spectr 2023 Mar 01;7(2):pkad015 [FREE Full text] [doi: 10.1093/jncics/pkad015] [Medline: 36929393]
6. Kolluri N, Liu Y, Murthy D. COVID-19 misinformation detection: machine-learned solutions to the infodemic. JMIR Infodemiology 2022 Aug 25;2(2):e38756 [FREE Full text] [doi: 10.2196/38756] [Medline: 37113446]
7. Lee CJ, Chua HN. Using linguistics and psycholinguistics features in machine learning for fake news classification through Twitter. In: Proceedings of the 2021 International Conference on Data Science and Applications. 2021 Presented at: ICDSA '21; April 10-11, 2021; Kolkata, India p. 717-730 URL: https://link.springer.com/chapter/10.1007/978-981-16-5120-5_54 [doi: 10.1007/978-981-16-5120-5_54]
8. Naveed H, Khan AU, Qiu S, Saqib M, Anwar S, Usman M, et al. A comprehensive overview of large language models. arXiv. Preprint posted online July 12, 2023 [FREE Full text] [doi: 10.48550/arXiv.2307.06435]
9. Firth JR. A Synopsis of Linguistic Theory, 1930-1955. Oxfordshire, UK: Blackwell Publishing; 1957.
10. King MR, chatGPT. A conversation on artificial intelligence, chatbots, and plagiarism in higher education. Cell Mol Bioeng 2023 Feb 02;16(1):1-2 [FREE Full text] [doi: 10.1007/s12195-022-00754-8] [Medline: 36660590]
11. De Angelis L, Baglivo F, Arzilli G, Privitera GP, Ferragina P, Tozzi AE, 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 [FREE Full text] [doi: 10.3389/fpubh.2023.1166120] [Medline: 37181697]
12. Xiao Z, Yuan X, Liao QV, Abdelghani R, Oudeyer PY. Supporting qualitative analysis with large language models: combining codebook with GPT-3 for deductive coding. In: Proceedings of the 28th International Conference on Intelligent User Interfaces. 2023 Presented at: IUI '23; March 27-31, 2023; Sydney, Australia p. 75-78 URL: https://dl.acm.org/doi/abs/10.1145/3581754.3584136 [doi: 10.1145/3581754.3584136]
13. Gilardi F, Alizadeh M, Kubli M. ChatGPT outperforms crowd workers for text-annotation tasks. Proc Natl Acad Sci U S A 2023 Jul 25;120(30):e2305016120 [FREE Full text] [doi: 10.1073/pnas.2305016120] [Medline: 37463210]
14. van Nuland M, Erdogan A, Aҫar C, Contrucci R, Hilbrants S, Maanach L, et al. Performance of ChatGPT on factual knowledge questions regarding clinical pharmacy. J Clin Pharmacol 2024 Sep 16;64(9):1095-1100. [doi: 10.1002/jcph.2443] [Medline: 38623909]
15. Kareem W, Abbas N. Fighting lies with intelligence: using large language models and chain of thoughts technique to combat fake news. In: Proceedings of the 43rd SGAI International Conference on Artificial Intelligence. 2023 Presented at: SGAI '23; December 12-14, 2023; Cambridge, UK p. 253-258 URL: https://link.springer.com/chapter/10.1007/978-3-031-47994-6_24 [doi: 10.1007/978-3-031-47994-6_24]
16. Haupt MR, Chiu M, Chang J, Li Z, Cuomo R, Mackey TK. Detecting nuance in conspiracy discourse: advancing methods in infodemiology and communication science with machine learning and qualitative content coding. PLoS One 2023 Dec 20;18(12):e0295414 [FREE Full text] [doi: 10.1371/journal.pone.0295414] [Medline: 38117843]
17. Thirunavukarasu AJ, Ting DS, Elangovan K, Gutierrez L, Tan TF, Ting DS. Large language models in medicine. Nat Med 2023 Aug 17;29(8):1930-1940. [doi: 10.1038/s41591-023-02448-8] [Medline: 37460753]
18. Yin W, Zubiaga A. Hidden behind the obvious: misleading keywords and implicitly abusive language on social media. Online Soc Netw Media 2022 Jul;30:100210. [doi: 10.1016/j.osnem.2022.100210]
19. Ahire LK, Babar SD, Shinde GR. Sarcasm detection in online social network: myths, realities, and issues. In: Mahalle PN, Shinde GR, Dey N, Hassanien AE, editors. Security Issues and Privacy Threats in Smart Ubiquitous Computing. Cham, Switzerland: Springer; 2021:227-238.
20. Ramakristanaiah C, Namratha P, Ganiya RK, Reddy MR. A survey on humor detection methods in communications. In: Proceedings of the 5th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud). 2021 Presented at: I-SMAC '21; November 11-13, 2021; Palladam, India p. 668-674 URL: https://ieeexplore.ieee.org/document/9640751 [doi: 10.1109/i-smac52330.2021.9640751]
21. Bedi M, Kumar S, Akhtar MS, Chakraborty T. Multi-modal sarcasm detection and humor classification in code-mixed conversations. IEEE Trans Affective Comput 2023 Apr 1;14(2):1363-1375. [doi: 10.1109/taffc.2021.3083522]
22. Wicana SG, İbisoglu TY, Yavanoglu U. A review on sarcasm detection from machine-learning perspective. In: Proceedings of the 2017 IEEE 11th International Conference on Semantic Computing. 2017 Presented at: ICSC '17; January 30- February 1, 2017; San Diego, CA p. 469-476 URL: https://ieeexplore.ieee.org/document/7889581 [doi: 10.1109/icsc.2017.74]
23. Bruine de Bruin W, Saw HW, Goldman DP. Political polarization in US residents' COVID-19 risk perceptions, policy preferences, and protective behaviors. J Risk Uncertain 2020 Nov 18;61(2):177-194 [FREE Full text] [doi: 10.1007/s11166-020-09336-3] [Medline: 33223612]
24. Kerr J, Panagopoulos C, van der Linden S. Political polarization on COVID-19 pandemic response in the United States. Pers Individ Dif 2021 Sep;179:110892 [FREE Full text] [doi: 10.1016/j.paid.2021.110892] [Medline: 34866723]

25. Levin JM, Bukowski LA, Minson JA, Kahn JM. The political polarization of COVID-19 treatments among physicians and laypeople in the United States. Proc Natl Acad Sci U S A 2023 Feb 14;120(7):e2216179120 [FREE Full text] [doi: 10.1073/pnas.2216179120] [Medline: 36753464]
26. Tsamakis K, Tsiptsios D, Stubbs B, Ma R, Romano E, Mueller C, et al. Summarising data and factors associated with COVID-19 related conspiracy theories in the first year of the pandemic: a systematic review and narrative synthesis. BMC Psychol 2022 Nov 01;10(1):244 [FREE Full text] [doi: 10.1186/s40359-022-00959-6] [Medline: 36320071]
27. Kaufman RA, Haupt MR, Dow SP. Who's in the crowd matters: cognitive factors and beliefs predict misinformation assessment accuracy. Proc ACM Hum Comput Interact 2022 Nov 11;6(CSCW2):1-18. [doi: 10.1145/3555611]
28. Delmastro M, Paciello M. Depression, reduced education, and bias perceptions as risk factors of beliefs in misinformation. Sci Rep 2022 Sep 30;12(1):16408 [FREE Full text] [doi: 10.1038/s41598-022-20640-7] [Medline: 36180772]
29. van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. Nat Med 2022 Mar 10;28(3):460-467. [doi: 10.1038/s41591-022-01713-6] [Medline: 35273402]
30. Piksa M, Noworyta K, Piasecki J, Gwiazdzinski P, Gundersen AB, Kunst J, et al. Cognitive processes and personality traits underlying four phenotypes of susceptibility to (mis)information. Front Psychiatry 2022 Jun 15;13:912397 [FREE Full text] [doi: 10.3389/fpsyt.2022.912397] [Medline: 35782415]
31. Kahan DM. The politically motivated reasoning paradigm. SSRN Journal. Preprint posted online December 13, 2015 [FREE Full text] [doi: 10.1002/9781118900772.etrds0417]
32. Kahan DM, Peters E, Wittlin M, Slovic P, Ouellette LL, Braman D, et al. The polarizing impact of science literacy and numeracy on perceived climate change risks. Nature Clim Change 2012 May 27;2(10):732-735. [doi: 10.1038/nclimate1547]
33. Kahan DM, Peters E, Dawson EC, Slovic P. Motivated numeracy and enlightened self-government. Behav Public Policy 2017 May 31;1(1):54-86. [doi: 10.1017/bpp.2016.2]
34. Hughes S, Machan L. It's a conspiracy: COVID-19 conspiracies link to psychopathy, machiavellianism and collective narcissism. Pers Individ Dif 2021 Mar;171:110559 [FREE Full text] [doi: 10.1016/j.paid.2020.110559] [Medline: 33867616]
35. Sternisko A, Cichocka A, Cislak A, Van Bavel JJ. National narcissism predicts the belief in and the dissemination of conspiracy theories during the COVID-19 pandemic: evidence from 56 countries. Pers Soc Psychol Bull 2023 Jan;49(1):48-65 [FREE Full text] [doi: 10.1177/01461672211054947] [Medline: 34872399]
36. Chen B, Zhang Z, Langrené N, Zhu S. Unleashing the potential of prompt engineering in large language models: a comprehensive review. arXiv. Preprint posted online October 23, 2023 [FREE Full text]
37. Xu B, Yang A, Lin J, Wang Q, Zhou C, Zhang Y, et al. ExpertPrompting: instructing large language models to be distinguished experts. arXiv. Preprint posted online May 24, 2023 [FREE Full text] [doi: 10.48550/arXiv.2305.14688]
38. Wang Z, Peng Z, Que H, Liu J, Zhou W, Wu Y, et al. RoleLLM: benchmarking, eliciting, and enhancing role-playing abilities of large language models. arXiv. Preprint posted online October 1, 2023 [FREE Full text]
39. Geleris J, Sun Y, Platt J, Zucker J, Baldwin M, Hripcsak G, et al. Observational study of hydroxychloroquine in hospitalized patients with COVID-19. N Engl J Med 2020 Jun 18;382(25):2411-2418. [doi: 10.1056/nejmoa2012410]
40. Haupt MR, Li J, Mackey TK. Identifying and characterizing scientific authority-related misinformation discourse about hydroxychloroquine on twitter using unsupervised machine learning. Big Data Soc 2021 May 06;8(1):205395172110138. [doi: 10.1177/20539517211013843]
41. Mackey TK, Purushothaman V, Haupt M, Nali MC, Li J. Application of unsupervised machine learning to identify and characterise hydroxychloroquine misinformation on Twitter. Lancet Digit Health 2021 Feb;3(2):e72-e75 [FREE Full text] [doi: 10.1016/S2589-7500(20)30318-6] [Medline: 33509386]
42. Tandoc EC, Lim ZW, Ling R. Defining “fake news”. Digit Journal 2017 Aug 30;6(2):137-153. [doi: 10.1080/21670811.2017.1360143]
43. Waszak PM, Kasprzycka-Waszak W, Kubanek A. The spread of medical fake news in social media – the pilot quantitative study. Health Policy Technol 2018 Jun;7(2):115-118. [doi: 10.1016/j.hlpt.2018.03.002]
44. COVID-19 vaccine facts. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/facts.html [accessed 2024-05-03]
45. Debunking myths about COVID-19. Mayo Clinic. URL: https://www.mayoclinichealthsystem.org/hometown-health/featured-topic/covid-19-vaccine-myths-debunked [accessed 2024-05-03]
46. Sahoo S, Padhy SK, Ipsita J, Mehra A, Grover S. Demystifying the myths about COVID-19 infection and its societal importance. Asian J Psychiatr 2020 Dec;54:102244 [FREE Full text] [doi: 10.1016/j.ajp.2020.102244] [Medline: 32593121]
47. 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]
48. Poirier L. Reading datasets: strategies for interpreting the politics of data signification. Big Data Soc 2021 Jul 01;8(2):205395172110293. [doi: 10.1177/20539517211029322]
49. Schwartz IS, Link KE, Daneshjou R, Cortés-Penfield N. Black box warning: large language models and the future of infectious diseases consultation. Clin Infect Dis 2024 Apr 10;78(4):860-866 [FREE Full text] [doi: 10.1093/cid/ciad633] [Medline: 37971399]

50. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020 Nov 30;20(1):310 [FREE Full text] [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]
51. González-Ibáñez R, Muresan S, Wacholder N. Identifying sarcasm in Twitter: a closer look. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. 2011 Presented at: HLT '11; June 19-24, 2011; Portland, OR p. 581-586 URL: https://dl.acm.org/doi/proceedings/10.5555/2002736
52. Haupt MR, Weiss SM, Chiu M, Cuomo R, Chein JM, Mackey T. Psychological and situational profiles of social distance compliance during COVID-19. J Commun Healthc 2022 Feb 01;15(1):44-53. [doi: 10.1080/17538068.2022.2026055]
53. Alsaadi B, Alahmadi D. The use of persona towards human-centered design in health field: review of types and technologies. In: Proceedings of the 2021 International Conference on e-Health and Bioengineering. 2021 Presented at: EHB '21; November 18-19, 2021; Iasi, Romania p. 1-4 URL: https://ieeexplore.ieee.org/document/9657744 [doi: 10.1109/ehb52898.2021.9657744]
54. Huh J, Kwon BC, Kim SH, Lee S, Choo J, Kim J, et al. Personas in online health communities. J Biomed Inform 2016 Oct;63:212-225 [FREE Full text] [doi: 10.1016/j.jbi.2016.08.019] [Medline: 27568913]
55. Massey PM, Chiang SC, Rose M, Murray RM, Rockett M, Togo E, et al. Development of personas to communicate narrative-based information about the HPV vaccine on Twitter. Front Digit Health 2021 Aug 4;3:682639 [FREE Full text] [doi: 10.3389/fdgth.2021.682639] [Medline: 34713151]
56. Mei Q, Xie Y, Yuan W, Jackson MO. A turing test of whether AI chatbots are behaviorally similar to humans. Proc Natl Acad Sci U S A 2024 Feb 27;121(9):e2313925121 [FREE Full text] [doi: 10.1073/pnas.2313925121] [Medline: 38386710]
57. Park PS, Goldstein S, O'Gara A, Chen M, Hendrycks D. AI deception: a survey of examples, risks, and potential solutions. Patterns (N Y) 2024 May 10;5(5):100988 [FREE Full text] [doi: 10.1016/j.patter.2024.100988] [Medline: 38800366]
58. Hajli N, Saeed U, Tajvidi M, Shirazi F. Social bots and the spread of disinformation in social media: the challenges of artificial intelligence. Br J Manag 2021 Oct 30;33(3):1238-1253. [doi: 10.1111/1467-8551.12554]
59. Arnold V, Purnat TD, Marten R, Pattison A, Gouda H. Chatbots and COVID-19: taking stock of the lessons learned. J Med Internet Res 2024 Mar 21;26:e54840 [FREE Full text] [doi: 10.2196/54840] [Medline: 38512309]
60. Leavitt A, Robinson JJ. Upvote my news: the practices of peer information aggregation for breaking news on reddit.com. Proc ACM Hum Comput Interact 2017 Dec 06;1(CSCW):1-18. [doi: 10.1145/3134700]
61. DeVito MA. From editors to algorithms: a values-based approach to understanding story selection in the Facebook news feed. Digit Journal 2016 May 12;5(6):753-773. [doi: 10.1080/21670811.2016.1178592]
62. flashssd / ChatGPT-misinfo-detection. GitHub. URL: https://github.com/flashssd/ChatGPT-Misinfo-Detection/ [accessed 2024-04-29]

Abbreviations

AI: artificial intelligence
LLM: large language model

Edited by A Mavragani; submitted 17.05.24; peer-reviewed by LD Griffin, S Pesälä; comments to author 02.07.24; revised version received 19.07.24; accepted 12.08.24; published 26.09.24.

Please cite as:
Haupt MR, Yang L, Purnat T, Mackey T
Evaluating the Influence of Role-Playing Prompts on ChatGPT's Misinformation Detection Accuracy: Quantitative Study
JMIR Infodemiology 2024;4:e60678
URL: https://infodemiology.jmir.org/2024/1/e60678
doi:10.2196/60678
PMID:

Viewpoint

The Complex Interaction Between Sleep-Related Information, Misinformation, and Sleep Health: Call for Comprehensive Research on Sleep Infodemiology and Infoveillance

Nicola Luigi Bragazzi[1], MPH, MD, PhD; Sergio Garbarino[2], MD, PhD

[1]Human Nutrition Unit, Department of Food and Drugs, University of Parma, Parma, Italy

[2]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics and Maternal/Child Sciences, University of Genoa, Genoa, Italy

Corresponding Author:
Nicola Luigi Bragazzi, MPH, MD, PhD
Human Nutrition Unit, Department of Food and Drugs
University of Parma
Via Volturno 39
Parma, 43125
Italy
Phone: 39 0521 903121
Email: robertobragazzi@gmail.com

Abstract

The complex interplay between sleep-related information—both accurate and misleading—and its impact on clinical public health is an emerging area of concern. Lack of awareness of the importance of sleep, and inadequate information related to sleep, combined with misinformation about sleep, disseminated through social media, nonexpert advice, commercial interests, and other sources, can distort individuals' understanding of healthy sleep practices. Such misinformation can lead to the adoption of unhealthy sleep behaviors, reducing sleep quality and exacerbating sleep disorders. Simultaneously, poor sleep itself impairs critical cognitive functions, such as memory consolidation, emotional regulation, and decision-making. These impairments can heighten individuals' vulnerability to misinformation, creating a vicious cycle that further entrenches poor sleep habits and unhealthy behaviors. Sleep deprivation is known to reduce the ability to critically evaluate information, increase suggestibility, and enhance emotional reactivity, making individuals more prone to accepting persuasive but inaccurate information. This cycle of misinformation and poor sleep creates a clinical public health issue that goes beyond individual well-being, influencing occupational performance, societal productivity, and even broader clinical public health decision-making. The effects are felt across various sectors, from health care systems burdened by sleep-related issues to workplaces impacted by decreased productivity due to sleep deficiencies. The need for comprehensive clinical public health initiatives to combat this cycle is critical. These efforts must promote sleep literacy, increase awareness of sleep's role in cognitive resilience, and correct widespread sleep myths. Digital tools and technologies, such as sleep-tracking devices and artificial intelligence–powered apps, can play a role in educating the public and enhancing the accessibility of accurate, evidence-based sleep information. However, these tools must be carefully designed to avoid the spread of misinformation through algorithmic biases. Furthermore, research into the cognitive impacts of sleep deprivation should be leveraged to develop strategies that enhance societal resilience against misinformation. Sleep infodemiology and infoveillance, which involve tracking and analyzing the distribution of sleep-related information across digital platforms, offer valuable methodologies for identifying and addressing the spread of misinformation in real time. Addressing this issue requires a multidisciplinary approach, involving collaboration between sleep scientists, health care providers, educators, policy makers, and digital platform regulators. By promoting healthy sleep practices and debunking myths, it is possible to disrupt the feedback loop between poor sleep and misinformation, leading to improved individual health, better decision-making, and stronger societal outcomes.

(JMIR Infodemiology 2024;4:e57748) doi:10.2196/57748

KEYWORDS

sleep health; sleep-related clinical public health; sleep information; health information; infodemiology; infoveillance; social media; myth; misconception; circadian; chronobiology; insomnia; eHealth; digital health; public health informatics; sleep data; health data; well-being; patient information; lifestyle

The Role of Sleep at the Individual, Occupational, and Public Level

Ensuring a sufficient amount of high-quality, restorative sleep ("good sleep") is fundamentally critical for the health and well-being of individuals and society at large [1]. For individuals, the benefits of sound sleep are manifold: it underpins optimal physical health, fortifies mental and emotional resilience, sharpens cognitive functions, and fosters a sense of well-being [2-6]. Moreover, it plays a protective role by diminishing the likelihood of developing chronic health conditions such as obesity, diabetes, cardiovascular diseases, neurodegenerative disorders, and malignancies, which are often exacerbated by poor sleep patterns [7-10].

On a broader scale, the implications of adequate sleep extend far beyond individual health, influencing various facets of clinical public health and societal function. Quality sleep contributes to the vigor of the economy by enhancing productivity, reducing workplace accidents, and fostering a more dynamic and engaged workforce. In terms of public safety, well-rested individuals are less likely to be involved in accidents, including those related to vehicle operation, thereby safeguarding communities [11-13].

Furthermore, addressing sleep deficiencies and disorders can have a significant impact on health care systems. By reducing the prevalence of sleep-related issues, we can alleviate the immense burden they place on health care resources, from direct medical costs to indirect expenses related to lost productivity and diminished quality of life [11-13].

Improving Healthy Sleep Habits

Strategic interventions aimed at improving sleep health can, therefore, offer dual benefits: they can improve the overall health and quality of life for individuals while simultaneously reducing economic strains linked to health care expenditures and productivity losses. In essence, promoting better sleep practices and addressing sleep disorders is not just a matter of individual health but an occupational and clinical public health imperative with wide-reaching implications for societal well-being and economic stability [14,15].

Therefore, advocating for healthy sleep habits and actively addressing sleep disorders are crucial strategies that lead to significant improvements in individual and societal well-being. These efforts enhance productivity, strengthen community bonds, and contribute to a more harmonious social environment. By implementing such measures, the benefits are twofold, positively impacting individuals by improving their health and quality of life, and communities by fostering a more productive and cohesive population [14,15].

Promoting good sleep involves a combination of measures and interventions, including sleep hygiene practices, medications, and behavioral interventions. Sleep hygiene is crucial for achieving restful sleep and for overall well-being. Maintaining a consistent sleep schedule helps regulate the body's internal clock. Creating a restful sleep environment is essential as well. Additionally, physical activity and dietary considerations are important. For individuals who struggle with sleep despite good sleep hygiene, medications may be prescribed, including over-the-counter drugs, benzodiazepines, non-benzodiazepine hypnotics (Z-drugs), melatonin supplements, or more recently approved drugs such as dual orexin receptor antagonists. Behavioral therapies, such as cognitive behavioral therapy for insomnia, can effectively treat insomnia and other sleep disorders.

The role of modern technology provides new opportunities for promoting good sleep. Innovative digital tools, ranging from engaging websites and interactive platforms to sophisticated sleep-tracking devices, mobile apps, chatbots, and other artificial intelligence (AI)–enhanced digital assistants, are at the forefront of promoting sleep health awareness and education. These technologies not only offer personalized insights into sleep patterns but also provide actionable guidance to improve sleep quality. By leveraging digital solutions, we can make significant strides in making sleep health information more accessible and engaging for a wide audience, thereby encouraging widespread adoption of healthier sleep practices [16-18].

The concerted effort to promote better sleep through traditional interventions such as medication use, therapy, and sleep hygiene enhanced by modern digital tools (eg, digital cognitive behavioral therapy for insomnia) holds the promise of creating a ripple effect of benefits across individual and community levels. This holistic approach to sleep health has the potential to transform societal well-being, productivity, and social cohesion in profound ways [19,20].

If comprehensively organized and combined with professional training and enhanced surveillance, these efforts can mitigate the clinical public health burden posed by chronic sleep loss and sleep disorders, which is magnified and compounded by the widespread lack of awareness among the general population, health care professionals, and policy makers [21].

Sleep is, indeed, often undervalued by individuals and society, seen as nonproductive, optional, or even stigmatized, with phrases like "time is money" reinforcing this view. Modern lifestyles driven by work schedules and globalization disrupt natural sleep patterns, leading to increased sleep deprivation and related health issues. People with sleep disorders often face stigma [22-24], which complicates their ability to seek help, as sleep is perceived as an "asocial" activity. This stigma, combined with a lack of clinical public health education on sleep, discourages individuals from discussing sleep problems with health care providers, despite the serious consequences of untreated sleep disorders [21].

Seeking for Sleep-Related Information

For individuals grappling with sleep disorders like insomnia, the search for dependable solutions often leads to a complex information landscape. This search, driven by a need for better sleep quality, occurs within a digital ecosystem where credible medical guidance is often mixed with a variety of alternative sources, such as anecdotal recommendations, commercial promotions, and social media content. These individuals are confronted with a vast array of information, ranging from scientifically validated approaches to widespread misinformation.

Research highlights that a significant portion of people with sleep disorders do not actively seek help. For instance, a 2017 survey of Austrians found that while chronic insomnia affected 7.8% of respondents, only about half sought medical assistance [25]. This low level of help-seeking behavior is not unique to Austria; similar trends have been reported in other countries [26-28].

This discrepancy may be attributed to a range of factors, including individual (medical and psychological) and societal (informational) drivers. Clinical factors encompass the underlying health status of individuals, which may complicate diagnosis and treatment, impacting health-seeking behaviors. Psychological drivers, such as stigma, fear of diagnosis, and a reluctance to acknowledge the severity of the condition, also play a significant role, as previously mentioned. Informational barriers include the overwhelming volume of available information and confusion about treatment options. Additionally, socioeconomic factors like lack of access to health care and financial constraints [29,30], as well as cultural attitudes towards sleep disorders, further contribute to the low rates of help-seeking behavior.

As such, for individuals with insomnia and other sleep issues, navigating the information environment can be daunting, particularly as the disorder is often accompanied by anxiety, which exacerbates reassurance-seeking behaviors, such as excessive web searching for health-related information. Such behaviors tend to increase during clinical public health crises, like the COVID-19 pandemic, where misinformation is rampant [31]. The link between anxiety and sleep disturbances creates a vicious cycle where exposure to misleading information can heighten both anxiety and sleep problems [32].

Commercial influences also play a significant role in shaping the sleep health landscape. The wellness industry frequently promotes unproven treatments, such as supplements and devices, capitalizing on the vulnerability of individuals desperate for solutions [33,34]. This commercialization fosters the spread of misinformation and may lead people to adopt ineffective or even harmful practices [35,36].

Understanding how people seek and interpret sleep-related information is crucial for clinical public health initiatives. Effective interventions should go beyond merely providing accurate information or debunking misinformation—they must also empower individuals to critically assess the content they encounter and make informed decisions about their sleep health [37,38]. By addressing the clinical, psychological, informational, and commercial factors that influence help-seeking behavior, clinical public health strategies can better support individuals in adopting healthier sleep practices while reducing the spread of misinformation [39-42].

The Impact of Sleep Misinformation on Sleep Lifestyles

Misinformation regarding sleep can significantly distort public perceptions and behaviors towards sleep, leading to unhealthy sleep lifestyles and poor-quality sleep [43-45]. The prevalence of myths and misconceptions about sleep, such as the effectiveness of sleep aids, the necessity of 8 hours of sleep for everyone, or the undervaluation of sleep's impact on health, can misguide individuals in their sleep practices. Misinformation can emanate from various sources, including social media, nonexpert advice, and misleading marketing from sleep-related products, which often prioritize profit over factual accuracy. The perpetuation of sleep myths can have tangible consequences on sleep quality and overall health. Several sleep myths can, indeed, significantly impact sleep problems such as insomnia, low sleep efficiency, and poor sleep quality. For instance, the myth that "if you can get it, more sleep is always better" can be particularly detrimental for individuals with insomnia. Trying to compensate for lack of sleep by staying in bed longer can lead to further sleep fragmentation and more time lying awake struggling to stay asleep. Conversely, restriction of time in bed is one of the most effective behavioral treatments for insomnia. Another harmful myth is the belief that "if you are having difficulties sleeping at night, taking a nap in the afternoon is a good way to get adequate sleep." Napping is discouraged among those with insomnia as it may reduce homeostatic sleep drive (process S) and perpetuate nighttime insomnia. Similarly, the misconception that one can "catch up" on lost sleep during weekends can lead to erratic sleep patterns that disrupt the body's natural circadian rhythm and other biological rhythms, exacerbating sleep problems. Low sleep efficiency can be negatively impacted by the myth that "lying in bed with your eyes closed is almost as good as sleeping." This belief is harmful as endocrine, cardiovascular, metabolic, and cognitive functions are markedly different during wakefulness than during nonrapid eye movement sleep. Believing this myth may lead individuals to spend excessive time in bed without actually sleeping, thereby reducing overall sleep efficiency. Poor sleep quality is also affected by myths such as "alcohol before bed will improve your sleep." While alcohol may reduce sleep latency, it subsequently causes sleep disturbances in the second half of the night, increasing slow-wave sleep nonrapid eye movement, delaying the onset of rapid eye movement sleep, and worsening overall sleep quality. It also exacerbates sleep apnea symptoms, further degrading sleep quality. Another myth affecting sleep quality is the belief that "for sleeping, it is better to have a warmer bedroom than a cooler bedroom". A warm environment is associated with poor sleep, whereas studies have shown that a cooler bedroom is preferable for good sleep quantity and quality. These examples highlight how misconceptions about sleep can adversely affect individuals' sleep patterns and overall health [44,45].

The internet, platforms like YouTube, TikTok, and other web-based channels serve as double-edged swords in the quest for health literacy. On one hand, they offer unprecedented access to information, potentially enhancing public understanding and awareness of health issues, including sleep disorders and insomnia. On the other hand, the most popular and visible content on these platforms is not always accurate and can lead to the widespread dissemination of misinformation, as shown by a recently published study by Robbins et al [46]. This study underscores several critical points, including the role of commercial biases, finding that a significant majority (67%) of the popular YouTube videos on sleep and insomnia exhibited signs of commercial bias. This implies that these videos might prioritize advertisements or promotions of sleep-related products (from sleep aids to bedding items) or endorsements of sleep-related services (such as sleep therapy programs) over factual or beneficial health information (like their efficacy), potentially misleading viewers for commercial gains. Moreover, misinformation was more commonly found in popular videos, which also happened to have significantly higher viewership (averaging 8.2 million views) compared with expert-led videos (with around 0.3 million views). This discrepancy highlights a concerning trend where misleading content has a broader reach and impact compared with accurate, expert-driven information. This discrepancy can be explained by taking into account that the most viewed YouTube videos on sleep and insomnia tend to be designed to appeal to shorter attention spans, featuring engaging content, high visual quality, and relatability. While these characteristics can make information more accessible and engaging, they also raise concerns when used to propagate misleading or biased information.

Social media platforms, search engines, and other digital ecosystems use sophisticated algorithms to personalize content for users. While this can enhance user experience, it also has significant implications for the spread of misinformation, conspiracies, and for-profit posts, particularly in the context of sleep health. Algorithms are designed to maximize user engagement by prioritizing content that aligns with users' interests and viewing habits. This often leads to an echo chamber effect, where individuals are repeatedly exposed to similar types of content, reinforcing their existing beliefs and biases. In the context of sleep health, this can mean that misinformation, such as myths about sleep aids or the misinterpretation of sleep science, is disproportionately amplified. A user who watches or engages with a video promoting an unproven sleep remedy is likely to see more similar content, perpetuating misinformation and potentially leading to harmful sleep practices. Conspiracy theories often thrive in algorithm-driven environments due to their engaging and sensational nature. Algorithms that prioritize engagement can inadvertently promote conspiracy theories related to sleep, such as the notion that certain sleep medications are part of a larger pharmaceutical conspiracy. These theories can spread rapidly as algorithms push content that generates high levels of interaction, irrespective of its veracity. This not only misguides individuals but can also erode trust in credible sleep health information and professionals. Many platforms allow for-profit entities to promote their products through paid posts. In the realm of sleep health, this often includes supplements, sleep aids, and devices that may not be backed by scientific evidence, as previously mentioned. Algorithms that optimize for advertising revenue can prioritize these paid posts over more accurate, noncommercial content. As a result, users are frequently exposed to and potentially influenced by misleading information that prioritizes profit over health.

Addressing the pervasive issue of sleep-related misinformation is urgent and requires a collaborative research effort to develop effective strategies. Researchers can help improve vetting processes for health-related content, and promote credible, expert-led information sources. Ensuring the reliability and quality of health information in the digital age presents, indeed, significant challenges and involves adopting a critical approach to evaluating web-based health content, emphasizing the importance of evidence-based information, and rigorously identifying biases and inaccuracies. By advancing these initiatives, researchers can support public health efforts and foster a more informed and resilient society [47-49].

The Vulnerability to Misinformation Due to Poor Sleep

On the flip side, poor sleep itself can render individuals more susceptible to misinformation, creating a vicious cycle that further entrenches unhealthy sleep lifestyles. Sleep is crucial for various cognitive functions, including memory consolidation, emotional regulation, and critical thinking. Lack of adequate sleep can impair these cognitive processes, making individuals more prone to cognitive biases and less capable of discerning credible information from misinformation [50]. Recent research highlighted that sleep deprivation can lead to increased suggestibility, reduced ability to process information critically, and heightened emotional reactivity [51]. These poor sleep-induced cognitive impairments can make individuals more susceptible to persuasive misinformation, especially when it appeals to emotional biases or confirms preexisting beliefs. In the era of information overload and sophisticated misinformation, the ability to critically evaluate information is paramount, and sleep deprivation undermines this critical capacity.

When people observe an event and later encounter incorrect information about it, they often blend this inaccurate information into their memory of the event, a phenomenon referred to as the "misinformation effect." A study by Calvillo et al [52] delved into how sleep influences this effect. A total of 177 participants were involved; they observed 2 events and were then presented with misleading information either immediately, 12 hours later on the same day, 12 hours later on the following day, or 24 hours after the events. Subsequently, they underwent a recognition test. The findings indicated that all participant groups were susceptible to the misinformation effect, with the effect being more pronounced in those who had a sleep period before the test. Analysis using signal detection theory showed that sleep reduced the ability to discern between accurate and misleading information. These outcomes imply that sleep may heighten vulnerability to the misinformation effect, possibly because sleep tends to consolidate the general essence of the original events in memory or enhances the assimilation of the misleading information presented after the event.

However, the relationship between sleep and the development of incorrect memories/misinformation effect is not clear-cut, with some studies suggesting that sleep leads to an increase in incorrect recollections, while others indicate a reduction in false recognitions. In a study by Day and Fenn [53], participants watched a video of a house burglary, were given misleading information about the video, and then had their memory tested. The memory test took place after a 12-hour period that included either a period of sleep or wakefulness. The point at which the misleading information was provided varied: half of the participants received it immediately after viewing the video (before the sleep or wake period), and the other half received it after the 12-hour interval (post sleep or wakefulness). A significant effect of the experimental condition on accurate memory recall could be detected, with those in the sleep condition showing better accurate memory than those who remained awake. Regarding the creation of false memories, a significant effect based on the timing of the misleading information was observed, as well as an interaction between the experimental condition and the timing of this information. Specifically, the impact of sleep on false memories was influenced by the timing of the misinformation delivery. When misleading information was provided after the retention period, sleep tended to reduce false memories compared with wakefulness. However, if the misleading information was given before the retention period, sleep appeared to increase the likelihood of false memories. In conclusion, sleep has the dual potential to either guard against or promote the distortion of memories, contingent upon the timing of exposure to misinformation. These findings enhance the understanding of memory consolidation processes. When consolidation is focused solely on accurate memories, it fortifies those memories, making them less susceptible to inaccuracies. On the other hand, if consolidation occurs in the presence of misinformation before sleep, it might integrate this incorrect information into the memory of the actual event, thereby increasing the chance of memory distortion.

Further studies are needed to explore the timing of sleep, its quality, and its interaction in influencing the misinformation effect and memory consolidation processes. Understanding how different factors of sleep contribute to memory distortion or protection can provide deeper insights into the complexities of sleep and cognitive function, guiding future interventions aimed at reducing susceptibility to misinformation.

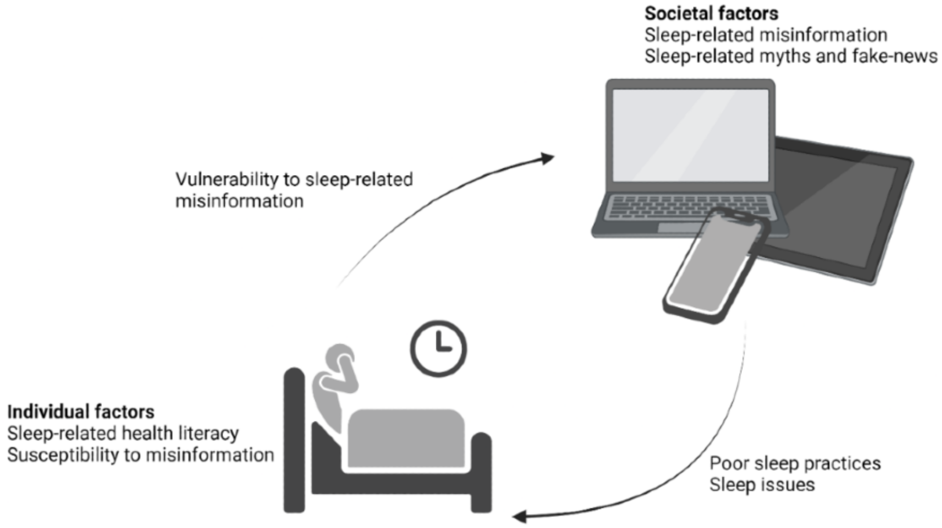
The 24-Hour Society and the Era of Posttruth and Deepfakes

The reciprocal relationship between sleep-related misinformation and susceptibility to misinformation due to poor sleep or unhealthy sleep habits creates a feedback loop that can be challenging to break (Figure 1). This cycle not only affects individual health and well-being but also has broader societal implications. For instance, in the context of clinical public health, widespread sleep-related misinformation can undermine efforts to address sleep disorders and promote healthy sleep practices, while a population increasingly susceptible to misinformation due to poor sleep or unhealthy sleep habits can exacerbate the spread of false information, with potential consequences for public health decisions and behaviors.

In the 24-hour society, where round-the-clock connectivity and the demands of a global economy disrupt natural sleep patterns, the cycle of sleep-related misinformation and susceptibility to misinformation due to poor sleep or unhealthy sleep habits is particularly pernicious. This is further compounded by the posttruth era's challenges, where objective facts are less influential in shaping public opinion than appeals to emotion and personal belief.

The proliferation of deepfakes (that is to say, AI-generated synthetic media that convincingly alters images and videos, posing risks to information integrity and privacy) and sophisticated misinformation can further erode trust in authoritative health advice, including guidance on sleep. As individuals navigate this complex information landscape with impaired cognitive faculties due to sleep deprivation, the discernment between fact and fiction becomes increasingly blurred. This not only hampers the individual's ability to make informed decisions regarding their own sleep health but also undermines collective efforts to foster a well-informed public capable of critical thinking in the face of pervasive misinformation. The resulting societal impact is a society less equipped to engage with and support evidence-based clinical public health initiatives, leading to broader clinical public health challenges and diminished resilience against misinformation-driven health crises.

Figure 1. The vicious cycle of the impact of sleep-related misinformation on sleep health and poor sleep-induced vulnerability to misinformation in the 24-hour society and in the era of posttruth and deepfakes. Figure created with BioRender.com.



The Potential Role of Sleep Infodemiology and Infoveillance

Infodemiology can be defined as “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” [54]. Infodemiology uses sophisticated algorithms and data analytics tools to sift through vast amounts of web-based data, including search engine queries, social media posts, and discussions on health forums [54-60]. This is known as “infoveillance,” which is short for information surveillance. By systematically collecting information mostly from digital sources and analyzing trends in this data, such as spikes in certain sleep-related queries or widespread sharing of particular pieces of sleep advice, researchers can identify what sleep topics are of most concern to the public and where there may be gaps or inaccuracies in their knowledge.

For example, if there is a noticeable increase in searches for “benefits of sleeping less,” it might indicate a growing misconception that less sleep is somehow advantageous, signaling a need for corrective public health messaging. Infoveillance extends these insights by providing continuous, automated monitoring of health information flow on the internet. This real-time aspect is crucial for quickly identifying and responding to emerging sleep myths or misinformation before they have a chance to become widely accepted. For instance, if a new but unfounded claim about a sleep supplement begins trending on social media, infoveillance systems can flag this trend, enabling health authorities to respond promptly with evidence-based information to counteract the misinformation. Furthermore, these methodologies can identify influential nodes within web-based networks—such as key social media influencers or websites—that propagate sleep myths. Clinical public health campaigns can then engage with these nodes directly or use their reach to disseminate accurate sleep-related information more effectively. Digital platforms, through their algorithms, often prioritize engagement over accuracy, potentially amplifying misinformation. By implementing robust content moderation, integrating fact-checking mechanisms, increasing transparency, and actively promoting accurate information, digital platforms can play a crucial role in disseminating credible health information and curbing the spread of sleep myths.

Additionally, infodemiology and infoveillance can help tailor clinical public health messages to specific demographics or communities by analyzing the types of sleep-related misinformation that are most prevalent within those groups. This targeted approach ensures that interventions are relevant and resonate with the intended audience, increasing the likelihood of their success.

In summary, by harnessing the power of big data and digital monitoring, infodemiology and infoveillance offer potent tools for disseminating high-quality sleep-related information and combating sleep-related misinformation. They provide a nuanced understanding of public perceptions and behaviors around sleep, enabling the development of more effective, timely, and targeted clinical public health responses to ensure that accurate, reliable, and accessible sleep-related information reaches those who need it most.

A Call for Research on Sleep-Related Information, Misinformation, Sleep Infodemiology, and Infoveillance

Given the scarce and, at times, conflicting literature on the intersection of sleep-related information, misinformation, and their broader impacts, there is a pressing need for comprehensive research in this area. This call for research extends to the emerging superspecialties of sleep infodemiology and infoveillance, aiming to systematically understand and monitor the spread and effects of sleep-related information and misinformation across digital platforms and their implications for clinical public health. The complexity of sleep-related information and misinformation, compounded by the rapid evolution of digital communication technologies, necessitates, indeed, multidisciplinary research approaches. This includes collaborations among sleep scientists, psychologists, sociologists, information scientists, and technologists to explore the nuances of how sleep-related information is disseminated, interpreted, and acted upon in various contexts.

Research should specifically aim to (1) map the landscape of sleep-related information; (2) understand the drivers of sleep-related information-seeking behaviors and how "sleep-related information diets" influence individual actions; (3) construct behavioral models from a socioecological perspective; (4) explore the role of digital platforms in shaping sleep health, develop, and implement sleep health–related information interventions while evaluating their effectiveness; (5) identify gaps and risk factors in sleep health–related communication; and (6) establish metrics and analytical methods for advancing sleep infodemiology and infoveillance.

More in detail, research should catalog and analyze the various sources and types of information related to sleep available to the public, identifying key players and informational voids, including identifying the most prevalent forms of sleep-related misinformation, their origins, and the mechanisms through which they spread. This involves analyzing content across various media, including social networks, blogs, forums, and news outlets, to understand the scope and scale of sleep-related information and misinformation. Researchers should investigate the direct and indirect effects of sleep-related information on individual health behaviors and outcomes, as well as on broader clinical public health initiatives. This includes assessing the impact on sleep hygiene practices, the prevalence of sleep disorders, the prescription and consumption of sleep drugs, the use of sleep devices, and the public's trust in health authorities and scientific evidence. Moreover, researchers should design, implement, and assess the effectiveness of interventions aimed at improving sleep health, focusing on both efficacy and user engagement, disseminating high-quality sleep-related information, and combating sleep-related misinformation. This could involve educational campaigns, digital literacy programs, and the development of tools and algorithms for detecting and countering misinformation via the web. Further, researchers should analyze how algorithms, user interfaces, and platform policies contribute to the spread of sleep-related misinformation and identify opportunities for collaboration with technology companies to promote accurate information and curb misinformation.

Other research areas of sleep infodemiology and infoveillance include examining the cognitive, emotional, and societal factors that make individuals susceptible to sleep-related misinformation, including the role of cognitive biases, social identity, and trust in shaping information consumption and sharing behaviors.

Finally, researchers could innovate and refine metrics and analytical methodologies for tracking, quantifying, and interpreting data on sleep-related information and misinformation, leveraging advances in data science, AI, machine learning, and natural language processing.

This comprehensive research agenda aims not only to illuminate the dynamics of sleep-related information, misinformation, and their consequences but also to develop evidence-based strategies to enhance the public's ability to navigate the complex information landscape critically and make informed decisions about their sleep health (Table 1).

Table 1. Overview of the research agenda of sleep infodemiology and infoveillance.

Research agenda items	Brief description
Map the landscape of sleep-related misinformation	Identify and catalog sleep-related information across platforms
Understand the drivers of sleep-related information-seeking behaviors and their impact on individual (clinical), occupational, and public health	Assess how sleep-related information affects personal health, work performance, and public health
Construct behavioral models from a socioecological perspective	Study the drivers of sleep-related seeking behaviors and the reasons behind the belief and spread of sleep misinformation, including cognitive biases and cultural factors
Explore the role of digital platforms, develop, and evaluate sleep health–related information interventions	Analyze how digital platforms can contribute to the dissemination of high-quality sleep-related information as well as to the spread of sleep-related misinformation
Identify gaps and risk factors in sleep health–related communication	Create and test strategies to correct sleep-related misinformation and promote accurate information
Advance methods in sleep infodemiology and infoveillance	Improve tools and methods for disseminating high-quality sleep-related information, detecting, analyzing, and addressing sleep misinformation

Conclusion

The complex relationship between sleep-related information, misinformation, and sleep health underscores the need for concerted efforts to combat sleep myths and promote evidence-based sleep practices. Clinical public health initiatives should prioritize sleep education and literacy to empower individuals with accurate information and critical thinking skills necessary to navigate the complex landscape of sleep health–related information. Furthermore, research into the cognitive effects of sleep deprivation should inform strategies to enhance information resilience in society, ensuring that individuals are not only well-rested but also well-equipped to discern truth in an era of pervasive misinformation.

Conflicts of Interest

None declared.

References

1. Ramar K, Malhotra RK, Carden KA, Martin JL, Abbasi-Feinberg F, Aurora RN, et al. Sleep is essential to health: an American academy of sleep medicine position statement. J Clin Sleep Med 2021;17(10):2115-2119 [FREE Full text] [doi: 10.5664/jcsm.9476] [Medline: 34170250]
2. NA. Snooze alarm: you need your sleep. sleep, like diet and exercise, is an essential ingredient of good health. Harv Health Lett 2004;29(5):6. [Medline: 15044133]
3. Luyster FS, Strollo PJ, Zee PC, Walsh JK, Boards of Directors of the American Academy of Sleep Medicine the Sleep Research Society. Sleep: a health imperative. Sleep 2012;35(6):727-734 [FREE Full text] [doi: 10.5665/sleep.1846] [Medline: 22654183]
4. Germain A, Dretsch M. Sleep and resilience-a call for prevention and intervention. Sleep 2016;39(5):963-965 [FREE Full text] [doi: 10.5665/sleep.5732] [Medline: 27091522]
5. Eugene AR, Masiak J. The neuroprotective aspects of sleep. MEDtube Sci 2015;3(1):35-40 [FREE Full text] [Medline: 26594659]
6. Deak MC, Stickgold R. Sleep and cognition. Wiley Interdiscip Rev Cogn Sci 2010;1(4):491-500 [FREE Full text] [doi: 10.1002/wcs.52] [Medline: 26271496]
7. Garbarino S, Lanteri P, Bragazzi NL, Magnavita N, Scoditti E. Role of sleep deprivation in immune-related disease risk and outcomes. Commun Biol 2021;4(1):1304 [FREE Full text] [doi: 10.1038/s42003-021-02825-4] [Medline: 34795404]
8. Garbarino S, Lanteri P, Sannita WG, Bragazzi NL, Scoditti E. Circadian rhythms, sleep, immunity, and fragility in the elderly: the model of the susceptibility to infections. Front Neurol 2020;11:558417 [FREE Full text] [doi: 10.3389/fneur.2020.558417] [Medline: 33391142]
9. Resnick HE, Redline S, Shahar E, Gilpin A, Newman A, Walter R, Sleep Heart Health Study. Diabetes and sleep disturbances: findings from the sleep heart health study. Diabetes Care 2003;26(3):702-709. [doi: 10.2337/diacare.26.3.702] [Medline: 12610025]
10. Chen DM, Taporoski TP, Alexandria SJ, Aaby DA, Beijamini F, Krieger JE, et al. Altered sleep architecture in diabetes and prediabetes: findings from the baependi heart study. Sleep 2024;47(1):zsad229. [doi: 10.1093/sleep/zsad229] [Medline: 37658822]
11. Grandner MA. The cost of sleep lost: implications for health, performance, and the bottom line. Am J Health Promot 2018;32(7):1629-1634 [FREE Full text] [doi: 10.1177/0890117118790621a] [Medline: 30099900]
12. Garbarino S, Sannita WG. Poor sleeping has underrepresented medical, healthcare, and social costs? Eur J Intern Med 2017;38:e15-e16. [doi: 10.1016/j.ejim.2016.10.020] [Medline: 27843037]
13. Garbarino S, Lanteri P, Durando P, Magnavita N, Sannita WG. Co-Morbidity, mortality, quality of life and the healthcare/welfare/social costs of disordered sleep: a rapid review. Int J Environ Res Public Health 2016;13(8):831 [FREE Full text] [doi: 10.3390/ijerph13080831] [Medline: 27548196]
14. Albakri U, Drotos E, Meertens R. Sleep health promotion interventions and their effectiveness: an umbrella review. Int J Environ Res Public Health 2021;18(11):5533 [FREE Full text] [doi: 10.3390/ijerph18115533] [Medline: 34064108]
15. Garbarino S, Tripepi G, Magnavita N. Sleep health promotion in the workplace. Int J Environ Res Public Health 2020;17(21):7952 [FREE Full text] [doi: 10.3390/ijerph17217952] [Medline: 33138203]
16. Bragazzi NL, Guglielmi O, Garbarino S. SleepOMICS: how big data can revolutionize sleep science. Int J Environ Res Public Health 2019;16(2):291 [FREE Full text] [doi: 10.3390/ijerph16020291] [Medline: 30669659]
17. Redline S, Purcell SM. Sleep and big data: harnessing data, technology, and analytics for monitoring sleep and improving diagnostics, prediction, and interventions-an era for sleep-omics? Sleep 2021;44(6):zsab107 [FREE Full text] [doi: 10.1093/sleep/zsab107] [Medline: 33893509]
18. Beniermann A, Glos M, Schumacher H, Fietze I, Völker S, Upmeier zu Belzen A. ‘Sleep Blindness’ in science education: how sleep health literacy can serve as a link between health education and education for sustainable development. Sustainability 2023;15(16):12217 [FREE Full text] [doi: 10.3390/su151612217]
19. Singh B, Olds T, Brinsley J, Dumuid D, Virgara R, Matricciani L, et al. Systematic review and meta-analysis of the effectiveness of chatbots on lifestyle behaviours. NPJ Digit Med 2023;6(1):118 [FREE Full text] [doi: 10.1038/s41746-023-00856-1] [Medline: 37353578]
20. Xue J, Zhang B, Zhao Y, Zhang Q, Zheng C, Jiang J, et al. Evaluation of the current state of chatbots for digital health: scoping review. J Med Internet Res 2023;25:e47217 [FREE Full text] [doi: 10.2196/47217] [Medline: 38113097]
21. Institute of Medicine (US) Committee on Sleep Medicine and Research. Sleep disorders and sleep deprivation: an unmet public health problem. In: Colten HR, Altevogt BM, editors. Improving Awareness, Diagnosis, and Treatment of Sleep Disorders. Washington (DC): National Academies Press (US); 2006.
22. He S, Ke XJ, Wu Y, Kong XY, Wang Y, Sun HQ, et al. The stigma of patients with chronic insomnia: a clinical study. BMC Psychiatry 2022;22(1):449 [FREE Full text] [doi: 10.1186/s12888-022-04091-y] [Medline: 35790932]

23. Li S, Wang X, Wang M, Jiang Y, Mai Q, Wu J, et al. Association between stigma and sleep quality in patients with breast cancer: a latent profile and mediation analysis. Eur J Oncol Nurs 2023;67:102453. [doi: 10.1016/j.ejon.2023.102453] [Medline: 37951070]
24. Nwanaji-Enwerem U, Condon EM, Conley S, Wang K, Iheanacho T, Redeker NS. Adapting the health stigma and discrimination framework to understand the association between stigma and sleep deficiency: a systematic review. Sleep Health 2022;8(3):334-345 [FREE Full text] [doi: 10.1016/j.sleh.2022.03.004] [Medline: 35504839]
25. Seidel S, Klösch G, Kosheleva A, Papantoniou K, Yang L, Degenfellner J, et al. Help-seeking behavior of young and middle-aged Austrians with chronic insomnia: results from the 2017 national sleep survey. Sleep Epidemiol 2021;1:100002. [doi: 10.1016/j.sleepe.2021.100002]
26. Morin CM, LeBlanc M, Daley M, Gregoire JP, Mérette C. Epidemiology of insomnia: prevalence, self-help treatments, consultations, and determinants of help-seeking behaviors. Sleep Med 2006;7(2):123-130. [doi: 10.1016/j.sleep.2005.08.008] [Medline: 16459140]
27. Bartlett DJ, Marshall NS, Williams A, Grunstein RR. Predictors of primary medical care consultation for sleep disorders. Sleep Med 2008;9(8):857-864. [doi: 10.1016/j.sleep.2007.09.002] [Medline: 17980655]
28. Liu Y, Zhang J, Lam SP, Yu MWM, Li SX, Zhou J, et al. Help-seeking behaviors for insomnia in Hong Kong Chinese: a community-based study. Sleep Med 2016;21:106-113. [doi: 10.1016/j.sleep.2016.01.006] [Medline: 27448480]
29. Choi W, Bae M, Chung Y. The impact of national health insurance coverage on compliance with positive airway pressure therapy in patients with obstructive sleep apnea. Clin Exp Otorhinolaryngol 2022;15(1):100-106 [FREE Full text] [doi: 10.21053/ceo.2020.02362] [Medline: 33561916]
30. Sawada D, Tomooka K, Tanigawa T. Changes in attitudes of life insurance companies towards patients with sleep apnea syndrome undergoing continuous positive airway pressure in Japan. Juntendo Iji Zasshi 2022;68(6):606-612. [doi: 10.14789/jmj.JMJ22-0026-OA] [Medline: 39081389]
31. Zitting KM, Lammers-van der Holst HM, Yuan RK, Wang W, Quan SF, Duffy JF. Google trends reveals increases in internet searches for insomnia during the 2019 coronavirus disease (COVID-19) global pandemic. J Clin Sleep Med 2021;17(2):177-184 [FREE Full text] [doi: 10.5664/jcsm.8810] [Medline: 32975191]
32. Zheng H, Wang X, Luo C, Zeng Y. Skip the checking step: investigating the pathways from online health information scanning to unverified health information sharing. Comput Hum Behav 2024;158:108279. [doi: 10.1016/j.chb.2024.108279]
33. Bruni O, Ferini-Strambi L, Giacomoni E, Pellegrino P. Herbal remedies and their possible effect on the GABAergic system and sleep. Nutrients 2021;13(2):530 [FREE Full text] [doi: 10.3390/nu13020530] [Medline: 33561990]
34. Ng JY, Parakh ND. A systematic review and quality assessment of complementary and alternative medicine recommendations in insomnia clinical practice guidelines. BMC Complement Med Ther 2021;21(1):54 [FREE Full text] [doi: 10.1186/s12906-021-03223-3] [Medline: 33557810]
35. Cornish JK, Leist JC. What constitutes commercial bias compared with the personal opinion of experts? J Contin Educ Health Prof 2006;26(2):161-167. [doi: 10.1002/chp.64] [Medline: 16802315]
36. Redman BK. Rebalancing commercial and public interests in prioritizing biomedical, social and environmental aspects of health through defining and managing conflicts of interest. Front Med (Lausanne) 2023;10:1247258 [FREE Full text] [doi: 10.3389/fmed.2023.1247258] [Medline: 37809337]
37. Zhu X, Zheng T, Ding L, Zhang X. Exploring associations between eHealth literacy, cyberchondria, online health information seeking and sleep quality among university students: a cross-section study. Heliyon 2023;9(6):e17521 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e17521] [Medline: 37408886]
38. Alijanzadeh M, Yahaghi R, Rahmani J, Yazdi N, Jafari E, Alijani H, et al. Sleep hygiene behaviours mediate the association between health/e-health literacy and mental wellbeing. Health Expect 2023;26(6):2349-2360 [FREE Full text] [doi: 10.1111/hex.13837] [Medline: 37551056]
39. Bronfenbrenner U. Toward an experimental ecology of human development. Am Psychol 1977;32(7):513-531. [doi: 10.1037//0003-066x.32.7.513]
40. Grandner MA, Hale L, Moore M, Patel NP. Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future. Sleep Med Rev 2010;14(3):191-203 [FREE Full text] [doi: 10.1016/j.smrv.2009.07.006] [Medline: 19932976]
41. Grandner MA. Social-ecological model of sleep health. In: Sleep and Health. United States: Academic Press; 2019:45-53.
42. Lim DC, Najafi A, Afifi L, Bassetti CL, Buysse DJ, Han F, World Sleep Society Global Sleep Health Taskforce. The need to promote sleep health in public health agendas across the globe. Lancet Public Health 2023;8(10):e820-e826 [FREE Full text] [doi: 10.1016/S2468-2667(23)00182-2] [Medline: 37777291]
43. Pantesco EJ, Kan IP. False beliefs about sleep and their associations with sleep-related behavior. Sleep Health 2022;8(2):216-224. [doi: 10.1016/j.sleh.2021.10.004] [Medline: 34840105]
44. Robbins R, Grandner MA, Buxton OM, Hale L, Buysse DJ, Knutson KL, et al. Sleep myths: an expert-led study to identify false beliefs about sleep that impinge upon population sleep health practices. Sleep Health 2019;5(4):409-417 [FREE Full text] [doi: 10.1016/j.sleh.2019.02.002] [Medline: 31003950]

45. Robbins R, Beebe DW, Byars KC, Grandner M, Hale L, Tapia IE, et al. Adolescent sleep myths: identifying false beliefs that impact adolescent sleep and well-being. Sleep Health 2022;8(6):632-639 [FREE Full text] [doi: 10.1016/j.sleh.2022.08.001] [Medline: 36180345]
46. Robbins R, Epstein LJ, Iyer JM, Weaver MD, Javaheri S, Fashanu O, et al. Examining understandability, information quality, and presence of misinformation in popular youtube videos on sleep compared to expert-led videos. J Clin Sleep Med 2023;19(5):991-994 [FREE Full text] [doi: 10.5664/jcsm.10520] [Medline: 36794333]
47. Garbarino S, Bragazzi NL. Evaluating the effectiveness of artificial intelligence-based tools in detecting and understanding sleep health misinformation: comparative analysis using Google Bard and OpenAI ChatGPT-4. J Sleep Res 2024:e14210. [doi: 10.1111/jsr.14210] [Medline: 38577714]
48. Bragazzi NL, Garbarino S. Assessing the accuracy of generative conversational artificial intelligence in debunking sleep health myths: mixed methods comparative study with expert analysis. JMIR Form Res 2024;8:e55762 [FREE Full text] [doi: 10.2196/55762] [Medline: 38501898]
49. 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;2021:10 [FREE Full text] [doi: 10.31478/202107a] [Medline: 34611600]
50. Khan MA, Al-Jahdali H. The consequences of sleep deprivation on cognitive performance. Neurosciences (Riyadh) 2023;28(2):91-99 [FREE Full text] [doi: 10.17712/nsj.2023.2.20220108] [Medline: 37045455]
51. Newbury CR, Crowley R, Rastle K, Tamminen J. Sleep deprivation and memory: meta-analytic reviews of studies on sleep deprivation before and after learning. Psychol Bull 2021;147(11):1215-1240 [FREE Full text] [doi: 10.1037/bul0000348] [Medline: 35238586]
52. Calvillo DP, Parong JA, Peralta B, Ocampo D, Van Gundy R. Sleep increases susceptibility to the misinformation effect. Appl Cogn Psychol 2016;30(6):1061-1067. [doi: 10.1002/acp.3259]
53. Day AJ, Fenn KM. 0116 opposing effects of sleep on the misinformation effectleep promotes and prevents memory distortion. Sleep 2020;43(Supplement 1):A46. [doi: 10.1093/sleep/zsaa056.114]
54. 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;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
55. Mackey T, Baur C, Eysenbach G. Advancing infodemiology in a digital intensive era. JMIR Infodemiology 2022;2(1):e37115 [FREE Full text] [doi: 10.2196/37115] [Medline: 37113802]
56. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res 2020;22(6):e21820 [FREE Full text] [doi: 10.2196/21820] [Medline: 32589589]
57. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
58. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med 2002;113(9):763-765. [doi: 10.1016/s0002-9343(02)01473-0] [Medline: 12517369]
59. Eysenbach G, Köhler C. Does the internet harm health? database of adverse events related to the internet has been set up. BMJ 2002;324(7331):239. [Medline: 11831207]
60. WHO. WHO Public Health Research Agenda for Managing Infodemics. Geneva: World Health Organization; 2021. URL: https://www.who.int/publications/i/item/9789240019508

Abbreviations

AI: artificial intelligence

Edited by T Mackey; submitted 25.02.24; peer-reviewed by R Robbins, S Wei, J Sy, T Purnat; comments to author 05.07.24; revised version received 07.07.24; accepted 24.10.24; published 13.12.24.

Please cite as:
Bragazzi NL, Garbarino S
The Complex Interaction Between Sleep-Related Information, Misinformation, and Sleep Health: Call for Comprehensive Research on Sleep Infodemiology and Infoveillance
JMIR Infodemiology 2024;4:e57748
URL: https://infodemiology.jmir.org/2024/1/e57748
doi:10.2196/57748
PMID:39475424

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org