JMIR Infodemiology

Journal Impact Factor (JIF) (2023): 3.5
Volume 3 (2023) ISSN 2564-1891 Editor-in-Chief: Tim Ken Mackey, MAS, PhD

Contents

Editorial

Original Papers

Editorial

Effective Infodemic Management: A Substantive Article of the Pandemic Accord

Kazuho Taguchi[1*], MSc; Precious Matsoso[2*], LLM; Roland Driece[3*], LLM; Tovar da Silva Nunes[4*], LLM, MA; Ahmed Soliman[5*], MD, MS; Viroj Tangcharoensathien[6], MD, PhD

[1]Permanent Mission of Japan to the International Organizations in Geneva, Geneva, Switzerland
[2]University of the Witwatersrand, Johannesburg, South Africa
[3]Ministry of Health, Welfare and Sports, The Hague, Netherlands
[4]Mission Permanente du Brésil auprès de l'Office des Nations Unies, Geneva, Switzerland
[5]Permanent Mission of the Arab Republic of Egypt to the United Nations Office and Other International Organizations, Geneva, Switzerland
[6]International Health Policy Program, Ministry of Public Health, Nonthaburi, Thailand
[*]these authors contributed equally

Corresponding Author:
Viroj Tangcharoensathien, MD, PhD
International Health Policy Program
Ministry of Public Health
Tivanond Road
Nonthaburi, 11000
Thailand
Phone: 66 818480297
Email: viroj@ihpp.thaigov.net

Abstract

Social media has proven to be valuable for disseminating public health information during pandemics. However, the circulation of misinformation through social media during public health emergencies, such as the SARS (severe acute respiratory syndrome), Ebola, and COVID-19 pandemics, has seriously hampered effective responses, leading to negative consequences. Intentionally misleading and deceptive fake news aims to harm organizations and individuals. To effectively respond to misinformation, governments should strengthen the management of an "infodemic," which involves monitoring the impact of infodemics through social listening, detecting signals of infodemic spread, mitigating the harmful effects of infodemics, and strengthening the resilience of individuals and communities. The global spread of misinformation requires multisectoral collaboration, such as researchers identifying leading sources of misinformation and superspreaders, media agencies identifying and debunking misinformation, technology platforms reducing the distribution of false or misleading posts and guiding users to health information from credible sources, and governments disseminating clear public health information in partnership with trusted messengers. Additionally, fact-checking has room for improvement through the use of automated checks. Collaboration between governments and fact-checking agencies should also be strengthened via effective and timely debunking mechanisms. Though the Intergovernmental Negotiating Body (INB) has yet to define the term "infodemic," Article 18 of the INB Bureau's text, developed for the Pandemic Accord, encompasses a range of actions aimed at enhancing infodemic management. The INB Bureau continues to facilitate evidence-informed discussion for an implementable article on infodemic management.

(JMIR Infodemiology 2023;3:e51760) doi:10.2196/51760

KEYWORDS

Pandemic Accord; infodemic; infodemic management; COVID-19; social media; Intergovernmental Negotiating Body; INB; INB Bureau; World Health Organization; WHO; misinformation; disinformation; public health

While social media has proven valuable for public health officials to disseminate crucial information during the COVID-19 pandemic, it has also been misused by many internet users to spread misinformation and fake news. This has hindered the public health response, particularly concerning issues like antivaccine and antimask sentiments. Fake news, which is intentionally misleading and deceptive, aims to harm organizations and individuals [1].

Misinformation is commonly shared on platforms such as Twitter, YouTube, Facebook, messaging applications, and personal websites. This has exacerbated the severity of the pandemic, eroded trust in public health experts, and led to nonadherence to public health and social measures [2,3].

Health-related misinformation on social media ranges from 0.2% to 28.8% [4]. Twitter, Facebook, YouTube, and Instagram play a critical role in rapidly spreading far-reaching misinformation. The increase in unreliable health information has delayed care provision and contributed to the dissemination of hateful and divisive rhetoric. However, social media can also be a useful tool to combat misinformation during crises [4].

Misinformation during large-scale infectious disease outbreaks has been observed since 2000, for example, during the SARS (severe acute respiratory syndrome) and Ebola outbreaks. False information is often linked to prevention, treatment, risk factors, transmission modes, and complications, and propagated by vaccine conspiracy theories [5]. A content analysis of 6600 randomly selected English-language tweets revealed that 22% contained false or partially false information, with the political nature of vaccines being most prevalent [6].

The World Health Organization (WHO) proposes addressing the challenge of managing an “infodemic” effectively using these five interconnected areas: (1) closely monitoring and measuring the impact of infodemics during health emergencies through social listening [7]; (2) detecting signals and understanding the spread and risk of infodemics; (3) responding with interventions that mitigate and protect against the harmful effects of infodemics; (4) assessing infodemic interventions and strengthening the resilience of individuals and communities; and (5) promoting the development, adaptation, and application of interventions and toolkits for infodemic management [8].

Messages should be based on scientific evidence and actionable behavioral change. They should be tailored to different population groups to enable them to make informed decisions to protect themselves and their communities during a pandemic. Vulnerable groups, such as women and youth, are particularly susceptible to misinformation. Health authorities should engage in multisectoral, interdisciplinary, and multilevel collaborations, including with the private sector, especially social media corporations and influencers [9]. Integrating artificial intelligence into social listening can provide real-time information about public discourse and support informed infodemic responses [10].

Digital health literacy refers to the ability to critically analyze elements presented in social media and determine their accuracy or credibility. Strengthening digital health literacy in the population can be done through education and engagement with communities, health professionals, and decision makers. A study revealed that less than half of health professionals possess a high level of digital health literacy [11]. In assessing digital health literacy in the population, a scoping review showed that the most commonly used tool is the eHealth Literacy Scale [12]. This short questionnaire consists of 8 questions with a 5-point Likert scale [13]. The eHealth Literacy Scale can be modified and used to regularly monitor and strengthen digital health literacy in the population.

Effective response to misinformation necessitates multisectoral and multidisciplinary collaboration. Researchers can identify leading sources of misinformation, including superspreaders. Media organizations can identify and debunk misinformation. Technology platforms can monitor and address misinformation by reducing the distribution of false or misleading posts and guiding users to health information from credible sources. Governments can disseminate clear public health information in partnership with trusted messengers [14].

Fact-checking has room for improvement. One study indicated that disinformation incorporates features from both propaganda and mainstream news, making it more challenging to detect. The scientific community should develop automated fact-checking processes that consider the degree of veracity [15]. Countries can apply Factcheck, a powerful tool in the fight against misinformation [16], along with effective and timely debunking mechanisms.

In response to the COVID-19 pandemic, WHO member states decided to draft and negotiate a WHO convention, agreement, or other international instrument on pandemic prevention, preparedness, and response (referred to as “WHO CA+”). The goal is to adopt this instrument under Article 19 or other provisions of the WHO Constitution as deemed appropriate by the Intergovernmental Negotiating Body (INB). This decision was mandated during the World Health Assembly Special Session in December 2021 [17], with the envisaged adoption of WHO CA+ by the 77th session of the World Health Assembly in May 2024.

Infodemics is an emerging challenge brought about by the widespread use of social media. This is the first time in history that the INB is discussing how to address misinformation. Article 18 of the INB Bureau’s text [18], “Communication and Public Awareness,” encompasses a range of actions on public communication that use scientific knowledge to increase awareness through community engagement and health literacy. Infodemic management is crucial to combat false or misleading information and disinformation. This includes regular social listening to assess the prevalence and profiles of misinformation. Active participation of the INB in the discussion will lead to an implementable article. We, the members of the INB Bureau [19], will continue to facilitate evidence-informed negotiations toward effective infodemic management.

Conflicts of Interest

All authors are members of the Intergovernmental Negotiating Body (INB) Bureau. We declare no competing interests. There was no medical writer involved in developing this editorial. This editorial neither represents the position of each bureau member’s country of origin nor prejudges the ongoing INB discussion or conclusions.

References

1. Giotakos O. Fake news in the age of COVID-19: evolutional and psychobiological considerations. Psychiatriki 2022 Sep 19;33(3):183-186 [FREE Full text] [doi: 10.22365/jpsych.2022.087] [Medline: 35947862]
2. Joseph A, Fernandez V, Kritzman S, Eaddy I, Cook O, Lambros S, et al. COVID-19 misinformation on social media: a scoping review. Cureus 2022 Apr;14(4):e24601 [FREE Full text] [doi: 10.7759/cureus.24601] [Medline: 35664409]
3. Prandi L, Primiero G. Effects of misinformation diffusion during a pandemic. Appl Netw Sci 2020 Oct 31;5(1):82 [FREE Full text] [doi: 10.1007/s41109-020-00327-6] [Medline: 33163618]
4. Borges do Nascimento IJ, Beatriz Pizarro A, Almeida J, Azzopardi-Muscat N, André Gonçalves M, Björklund M, et al. Infodemics and health misinformation: a systematic review of reviews. Bull World Health Organ 2022 Sep 01;100(9):544-561. [doi: 10.2471/blt.21.287654]
5. Chowdhury N, Khalid A, Turin TC. Understanding misinformation infodemic during public health emergencies due to large-scale disease outbreaks: a rapid review. Z Gesundh Wiss 2023 May 01;31(4):553-573 [FREE Full text] [doi: 10.1007/s10389-021-01565-3] [Medline: 33968601]
6. Trotochaud M, Smith E, Hosangadi D, Sell TK. Analyzing social media messaging on masks and vaccines: a case study on misinformation during the COVID-19 pandemic. Disaster Med Public Health Prep 2023 Jan 10:1-9. [doi: 10.1017/dmp.2023.16] [Medline: 36624607]
7. Boender T, 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]
8. Calleja N, AbdAllah A, Abad N, Ahmed N, Albarracin D, Altieri E, et al. A public health research agenda for managing infodemics: methods and results of the first WHO Infodemiology Conference. JMIR Infodemiology 2021 Sep 15;1(1):e30979 [FREE Full text] [doi: 10.2196/30979] [Medline: 34604708]
9. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
10. Purnat TD, Wilson H, Nguyen T, Briand S. EARS - A WHO platform for AI-supported real-time online social listening of COVID-19 conversations. Stud Health Technol Inform 2021 May 27;281:1009-1010. [doi: 10.3233/SHTI210330] [Medline: 34042825]
11. Ahmed MH, Guadie HA, Ngusie HS, Teferi GH, Gullslett MK, Hailegebreal S, et al. Digital health literacy during the COVID-19 pandemic among health care providers in resource-limited settings: cross-sectional study. JMIR Nurs 2022 Nov 14;5(1):e39866 [FREE Full text] [doi: 10.2196/39866] [Medline: 36301671]
12. Faux-Nightingale A, Philp F, Chadwick D, Singh B, Pandyan A. Available tools to evaluate digital health literacy and engagement with eHealth resources: a scoping review. Heliyon 2022 Aug;8(8):e10380 [FREE Full text] [doi: 10.1016/j.heliyon.2022.e10380] [Medline: 36090207]
13. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
14. Office of the Surgeon General. We can take action. Confronting Health Misinformation: The U.S. Surgeon General's Advisory on Building a Healthy Information Environment. Washington DC: US Department of Health and Human Services; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK572168/ [accessed 2023-07-26]
15. Pathak A, Srihari RK, Natu N. Disinformation: analysis and identification. Comput Math Organ Theory 2021 Jun 18;27(3):357-375 [FREE Full text] [doi: 10.1007/s10588-021-09336-x] [Medline: 34177355]
16. 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 Sep 14;118(37) [FREE Full text] [doi: 10.1073/pnas.2104235118] [Medline: 34507996]
17. The world together: establishment of an intergovernmental negotiating body to strengthen pandemic prevention, preparedness and response. World Health Organization. 2021 Dec 01. URL: https://apps.who.int/gb/ebwha/pdf_files/WHASSA2/SSA2(5)-en.pdf [accessed 2023-07-25]
18. Bureau's text of the WHO convention, agreement or other international instrument on pandemic prevention, preparedness and response (WHO CA+). World Health Organization. 2023 Jun 23. URL: https://apps.who.int/gb/inb/pdf_files/inb5/A_INB5_6-en.pdf [accessed 2023-07-25]
19. First report of the first meeting of the Intergovernmental Negotiating Body to draft and negotiate a WHO convention, agreement or other international instrument on pandemic prevention, preparedness and response. World Health Organization. 2022 Mar 22. URL: https://apps.who.int/gb/inb/pdf_files/inb1/A_INB1_4Rev1-en.pdf [accessed 2023-07-25]

Abbreviations

INB: Intergovernmental Negotiating Body
SARS: severe acute respiratory syndrome
WHO: World Health Organization

Edited by T Mackey; submitted 12.08.23; this is a non–peer-reviewed article;accepted 23.08.23; published 20.09.23.

Please cite as:
Taguchi K, Matsoso P, Driece R, da Silva Nunes T, Soliman A, Tangcharoensathien V
Effective Infodemic Management: A Substantive Article of the Pandemic Accord
JMIR Infodemiology 2023;3:e51760
URL: https://infodemiology.jmir.org/2023/1/e51760
doi:10.2196/51760
PMID:37728969

Original Paper

Predicting Patient Satisfaction With Medications for Treating Opioid Use Disorder: Case Study Applying Natural Language Processing to Reviews of Methadone and Buprenorphine/Naloxone on Health-Related Social Media

Samaneh Omranian[1], MSc; Maryam Zolnoori[2], PhD; Ming Huang[3], PhD; Celeste Campos-Castillo[4], PhD; Susan McRoy[1], PhD

[1]Department of Electrical Engineering and Computer Science, College of Engineering & Applied Science, University of Wisconsin-Milwaukee, Milwaukee, WI, United States

[2]School of Nursing, Columbia University, New York, NY, United States

[3]Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, MN, United States

[4]Department of Media and Information, Michigan State University, East Lansing, MI, United States

Corresponding Author:
Samaneh Omranian, MSc
Department of Electrical Engineering and Computer Science
College of Engineering & Applied Science
University of Wisconsin-Milwaukee
3200 N Cramer St
Milwaukee, WI, 53211
United States
Phone: 1 414 229 4000
Email: omranian@uwm.edu

Abstract

Background: Medication-assisted treatment (MAT) is an effective method for treating opioid use disorder (OUD), which combines behavioral therapies with one of three Food and Drug Administration–approved medications: methadone, buprenorphine, and naloxone. While MAT has been shown to be effective initially, there is a need for more information from the patient perspective about the satisfaction with medications. Existing research focuses on patient satisfaction with the entirety of the treatment, making it difficult to determine the unique role of medication and overlooking the views of those who may lack access to treatment due to being uninsured or concerns over stigma. Studies focusing on patients' perspectives are also limited by the lack of scales that can efficiently collect self-reports across domains of concerns.

Objective: A broad survey of patients' viewpoints can be obtained through social media and drug review forums, which are then assessed using automated methods to discover factors associated with medication satisfaction. Because the text is unstructured, it may contain a mix of formal and informal language. The primary aim of this study was to use natural language processing methods on text posted on health-related social media to detect patients' satisfaction with two well-studied OUD medications: methadone and buprenorphine/naloxone.

Methods: We collected 4353 patient reviews of methadone and buprenorphine/naloxone from 2008 to 2021 posted on WebMD and Drugs.com. To build our predictive models for detecting patient satisfaction, we first employed different analyses to build four input feature sets using the vectorized text, topic models, duration of treatment, and biomedical concepts by applying MetaMap. We then developed six prediction models: logistic regression, Elastic Net, least absolute shrinkage and selection operator, random forest classifier, Ridge classifier, and extreme gradient boosting to predict patients' satisfaction. Lastly, we compared the prediction models' performance over different feature sets.

Results: Topics discovered included oral sensation, side effects, insurance, and doctor visits. Biomedical concepts included symptoms, drugs, and illnesses. The F-score of the predictive models across all methods ranged from 89.9% to 90.8%. The Ridge classifier model, a regression-based method, outperformed the other models.

Conclusions: Assessment of patients' satisfaction with opioid dependency treatment medication can be predicted using automated text analysis. Adding biomedical concepts such as symptoms, drug name, and illness, along with the duration of treatment and topic models, had the most benefits for improving the prediction performance of the Elastic Net model compared to other models. Some of the factors associated with patient satisfaction overlap with domains covered in medication satisfaction scales (eg, side effects) and qualitative patient reports (eg, doctors' visits), while others (insurance) are overlooked, thereby underscoring the value added from processing text on online health forums to better understand patient adherence.

(JMIR Infodemiology 2023;3:e37207) doi:10.2196/37207

KEYWORDS

machine learning; online forums; text classification; topic modeling; MetaMap; drug review; opioid treatment, opioid use disorder; patient-generated text

Introduction

Opioid Use Disorder

The 2018 National Survey on Drug Use and Health estimated that 10.3 million people over 12 years old misused opioids, including 9.9 million individuals who misused prescribed pain relievers and 808,000 heroin users [1]. Long-term misuse of opioids and heroin affects the brain's normal functionalities and results in opioid tolerance, dependence, or addiction [2]. The Centers for Disease Control and Prevention (CDC) has preferred the term "opioid use disorder" (OUD) over "opioid abuse or dependence" owing to the set of behavioral, cognitive, and physiological symptoms after repeated substance use [3]. In response to the opioid dependence crisis, the National Institute on Drug Abuse has invested in the implantation of science and patient care to increase access to medication-assisted treatment (MAT), which consists of medication and behavioral therapies to reduce OUD across health care and the justice system [4]. MAT with opioid agnostic medications such as buprenorphine and methadone helps patients with OUD reduce relapse rates of quitting opioids; lowers illicit opioid use; and results in an overall reduction of the burden of opioid dependency on patients, caregivers, and the health care system [5].

Based on the CDC reports, MAT is a practical, systematic approach that incorporates medications such as methadone, buprenorphine, or naloxone along with behavioral therapy to meet the needs of patients with OUD. Methadone, a full opioid agonist, has been the most generally recognized and well-researched among pharmacological and nonpharmacological treatments since its introduction in 1965 [6]. Methadone provokes cells in the same way as illicit opioids but does not invoke the same cellular response that leads to dependence on the drug [7]. Another well-tolerated MAT supervised by the medical profession is buprenorphine/naloxone, marketed under the brand name Suboxone [8]. Buprenorphine is a partial opioid agonist that binds to the same opioid receptors as the opioid drugs in the brain, decreasing craving and withdrawal symptoms [8]. Methadone was developed for oral applications, and buprenorphine/naloxone is formulated for sublingual applications [9].

In a controlled comparative randomized study, Saxon et al [10] assessed the retention rates of methadone and buprenorphine/naloxone in individuals with OUD (N=1269). The study found that 74% of patients taking methadone completed the 24-week treatment, while only 46% of patients taking buprenorphine/naloxone completed the 24-week treatment, suggesting that over the net of behavioral therapies and other client services offered, medications play a key role in experiences. These findings suggest that a methadone treatment course may produce a better retention rate (medication's overall effectiveness) than a buprenorphine/naloxone treatment course, yet patients may prefer the latter when given a choice [11].

Various factors play into patient preferences and overall satisfaction with medication, including financial barriers, ease of use, and side effects, particularly withdrawal symptoms [12]. Unlike MAT, withdrawal symptoms appear immediately after opioid discontinuation or lowering of the dosage of opioids [13,14]. Cicero et al [15] found that the fear of withdrawal symptoms is a compelling motivator to relapse a short time after OUD treatment. As a result, poor/nonadherence and treatment dropout are quite common in MAT [16]. Bastiaans et al [17] reported elevated pulse rate, piloerection, pains, nausea, and many other symptoms as signs of withdrawal. Therefore, understanding the patients' satisfaction with medications used in OUD treatment may help health care professionals make informed treatment decisions [18].

Unfortunately, existing data have several limitations. Studies examining patient experiences with OUD treatment often evaluate satisfaction with the entire treatment and do not disaggregate satisfaction with the medications used [19]. When studies do measure satisfaction with the medications specifically, they are limited by the lack of scales designed for OUD treatment and that are short enough to administer regularly [20]. Qualitative studies can provide opportunities for patients with OUD to volunteer a broad range of factors shaping their satisfaction with medication [19], but they share a limitation with quantitative studies in that they often sample from those who are enrolled in treatment [12]. This approach misses the perspectives of those who cannot access treatment because of concerns over stigma or lack insurance [21]. An alternative data source is therefore needed.

Online Health Forums

To better understand the patients' experiences and address limitations in existing data sources, online health forums have been proven to be useful resources, as patients are not biased by the presence of a medical professional [22]. Accordingly, patients seek external information sources such as health care forums or online health care communities, particularly reports of patients with similar health conditions and treatment [23]. Besides, these forums also provide valuable social support, encouragement, and friendship [24,25]. In research on the efficacy of online health discussion forums for prescription drug abuse, findings imply that an online health forum is useful for assisting users with physical detoxification and opiate withdrawal [26]. Another advantage of these online platforms compared to survey data is that people decide when to post a review compared with patient satisfaction surveys. In a study on bias in patient satisfaction surveys, Dunsch et al [27] demonstrated how assessments of patient satisfaction are extremely sensitive to how the questions are framed. They also found convincing evidence of the acceptance bias, or peoples' inclination to accept a statement regardless of its content, in particular [27].

Natural Language Processing

Analyzing data from online health forums is not without its own challenges. The text is unstructured and may contain a mix of formal and informal concepts. Moreover, across the reviews, there may be different terms used to refer to the same biomedical concepts. Natural language processing (NLP) techniques increasingly offer an alternative analytic strategy for addressing complicated interactions in large data sets, recognizing hidden patterns, and providing effective predictions in health-related texts [28,29]. Several prior studies have used NLP methods to predict opioid dependency [30,31], overdose [32], prolonged use of opioids after surgery [33], suicidality among opioid users on the online forum Reddit, or other related outcomes [34]. Moreover, in analyzing health-related online review posts, Lu et al [35] discovered health-related topics using text clustering algorithms on social media data.

Contributions

To address limitations in existing data on patient satisfaction with medications for opioid dependency, we examined online health forums. The aim of this study was to utilize NLP to detect patient satisfaction with opioid medication treatments from patient reviews in health forums that mention methadone and buprenorphine as targeted OUD medications. To the best of our knowledge, this is the first study to identify biomedical concepts influencing patients' satisfaction with opioid medication treatments and automatically detect patient satisfaction using those concepts as model features. To achieve our goal, we utilized patient reviews from two well-known health care forums, WebMD [36] and Drugs.com [37], on opioid treatment medications. We also used MetaMap (an NLP and computational-linguistic tool developed by the National Library of Medicine) to extract biomedical terms used in the patients' posts. We leveraged these terms along with the duration of treatments to train a stratified 10-fold cross-validation (CV) model to detect patient satisfaction with targeted medications.

Methods

Study Design

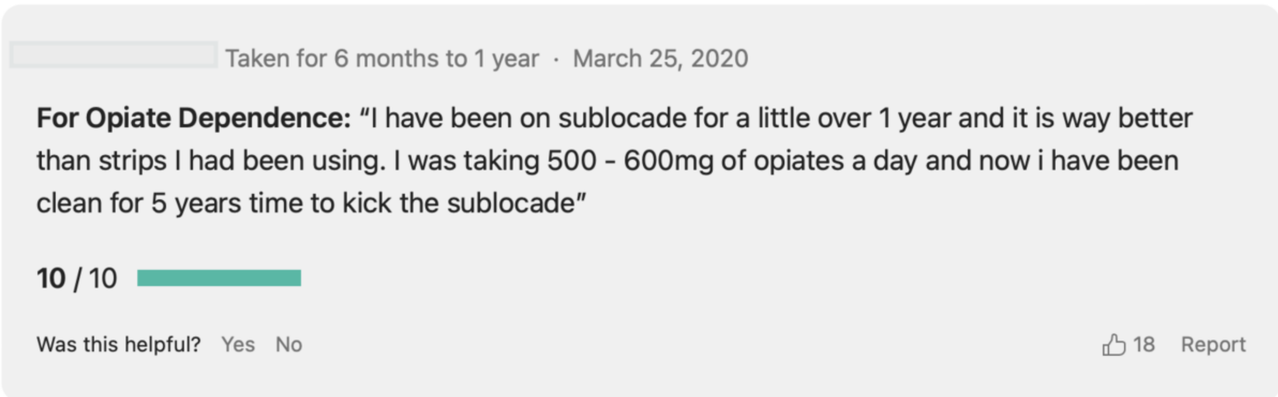
The methodology of this study consisted of four stages: (1) data collection and preprocessing, (2) identifying hidden topic models and duration of treatment, (3) identifying biomedical concepts by applying MetaMap, and (4) developing a predictive model to detect patient satisfaction with opioid medications from reviewers' posts. We describe each of the stages in detail below.

Data Sources

We used two health care forums, Drugs.com and WebMD, as our data source for this study. Both forums collect patients' self-reported experiences for a wide range of medications. In both forums, patients can report their experiences with medication in a field called "comments." In the WebMD forum, patients can enter their gender and age range, while the Drugs.com forum does not have an option for gender and age. In both forums, each review post includes a rating attribute for the reviewer to rate the treatment effectiveness experience as a number, which is in the range of 1-10 in Drugs.com and 1-5 in WebMD. In addition, in either forum, the reviewers can input the duration of their treatments into four categories: too short, less than 1 month, too long, and more than 10 years. WebMD also has options for collecting the "drug satisfaction" and "ease of use," while Drugs.com does not have these two rating features. The date of reports in both forums is recorded automatically using the system. The patient's ID is visible; however, the forums collect the patient consent to make the reported experience publicly available. Figure 1 shows a sample review post from the Drugs.com forum.

In this study, our targeted drugs were the two well-studied [38] OUD treatment medications methadone hydrochloride and buprenorphine/naloxone hydrochloride (Zubsolv, Suboxone, Subutex, and Bunavail). Methadone and naloxone (brand names: Methadose and Dolophine) are from a class of medications called opioid analgesics, whereas buprenorphine is from the partial agonist-antagonists class.

Figure 1. A sample review post from Drugs.com.



Data Collection

We collected 4353 drug reviews from the two online forums via an automatic web scraper. We used *beautiful soup* [39] in Python programming to develop the web scraper. The collected data included review posts that mentioned these two drugs (both generic and brand names for each) from 2008 to 2021, the duration of the treatment, and drug effectiveness rating. Henceforward, because the term "effectiveness" has a particular meaning in the medical literature, we refer to the numerical patient rating as "patient satisfaction."

Data Preprocessing

In this study, we used the whole review comment as the unit of our analysis to preserve the meaning of the patients' review comments. We removed all posts that did not provide any comment text for the post. We then used Natural Language Toolkit [40] to remove all stop words, punctuation, and non-ASCII characters. Subsequently, the words were stemmed and lemmatized for applying the topic modeling approach. This data-cleaning process improves the performance of topic modeling as it avoids repeated versions of a word in a topic and improves the detection rate significantly [41]. Finally, we rescaled patients' ratings for the two medication treatments (methadone and buprenorphine/naloxone) from the two forums for further developing the predictive model process. As stated earlier, in both forums, each review post includes a rating attribute for rating the treatment medication satisfaction by the reviewer. Drugs.com's rating ranges from 1 to 10, whereas this rating is from 1 to 5 on WebMD. To make the rating uniform, we employed the approach proposed by Dawes [42] to rescale 1-10 ratings on a scale from 1 to 5. In this approach, 1 remains as 1 and 10 is rescaled to 5, and then the midpoint of 5.5 on a 10-point scale is changed to be 3, the midpoint of 1 is changed to 5, and so on. We then rescaled the satisfaction ratings from a 5-score scale into a binary score, in which a score of 1, 2, or 3=unsatisfied and a score of 4 or 5=satisfied [43].

Sample Size Calculation

To identify the best features from users' posts, we needed to determine the best random sample for the collected review posts. We first identified the ideal sample size using the finite population correction factor when sampling without replacement [44]. The formula used for calculating the sample size n with the limited population factor formula in statistics is as follows:

$$n=n_0N/n_0+(N–1)$$

where N is the population size and n_0 is the size of the sample before the finite population factors are applied. We calculated n_0 with the following formula:

$$n_0=z^2p(1–p)/e^2$$

where e is the sampling error, p is the population standard deviation, and z is the confidence level. We then calculated the ideal size n for a sample with N=4353, e=0.007, p=0.15, and z=1.96 (95% confidence), yielding n=100. To find the best random sample of size 100, we used a stratified sample method. This approach divides the population into groups, and a proportionate number is randomly sampled from each group [44].

Content Analysis of Drug Reviews and NLP Tools for Feature Extraction

After determining the sample size, we reviewed 100 posts via the stratified sample method manually to get a better sense of the posts' content to determine the most suitable techniques for feature extraction. Our manual content analysis showed that the patients use both colloquial and formal medical language to express medical concepts presenting symptoms, adverse drug effects, drug effectiveness, and some social concepts (eg, social isolation or financial stress) with medication. Therefore, to identify the medical concepts, we used MetaMap to extract the medical and social concepts, particularly for formal expressions (see the Biomedical Concepts Extracted by MetaMap section below for more details). In addition, to identify the major themes in each drug review, we used the topic modeling approach. We also used vectorized text and n-grams as the baseline feature set. Furthermore, we conducted feature importance in Python to determine the contribution of each feature set to the model performance. The following sections provide more detail on each feature set.

Biomedical Concepts Extracted by MetaMap as Features

To identify the biomedical concepts such as symptoms, drugs, and illnesses mentioned in the review posts, we employed MetaMap, a publicly available program based on NLP and computational-linguistic methods developed by the National Library of Medicine. MetaMap is commonly used in information extraction, classification, biomedical and clinical literature analysis in natural language, and unified medical language system (UMLS) concept-based indexing and retrieval. MetaMap maps biomedical text to concepts in the UMLS Metathesaurus [45], which assists in organizing different vocabularies used to refer to the same biomedical concept. It takes the text and breaks it down into components that include terms, phrases, linguistic elements, and tokens through a series of modules. In a comprehensive study on MetaMap features, Aronson et al [46] reported that MetaMap has an extension of the NegEx algorithm [47] to detect negated concepts.

The number of biomedical concepts extracted by MetaMap for all collected reviews was 556. To improve the performance of machine-learning algorithms, sparse features and features with low frequency for an identified concept were removed. Thus, we primarily focused on concepts with higher frequency, leaving 424 biomedical concepts on three groups of symptoms, drugs, and illnesses. The detailed procedure, including MetaMap methods and the associated results summary, is provided in Multimedia Appendix 1.

Topic Models as Features

Topic modeling is a statistical technique that groups the words of a collection of documents based on their frequency of co-occurrence. Topic modeling's core assumption is that a document contains a mixture of themes. To identify the main underlying themes or "hidden topics" among the patients' posts, we utilized latent Dirichlet allocation (LDA) [48], one of the popular topic modeling algorithms in NLP. LDA is a three-level hierarchical Bayesian model that models each item of a collection as a finite mixture over an underlying list of topics. The main advantage of LDA is that it is a probabilistic model with an interpretable subject and different parameters [49]. Additionally, studies on online health forums dealing with breast cancer [50] and Chinese social media [51] revealed that LDA can be used as a feature for developing predictive models to detect postings that contain informational and emotional support automatically. However, the basic disadvantage of LDA is that it lacks objective metrics to justify hyperparameter selection [52].

For our analysis, we used the LDA algorithm implemented in Python's Gensim package [53]. We utilized the Mallet function from the Gensim package. Selecting the best number of topics is important to create a meaningful set of topics. Steyvers and Griffiths [54] observed that the best number of topics varies from task to task and needs to result in the best generalization performance. Their research concluded that picking too few topics causes a vast topic, limiting the ability to discriminate. In contrast, too many topics results in topics that tend to catch unusual word combinations [54]. To that extent, to determine the best number of topics and words per topic, we experimented with different numbers of topics (5, 10, 15, 20, 25, 30, and 35) and words (5, 10, 15, and 20) alongside manual tuning of the LDA parameters. We also assessed the coherence score [55] corresponding to each extracted topic model calculated by Mallet and confirmed their reasonableness by manual inspection. By this method, the most meaningful set included 20 topics with 10 words per topic.

Duration of Treatments

Fishbain et al [56] found that between 3.3% and 14.5% of long-term prescription opioid users developed an opioid dependency after an average of 22.1 months of exposure [57]. Therefore, we considered the duration of opioid treatment medication as one of the predictive model's features. On both online platforms, the users have seven choices to enter the time on the medication: less than 1 month, 1-6 months, 6 months to 1 year, 1-2 years, 2-5 years, 5-10 years, and 10 or more years.

Approximately 10% of the collected data had missing information about the treatment duration. To handle the missing data, we utilized maximum-likelihood estimation, a statistical strategy for estimating missing data based on the available data that have been seen [58].

As a final step, the features extracted using NLP techniques (Table 1) were combined with the duration of treatment for developing machine-learning models to predict patient satisfaction with the two targeted opioid treatment medications: methadone and buprenorphine/naloxone (see the Data Source section for details).

Table 1. Four different feature sets used for the predictive model.

Input features	Baseline feature set	Feature set 1	Feature set 2	Feature set 3
Vectorized text, unigrams, and bi-grams	✓	✓	✓	✓
Biomedical concepts extracted by MetaMap		✓		✓
Topic models			✓	✓
Duration of treatment			✓	✓

Machine-Learning Algorithms

In this study, we selected six machine-learning algorithms to predict the patients' satisfaction with OUD treatments. We chose six predictive approaches based on prior studies that have frequently produced the best prediction outcomes in classification talk [59,60]. These approaches are logistic regression, the elastic network model (Elastic Net), least absolute shrinkage and selection operator (LASSO) regression, ridge regression model (Ridge), and two decision tree models, random forest and extreme gradient boosting (XGBoost).

We used stratified k-fold CV, which automatically selects training and test sets for each iteration, to train and test the machine-learning models; none of the models tune hyperparameters shared across different iterations. The k-fold CV splits the data set into k folds randomly each time and uses one dedicated fold for the test set and the rest of the folds for the training set [61]. In a stratified k-fold CV, the folds are stratified to ensure that each fold of the data set has the same proportion of observations with a given label, particularly in the case of an imbalanced labeled data set [62]. Han et al [63] recommended stratified 10-fold CV owing to its low bias and variance for assessing the performance of machine-learning algorithms. Therefore, we used a stratified 10-fold CV to train and test the models, and the average of the folds was taken to compare the metrics.

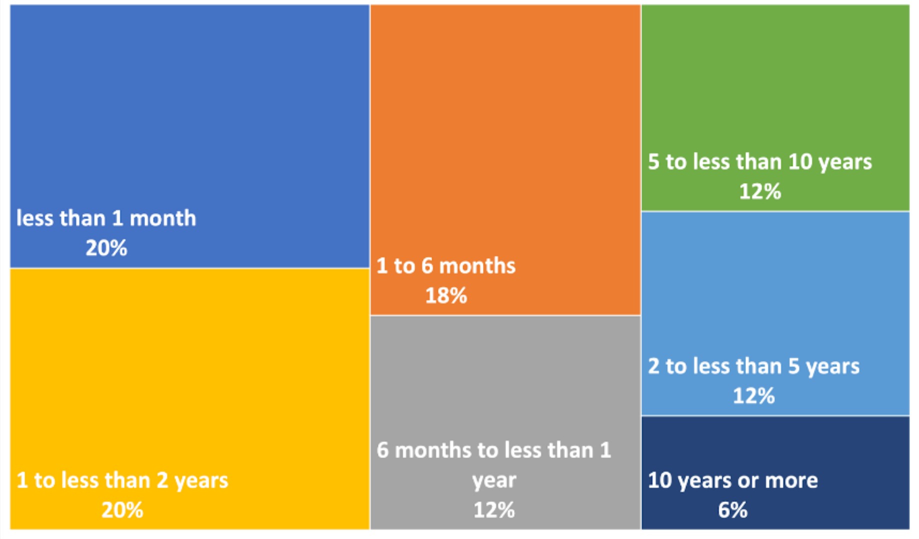
The above algorithms were fed with four novel combinations of input features (Table 1) as follows: vectorized text, which includes unigrams and bigrams (baseline feature set); vectorized text along with features from MetaMap (feature set 1); vectorized text along with features from topic models and duration of treatment (feature set 2); and lastly, vectorized text along with features from both MetaMap and topic models and duration of treatment (feature set 3). In the next section, we describe the details of the model features. We evaluated each model's performance using general metrics of accuracy, precision, recall, F-score, and area under the curve (AUC).

Results

Statistical Analysis

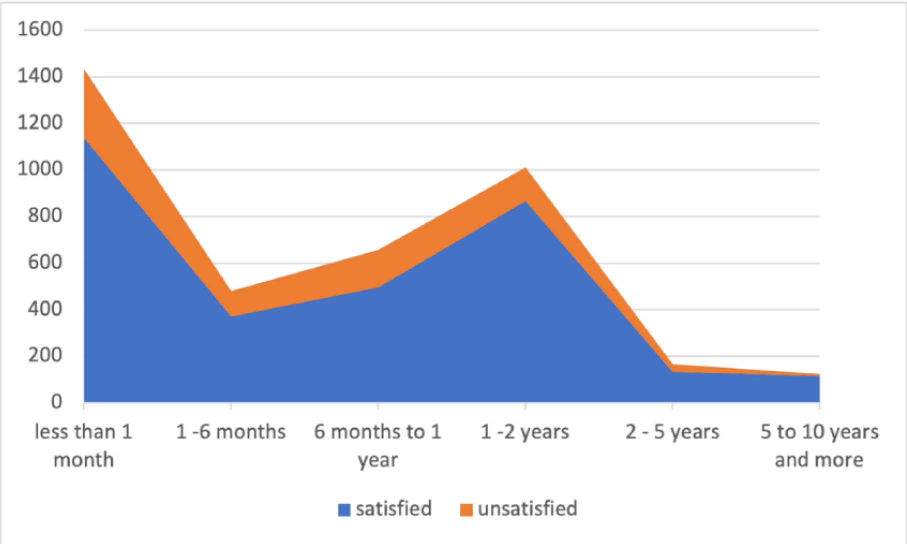
Removing the posts with empty comment texts reduced the sample of 4353 posts to 4048. The average number of words per post review was 93. Among all review posts, three quarters of reviewers utilized buprenorphine/naloxone (Suboxone) and the rest used methadone. Figure 2 demonstrates the distribution of treatment duration reported by patients among six categories of time on the medication.

Figure 2. Distribution of the duration of treatment reported by patients.



Our statistical analysis revealed that 18% of the patients were unsatisfied with the treatment medication and 82% reported satisfaction with targeted medications. As shown in Figure 3, 36% of satisfied patients with the medications reported using the medication for less than 1 month.

Figure 3. Distribution of "satisfied" and "unsatisfied" patients according to the duration of treatment.



After applying MetaMap to all review posts, we found that the patients mentioned symptoms such as breathing problems, dehydration, vomit, and confusion. The reviewers also mentioned illnesses such as adrenal crisis, delirium, and chronic headaches. The top 10 symptoms, other drug names, and illnesses extracted by MetaMap are summarized in Table 2.

Table 2. Top 10 biomedical concepts for symptoms, other drugs, and illnesses extracted by MetaMap.

Category	Biomedical concepts
Symptoms	Suffocate, breathing problem, dehydration, vomit, confusion, restlessness, disoriented, muscle weakness, mood changes, depressing
Other mentioned drugs	hydroxyzine (anxiety, nausea), luvox (OCD[a]), temazepam (insomnia), vitamins (body needs), miralax (constipation), fioricet (pain and fever reliever), Narcan (overdose), magnesium citrate (bowel movement), Adderall (ADHD[b]), Ambien (insomnia in adults)
Illness	adrenal crisis, delirium, reflex sympathetic dystrophy, cyst, hydrocephalus, chronic headaches, sciatica, fibromyalgia, herniated discs, degenerative joint disease

[a]OCD: obsessive compulsive disorder.

[b]ADHD: attention deficit and hyperactivity disorder.

As depicted in Table 2, reviewers mentioned other drugs in their posts, such as hydroxyzine, which is used for anxiety and nausea, and temazepam, which is helpful for insomnia. After applying topic modeling methods, the top four hidden topics extracted by LDA were an oral sensation, side effects, insurance, and doctor visit. Table 3 demonstrates the top 10 words associated with each cluster of topics.

Table 3. Four meaningful topics and associated words extracted by Mallet.

Topic	Top 10 associated words
Oral sensation	dissolve, experience, tongue, back, minute, mouth, cheek, form, considerable, night
Side effects	bad, anxiety, depression, sober, luck, meditation, write, panic, person, properly
Insurance	insurance, find, cover, anymore, strong, worry, company, doctor, couple, list
Doctor visit	doctor, thing, put, prescribe, amazing, blood, sublingual, leave, worried, shot

Prediction Model Performance

Table 4 shows the predictive model performance through four different feature sets: vectorized text, including unigrams and bigrams (baseline feature set); vectorized text along with biomedical concepts (feature set 1); vectorized text along with topic models and duration of treatment (feature set 2); and lastly, vectorized text along with biomedical concepts, topic models, and duration of treatment (feature set 3). Our feature importance analysis revealed that text alone had higher importance in the models' performance than features extracted by MetaMap, topic models, and the duration of treatments. After feeding each model with a different set of features, we found a slight improvement in the F-scores of the predictive models compared to the baseline model, except for the Ridge classifier with a small deterioration.

Table 4. Performance (F1-scores) of six classifiers for each combination of features.

Model	Baseline feature set	Feature set 1	Feature set 2	Feature set 3
Logistic regression	90.2	90.3	90.2	90.5
Elastic Net	90.2	90.2	90.2	90.6
LASSO[a]	90.3	90.1	90.3	90.5
Random forest	90.0	90.0	90.0	90.0
Ridge classifier	90.8	90.2	90.7	90.6
XGBoost[b]	89.9	90.2	90.1	90.2

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

[b]XGBoost: extreme gradient boosting.

The receiver operating characteristic curves in Figure 4 compare different classifiers for different feature sets. Comparing the AUC values shows that the logistic regression model (AUC=78.8) outperformed the other models. The complete curves for all classifiers and feature sets are available in Multimedia Appendix 2.

Figure 4. Receiver operating characteristic plots showing the performance comparison among six classifiers, including feature set 3. AUC: area under the curve; LASSO: least absolute shrinkage and selection operator; XGBoost: extreme gradient boosting.

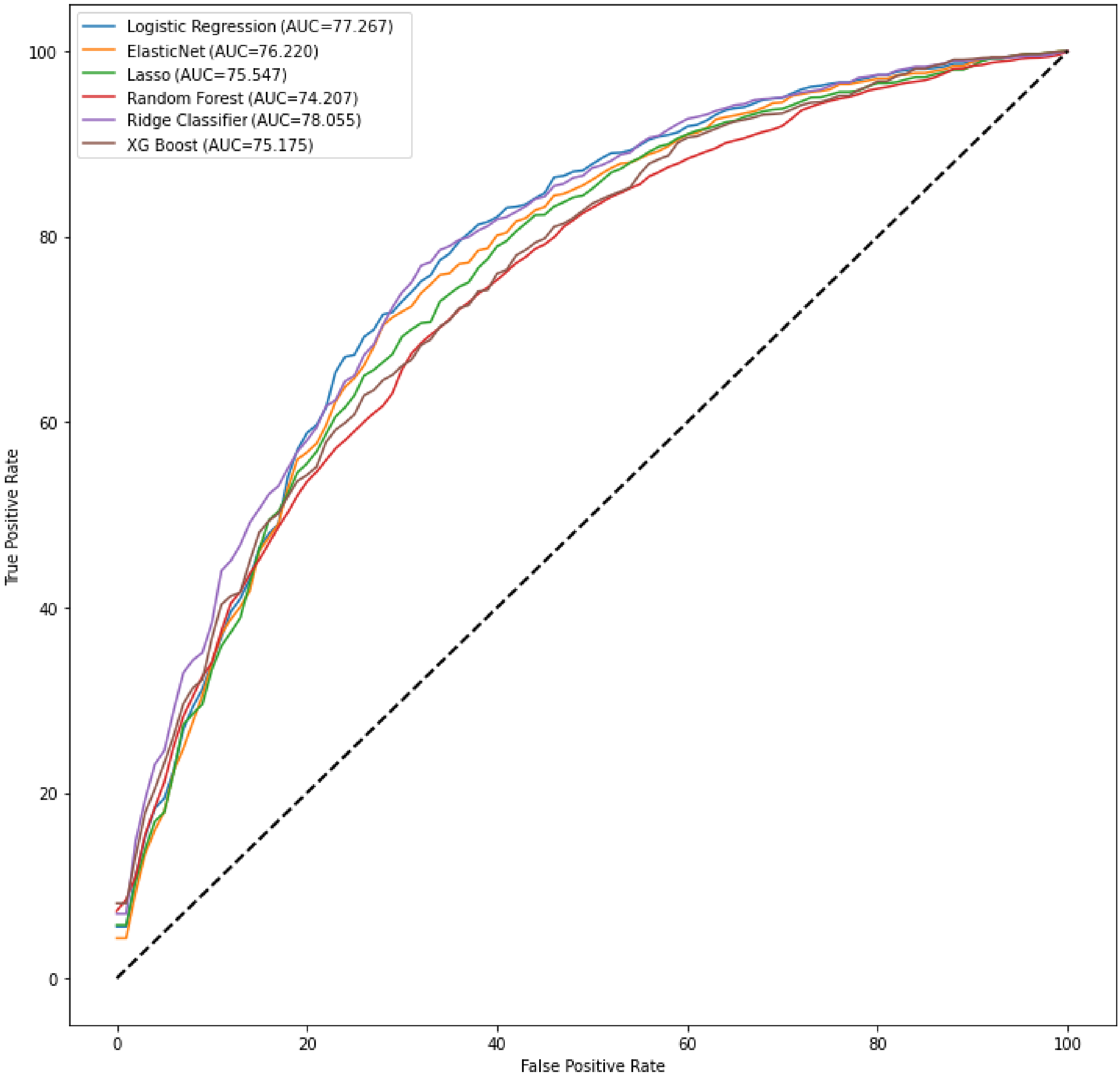
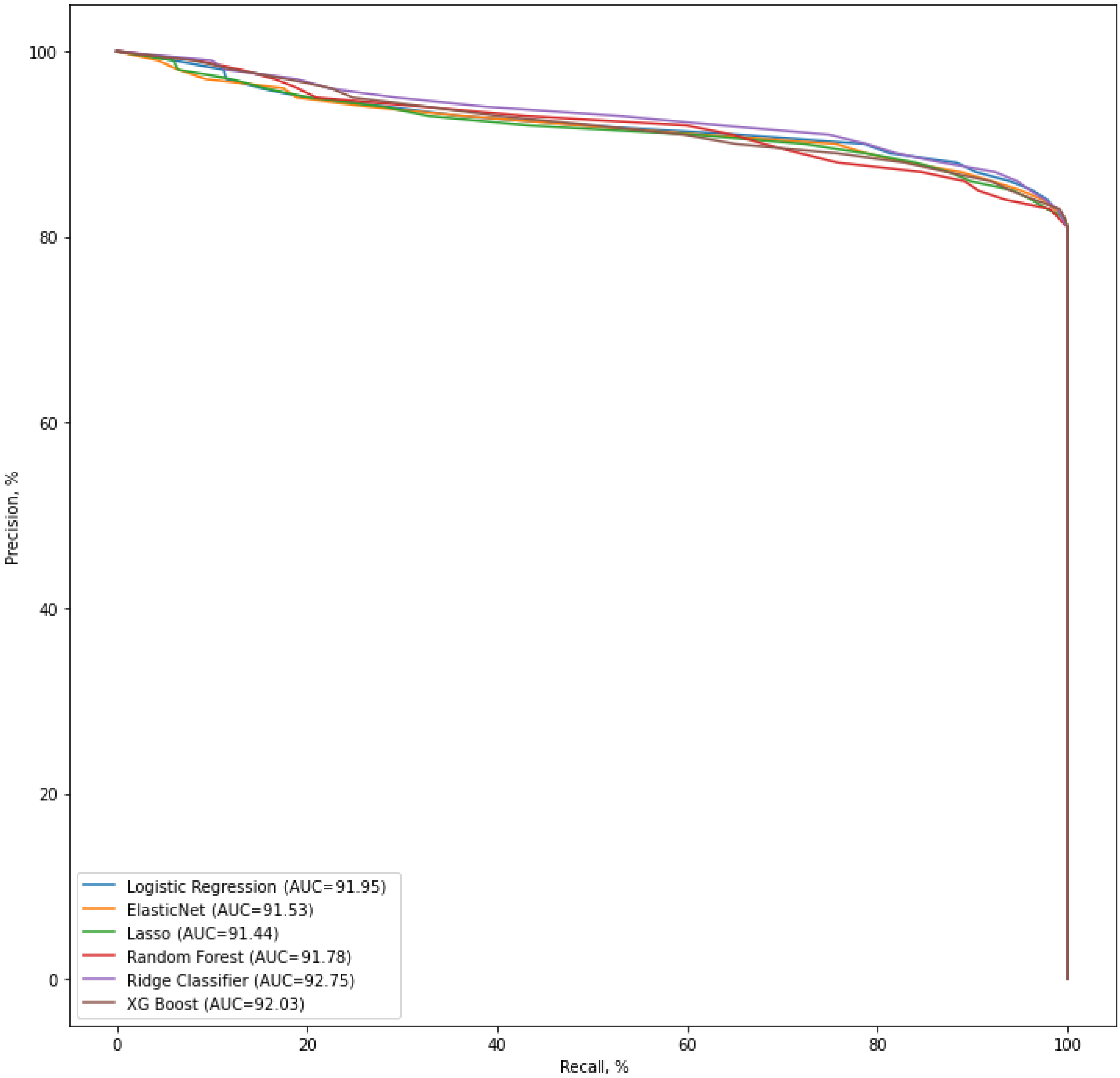


Figure 5 shows the precision-recall curves of all models on the different feature sets. The AUC values were calculated based on the average of the AUC of each curve. Comparing the prediction scores shows that all models had a close range of scores. As shown in Table 4, the F-score for all models ranged from 89.9% to 90.8% for the baseline feature set, biomedical concepts, topic models, and duration of treatment feature combination set. The Ridge classifier model scores, in general, were better than the other models' scores. Adding biomedical concepts, topic models, and duration of treatment as features individually and in combination improved the performance measures of the logistic regression, Elastic Net, LASSO, and XGBoost models, whereas there were no changes for the random forest model. The results also revealed that the Ridge classifier gained the highest performance (F-score=90.8) by having the baseline feature set as its input.

Figure 5. Comparison of the precision-recall curves of all models, including feature set 3. AUC: area under the curve; LASSO: least absolute shrinkage and selection operator; XGBoost: extreme gradient boosting.



Discussion

Principal Findings

In this study, we collected review posts from two online health-related forums, Drugs.com and WebMD, to investigate creating a predictive model to automatically identify patient satisfaction with OUD treatment medications. The data source presents challenges for analysis because it is unstructured, patient-generated text. We demonstrated how analysts can use MetaMap to detect biomedical concepts alongside an NLP rule-based algorithm and topic modeling (unsupervised NLP algorithm) to detect patient satisfaction. Our analysis of feature importance uncovered that the text alone as a baseline feature is a significant input variable to predict the output variable. This is aligned with the results coming from integrating biomedical concepts extracted from MetaMap and topic models with other features such as duration of treatment, which led to adding slight value to the predictive model performance. Our study also found that compared to other models, Elastic Net, a regularized regression method, improved the most upon the addition of biomedical features along with other features, which is in line with Marafino et al's [64] study on biomedical text classification on nurses' notes. The F-score ranged from 89.9 for the XGBoost model to 90.8 for the Ridge classifier model, including the baseline feature set. The AUC value ranged from 74.0 (random forest) to 78.8 (logistic regression). When training models with different machine-learning models, we manually considered some alternative values for the models' parameters, but this resulted in no significant improvements. For instance, we manually adjusted the number of iterations to the logistic regression, Elastic Net, and LASSO models from 1000 to 50,000, but we noticed no significant change.

To the best of our knowledge, this study is the first to identify the biomedical concepts from reviews of opioid medication treatments among patients who have been struggling with the issue of OUD treatment and to predict patient satisfaction with these medications. This is critical given previous research showing the importance of medications for shaping experiences with opioid dependency treatment [11].

This study aligns with the findings from other research, while also underscoring the added value of analyzing reviews from online health forums. Our study showed that patients who used different forms of buprenorphine/naloxone (Suboxone) and methadone for their OUD mentioned numerous symptoms, which is in line with the findings of Perlogizzi et al [65] on opioid withdrawal symptoms, who showed that these symptoms are both a motivator for continuing opioid usage and a barrier to stopping them [65]. Symptoms may also reflect side effects, which are commonly ascertained in self-reported surveys of medication satisfaction and volunteered during interviews and focus groups [12,20]. Oral sensation when ingesting the medication and frequency of doctors' visits also appear regularly in patient reports of their experiences [12]. Notably, insurance was revealed as a topic appearing in the reviews, and words related to this topic included the term "worry," possibly indicating concerns about having insurance to assist with financial barriers to treatment. Approximately one-fifth of those who experience opioid dependency lack health insurance coverage, which increases their risk of forgoing treatment [20]. Insurance coverage and concerns about other financial barriers are rarely considered in medication satisfaction scales, however, which highlights the contributions of monitoring online health forums to capture patient satisfaction more fully.

Limitations

This study has several limitations that may impact the results. The reviews we incorporated may not reflect the viewpoints of the population of patients with OUD fully because we only collected reviews from two websites (WebMD and Drugs.com) and we cannot determine the demographic or medical background of the reviewers. Moreover, because we were limited in only using the formal names of medications as keywords, we may have missed more colloquial discourse that refers to these using slang. Other platforms such as Twitter may provide sufficient information to infer background characteristics and capture more colloquial references to the medications, but previous work found that such data were not as well-structured or relevant as review text [43].

By incorporating reviews from two different websites, this imposed restrictions in how we structured and processed the data. Because patient satisfaction was measured on one site on a 5-point scale and measured on a 10-point scale in the second site, we had to rescale the ratings to make them uniform (on a scale from 1 to 5), which resulted in the loss of some information. Besides the text of the review, we also only had one feature present in both websites (duration of treatment) to use as one of the predictive model's features. To further assist in managing text from two different websites that may include a mix of biomedical and informal language, we used a combination of NLP techniques. We used MetaMap, which is useful for identifying biomedical concepts because it leverages the UML Metathesaurus, but it may still fail in recognizing and mapping a disease name effectively [66]. We also used topic modeling but used it on the entire review, and each review may contain different sentences with different sentiments and topics, as people reflect on their lives before or during treatment. Despite these limitations, we achieved high accuracy, and the resulting algorithm may still help address a complex crisis entangled with public health as well as with social and economic welfare, especially in the treatment of pain, a major health issue [67].

Future Work

Based on the current methods of this study and the limitations mentioned above, several future directions are suggested to build on this research. Foremost, adding demographics such as the gender of the reviewer and evaluating whether these interact with treatment can play an essential role in testing whether there are demographic disparities in responses to opioid treatments. Furthermore, future work could extend the analysis to explore the relationship between online opioid treatment reviews by patients and clinical notes by health care providers. In addition, sentiment analysis could be performed and added as a feature in addition to hidden topics. Moreover, applying advanced filtering techniques to the reviews may improve retrieving text more relevant to the subject of study and refining contextual polarity to better grasp what a word or phrase implies in a given context. Lastly, word embeddings and deep-learning methods are other suggestions for future work to investigate the improvement in the model's performance.

Conclusions

To address the need to more fully capture patients' experiences with medications for OUD treatments, this study used different models and classifiers to predict patient satisfaction using reviews from two online health forums. As a part of this research, we performed topic modeling and found that patients' main concerns regarding OUD treatments are insurance, anxiety/depression, doctor visits, and types of medications. Insurance is a topic rarely covered in scales to measure medication satisfaction during OUD treatments, despite one-fifth of those with OUD lacking health insurance. We also found that including treatment duration, hidden topics, and biomedical concepts such as symptoms, drug names, and illnesses was beneficial in developing some of the predictive models, specifically the Elastic Net model, for this study. Despite the data source comprising unstructured patient-generated text, these methods showed that we could analyze patient reviews and predict patient satisfaction with an opioid dependency with an F-score of approximately 90%. This result offers a promising method for automatically extracting information from patients' comments on health care web forums.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MetaMap methods and results of its application to the data of this study.
[PDF File (Adobe PDF File), 293 KB - infodemiology_v3i1e37207_app1.pdf]

Multimedia Appendix 2
Area under the receiver operating characteristic (ROC) curve (AUC) results for different classifiers with four feature sets.
[PDF File (Adobe PDF File), 497 KB - infodemiology_v3i1e37207_app2.pdf]

References

1. Key Substance Use and Mental Health Indicators in the United States: Results from the 2018 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration. 2019. URL: https://store.samhsa.gov/product/key-substance-use-and-mental-health-indicators-in-the-united-states-results-from-the-2018-national-survey-on-Drug-Use-and-Health/PEP19-5068 [accessed 2023-01-06]
2. Kosten TR, George TP. The neurobiology of opioid dependence: implications for treatment. Sci Pract Perspect 2002 Jul;1(1):13-20 [FREE Full text] [doi: 10.1151/spp021113] [Medline: 18567959]
3. Commonly used terms: opioids. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/opioids/basics/terms.html [accessed 2023-01-06]
4. National Academies of Sciences, Engineering, and Medicine. Medication-Assisted Treatment for Opioid Use Disorder: Proceedings of a Workshop—in Brief. Washington, DC: National Academies Press; 2018.
5. Au VYO, Rosic T, Sanger N, Hillmer A, Chawar C, Worster A, et al. Factors associated with opioid overdose during medication-assisted treatment: How can we identify individuals at risk? Harm Reduct J 2021 Jul 08;18(1):71 [FREE Full text] [doi: 10.1186/s12954-021-00521-4] [Medline: 34238301]
6. Ali S, Tahir B, Jabeen S, Malik M. Methadone treatment of opiate addiction: a systematic review of comparative studies. Innov Clin Neurosci 2017;14(7-8):8-19 [FREE Full text] [Medline: 29616150]
7. Kosten T, George TP. The neurobiology of opioid dependence: implications for treatment. Sci Pract Perspect 2002 Jul;1(1):13-20 [FREE Full text] [doi: 10.1151/spp021113] [Medline: 18567959]
8. Orman J, Keating GM. Buprenorphine/naloxone: a review of its use in the treatment of opioid dependence. Drugs 2009;69(5):577-607. [doi: 10.2165/00003495-200969050-00006] [Medline: 19368419]
9. Shulman M, Wai JM, Nunes EV. Buprenorphine treatment for opioid use disorder: an overview. CNS Drugs 2019 Jun;33(6):567-580 [FREE Full text] [doi: 10.1007/s40263-019-00637-z] [Medline: 31062259]
10. Saxon A, Hser YI, Woody G, Ling W. Medication-assisted treatment for opioid addiction: methadone and buprenorphine. J Food Drug Anal 2013 Dec;21(4):S69-S72 [FREE Full text] [doi: 10.1016/j.jfda.2013.09.037] [Medline: 24436573]
11. Checkley L, Steiger S, Knight KR. "I wanted something that was more flexible": a qualitative study of patient preferences on choosing buprenorphine over methadone in a large, safety-net hospital opioid treatment program. Subst Abus 2022 Feb 03;43(1):767-773 [FREE Full text] [doi: 10.1080/08897077.2021.2010251] [Medline: 35112998]
12. Cioe K, Biondi BE, Easly R, Simard A, Zheng X, Springer SA. A systematic review of patients' and providers' perspectives of medications for treatment of opioid use disorder. J Subst Abuse Treat 2020 Dec;119:108146. [doi: 10.1016/j.jsat.2020.108146] [Medline: 33138929]
13. Cicero T, Ellis MS. The prescription opioid epidemic: a review of qualitative studies on the progression from initial use to abuse. Dial Clin Neurosc 2022 Apr 01;19(3):259-269. [doi: 10.31887/dcns.2017.19.3/tcicero]
14. Weiss R, Potter JS, Griffin ML, McHugh RK, Haller D, Jacobs P, et al. Reasons for opioid use among patients with dependence on prescription opioids: the role of chronic pain. J Subst Abuse Treat 2014 Aug;47(2):140-145. [doi: 10.1016/j.jsat.2014.03.004] [Medline: 24814051]
15. Cicero TJ, Ellis MS. Understanding the demand side of the prescription opioid epidemic: Does the initial source of opioids matter? Drug Alcohol Depend 2017 Apr 01;173(Suppl 1):S4-S10 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.03.014] [Medline: 28363319]
16. Strang J, Reed K, Bogdanowicz K, Bell J, van der Waal R, Keen J, et al. Randomised comparison of a novel buprenorphine oral lyophilisate versus existing buprenorphine sublingual tablets in opioid-dependent patients: a first-in-patient phase II randomised open label safety study. Eur Addict Res 2017;23(2):61-70. [doi: 10.1159/000456612] [Medline: 28268215]
17. Alper KR, Lotsof HS, Frenken GM, Luciano DJ, Bastiaans J. Treatment of acute opioid withdrawal with ibogaine. Am J Addict 1999;8(3):234-242. [doi: 10.1080/105504999305848] [Medline: 10506904]
18. Raynor D, Savage I, Knapp P, Henley J. We are the experts: people with asthma talk about their medicine information needs. Patient Educ Couns 2004 May;53(2):167-174. [doi: 10.1016/S0738-3991(03)00126-5] [Medline: 15140456]

19. Trujols J, Iraurgi I, Oviedo-Joekes E, Guàrdia-Olmos J. A critical analysis of user satisfaction surveys in addiction services: opioid maintenance treatment as a representative case study. Patient Prefer Adherence 2014;8:107-117 [FREE Full text] [doi: 10.2147/PPA.S52060] [Medline: 24482571]
20. Lavsa S, Holzworth A, Ansani N. Selection of a validated scale for measuring medication adherence. J Am Pharm Assoc 2011;51(1):90-94. [doi: 10.1331/JAPhA.2011.09154] [Medline: 21247831]
21. Key facts about uninsured adults with opioid use disorder. KFF. 2019 Jul 15. URL: https://www.kff.org/uninsured/issue-brief/key-facts-about-uninsured-adults-with-opioid-use-disorder/ [accessed 2020-02-19]
22. McRoy S, Rastegar-Mojarad M, Wang Y, Ruddy KJ, Haddad TC, Liu H. Assessing unmet information needs of breast cancer survivors: exploratory study of online health forums using text classification and retrieval. JMIR Cancer 2018 May 15;4(1):e10 [FREE Full text] [doi: 10.2196/cancer.9050] [Medline: 29764801]
23. Sinha A, Porter T, Wilson A. The use of online health forums by patients with chronic cough: qualitative study. J Med Internet Res 2018 Jan 24;20(1):e19 [FREE Full text] [doi: 10.2196/jmir.7975] [Medline: 29367181]
24. Kendal S, Kirk S, Elvey R, Catchpole R, Pryjmachuk S. How a moderated online discussion forum facilitates support for young people with eating disorders. Health Expect 2017 Feb 03;20(1):98-111 [FREE Full text] [doi: 10.1111/hex.12439] [Medline: 26725547]
25. Barney LJ, Griffiths KM, Banfield MA. Explicit and implicit information needs of people with depression: a qualitative investigation of problems reported on an online depression support forum. BMC Psychiatry 2011 May 18;11(1):88 [FREE Full text] [doi: 10.1186/1471-244x-11-88]
26. Mulveen R, Hepworth J. An interpretative phenomenological analysis of participation in a pro-anorexia internet site and its relationship with disordered eating. J Health Psychol 2006 Mar;11(2):283-296. [doi: 10.1177/1359105306061187] [Medline: 16464925]
27. Dunsch F, Evans DK, Macis M, Wang Q. Bias in patient satisfaction surveys: a threat to measuring healthcare quality. BMJ Glob Health 2018;3(2):e000694 [FREE Full text] [doi: 10.1136/bmjgh-2017-000694] [Medline: 29662696]
28. Gonzalez-Hernandez G, Sarker A, O'Connor K, Savova G. Capturing the patient's perspective: a review of advances in natural language processing of health-related text. Yearb Med Inform 2017 Aug;26(1):214-227 [FREE Full text] [doi: 10.15265/IY-2017-029] [Medline: 29063568]
29. Lo-Ciganic W, Huang JL, Zhang HH, Weiss JC, Wu Y, Kwoh CK, et al. Evaluation of machine-learning algorithms for predicting opioid overdose risk among Medicare beneficiaries with opioid prescriptions. JAMA Netw Open 2019 Mar 01;2(3):e190968 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.0968] [Medline: 30901048]
30. Ellis R, Wang Z, Genes N, Ma'ayan A. Predicting opioid dependence from electronic health records with machine learning. BioData Min 2019;12(1):3 [FREE Full text] [doi: 10.1186/s13040-019-0193-0] [Medline: 30728857]
31. Hasan M, Young GJ, Patel MR, Modestino AS, Sanchez LD, Noor-E-Alam M. A machine learning framework to predict the risk of opioid use disorder. Mach Learn Appl 2021 Dec;6:100144. [doi: 10.1016/j.mlwa.2021.100144]
32. Ward A, Jani T, De Souza E, Scheinker D, Bambos N, Anderson TA. Prediction of prolonged opioid use after surgery in adolescents: insights from machine learning. Anesth Analg 2021 Aug 01;133(2):304-313. [doi: 10.1213/ANE.0000000000005527] [Medline: 33939656]
33. 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]
34. Prieto JT, Scott K, McEwen D, Podewils LJ, Al-Tayyib A, Robinson J, et al. The detection of opioid misuse and heroin use from paramedic response documentation: machine learning for improved surveillance. J Med Internet Res 2020 Jan 03;22(1):e15645 [FREE Full text] [doi: 10.2196/15645] [Medline: 31899451]
35. Hedegaard H, Miniño AM, Warner M. Drug overdose deaths in the United States, 1999-2018. NCHS Data Brief No. 356. Centers for Disease Control and Prevention. 2020 Jan. URL: http://www.cdc.gov/nchs/data/databriefs/db356-h.pdf [accessed 2023-01-06]
36. Better information. Better health. WebMD. 2005. URL: https://www.webmd.com/ [accessed 2023-01-06]
37. Drugs.com. 2000. URL: https://www.drugs.com/ [accessed 2023-01-06]
38. Lu Y, Wu Y, Liu J, Li J, Zhang P. Understanding health care social media use from different stakeholder perspectives: a content analysis of an online health community. J Med Internet Res 2017 Apr 07;19(4):e109 [FREE Full text] [doi: 10.2196/jmir.7087] [Medline: 28389418]
39. Meschenmoser P, Meuschke N, Hotz M, Gipp B. Scraping scientific web repositories: challenges and solutions for automated content extraction. D-Lib 2016;22(9/10):1 [FREE Full text] [doi: 10.1045/september2016-meschenmoser]
40. Natural Language Toolkit. 2021. URL: https://www.nltk.org [accessed 2023-01-06]
41. May C, Cotterell R, Van Durme B. An analysis of lemmatization on topic models of morphologically rich language. arXiv. 2019 May 10. URL: https://arxiv.org/abs/1608.03995 [accessed 2023-01-06]
42. Prakash B. Patient satisfaction. J Cutan Aesthet Surg 2010 Sep;3(3):151-155 [FREE Full text] [doi: 10.4103/0974-2077.74491] [Medline: 21430827]

43. Omranian S. Information retrieval of opioid dependence medications reviews from health-related social media. Masters thesis. University of Wisconsin-Milwaukee Digital Commons. 2020. URL: https://dc.uwm.edu/cgi/viewcontent.cgi?article=3575&context=etd [accessed 2023-01-06]
44. Ramachandran KM, Tsokos CP. Mathematical Statistics with Applications in R. 2nd ed. Cambridge, MA: Academic Press; 2014.
45. Unified Medical Language System® (UMLS®). Metamap. National Library of Medicine. 2021. URL: https://www.nlm.nih.gov/research/umls/implementation_resources/metamap.html [accessed 2023-01-06]
46. Aronson A, Lang FM. An overview of MetaMap: historical perspective and recent advances. J Am Med Inform Assoc 2010;17(3):229-236 [FREE Full text] [doi: 10.1136/jamia.2009.002733] [Medline: 20442139]
47. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 2001 Oct;34(5):301-310 [FREE Full text] [doi: 10.1006/jbin.2001.1029] [Medline: 12123149]
48. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text] [doi: 10.1162/jmlr.2003.3.4-5.993]
49. Binkley D, Heinz D, Lawrie D, Overfelt J. Understanding LDA in source code analysis. 2014 Presented at: 22nd International Conference on Program Comprehension (ICPC). ICSE '14: 36th International Conference on Software Engineering; June 2-3, 2014; Hyderabad, India. [doi: 10.1145/2597008.2597150]
50. Wang YC, Kraut R, Levine J. To stay or leave?: The relationship of emotional and informational support to commitment in online health support groups. 2012 Presented at: Conference on Computer Supported Cooperative Work; February 2012; Seattle p. 11-15. [doi: 10.1145/2145204.2145329]
51. Wang S, Paul M, Dredze M. Exploring health topics in Chinese social media: an analysis of Sina Weibo. 2014 Presented at: The AAAI Workshop on the World Wide Web and Public Health Intelligence; July 27-28, 2014; Quebec, QC, Canada p. 27-28 URL: https://www.cs.jhu.edu/~mdredze/publications/2014_w3phi_weibo.pdf
52. 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]
53. Latent Dirichlet allocation. Gensim: topic modelling for humans. URL: https://radimrehurek.com/gensim/models/ldamodel.html [accessed 2023-01-06]
54. Steyvers T, Griffiths T. Probabilistic topic models. In: Landauer T, McNamara D, Dennis S, Kintsch W, editors. Latent semantic analysis: a road to meaning. Mahwah, NJ: Lawrence Erlbaum; 2007:1-15.
55. Rosner F, Hinneburg A, Röder M, Nettling M, Both A. Evaluating topic coherence measures. arXiv. 2014 Mar 25. URL: https://arxiv.org/abs/1403.6397 [accessed 2023-01-06]
56. Fishbain D, Cole B, Lewis J, Rosomoff HL, Rosomoff RS. What percentage of chronic nonmalignant pain patients exposed to chronic opioid analgesic therapy develop abuse/addiction and/or aberrant drug-related behaviors? A structured evidence-based review. Pain Med 2008;9(4):444-459. [doi: 10.1111/j.1526-4637.2007.00370.x] [Medline: 18489635]
57. Cragg A, Hau JP, Woo SA, Liu C, Doyle-Waters MM, Hohl CM. Risk factors for addiction among patients receiving prescribed opioids: a systematic review protocol. Syst Rev 2017 Dec 28;6(1):265 [FREE Full text] [doi: 10.1186/s13643-017-0642-0] [Medline: 29282114]
58. Kang H. The prevention and handling of the missing data. Korean J Anesthesiol 2013 May;64(5):402-406 [FREE Full text] [doi: 10.4097/kjae.2013.64.5.402] [Medline: 23741561]
59. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, NY: Springer; 2008.
60. Ren Y, Wu D, Singh A, Kasson E, Huang M, Cavazos-Rehg P. Automated detection of vaping-related tweets on Twitter during the 2019 EVALI outbreak using machine learning classification. Front Big Data 2022;5:770585 [FREE Full text] [doi: 10.3389/fdata.2022.770585] [Medline: 35224484]
61. Jung Y, Hu J. AK-fold Averaging cross-validation procedure. J Nonparametr Stat 2015;27(2):167-179. [doi: 10.1080/10485252.2015.1010532] [Medline: 27630515]
62. Hei H, Ma Y. Imbalanced learning: foundations, algorithms, and applications. New York, NY: The Institute of Electrical and Electronics Engineers, Inc; 2013.
63. Han J, Kamber M, Pei J. Data mining concepts and techniques 3rd Edition. Waltham: Morgan Kaufmann Publishers; 2012.
64. Marafino B, Boscardin WJ, Dudley RA. Efficient and sparse feature selection for biomedical text classification via the elastic net: Application to ICU risk stratification from nursing notes. J Biomed Inform 2015 Apr;54:114-120 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.003] [Medline: 25700665]
65. Pergolizzi JV, Raffa RB, Rosenblatt MH. Opioid withdrawal symptoms, a consequence of chronic opioid use and opioid use disorder: Current understanding and approaches to management. J Clin Pharm Ther 2020 Oct 27;45(5):892-903. [doi: 10.1111/jcpt.13114] [Medline: 31986228]
66. Huang M, Zolnoori M, Balls-Berry JE, Brockman TA, Patten CA, Yao L. Technological innovations in disease management: text mining US Patent data From 1995 to 2017. J Med Internet Res 2019 Apr 30;21(4):e13316 [FREE Full text] [doi: 10.2196/13316] [Medline: 31038462]

67. Compton W, Jones CM. Epidemiology of the U.S. opioid crisis: the importance of the vector. Ann N Y Acad Sci 2019 Sep;1451(1):130-143. [doi: 10.1111/nyas.14209] [Medline: 31378974]

Abbreviations

AUC: area under the curve
CDC: Center for Disease Control
CV: cross-validation
LASSO: least absolute shrinkage and selection operator
LDA: latent Dirichlet allocation
MAT: medication-assisted treatment
NLP: natural language processing
OUD: opioid use disorder
UMLS: Unified Medical Language System
XGBoost: extreme gradient boosting

Edited by M Meacham; submitted 11.02.22; peer-reviewed by R Pryss, H Xu; comments to author 11.05.22; revised version received 06.07.22; accepted 30.12.22; published 23.01.23.

Please cite as:
Omranian S, Zolnoori M, Huang M, Campos-Castillo C, McRoy S
Predicting Patient Satisfaction With Medications for Treating Opioid Use Disorder: Case Study Applying Natural Language Processing to Reviews of Methadone and Buprenorphine/Naloxone on Health-Related Social Media
JMIR Infodemiology 2023;3:e37207
URL: https://infodemiology.jmir.org/2023/1/e37207
doi:10.2196/37207
PMID:37113381

Original Paper

Influence of User Profile Attributes on e-Cigarette–Related Searches on YouTube: Machine Learning Clustering and Classification

Dhiraj Murthy[1], PhD; Juhan Lee[2], PhD; Hassan Dashtian[1], PhD; Grace Kong[2], PhD

[1]Computational Media Lab, School of Journalism and Media, Moody College of Communication, The University of Texas at Austin, Austin, TX, United States

[2]Department of Psychiatry, Yale School of Medicine, New Haven, CT, United States

Corresponding Author:
Dhiraj Murthy, PhD
Computational Media Lab, School of Journalism and Media
Moody College of Communication
The University of Texas at Austin
300 W Dean Keeton (A0900)
Austin, TX, 78712
United States
Phone: 1 512 471 5775
Email: dhiraj.murthy@austin.utexas.edu

Abstract

Background: The proliferation of e-cigarette content on YouTube is concerning because of its possible effect on youth use behaviors. YouTube has a personalized search and recommendation algorithm that derives attributes from a user's profile, such as age and sex. However, little is known about whether e-cigarette content is shown differently based on user characteristics.

Objective: The aim of this study was to understand the influence of age and sex attributes of user profiles on e-cigarette–related YouTube search results.

Methods: We created 16 fictitious YouTube profiles with ages of 16 and 24 years, sex (female and male), and ethnicity/race to search for 18 e-cigarette–related search terms. We used unsupervised (k-means clustering and classification) and supervised (graph convolutional network) machine learning and network analysis to characterize the variation in the search results of each profile. We further examined whether user attributes may play a role in e-cigarette–related content exposure by using networks and degree centrality.

Results: We analyzed 4201 nonduplicate videos. Our k-means clustering suggested that the videos could be clustered into 3 categories. The graph convolutional network achieved high accuracy (0.72). Videos were classified based on content into 4 categories: product review (49.3%), health information (15.1%), instruction (26.9%), and other (8.5%). Underage users were exposed mostly to instructional videos (37.5%), with some indication that more female 16-year-old profiles were exposed to this content, while young adult age groups (24 years) were exposed mostly to product review videos (39.2%).

Conclusions: Our results indicate that demographic attributes factor into YouTube's algorithmic systems in the context of e-cigarette–related queries on YouTube. Specifically, differences in the age and sex attributes of user profiles do result in variance in both the videos presented in YouTube search results as well as in the types of these videos. We find that underage profiles were exposed to e-cigarette content despite YouTube's age-restriction policy that ostensibly prohibits certain e-cigarette content. Greater enforcement of policies to restrict youth access to e-cigarette content is needed.

(JMIR Infodemiology 2023;3:e42218) doi:10.2196/42218

KEYWORDS

electronic cigarettes; electronic nicotine delivery systems; ENDS; tobacco products; YouTube; social media; minority groups; exposure; youth; behavior; user; machine learning; policy

Introduction

Nicotine exposure through e-cigarettes, particularly during adolescence, poses negative health outcomes such as brain maldevelopment and subsequent substance use [1]. In 2022, 9.4% (representing 2,550,000 students) of US middle and high school students reported using e-cigarettes in the past 30 days [1]. e-Cigarettes are also popular among adults (5.1% used them in the past 30 days in 2020), but these are most often used by young adults (15.6%) [2,3]. e-Cigarette use among adolescents and young adults (referred to as “youth” from here onward) may be driven, in part, by its heavy presence and positive portrayal on social media [4,5]. There is accumulating literature documenting e-cigarette promotion on social media. e-Cigarettes are portrayed on social media as fashionable, acceptable, and cool [6]. There are also themes that specifically appeal to youth, such as cartoon-based images on Instagram [7] and vape tricks (ie, blowing large vapor clouds or shapes like rings) on YouTube [8]. Studies have suggested that positive perceptions such as e-cigarette use being socially acceptable is related to its use among youth [9,10]. Studies have also shown that such positive portrayals of e-cigarettes on social media platforms have contributed to youth appeal and use behaviors [11]. For example, Lee et al [12] used state-level population data and found that the daily use of social media platforms, namely, Instagram, Snapchat, Facebook, and Twitter, was associated with e-cigarette use among adolescents, suggesting that youth may be exposed to e-cigarette–related information on social media. Given the high rate of social media usage by youth [13] and the unregulated environment [14], surveillance of e-cigarette–related content on social media platforms is warranted.

Social media platforms custom-tailor content to user characteristics [15]. However, these algorithms are proprietary, and it is unclear how information regarding e-cigarettes is featured to youth users. In this study, we examined how user profile attributes (ie, age and sex) influence the e-cigarette content being shown to youth users on YouTube—an online video streaming social media platform that has more than 2 billion users and is viewed more than 1 billion hours/day [16]. Users can upload and watch videos on YouTube and interact with other users by posting comments, reacting to videos (like/dislike), sharing content, and subscribing to YouTube channels. YouTube was the most frequently used social media platform in 2021, with 81% of the social media users reporting having used the platform [17]. Moreover, YouTube is the most popular platform among youth [8].

e-Cigarette content is prolific on YouTube. For instance, Huang et al [18] analyzed 28,000 e-cigarette–related YouTube videos and found that those videos had received more than 100 million views, indicating high engagement by users [18]. Further, e-cigarettes are frequently positively portrayed on YouTube and pro–e-cigarette videos are commonly sponsored by the e-cigarette industry [19]; 85% of the e-cigarette–related videos on YouTube are sponsored by e-cigarette marketers, including e-cigarette companies or people endorsing e-cigarette companies [20]. Pro–e-cigarette videos include portrayals of e-cigarettes as safer, cleaner, and less malodorous than combustible cigarettes [21]; videos showcasing or teaching how to conduct vape tricks (ie, using e-cigarettes to blow large, thick amounts of exhaled aerosol or shapes) [8]; modification of e-cigarette devices for unintended purposes such as increasing the temperature and using other substances in it [19,22,23]; instructions on how to use e-cigarettes (eg, how to puff) [24]; product reviews [25]; and health information or misinformation about e-cigarette use [26]. Concerningly, these e-cigarette contents are readily available on YouTube without a warning label/statement [27], and these videos are easily accessible to youth [4]. In sum, there are abundant e-cigarette–related videos on YouTube. However, less known is specifically what content youth are exposed to. All users do not receive the same results when they search for the same terms on YouTube. This is partially due to YouTube’s personalized search and recommendation algorithms, which consider, to some extent, a user’s age, sex, and the history of the searched items by that specific user [28,29].

YouTube’s search and recommendation algorithms are responsible for creating personalized content for users from an ever-growing collection of videos. Similar to other social networks, YouTube has undergone a paradigm shift toward using deep machine learning—systems based on artificial neural networks—as a solution for scaling the systems used by YouTube’s search and recommendation algorithms [30]. However, the opaque nature of the search and recommendation algorithms poses questions concerning whether algorithmic visibility can be evaluated. Search and recommendation algorithms may be developed to take viewers’ demographic profiles (eg, age, sex) as inputs in determining what search results users receive. Therefore, YouTube’s search and recommendation algorithms have important public health implications. For instance, males have consistently shown a higher level of e-cigarette use among adolescents and adults [31], and data suggest that e-cigarette–related videos such as vape tricks videos mostly feature young men and seem to be targeting this population [8]. A recent study identified that e-cigarette content on YouTube contained themes related to product reviews (provide reviews of e-cigarette products), instructional videos (teach viewers how to use, modify, or create e-cigarette products), health information (provide health information related to e-cigarettes), vape tricks (feature different vape tricks), cannabis (cannabis vaping–related topics), and other (a variety of other themes such as news clips related to e-cigarette use) [19]. However, less known is whether these video themes are differentially exposed to users by their demographic attributes. Such information is important to inform tobacco regulatory actions in restricting marketing that targets at-risk populations such as underage youth, and it can be used to inform how prevention strategies such as countermarketing can be targeted to these populations.

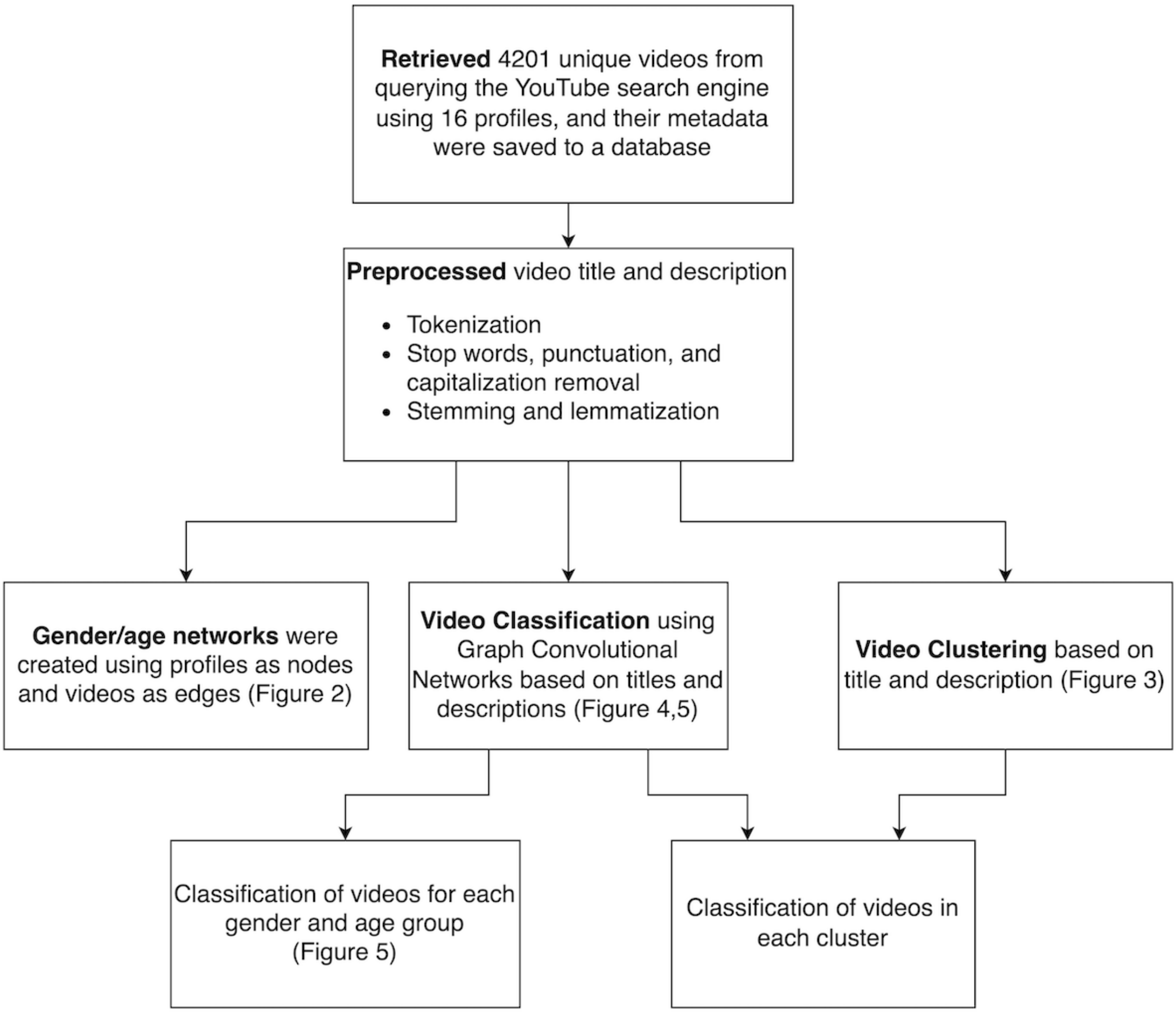
Methods

Overview

The goal of this study was to understand the role of the demographic factors (ie, sex, age) of YouTube users’ profiles in influencing the variations in e-cigarette–related search results presented to users. To accomplish this goal, we developed a

3-step framework, which combined computational methods and human labeling (Figure 1). First, we used an unsupervised machine learning method, the k-means method, which used the distribution of words in the video data (ie, titles and descriptions) to cluster the videos into themes. Human-labeled data sets using titles and descriptions of the YouTube videos were then used to confirm the themes identified in our k-means clustering results. Second, we used this human-labeled data set to train a supervised machine learning method, that is, the graph convolutional network (GCN), to classify all the videos in our data set based on their identified themes. Finally, we performed unsupervised network analysis to measure how YouTube video results varied by user attributes (ie, age and sex). We examined whether there were differences in the video themes between different age and sex profiles. The application of these machine learning–based methods is novel in tobacco regulatory science work using social media data. Our approach is also scalable to large volumes of data and can be extended to a variety of social media platforms.

Figure 1. Overall framework of data collection, preprocessing, and analytics.



Ethical Considerations

This research is not deemed as human subjects research according to the definition provided by the Office of Human Research Protections, US Department of Health and Human Services. We examined publicly available data, and we did not report any identifying information of the content observed on social media. Additionally, this observational study was deemed exempt as human subjects research by the Yale Institutional Review Board (HIC 2000028350).

Search Methods

We created 16 fictitious profiles on YouTube that sought to vary and reflect particular demographic attributes (ie, age, sex, and race) [32]. Profile photos were not added. To attempt to reflect particular racial and ethnic attributes, we created profiles by using common African American, Hispanic, and White first and last names by using existing name data [33]. The profiles consisted of African American females and males aged 16 and 24 years (4 profiles), Hispanic females and males aged 16 and 24 years (4 profiles), and 2 sets of White females and males aged 16 and 24 years (8 profiles). We oversampled White users to be more reflective of the e-cigarette use population. To create each fictitious profile on YouTube, we used a new SIM card and phone number and performed a factory reset of an Android phone. Sex and age were entered during each fictitious account creation. No other demographic metadata were included during account creation. We used a mobile phone rather than a web browser to conduct our searches to best replicate how youth access YouTube content [13].

During the course of a week in June 2020, we collected data for 2-3 profiles per day. Once we collected 140 videos per profile per search term, we factory reset the Android phone and moved to the next profile. For each profile, the following keywords were searched for each profile by using Orbot, a mobile app that allows one to use an anonymized Tor bridge (to avoid location or IP address personalization): box mods, cigalikes, disposable e-cigs, disposables, disposable vape, e-cig, e-cigarette, e-juice, electronic cigarette, e-liquid, ENDS, pod mods, vape, vaping, vape juice, vape mods, vape pens, vape pods. Studies typically examine the first page [18,20,34] of the search results on YouTube, which has 20 videos, or the first 2 pages, which has 40 videos. However, some users may search through multiple pages if they do not find what they are looking for in the first few pages. Thus, for each of our 16 fictitious YouTube profiles, we searched through 7 pages (140 videos) for each of our keywords (n=5875). This approach is therefore far more aggressive than previous work [35]. After removing duplicates (n=1674), we arrived at the final sample (N=4201) of unique videos. We collected video metadata such as title, description, transcript, view counts, likes/dislikes, comments, date published, channel name, and category. The methods are further explained in Dashtian et al [32].

Preprocessing Data

We converted the text into numerical form so that we can apply machine learning algorithms such as clustering and classification to the data. The preprocessing steps included tokenization, stop words removal, stemming, and lemmatization. Tokenization is the process of splitting a set of texts into words (also called tokens) and then removing certain characters such as blank sequences and punctuation. Stop words are usually frequent in English text (eg, a, an, the, that, I, be, other, with). The goal of both stemming and lemmatization is to find the base form of a word from its inflectional forms and derivatives (eg, vaped, vaping have a base of vape). We used Porter stemmer, an algorithm which has been successfully used by others for the stemming of health-related texts for machine learning purposes [36].

Video Clustering (Unsupervised Machine Learning)

K-means automatically arranges texts into clusters such that text data within clusters are relatively similar in terms of content when compared to text data in other clusters [37]. Another health-related work [38] has successfully used the k-means clustering algorithm for automated text classification. We therefore chose to use k-means to categorize video types. In our case, the input to the k-means clustering is preprocessed text (video title and the description provided by the uploader to describe the video). We used the elbow method to find the optimum value for the number of clusters (k). The elbow method provides a good indication that the underlying model and number of (k) fits best at that point and has been successfully used in other health-related machine learning studies [39]. We examined the results visually to discern the point at which diminishing returns are observed (ie, an elbow appears). K-means seeks to cluster around optimal centroids (ie, cluster centers). The best placement of initial centroid positions is a standard method for maximizing the k-means clustering process. To avoid any bias, we randomly selected initial centroids and iterated the algorithm several times for each k to confirm that the initial centroids do not change our optimized clustering results. We measured cosine similarity to generate a measure of similarity between each video and the other videos in the search results. Cosine similarity is a measure mostly used for k-means clustering of text documents. The distance matrix was then converted into a 2D array by using multidimensional scaling.

Video Classification (Human Labeling)

Members of the research team with expertise in e-cigarettes randomly selected videos from the full corpus of the collected videos (n=1000) [19] and labeled the videos by the following classes: (1) product review (ie, an individual(s) providing a review of an e-cigarette product), (2) health information (ie, health information related to e-cigarette use), (3) instructional (ie, a tutorial on how to use an e-cigarette or how to modify an e-cigarette), and (4) other (which consists of a variety of other themes, including cannabis, television/news clips, vape tricks). Interrater reliability (Cohen κ) was 0.93, indicating “almost perfect” agreement between coders. These categories were used in previous research [32]. Please refer to Kong et al [19] for more information on how these themes were determined and labeled.

Text Classification Using GCNs (Supervised Machine Learning)

We used GCN, which is a supervised machine learning method, to classify data (ie, titles and descriptions) by theme to better understand the unique clusters identified through k-means clustering. In GCN, word frequency and word co-occurrence information are used to build the word-to-word and word-to-video edges (ie, as common videos between pairs), respectively. We also classified the nodes (ie, entities in the network) instead of the actual videos. The entities in the network represented just the nodes in the graph. These do not refer to the themes. GCN has shown strong performance for classification with a small portion of labeled data similar to the data used in our study [40].

To model the global word co-occurrence, we built a large 2-mode graph (ie, 2 types of nodes). Our graph contains word nodes (which represent single words) and document nodes (which represent whole documents with many words). See Multimedia Appendix 1 for a visual rendering of the relationship between the document nodes and word nodes. Specifically, the first mode of nodes consists of words and the second mode of nodes consists of documents with titles and descriptions (ie, with many words). One document represents 1 video (title and description together). Document nodes and word nodes are interconnected and intraconnected. The number of nodes in the text graph |V| is the number of documents (document nodes) plus the number of unique words in the documents (word nodes). We set feature matrix X = I as an identity matrix, which means every word or document is represented as a 1-hot vector as input to text GCN. One-hot encoding converts categorical data into binary values suitable for machine learning algorithms. We build edges (ie, connections) between nodes based on word occurrence in documents (document-word edges) and word co-occurrence in the whole corpus (word-word edges). The weight of the edge between a document node and a word node is the term frequency-inverse document frequency of the word in the document. Term frequency is the number of times the word appears in the document, and inverse document frequency is the logarithmically scaled inverse fraction of the number of documents that contain the word. After performing clustering and classification on preprocessed data, we calculated the percentage of each video type (derived from classification) in each category (derived from clustering).

Profiles Network

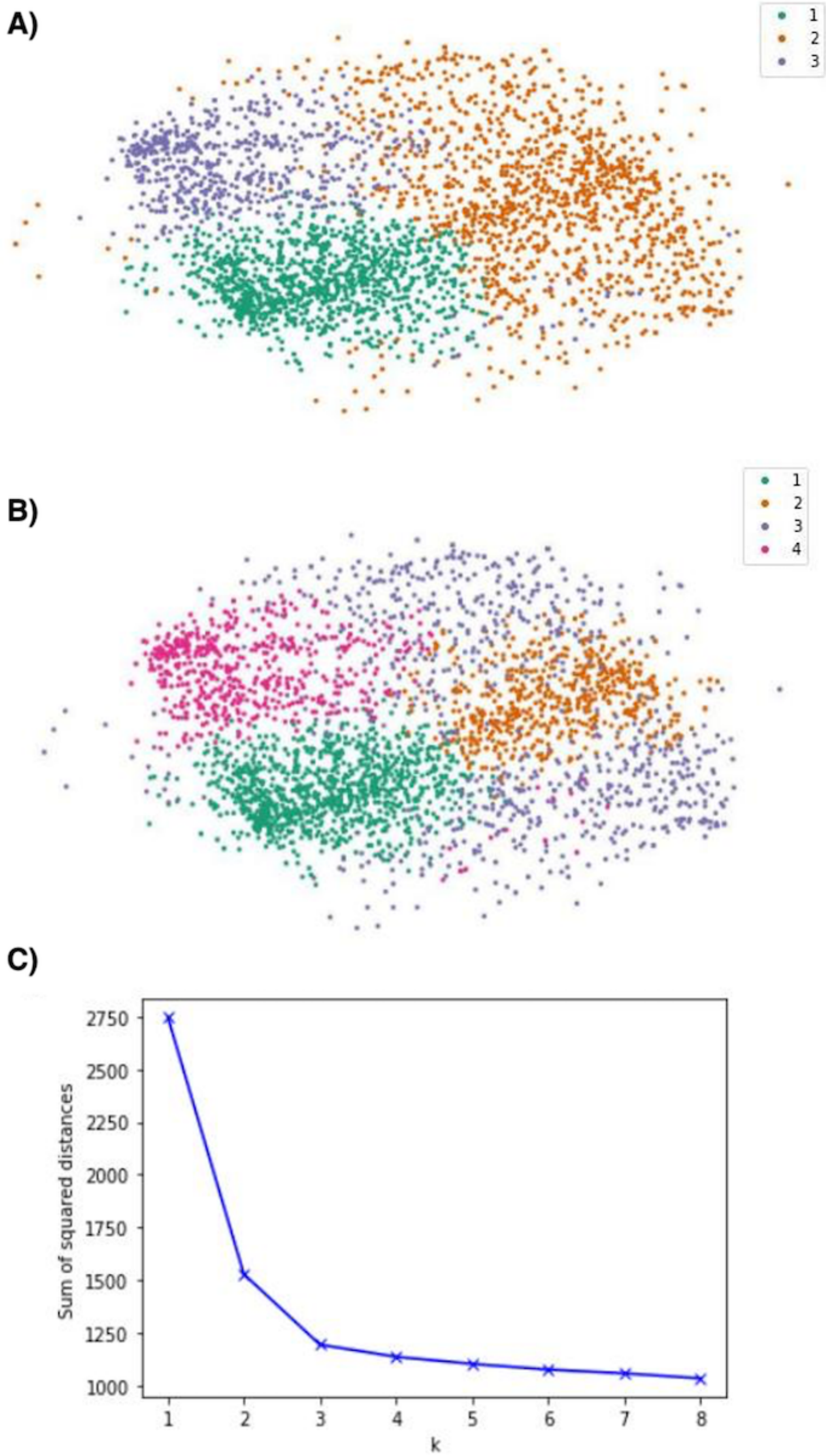
The frequency of common videos between different ages and sexes can be used as a measure to quantify the strength of the relationships between these variables. For example, the overlap of videos among the same sex and age profiles can be used to discern whether users with these attributes (eg, both female and male, adolescents or young adults) receive similar information from YouTube's search engine. Furthermore, the connections between nodes in a network provide information about the structure of the network. We can also use the number of connections of a node in each demographic group to identify the most influential nodes in the network. Specifically, the network of 4 demographic groups can be represented as nodes with their edges representing common videos between pairs of groups. To show the connections, we plotted a line between two groups and calculated the number of common videos between them. Lines with a larger value represent more common videos between a pair than lines with smaller values. We assessed 2 separate networks: one with common videos between age and sex and another that assessed a combination of the two.

Results

Video Clustering (Unsupervised Machine Learning)

To better understand which content shows up for different demographic profiles, we identified the types of videos in our data set by using k-means to cluster videos. Figures 2A and B illustrate the video clusters as 3 clusters and 4 clusters, respectively. The former had 3 distinct topical clusters, whereas the latter had 3 distinct topical clusters and 1 diffuse cluster (that likely represents the "other" content cluster). The elbow method indicated that the plateau (ie, the first stable k value in the sum of squared distances) is at k=3 (Figure 2C). In some cases, the elbow method has ambiguity [41]. However, in our case, we had a clear result that videos can be automatically clustered into 3 main clusters.

Figure 2. Each dot in (A)/(B) indicates a video and each color represents a cluster. (A) Videos categorized into 3 clusters. (B) Videos categorized into 4 clusters. (C) Elbow method results, which show the sum of the squared distances as a function of the number of clusters (k).



Video Classification (Human Labeling)

Human labeling identified 3 distinct classes: (1) product reviews, (2) instructional, and (3) health information. We also included a fourth catchall class of "others" for any videos that did not fit into the other 3 distinct classes. Product reviews are videos that provide reviews of e-cigarette products, instructional videos provide instructions on how to use/modify/create e-cigarette products, health information videos provide information on the health risk of e-cigarettes, and other videos are topics that do not fall into these 3 classes and include a range of topics such as cannabis and vape tricks. We found that GCN was able to successfully classify videos based on these 3 distinct classes as well as a separate "other" class. Overall, product review was the most common type of videos identified (49.3%), followed by instructional (26.9%), health information (15.1%), and other (eg, cannabis, television/news clip) (8.5%). We further estimated the prevalence of each video type exposure by demographic attributes (Figure 3). For all demographic groups, except the 16-year-old group, product review videos showed the highest percentage in the search results, followed by instructional videos. Instructional videos showed the highest percentage in the search results of 16-year-old students. We estimated the prevalence of video themes separated by age and sex (Figure 4). The product review label was the dominant class for 24-year-old male (39.4%) and 24-year-old female (38%) profiles. Instructional videos showed the highest percentage in the search results of 16-year-old female (42.5%) and 16-year-old male (30.9%) profiles; notably, the 16-year-old female profile had the highest percentage of search results for this label. All profiles were least exposed to health information videos.

Figure 3. Prevalence of video type shown, split by demographic variables. The percentage of each label (class) is shown based on the results from graph convolutional networks. TV: television.

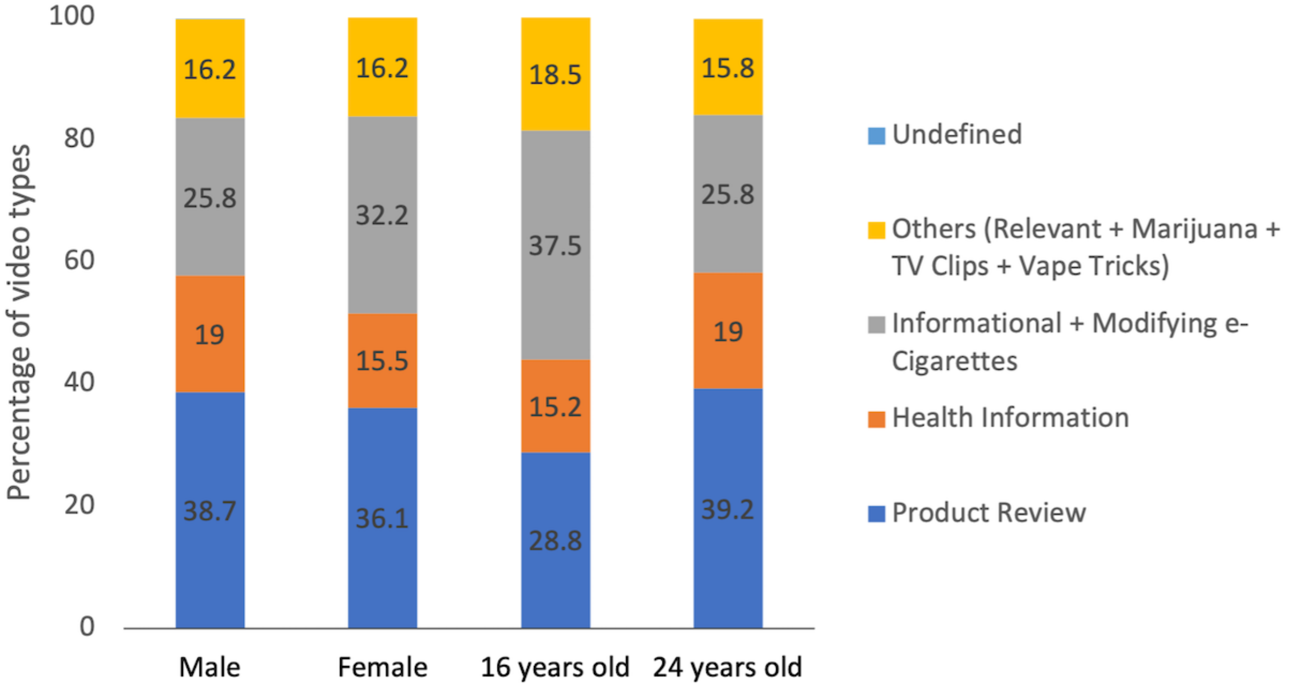
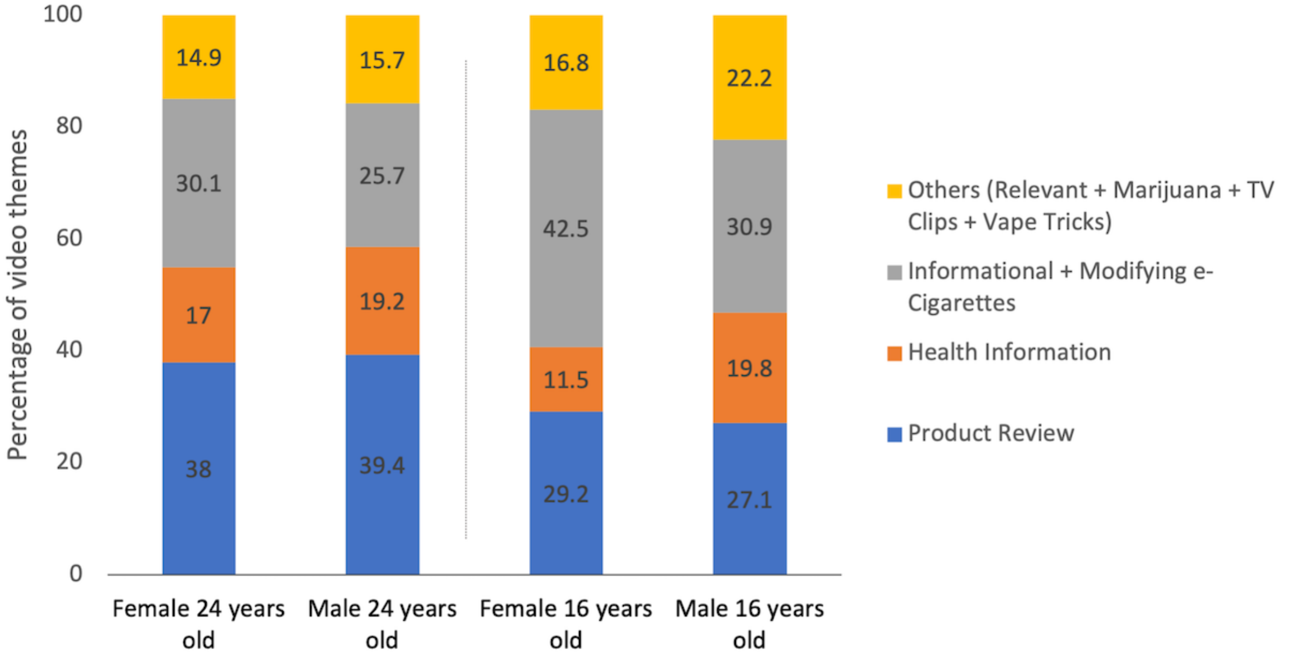


Figure 4. Results of the classification of videos in each demographic group. We grouped YouTube profiles based on age (24 or 16 years old) and sex (male and female). TV: television.



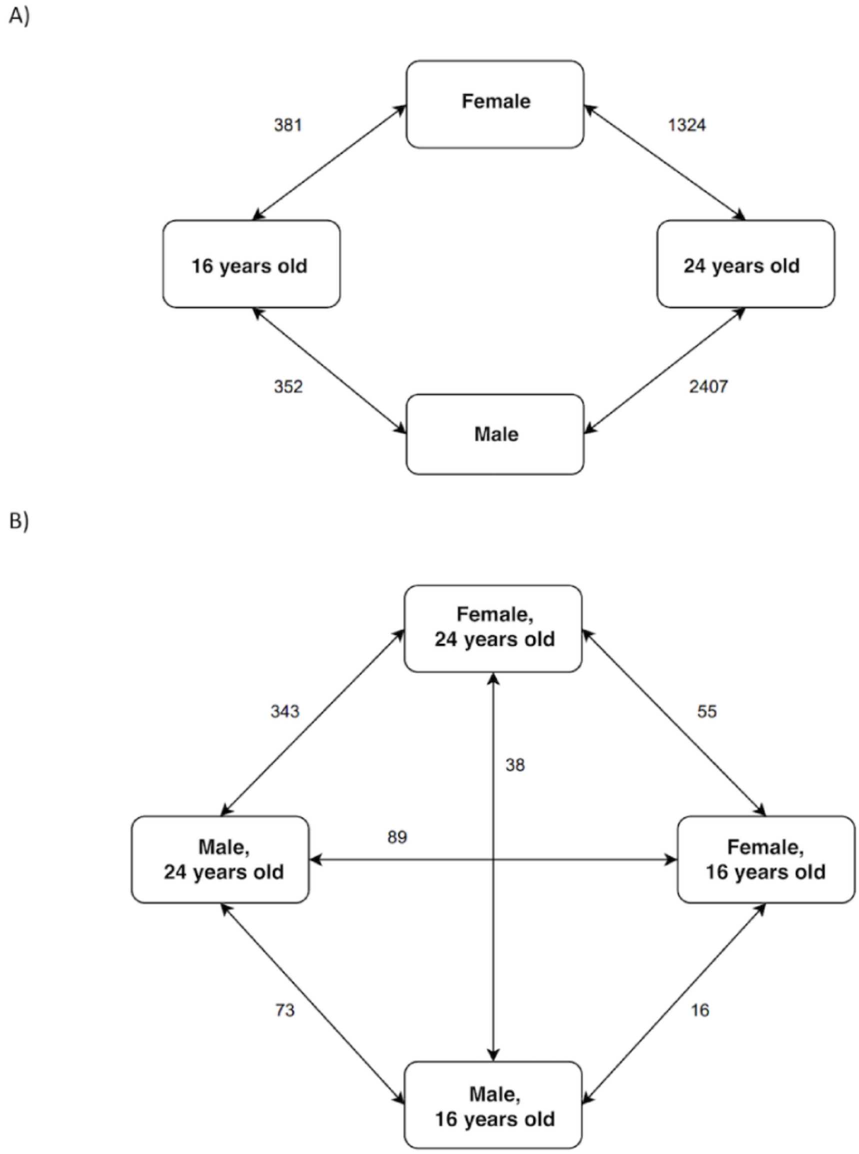
Text Classification Using GCNs (Supervised Machine Learning)

We used text classification using GCN, a supervised machine learning technique, to classify the text of video titles and description into human-labeled classes (ie, product review, health information, instructional, other). We found that the accuracy of the GCN model for the classification of e-cigarette–related YouTube videos is 0.72 for the parameters that we set. The precision, recall, and F1-score values were 0.70, 0.78, and 0.74, respectively.

Profiles Network

The connections between the profile groups based on the common videos that were retrieved from the YouTube search are shown in Figure 5. The number of common edges between 16-year-old and 24-year-old pairs was the lowest among the other pairs. As shown in Figure 5A, the connection between the nodes of 24 years old and male is very strong, as indicated by the edge weight of n=2407 (ie, the number of common videos). We also constructed another network by using a combination of age and sex. The videos of all the profiles were grouped into 4 subsets: 24-year-old male, 16-year-old male, 24-year-old female, and 16-year-old female. Similar to that in the previous network, each node in the network represents one of these groups, and common videos between pairs of groups are shown as an edge. Compared to the previous network (Figure 5A), the network of combined age and sex (Figure 5B) had fewer edges (connections). When we examined the network of age and sex together, we imposed further restrictions on the videos that belonged to a specific node. Thus, the number of videos and therefore, the number of connections between nodes in the network of age and sex was smaller than that of age or sex alone. Figure 5B shows that 24-year-old male and 24-year-old female profiles have the highest number of common edges, while 16-year-old male and 16-year-old female profiles have the lowest number of common edges.

Figure 5. Network of demographic attributes and videos. Edge weights are provided next to the edge line between 2 pairs, and these edge weight values indicate the number of common videos between 2 corresponding nodes (ie, between the demographic attributes of sex and age).



Discussion

Principal Findings

In this study, we examined how YouTube profile attributes, specifically age and sex, affected e-cigarette–related YouTube search results. Our profile network analysis indicated that there were more common videos between male and female 24-year-old profiles relative to other demographic groupings. Using our own human-labeled data, we developed a GCN machine learning model that was able to classify the videos into 4 main classes. We found that the highest proportion of younger age groups (16 years old) was exposed to instructional videos (37.5%), while the highest proportion of young adult age groups (24 years old) were exposed to product review videos (39.2%). Additionally, the group with the highest proportion of exposure to instructional videos was 16-year-old females relative to other age/sex pairs. Our findings are consistent with prior studies that observed that common video themes related to e-cigarettes on YouTube were product reviews and instructional videos on how to use/modify/create videos [8,19,22,25]. However, our results uniquely contribute to the literature by demonstrating that demographic attributes factor into YouTube's algorithmic systems in how video themes are differentially shown to profiles with different age/gender attributes.

It is unclear what drives the differences in exposure to e-cigarette content and the volume of this content among different demographic profiles. Previous studies have shown that age and sex affect the results delivered in search engines (eg, Google) [15]. Our findings are consistent with other research that indicate that YouTube also may use demographic information to provide the most relevant information to users [29]. Specifically, Hussein et al [29] found that once a user develops a watch history in the YouTube search engine, the demographic attributes do affect the extent of content recommended to them. However, in this study, we used the same search words between each profile and used a mobile phone that was factory reset after each profile's searches were conducted to prevent tailoring of search results. It is therefore unlikely that these factors account for differences in exposure to e-cigarette content. It appears that YouTube's search engines and recommendation algorithms are driven by the demographic factors of its users. Personalization of search engines, where individual users receive distinct results for the same search query, has also led to public concerns about the so-called "filter bubble" effects [42], where users are unable to access diverse information that a search engine's algorithm decides is irrelevant to a user [43]. Our results indicate that there might be differences in the type of exposure specific to e-cigarettes that are provided to different demographic groups. We further break down these differences in terms of age and sex attributes.

Our network of search results, which shows the influence of age and sex on search results, indicates a noteworthy difference between the number of edges (common videos) for various pairs of nodes (common videos between 2 groups) in the network, including male/female and 16-year-old and 24-year-old profiles. For example, the videos common to both 16-year-old and 24-year-old groups are the lowest. However, the second network analysis showed that 24-year-old male group and 24-year-old female group pairs have the highest number of common videos. There is a greater number of edges between the male group and 24-year-old group than between the female and the 24-year-old group, indicating that males and 24-year-old groups have more common videos than females and 24-year-old groups. These results indicate that 24-year-old profiles are most exposed to e-cigarette content, and this exposure is greater among 24-year-old male groups compared to their female counterparts.

Our finding that e-cigarette content is mostly available to male young adult groups is consistent with research findings that show that e-cigarette–related videos on YouTube feature more males. For instance, an examination of vape tricks on YouTube showed that 80% of the vape tricks videos featured young adult males [8]. There is also research showing that males are more engaged with YouTube content than females. Khan [44] found that male users are more likely to read comments on YouTube; Molyneaux et al [45] found that there was a greater number of comments posted by male users. Perhaps, the high engagement of males on social media platforms such as YouTube can explain the higher e-cigarette use rates among males. A review on e-cigarette use behaviors among adolescents showed that e-cigarettes are used more by male adolescents than by female adolescents [46], and national data also show that e-cigarette use is higher among male adolescents and young adults [47]. However, it is important to also highlight that e-cigarette use among females is also high: up to 20% of females in middle and high school surveyed in a study in 2020 were found to use e-cigarettes [47]. It is possible that females are engaging with e-cigarette–related social media content but doing so differently from males. For instance, there was no difference between males and females in viewing YouTube videos [44] or in the rating of YouTube videos [45], suggesting that females are engaging with YouTube content similarly as males.

The lower number of e-cigarette–related videos shown to 16-year-old profiles than 24-year-old profiles may be due, in part, to the age-restriction process of e-cigarette–related content by YouTube. YouTube's current policy prohibits tobacco-related advertisements. YouTube considers content that "promotes a product that contains drugs, nicotine …" as age-restricted content [16]. They exemplified "a video reviewing brands of nicotine e-liquid" as an example of age-restricted content. This rule may explain why we observed more product review videos in the 24-year-old group (39.2%) compared to those in the 16-year-old group (28.8%). This finding also suggests that despite these self-imposed limits on e-cigarette promotional content on YouTube, there is evidence that these restrictions may be loosely implemented and content that are restricted may be shown to underage minors on this and other social media platforms [14,19,48]. It is noteworthy that in our study, the 16-year-old profiles were exposed to e-cigarette content despite YouTube's age-restriction policy that prohibits certain e-cigarette content such as product reviews. This finding is consistent with that in other studies that found that e-cigarette content such as vape tricks were readily available using non–age-verified accounts [8]. This study highlights the importance of strong policies and the enforcement of these policies to prohibit the exposure of e-cigarette–related videos to youth on YouTube. This finding also suggests that young adults are the highest consumers of e-cigarettes among adults [49]; they may search for more information about e-cigarette products to purchase through product reviews.

Concerningly, the instruction label was observed in the highest percentage (37.5%) of search results of the 16-year-old group, and exposure to instructional videos among 16-year-old female profiles was particularly high (42.5%), suggesting that underage youth are more exposed to instructional videos, which may provide tutorials on e-cigarette use. Further, instructional videos include other content such as how to hack or modify the device to use for unintended purpose as well as to use cannabis [22,23]. The high prevalence of modification of e-cigarette content on YouTube has been shown in other studies. For instance, Massey et al [23] analyzed 168 e-cigarette–related YouTube videos and found that 20.2% of the videos were modifications of e-liquids to using cannabis. Future studies should identify whether youth modify/hack e-cigarettes and the health implications of engaging in such behaviors.

Future Work and Limitations

Several limitations in this study are noteworthy. First, we might have missed potential search terms related to e-cigarettes. For example, these may include brand-specific terms (eg, Juuling) and e-cigarette use–related slang (eg, stick). Thus, our collected videos may not represent an exhaustive list of e-cigarette–related videos. However, our study uses 18 e-cigarette–related search terms that were successfully tested and used to collect a broad range of e-cigarette–related YouTube videos [32]. Second, due to a limited number of fictitious profiles, our findings do have limits in terms of generalizability. Third, we included race/ethnicity as an element when creating profiles (ie, White, African American, and Hispanic) to be inclusive of diverse racial backgrounds. The first and last names of each profile were randomly selected by choosing names from existing data sets that were shown to be most commonly associated with a specific race/ethnicity [33,50]. However, as we created a limited number of fictitious profiles, we did not have enough data points for each race/ethnicity to incorporate machine learning to determine whether search results differed by race/ethnicity. Fourth, we only used 2 age groups (ie, 16 and 24 years), and it is possible that the search results may be different if younger or older age groups were used. Future research should therefore place an emphasis on assessing whether race/ethnicity as well as other factors (eg, viewing history, age) has an effect on search results related to e-cigarettes on YouTube. Fifth, anonymous Tor-based IP addresses may have influenced our search results; therefore, results may differ if searches were to be conducted using nonanonymized IP addresses. There may be other factors that may drive results, such as the date/time of searches as well as what content is popular on YouTube at a given time. Sixth, we cannot confirm how, whether, or to what extent YouTube's personalized search parameters read the demographic attributes (ie, age, sex, and race) that we populated our fictitious profiles with because the algorithm is proprietary. However, we used a factory-reset Android device without any search history or cookies to avoid any implicit bias in the results. The searches were conducted using the same terms to ensure that the differences between profiles, from our vantage point, are only the demographic characteristics. Nevertheless, as we only used search results collected from a mobile device, future work can explore whether web-based results are different. Seventh, we applied our methods, that is, natural language processing, video classification, and network modeling to only a single platform, that is, YouTube. Future studies would therefore benefit from extending our methodological framework to other social media platforms. Eighth, given that after we collected 140 videos per profile per search term, we factory reset our phone and moved to the next profile; our approach does not emulate or reflect the high levels of personalization that a user who uses YouTube everyday might experience. Future studies would therefore benefit by comparing our results from collecting data from YouTube in 1 setting with fictitious profile data collection done over a longer period and with some levels of variation. Ninth, we did not undertake statistical tests comparing the proportions of content classification by profile demographics nor were we able to determine how each theme was manifested by demographic attributes (eg, was health information present more for male profiles than female profiles?). Future work could make these comparisons based on the classes identified by the GCN analysis and determine how and why content themes vary by different profile attributes. Lastly, as we did not have a control group in our data collection methods, future work would benefit from the use of a control group and the examination of some of these variables.

Conclusion

Our findings underscore the value of machine learning methods in studying how profile attributes on YouTube may influence e-cigarette–related content and move the field forward by highlighting the critical need to take into consideration how social media algorithms work in practice. We used unsupervised (k-means clustering) and supervised (GCN classification) machine learning models in combination with network models to study the variation of e-cigarette–related videos on YouTube. Our methods were designed to specifically identify the similarities and differences in the videos by using selected demographic attributes, that is, age and sex. Collectively, our results suggest that advanced computational methods can be used to help understand how YouTube's current search and recommendation algorithm customizes e-cigarette–related content based on demographic attributes such as sex and age. This suggests an urgent need for surveillance and prohibition of e-cigarette–related content on social media such as YouTube to prevent e-cigarette use among youth.

Acknowledgments

This study was supported by the National Institutes of Health's National Institute on Drug Abuse (award R01DA049878).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Document node and word node composition and relationship.

[DOCX File , 483 KB - infodemiology_v3i1e42218_app1.docx]

References

1. Yuan M, Cross SJ, Loughlin SE, Leslie FM. Nicotine and the adolescent brain. J Physiol 2015 Jun 23;593(16):3397-3412. [doi: 10.1113/jp270492]
2. Gentzke AS, Wang TW, Cornelius M, Park-Lee E, Ren C, Sawdey MD, et al. Tobacco Product Use and Associated Factors Among Middle and High School Students - National Youth Tobacco Survey, United States, 2021. MMWR Surveill Summ 2022 Mar 11;71(5):1-29 [FREE Full text] [doi: 10.15585/mmwr.ss7105a1] [Medline: 35271557]
3. Boakye E, Osuji N, Erhabor J, Obisesan O, Osei AD, Mirbolouk M, et al. Assessment of Patterns in e-Cigarette Use Among Adults in the US, 2017-2020. JAMA Netw Open 2022 Jul 01;5(7):e2223266 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.23266] [Medline: 35867060]
4. O'Brien EK, Hoffman L, Navarro MA, Ganz O. Social media use by leading US e-cigarette, cigarette, smokeless tobacco, cigar and hookah brands. Tob Control 2020 Dec;29(e1):e87-e97. [doi: 10.1136/tobaccocontrol-2019-055406] [Medline: 32217772]
5. Amin S, Dunn AG, Laranjo L. Social Influence in the Uptake and Use of Electronic Cigarettes: A Systematic Review. Am J Prev Med 2020 Jan;58(1):129-141. [doi: 10.1016/j.amepre.2019.08.023] [Medline: 31761515]
6. Sapru S, Vardhan M, Li Q, Guo Y, Li X, Saxena D. E-cigarettes use in the United States: reasons for use, perceptions, and effects on health. BMC Public Health 2020 Oct 09;20(1):1518 [FREE Full text] [doi: 10.1186/s12889-020-09572-x] [Medline: 33032554]
7. Dormanesh A, Kirkpatrick MG, Allem J. Content Analysis of Instagram Posts From 2019 With Cartoon-Based Marketing of e-Cigarette-Associated Products. JAMA Pediatr 2020 Nov 01;174(11):1110-1112 [FREE Full text] [doi: 10.1001/jamapediatrics.2020.1987] [Medline: 32687566]
8. Kong G, LaVallee H, Rams A, Ramamurthi D, Krishnan-Sarin S. Promotion of Vape Tricks on YouTube: Content Analysis. J Med Internet Res 2019 Jun 18;21(6):e12709 [FREE Full text] [doi: 10.2196/12709] [Medline: 31215510]
9. Romijnders KAGJ, van Osch L, de Vries H, Talhout R. Perceptions and Reasons Regarding E-Cigarette Use among Users and Non-Users: A Narrative Literature Review. Int J Environ Res Public Health 2018 Jun 06;15(6):1190 [FREE Full text] [doi: 10.3390/ijerph15061190] [Medline: 29882828]
10. Kong G, Morean ME, Cavallo DA, Camenga DR, Krishnan-Sarin S. Reasons for Electronic Cigarette Experimentation and Discontinuation Among Adolescents and Young Adults. Nicotine Tob Res 2015 Jul;17(7):847-854 [FREE Full text] [doi: 10.1093/ntr/ntu257] [Medline: 25481917]
11. Donaldson SI, Dormanesh A, Perez C, Majmundar A, Allem J. Association Between Exposure to Tobacco Content on Social Media and Tobacco Use: A Systematic Review and Meta-analysis. JAMA Pediatr 2022 Sep 01;176(9):878-885 [FREE Full text] [doi: 10.1001/jamapediatrics.2022.2223] [Medline: 35816331]
12. Lee J, Tan AS, Porter L, Young-Wolff KC, Carter-Harris L, Salloum RG. Association Between Social Media Use and Vaping Among Florida Adolescents, 2019. Prev Chronic Dis 2021 May 13;18:E49 [FREE Full text] [doi: 10.5888/pcd18.200550] [Medline: 33988495]
13. Vogels EA, Risa GW, Navid M. Teens, social media and technology 2022. Pew Research Center. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2022/08/PI_2022.08.10_Teens-and-Tech_FINAL.pdf [accessed 2022-09-05]
14. Jackler RK, Li VY, Cardiff RAL, Ramamurthi D. Promotion of tobacco products on Facebook: policy versus practice. Tob Control 2019 Jan;28(1):67-73. [doi: 10.1136/tobaccocontrol-2017-054175] [Medline: 29622602]
15. Lorigo L, Pan B, Hembrooke H, Joachims T, Granka L, Gay G. The influence of task and gender on search and evaluation behavior using Google. Information Processing & Management 2006 Jul;42(4):1123-1131. [doi: 10.1016/j.ipm.2005.10.001]
16. YouTube. Official Blog and Support Pages. URL: https://www.youtube.com/ [accessed 2022-10-05]
17. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf [accessed 2023-09-04]
18. Huang J, Kornfield R, Emery SL. 100 Million Views of Electronic Cigarette YouTube Videos and Counting: Quantification, Content Evaluation, and Engagement Levels of Videos. J Med Internet Res 2016 Mar 18;18(3):e67 [FREE Full text] [doi: 10.2196/jmir.4265] [Medline: 26993213]
19. Kong G, Schott AS, Lee J, Dashtian H, Murthy D. Understanding e-cigarette content and promotion on YouTube through machine learning. Tob Control. Online ahead of print . [doi: 10.1136/tobaccocontrol-2021-057243] [Medline: 35504690]

20. Paek H, Kim S, Hove T, Huh JY. Reduced harm or another gateway to smoking? source, message, and information characteristics of E-cigarette videos on YouTube. J Health Commun 2014;19(5):545-560. [doi: 10.1080/10810730.2013.821560] [Medline: 24117370]
21. Sears C, Walker K, Hart J, Lee A, Siu A, Smith C. Clean, cheap, convenient: promotion of Electronic cigarettes on YouTube. Tob Prev Cessat 2017 Apr;3:10 [FREE Full text] [doi: 10.18332/tpc/69393] [Medline: 28725876]
22. Guy M, Helt J, Palafox S, Green K, Soule E, Maloney S, et al. Orthodox and Unorthodox Uses of Electronic Cigarettes: A Surveillance of YouTube Video Content. Nicotine Tob Res 2019 Sep 19;21(10):1378-1384 [FREE Full text] [doi: 10.1093/ntr/nty132] [Medline: 29961828]
23. Massey ZB, Li Y, Holli J, Churchill V, Yang B, Henderson K, et al. Modifications to Electronic Nicotine Delivery Systems: Content Analysis of YouTube Videos. J Med Internet Res 2020 Jun 02;22(6):e17104 [FREE Full text] [doi: 10.2196/17104] [Medline: 32348288]
24. Hua M, Alfi M, Talbot P. Health-related effects reported by electronic cigarette users in online forums. J Med Internet Res 2013 Apr 08;15(4):e59 [FREE Full text] [doi: 10.2196/jmir.2324] [Medline: 23567935]
25. Luo C, Zheng X, Zeng DD, Leischow S. Portrayal of electronic cigarettes on YouTube. BMC Public Health 2014 Oct 03;14:1028 [FREE Full text] [doi: 10.1186/1471-2458-14-1028] [Medline: 25277872]
26. Tan AS, Bigman CA. Misinformation About Commercial Tobacco Products on Social Media—Implications and Research Opportunities for Reducing Tobacco-Related Health Disparities. Am J Public Health 2020 Oct;110(S3):S281-S283. [doi: 10.2105/ajph.2020.305910]
27. Jones DM, Guy MC, Soule E, Sakuma KK, Pokhrel P, Orloff M, et al. Characterization of Electronic Cigarette Warning Statements Portrayed in YouTube Videos. Nicotine Tob Res 2021 Aug 04;23(8):1358-1366 [FREE Full text] [doi: 10.1093/ntr/ntaa272] [Medline: 33400781]
28. Davidson J, Liebald B, Liu J, Nandy P, Van Vleet T, Gargi U, et al. The YouTube video recommendation system. 2010 Presented at: RecSys'10 - Proceedings of the 4th ACM Conference on Recommender Systems; September 26; Barcelona, Spain p. 293-296. [doi: 10.1145/1864708.1864770]
29. Hussein E, Juneja P, Mitra T. Measuring misinformation in video search platforms: an audit study on YouTube. 2020 Presented at: Proceedings of the ACM on Human-Computer Interaction; May 28; New York, USA p. 1-27. [doi: 10.1145/3392854]
30. Covington P, Adams J, Sargin E. Deep neural networks for YouTube recommendations. 2016 Presented at: RecSys - Proceedings of the 10th ACM Conference on Recommender Systems; September 7; Boston, MA p. 191-198. [doi: 10.1145/2959100.2959190]
31. 2016 Surgeon General's report: e-Cigarette use among youth and young adults. Centers for Disease Control. URL: https://www.cdc.gov/tobacco/sgr/e-cigarettes/index.htm [accessed 2023-03-17]
32. Dashtian H, Murthy D, Kong G. An Exploration of e-Cigarette-Related Search Items on YouTube: Network Analysis. J Med Internet Res 2022 Jan 27;24(1):e30679 [FREE Full text] [doi: 10.2196/30679] [Medline: 35084353]
33. Tzioumis K. Demographic aspects of first names. Sci Data 2018 Mar 06;5:180025 [FREE Full text] [doi: 10.1038/sdata.2018.25] [Medline: 29509189]
34. Romito LM, Hurwich RA, Eckert GJ. A Snapshot of the Depiction of Electronic Cigarettes in YouTube Videos. Am J Hlth Behav 2015 Nov 01;39(6):823-831. [doi: 10.5993/ajhb.39.6.10]
35. Stellefson M, Chaney B, Ochipa K, Chaney D, Haider Z, Hanik B, et al. YouTube as a source of chronic obstructive pulmonary disease patient education: a social media content analysis. Chron Respir Dis 2014 May;11(2):61-71 [FREE Full text] [doi: 10.1177/1479972314525058] [Medline: 24659212]
36. Kloehn N, Leroy G, Kauchak D, Gu Y, Colina S, Yuan NP, et al. Improving Consumer Understanding of Medical Text: Development and Validation of a New SubSimplify Algorithm to Automatically Generate Term Explanations in English and Spanish. J Med Internet Res 2018 Aug 02;20(8):e10779 [FREE Full text] [doi: 10.2196/10779] [Medline: 30072361]
37. Curiskis SA, Drake B, Osborn TR, Kennedy PJ. An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. Information Processing & Management 2020 Mar;57(2):102034. [doi: 10.1016/j.ipm.2019.04.002]
38. Benis A, Barak Barkan R, Sela T, Harel N. Communication Behavior Changes Between Patients With Diabetes and Healthcare Providers Over 9 Years: Retrospective Cohort Study. J Med Internet Res 2020 Aug 11;22(8):e17186 [FREE Full text] [doi: 10.2196/17186] [Medline: 32648555]
39. Haque MM, Rabbani M, Dipal DD, Zarif MII, Iqbal A, Schwichtenberg A, et al. Informing Developmental Milestone Achievement for Children With Autism: Machine Learning Approach. JMIR Med Inform 2021 Jun 08;9(6):e29242 [FREE Full text] [doi: 10.2196/29242] [Medline: 33984830]
40. Zhang S, Tong H, Xu J, Maciejewski R. Graph convolutional networks: a comprehensive review. Comput Soc Netw 2019 Nov 10;6(1):1. [doi: 10.1186/s40649-019-0069-y]
41. Horne E, Tibble H, Sheikh A, Tsanas A. Challenges of Clustering Multimodal Clinical Data: Review of Applications in Asthma Subtyping. JMIR Med Inform 2020 May 28;8(5):e16452 [FREE Full text] [doi: 10.2196/16452] [Medline: 32463370]
42. Pariser E. The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think. New York, NY: Penguin Books; Apr 24, 2014.

43. Hannak A, Sapiezynski P, Molavi Kakhki A, Krishnamurthy B, Lazer D, Wilson C. Measuring personalization of web search. 2013 Presented at: Proceedings of the 22nd International Conference on World Wide Web; May 13; Rio de Janeiro, Brazil. [doi: 10.1145/2488388.2488435]
44. Khan ML. Social media engagement: What motivates user participation and consumption on YouTube? Computers in Human Behavior 2017 Jan;66:236-247. [doi: 10.1016/j.chb.2016.09.024]
45. Molyneaux H, O'Donnell S, Gibson K, Singer J. Exploring the gender divide on YouTube: an analysis of the creation and reception of vlogs. American Communication Journal. 2008. URL: https://ac-journal.org/journal/2008/Spring/3GenderandYoutube.pdf?origin=publication_detail [accessed 2023-03-17]
46. Kong G, Kuguru KE, Krishnan-Sarin S. Gender Differences in U.S. Adolescent E-Cigarette Use. Curr Addict Rep 2017 Dec;4(4):422-430 [FREE Full text] [doi: 10.1007/s40429-017-0176-5] [Medline: 29545987]
47. Wang TW, Gentzke AS, Neff LJ, Glidden EV, Jamal A, Park-Lee E, et al. Characteristics of e-Cigarette Use Behaviors Among US Youth, 2020. JAMA Netw Open 2021 Jun 01;4(6):e2111336 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.11336] [Medline: 34097049]
48. Kong G, Laestadius L, Vassey J, Majmundar A, Stroup AM, Meissner HI, et al. Tobacco promotion restriction policies on social media. Tob Control. Online ahead of print [FREE Full text] [doi: 10.1136/tc-2022-057348] [Medline: 36328589]
49. Olfson M, Wall MM, Liu S, Sultan RS, Blanco C. E-cigarette Use Among Young Adults in the U.S. Am J Prev Med 2019 May;56(5):655-663 [FREE Full text] [doi: 10.1016/j.amepre.2018.12.004] [Medline: 30885517]
50. Comenetz J. Frequently occurring surnames from the 2010 census. US Census Bureau. URL: https://www.census.gov/topics/population/genealogy/data/2010_surnames.html [accessed 2021-05-05]

Abbreviations

GCN: graph convolutional network

Edited by T Mackey; submitted 26.08.22; peer-reviewed by J Fetterman, T Hong, J Vassey; comments to author 05.11.22; revised version received 30.12.22; accepted 06.02.23; published 12.04.23.

Please cite as:
Murthy D, Lee J, Dashtian H, Kong G
Influence of User Profile Attributes on e-Cigarette–Related Searches on YouTube: Machine Learning Clustering and Classification
JMIR Infodemiology 2023;3:e42218
URL: https://infodemiology.jmir.org/2023/1/e42218
doi:10.2196/42218
PMID:37124246

Original Paper

State and Federal Legislators' Responses on Social Media to the Mental Health and Burnout of Health Care Workers Throughout the COVID-19 Pandemic: Natural Language Processing and Sentiment Analysis

Matthew P Abrams[1,2,3], BA; Arthur P Pelullo[2], MA, MS; Zachary F Meisel[1,3,4], MD, MPH, MSHP; Raina M Merchant[1,2,3,4], MD, MSHP; Jonathan Purtle[5], MSc, DrPH; Anish K Agarwal[1,2,3,4], MD, MPH, MS

[1]Department of Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Center for Digital Health, University of Pennsylvania, Philadelphia, PA, United States

[3]Center for Emergency Care Policy and Research, Department of Emergency Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[4]Leonard Davis Institute of Health Care Economics, University of Pennsylvania, Philadelphia, PA, United States

[5]Department of Public Health Policy & Management, School of Global Public Health, New York University, New York, NY, United States

Corresponding Author:
Matthew P Abrams, BA
Department of Emergency Medicine
Perelman School of Medicine
University of Pennsylvania
8th Fl
3600 Civic Center Blvd
Philadelphia, PA, 19104
United States
Phone: 1 619 227 9228
Email: mattpabrams@knights.ucf.edu

Abstract

Background: Burnout and the mental health burden of the COVID-19 pandemic have disproportionately impacted health care workers. The links between state policies, federal regulations, COVID-19 case counts, strains on health care systems, and the mental health of health care workers continue to evolve. The language used by state and federal legislators in public-facing venues such as social media is important, as it impacts public opinion and behavior, and it also reflects current policy-leader opinions and planned legislation.

Objective: The objective of this study was to examine legislators' social media content on Twitter and Facebook throughout the COVID-19 pandemic to thematically characterize policy makers' attitudes and perspectives related to mental health and burnout in the health care workforce.

Methods: Legislators' social media posts about mental health and burnout in the health care workforce were collected from January 2020 to November 2021 using Quorum, a digital database of policy-related documents. The total number of relevant social media posts per state legislator per calendar month was calculated and compared with COVID-19 case volume. Differences between themes expressed in Democratic and Republican posts were estimated using the Pearson chi-square test. Words within social media posts most associated with each political party were determined. Machine-learning was used to evaluate naturally occurring themes in the burnout- and mental health–related social media posts.

Results: A total of 4165 social media posts (1400 tweets and 2765 Facebook posts) were generated by 2047 unique state and federal legislators and 38 government entities. The majority of posts (n=2319, 55.68%) were generated by Democrats, followed by Republicans (n=1600, 40.34%). Among both parties, the volume of burnout-related posts was greatest during the initial COVID-19 surge. However, there was significant variation in the themes expressed by the 2 major political parties. Themes most correlated with Democratic posts were (1) frontline care and burnout, (2) vaccines, (3) COVID-19 outbreaks, and (4) mental health services. Themes most correlated with Republican social media posts were (1) legislation, (2) call for local action, (3) government support, and (4) health care worker testing and mental health.

Conclusions: State and federal legislators use social media to share opinions and thoughts on key topics, including burnout and mental health strain among health care workers. Variations in the volume of posts indicated that a focus on burnout and the mental health of the health care workforce existed early in the pandemic but has waned. Significant differences emerged in the content posted by the 2 major US political parties, underscoring how each prioritized different aspects of the crisis.

(JMIR Infodemiology 2023;3:e38676) doi:10.2196/38676

KEYWORDS

burnout; wellness; mental health; social media; policy; health care workforce; COVID-19; infodemiology; healthcare worker; mental well-being; psychological distress; Twitter; content analysis; thematic analysis; policy maker; healthcare workforce; legislator

Introduction

Health care workers have been disproportionately affected by burnout and mental health symptoms, including depression and anxiety [1-3]. The COVID-19 pandemic has exacerbated mental health symptoms, disorders, and burnout across the workforce [4-14]. Health care workers continue to experience rapid shifts in case volume, critical supply shortages (eg, of personal protective equipment), vaccination rates, death rates, and public health measures [12,15-17]. The emotions and mental health symptoms experienced by workers continue to fluctuate dramatically [14,17,18]. Depression, anxiety, and burnout continue to rise at alarming rates across the health care workforce [19] and have public-facing consequences, such as worse patient outcomes and higher costs [20,21].

State and federal policy responses to the pandemic continue to change across the nation [22,23]. These policy changes have been debated in the public forum by health experts, physicians, and politicians [24]. The link between these policies and case count may lead to hospital-based capacity strain and impact the mental health of the workforce. Ultimately, state COVID-19 policies and political trends are shaping national legislation. For example, President Joe Biden recently signed the Dr. Lorna Breen Health Care Provider Protection Act, inspired by Dr Breen's death by suicide from the strain of providing care during the COVID-19 pandemic. This reflects how national legislators are starting to recognize the urgent need for improved behavioral health among health care providers.

Social media provides state and federal legislators the opportunity to directly communicate health-related information—including mental health information—to the public and to gauge public interest in a topic [25]. A recent systematic review identified that Twitter can be used to promote public health in 6 main ways, including analysis of shared content and public engagement [26], ultimately informing how governments and health care organizations shape appropriate responses to the COVID-19 pandemic [27]. Social media has also been analyzed to provide insights about the mental health of the general public during the COVID-19 pandemic [28].

The content and language used by state legislators in public-facing venues such as social media reflect their opinions and priorities [23,24,29]. Legislators' social media posts may also signal attention toward legislation and policy engagement in real time, in addition to their priorities [29-31]. Understanding what policy makers and legislators are saying in these forums is also important, as they have influence over public opinion and impact behavior [32,33]. This may be of particular interest during the COVID-19 pandemic, as US legislators connect with their constituents and influence behaviors related to COVID-19 prevention, safety, and exposure [33,34].

As burnout and mental health symptoms increase among health care workers, the support and opinions of legislators displayed on social media are also important in understanding the message being relayed to the public. Legislators interact on social media broadly, to a greater extent than they share legislative votes or cosponsorship [35]. The growing body of social media exposure on platforms such as Twitter and Facebook between legislators and the general public creates a repository of political opinion and indicators of key policy shifts and messaging. Further, prior studies have found differences reflect a growing divide between Republican and Democratic legislators' priorities regarding COVID-19 policies [34,36] and, overall, more partisanship than cosponsorship among online interactions between legislators [35]. However, no studies, to our knowledge, have examined possible differences in the views legislators have expressed online regarding the mental health and burnout of the health care workforce.

The objective of this study was to examine state and federal legislators' social media posts on Twitter and Facebook throughout the COVID-19 pandemic to identify and understand themes related to mental health and burnout of the health care workforce and look for indicators of temporal shifts in political priorities regarding mental health. Specifically, we sought to describe variations in content over time, differences in language and sentiment used across parties, and party-specific theme prevalence. This content is important to analyze in order to understand the public discourse, opinions of the legislature, and the overall response from legislators to burnout and the mental health of the health care workforce.

Methods

Data Source

We identified state legislators' Facebook and Twitter posts related to mental health and burnout in the health care workforce from January 2020 to November 2021 using Quorum (Quorum Analytics) [37], a software platform that collects policy-related documents, including social media content, from politicians during their time in office. For context, there are about 7312 state legislators [38] and 600 federal legislators [39] in the United States. Posts from all members of the upper and lower houses, as applicable, of the 50 US state legislature with 1 or more terms from each of the following keyword groups were selected for analysis: ["healthcare worker," "doctor," "physician," "nurse"] AND ["wellness," "wellbeing," "burnout," "resilience," "compassion," "fatigue," "depression," "suicide," "mental health," "anxiety," "sad," "depressed," "stress," "stressed," "tired," "frustrated," "frustration"]. Of note, the 4 keywords in the first string of search terms were carefully selected by the research team to capture the health care workers perceived to be most discussed by legislators online and were not inclusive of all frontline workers. Retweets and other posts duplicating the content of another user were also included in the analysis, as these posts indicate the significance of the original content and intent to propagate to a larger audience. This study was conducted in partnership with the Research-to-Policy Collaboration, which is affiliated with Pennsylvania State University's Edna Bennett Pierce Prevention Research Center.

Descriptive Analysis

Summary statistics were used to describe the volume of relevant burnout-related posts on each social media platform and across parties and legislative bodies. The monthly volume of social media posts related to mental health and burnout between January 2020 and November 2021, stratified by social media platform and political party, was compared with monthly COVID-19 case volume during the same time period. Differences between themes expressed in Democratic and Republican posts were estimated using the Pearson chi-square test. Themes expressed by legislators with independent or unknown affiliations were excluded from the analyses and assessments due to small sample size. Likewise, social media posts from government entities (rather than individual legislators) were excluded from the analyses due to small sample size.

Natural Language Processing

Preprocessing

Post text was converted to lowercase, extraneous white space was stripped, and link URLs, email addresses, user mentions, hashtags, and stop words were removed. Remaining terms were lemmatized to group-inflected forms with the same word stem, and the relative frequency of single words and phrases was extracted to build a baseline set of language features (rows indicated posts, and columns indicated word/phrase frequency), from which the top 50 most frequent words posted by Republicans and Democrats were identified. These methods have been used in prior work characterizing legislator discourse on social media [36,40,41].

Theme Modeling

We applied latent Dirichlet allocation (LDA), an unsupervised clustering algorithm, to the baseline set of language features to identify 20 data-driven word clusters (ie, topics) and constructed a topic feature set (rows indicated posts, and columns indicated topic prevalence); LDA assumes that posts have a small number of topics (ie, themes) and that topics are composed of groups of frequently co-occurring words and phrases across posts [42,43]. The topic model was trained using the Machine Learning for Language Toolkit 2.0 [44], and the optimal number of themes was selected via analysis of model coherence scores, visual inspection of topic separation with principal component analysis, and manual evaluation of topic interpretability.

Topic features were correlated (Pearson *r*) with political party (coded as a binary variable, where 0 indicated a Democratic post and 1 indicated a Republican post) to further distinguish linguistic differences across political parties in social media posts about mental health and burnout in the health care workforce. Significant correlations with a Benjamini-Hochberg–corrected *P* value of <.001 and their 95% CIs are reported. Authors AKA and MPA independently evaluated each topic for thematic meaning by reviewing the 10 words and 10 social media posts most associated with each topic [45,46].

Sentiment

We applied the Valence Aware Dictionary and Sentiment Reasoner (VADER) [47], a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media, to the baseline set of language features to identify weekly changes in post sentiment over time across political parties. Post sentiment scores were calculated as the mean sentence sentiment in each post (as suggested in the VADER documentation), and weekly sentiment scores were calculated as the mean post sentiment for all posts in a given week stratified by party. Sentiment data were visualized via weekly sentiment means overlaid with the exponentially weighted mean of weekly sentiment means. This was repeated to identify monthly changes in sentiment.

All statistical analyses were performed using Python (version 3.7.7).

Ethical Considerations

This study is exempt from ethical review under University of Pennsylvania Institutional Review Board guidelines, as it does not meet the criteria for human-subject research and utilizes publicly available social media posts.

Results

The search criteria resulted in 4165 health care workforce burnout–related social media posts, including 1400 tweets and 2765 Facebook posts, that met the inclusion criteria (Table 1).

These posts were generated by 2047 unique social media accounts, consisting of 2009 state and federal legislator accounts (1257 Facebook accounts and 752 Twitter accounts owned by 1685 unique individuals) and 38 government entity accounts, such as state health departments (n=38 Twitter accounts). The majority of the social media posts (2319/4165, 55.68%) were generated by Democrats. Republicans were responsible for 40.34% (1600/4165) of health care–associated burnout-related social media posts and all other legislators were responsible for 3.58% of posts (166/4165). The most common legislators were representatives (2139/4165, 51.36%) followed by senators (1259/4165, 29.52%). The mean word count was 43.47 (SD 18.87) words for Twitter posts and 422.72 (SD 277.15) words for Facebook posts. Variation in volume of posts generated varied over time, with the majority occurring during the initial surge (Figure 1). This general waning of the volume of burnout-related posts as the pandemic progressed was similar among legislators from both major political parties.

Notable differences were observed between platform use and political party affiliation. Democrats made the majority of Twitter posts (1033/1400, 73.79%) and Republicans made the slight majority of Facebook posts (1425/2765, 51.54%). Additionally, there were notable geographic differences along party lines in the volume of Facebook posts, with Democrats posting more often than Republicans in the Northeast (n=505 vs n=378), and Republicans posting more often from the South (n=727 vs n=418) and Midwest (n=246 vs n=91). However, these regional differences may partially reflect differences in the size and partisan composition of state legislatures across these geographies.

Thematic content generated from the natural-language processing and LDA approaches revealed varying content themes between the 2 major political parties (Figures 2 and 3).

The top 4 themes from social media posts most significantly correlated with the Democratic Party were (1) frontline care and burnout, (2) vaccines, (3) COVID outbreaks, and (4) mental health services. The top 4 themes associated with the Republican Party social media posts were (1) legislation, (2) call for local action, (3) government support, and (4) health care worker testing and mental health. Table 2 shows themes, words, and correlation strength with party.

Figures 4 and 5 show word clouds for each of the top 4 themes across party affiliation. Full post content and the list of themes are available in Multimedia Appendix 1, Table S1.

Sentence-level sentiment analyses also revealed differential sentiment patterns by political party throughout the timeline of the study (Multimedia Appendix 1, Figure S1A). The mean monthly post sentiment analysis found that both parties' posts remained within the slightly positive to positive sentiment range when mean sentiment scores were averaged per month and exponentially weighted. However, the more granular weekly post sentiment analysis by party revealed that during most spikes in COVID-19 case counts, the weekly exponentially weighted mean sentiment scores of Democratic posts more often entered the neutral or negative range compared to Republican posts (Multimedia Appendix 1, Figure S1B).

Table 1. Characteristics of social media posts.

Characteristic	Twitter (n=1400), n (%)	Facebook (n=2765), n (%)
Party		
Democratic	1033 (73.79)	1286 (46.51)
Republican	255 (18.21)	1425 (51.54)
Independent	5 (0.36)	3 (0.11)
Unknown	107 (7.64)	51 (1.84)
Region[a]		
Northeast	445 (31.79)	897 (32.64)
South	365 (26.07)	1170 (42.58)
Midwest	285 (20.36)	446 (16.23)
West	304 (21.71)	235 (8.55)
Status		
Current	1338 (95.57)	2535 (91.68)
Designate	1 (0.07)	3 (0.11)
Former	61 (4.36)	227 (8.21)
Government title[b]		
Representative	623 (22.53)	1516 (54.83)
Senator	439 (15.88)	820 (29.66)
Assembly	115 (4.16)	158 (5.71)
Delegate	54 (1.95)	118 (4.27)
Governor	54 (1.95)	73 (2.64)
Speaker	9 (0.33)	40 (1.45)
Member	0 (0)	14 (0.51)
Other	18 (0.65)	26 (0.94)

[a]Posts from Guam, the Virgin Islands, and the Northern Mariana Islands were not included.

[b]Posts from government entities (n=88 Twitter posts) were not included.

Figure 1. COVID-19 case counts and volume of social media posts by party over time. D: Democratic; I: independent; R: Republican; U: unknown.

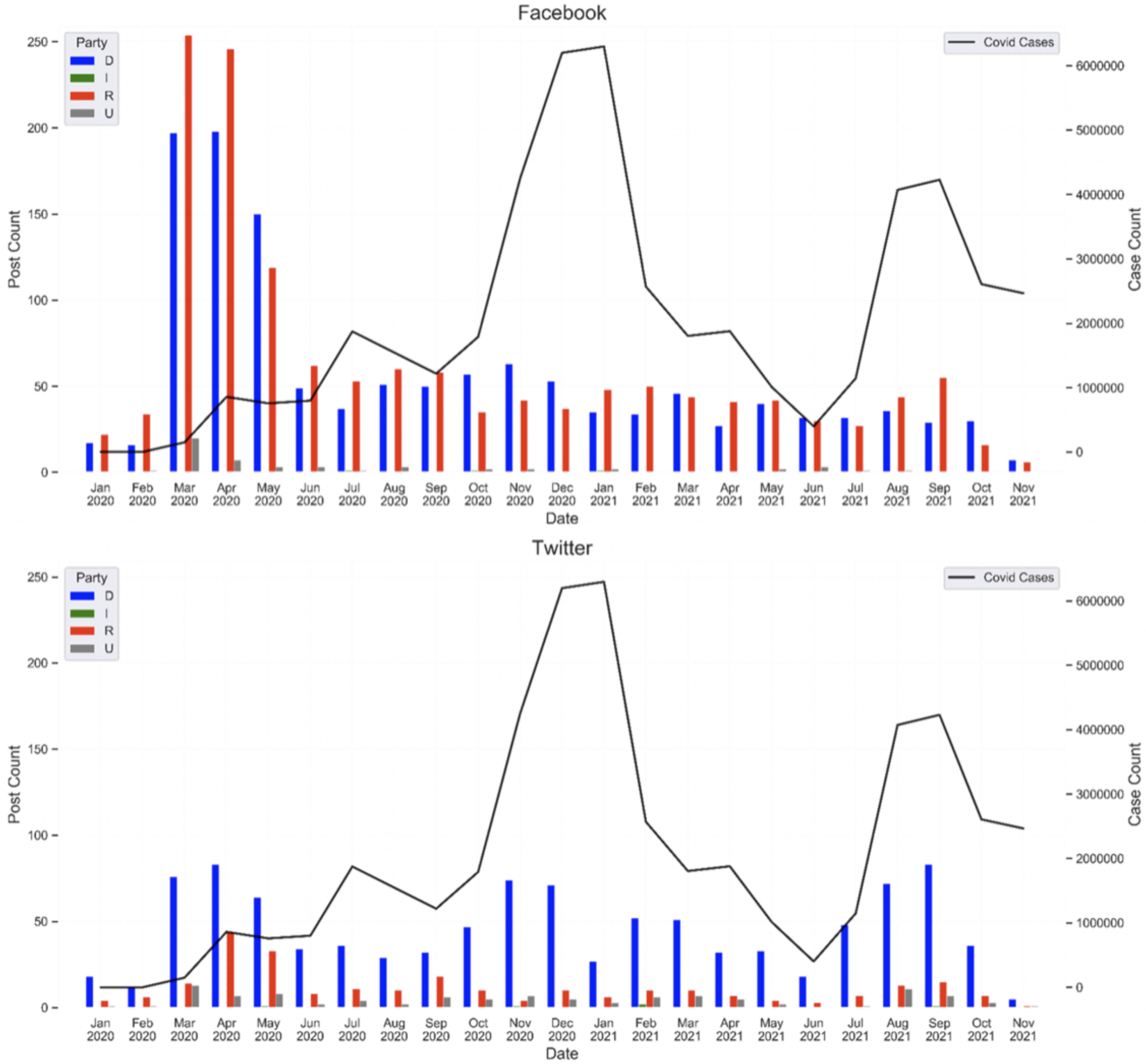


Figure 2. Words most frequently used in Democratic social media posts.

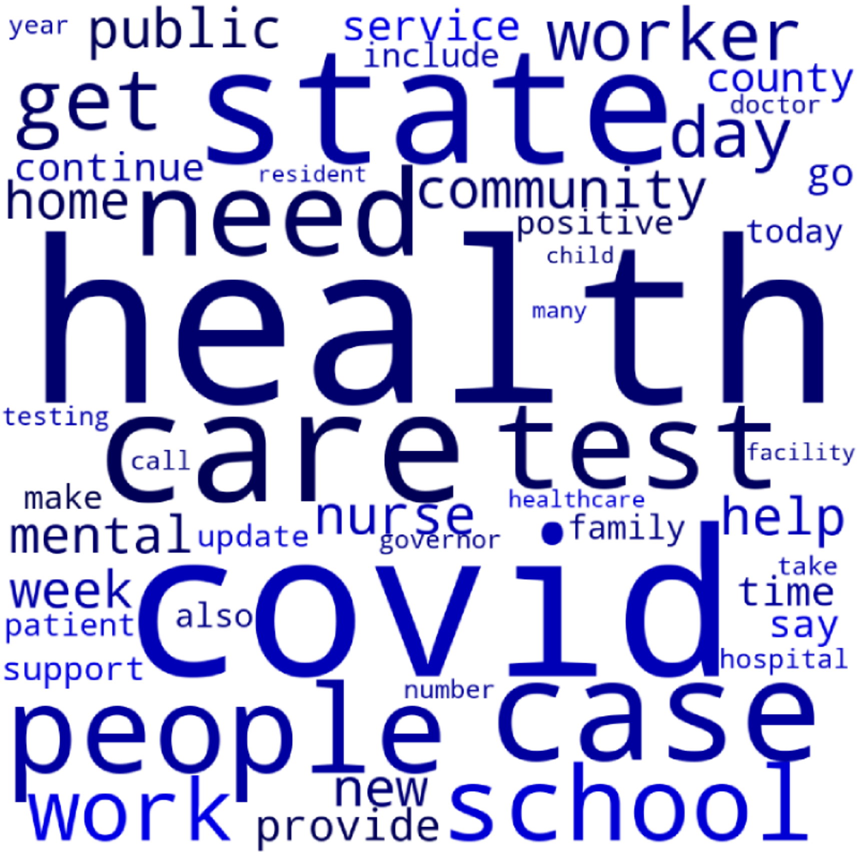


Figure 3. Words most frequently used in Republican social media posts.

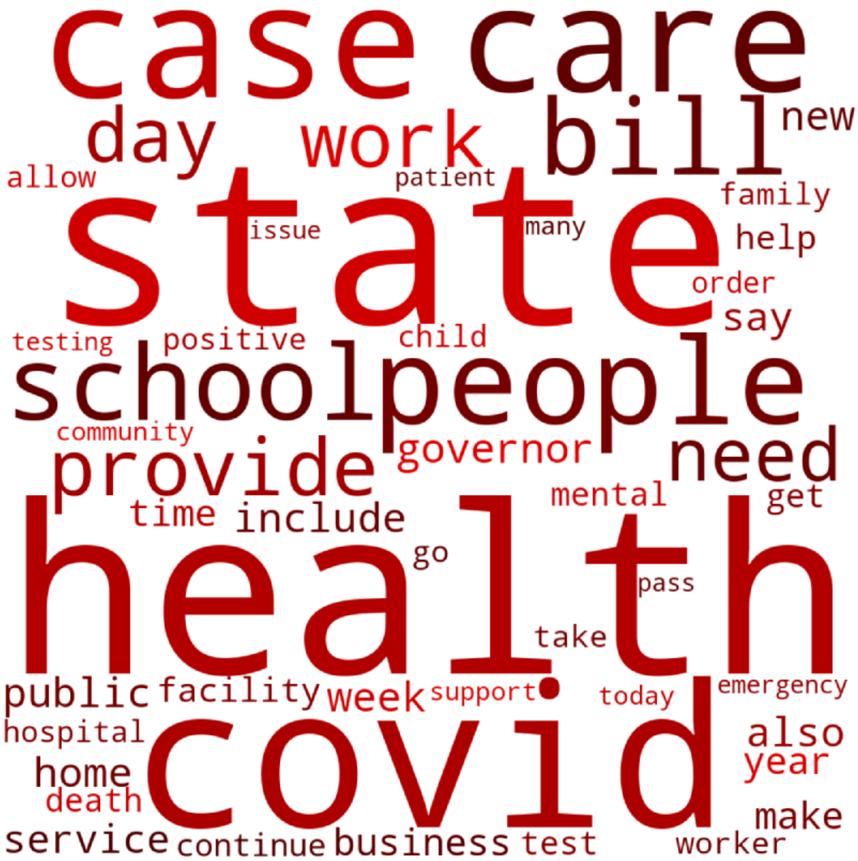


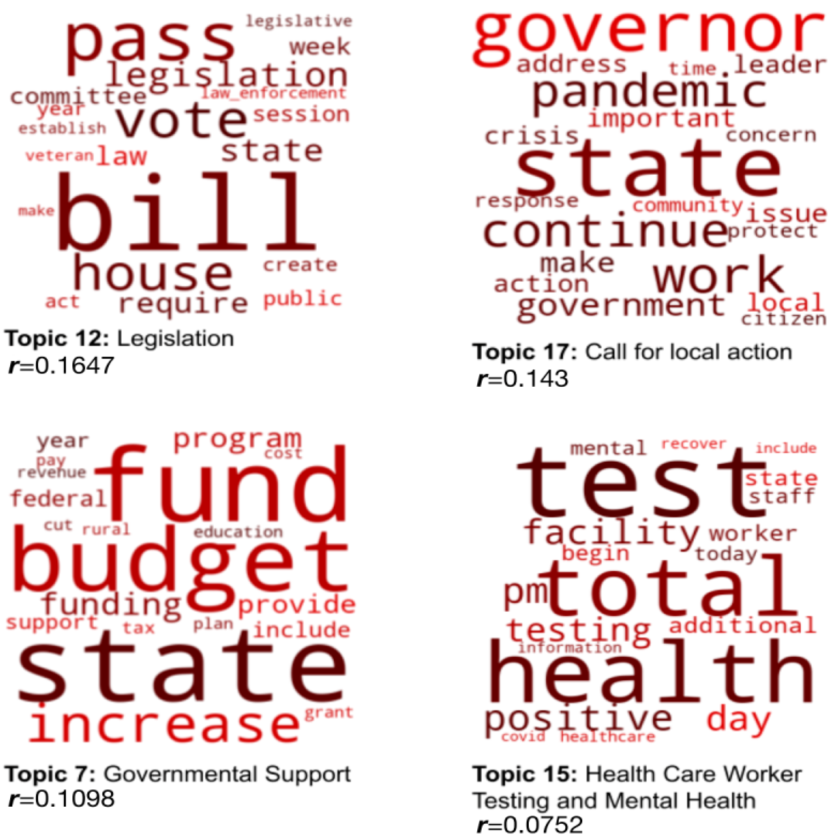
Table 2. Themes associated with Democratic or Republican posts.

Theme	Top words	Pearson *r* (95% CI)
Themes associated with Democratic posts		
Frontline care burnout and stress	worker, nurse, healthcare, work, doctor, community, fight, stress, pandemic, frontline, first_responder, social, tired, year, life, month, support, serve, front_line, proud	–0.2615[a] (–0.31 to –0.21)
Vaccine	vaccine, call, covid, vaccination, receive, week, information, vaccinate, appointment, online, local, office, meal, website, free, visit, find, pm, question, age	–0.1278[a] (–0.18 to –0.08)
COVID outbreaks	county, people, governor, test, work, state, back, continue, school, number, outbreak, rate, testing, day, make, positive, move, system, good, lot	–0.1118[a] (–0.16 to –0.06)
Mental health services	health, mental, care, service, access, support, patient, provider, treatment, physician, professional, insurance, crisis, practice, resource, behavioral, medical, provide, system, network	–0.1102[a] (–0.16 to –0.06)
COVID testing	covid, testing, health, information, state, include, public, update, test, site, department, community, today, member, resident, contact, resource, day, announce, visit	–0.0582 (–0.11 to –0.01)
State information	case, covid, county, health, statewide, update, coronavirus, individual, state, home, death, total, patient, provide, stay, information, report, number, continue, resident	–0.0398 (–0.09 to 0.01)
Schools and education	school, child, student, education, year, district, teacher, high, parent, family, plan, learn, person, work, board, adult, college, staff, opportunity, ensure	–0.0348 (–0.09 to 0.02)
Masking to slow spread	virus, people, spread, mask, risk, coronavirus, medical, disease, sick, doctor, prevent, stay, condition, symptom, show, time, flu, avoid, slow, wear	–0.0243 (–0.08 to 0.03)
Frontline/essential service support and volunteers	service, include, provide, medical, support, public, community, food, individual, provider, essential, work, worker, center, company, care, supply, volunteer, health, equipment	–0.0116 (–0.06 to 0.04)
Family/support systems	worker, nurse, healthcare, work, doctor, community, fight, stress, pandemic, frontline, first_responder, social, tired, year, life, month, support, serve, front_line, proud	–0.0019 (–0.05 to 0.05)
Themes associated with Republican posts		
Legislation	bill, pass, vote, house, legislation, state, require, law, week, committee, session, public, year, create, act, law_enforcement, veteran, legislative, establish, make	0.1647[a] (0.11 to 0.21)
Call for local action	state, governor, work, continue, pandemic, government, make, issue, important, address, action, local, crisis, leader, response, concern, protect, community, citizen, time	0.1430[a] (0.09 to 0.19)
Governmental support	state, fund, budget, increase, funding, program, provide, federal, include, year, support, tax, grant, cut, plan, education, pay, revenue, cost, rural	0.1098[a] (0.06 to 0.16)
Health care worker testing and mental health	test, health, total, positive, pm, facility, day, testing, additional, worker, state, begin, staff, mental, today, healthcare, recover, information, include, covid	0.0752[a] (0.02 to 0.13)
Business/economy	business, order, home, public, health, stay, guidance, follow, close, guideline, essential, open, social_distance, employee, issue, reopen, continue, activity, remain, limit	0.0676[a] (0.02 to 0.12)
Pandemic time course	time, day, people, make, work, give, place, put, today, good, week, call, happen, start, month, understand, long, point, post, end	0.0675[a] (0.02 to 0.12)
Emergency public health measures	emergency, state, provide, program, public, benefit, federal, assistance, business, payment, requirement, covid, extend, pay, department, governor, sign, receive, apply, require	0.0379 (–0.01 to 0.09)
Debate surrounding public policies	woman, mandate, decision, government, protect, policy, force, doctor, fail, power, lead, lose, sadly, life, abortion, hearing, drug, speak, freedom, science	0.0302 (–0.02 to 0.08)
Case counts	case, death, positive, covid, active, test, change, report, yesterday, number, week, hospital, total, bed, patient, update, day, increase, confirm, rate	0.0268 (–0.03 to 0.08)
Long-term care facilities	care, facility, home, family, nursing, health, hospital, resident, nursing home, staff, visit, long-term, member, patient, visitation, vulnerable, person, senior, hour, individual	0.0197 (–0.03 to 0.07)

[a]These values were significant at the $P<.001$ level after applying the Benjamini-Hochberg correction for multiple tests.

Figure 4. Word clouds representing the top 20 most representative words for each of the 4 themes most correlated with Democratic social media posts.

Figure 5. Word clouds representing the top 20 most representative words for each of the 4 themes most correlated with Republican social media posts.



Discussion

Principal Findings

This study investigated the social-media posts of US legislators throughout the COVID-19 pandemic with a focus on content related to health care–associated burnout and the mental health of the workforce. It has 3 key findings. First, state and federal legislators are actively using social media to discuss the pandemic and burnout. Second, the focus on burnout and the mental health of the health care workforce was primarily seen in the early surge of the pandemic and then dramatically waned. Third, key differences emerged in the social media content posted by the 2 major US political parties. Addressing the overlapping nature of the COVID-19 pandemic and health care–associated burnout is a national priority for health systems, payers, clinicians, and patients [7], yet the 2 parties appear to highlight and prioritize different aspects of the crisis.

State and federal legislators are increasingly using social media as a platform to discuss health care and medicine [31,35,36,40]. Previous literature has investigated the relationship between Democrats' and Republicans' social media content within the context of the opioid epidemic, showing that overall partisanship across topics increased from 2016 to 2019 [40]. In the setting of the COVID-19 pandemic, a recent study also showed that Republican legislators who were previously less engaged in discussion of vaccination on social media became significantly more publicly engaged following the arrival of COVID-19 compared to their Democratic counterparts, suggesting a possible convergence of priorities in light of the COVID-19 pandemic [41]. The content posted on Twitter and Facebook is public facing, and given the rise of digital technology and social media, the content posted by legislators in the United States provides a window into political thoughts, agendas, and priorities. The pandemic has certainly worsened the mental health strain and burnout faced by health care providers and is projected to continue despite improvements in case volume [7]. This is among the first studies to discover and investigate the social media content from US legislators specific to burnout and mental health of the workforce. Perhaps less surprising is the rise in these social media posts early in the pandemic, as attention was keenly focused on the workforce. Unfortunately, this data set shows that after the initial wave, there has been less attention over time despite recurrent surges (eg, Delta variants). In line with Kingdon's multiple streams model [48], this may indicate that the "policy window" for mental health–related legislation regarding the health care workforce was open early in the pandemic. That said, there remains a persistent, yet small, discussion across parties, but ultimately it is low.

The themes and words that state and federal legislators used in these mental health–related social media posts were notably different between the 2 major political parties, including in their emphasis. This is consistent with another recent analysis of tweets from legislators that found differences in health care–related themes according to party lines [36]. In our study, Republican-affiliated legislator posts revealed a greater representation of themes central to public policies and legislation. The themes indicated a focus on local and federal action as seen through 2 of the top 4 most strongly correlated themes, "call for local action" and "governmental support." This may reflect support for implementing broader policies to help support health care workers. Republican posts also included a focus on COVID-19 testing for the workforce. In contrast, Democratic social media posts more specifically focused on the mental health services and acute strain on the workers themselves. The thematic analysis showed that 2 of the top 4 themes focused on "frontline care and burnout/stress" and "mental health services." In addition, Democratic posts were varied in their overall content, with other themes emerging related to capacity strain on health systems related to outbreaks and vaccines and vaccinations themselves. These themes appear to be much more granular and focused on the workers themselves and the stress and burnout they face throughout the pandemic surges.

This is among the first studies to use natural language processing of state and federal legislators' social media content to measure and describe trends in content and posting issues over time with specific attention to health care worker burnout and mental health. State and federal legislators' word choices on social media carry great influence, and their reach is broad. The posts generated by legislators reflect the immense initial concern and the seeming loss of focus as the public response evolved over the course of the pandemic. Discussing mental health and burnout in public forums is important in health care, where significant stigmas exist and the consequences are grave, as seen by the high relative rate of physician suicide [49-51]. State and federal legislators carry power in their voices, whether they are live or on social media, and their words can lead to important action to help support and sustain the workforce. Recognizing the urgent need for improved behavioral health among health care providers, President Joe Biden recently signed the Dr. Lorna Breen Health Care Provider Protection Act, inspired by Dr Breen's death by suicide from the strain of providing care during the COVID-19 pandemic. Highlighting the important role of legislators' social media, the post on the President's Instagram account (@Potus) about this new act's aim of "reducing and preventing suicide, burnout, and mental health and substance use conditions among healthcare professionals" received over 330,000 likes and 7200 comments, suggesting social media is an important tool for legislators to interface with constituents about the mental health of the workforce.

Limitations

This study has several limitations. Quorum does not report state or federal legislators' years in office, only whether they are a current or former legislator at the time of data download. We therefore were unable to stratify for years in office in our measures of legislators' number of social media posts related to burnout or mental health. Similarly, Quorum does not report the gender of legislators. It is possible that the content may be different based on the gender of legislators, so future studies should aim to analyze legislators' posts by gender. We also did not have access to the total number of social media posts for each legislator. We were therefore also unable to stratify for a legislator's general social media activity in our analysis. Another limitation is that social media posts from both state and federal legislators were aggregated and analyzed together. However, it is possible that variations in the content and sentiment of social media posts may differ based on whether a legislator works at the state or federal level.

Moreover, cross-party comparisons in post volume are impacted by the size and partisan composition of state and federal legislatures, which are often not evenly distributed along party lines; therefore, regional differences in attention to burnout within these geographical regions should be interpreted with caution, since there may be different numbers of Democratic versus Republican legislators in a given region. Another limitation is that changes in the content of social media posts in relation to major changes in pandemic prevention and control, such as lockdowns, the introduction of vaccines, vaccine mandates, and masking, were not considered in the analyses. Given it is possible that the content in posts may vary based on these major events, more granular analyses that look at how social media content was influenced by prevention efforts should be conducted in the future.

Finally, social media language does not necessarily lead to specific votes or policy decisions. Identifying relationships between state and federal legislator social media content and legislator voting patterns was beyond the scope of this project.

Conclusion

Health care–associated burnout and mental health strain has grown tremendously throughout the pandemic. Public and legislative response and attention is key to ensuring those working in health care are supported and cared for, as burnout impacts clinicians and the care they provide. Social media can provide valuable insight into trends in state and federal legislators' burnout and mental health–related content. We found an initial surge in the volume of posts that has diminished throughout the pandemic and, perhaps unsurprisingly, a divide in how Democrats and Republicans think about the issues. Democrats increasingly post content related to individuals and stress and Republicans increasingly post content related to legislation. As the pandemic case count diminished, we found an unfortunate similar decrease in attention from legislators to the issue of supporting the mental health of health care workers and combating burnout.

Acknowledgments

This study was conducted in partnership with the Research-to-Policy Collaboration, which is affiliated with Pennsylvania State University's Edna Bennett Pierce Prevention Research Center.

RMM receives funding from the National Institute of Mental Health (NIMH; grant number R01MH127686) and the National Heart, Lung, and Blood Institute (grant number K24 HL157621). JP receives funding from NIMH (grant number P50MH113662).

Data Availability

The data sets generated during and/or analyzed during the current study are available publicly from the Quorum database, as well as from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplemental Figure 1: Mean Social Media Post Sentiment Scores by Political Party Affiliation and Supplemental Table 1: Representative Social Media Posts Associated with Each Theme.

[DOCX File , 471 KB - infodemiology_v3i1e38676_app1.docx]

References

1. Slavin S. Mental health from medical school to medical practice: finding a path forward. Mo Med 2021;118(1):7-12 [FREE Full text] [Medline: 33551470]
2. Amir M, Dahye K, Duane C, Wendy L W. Medical student and resident burnout: a review of causes, effects, and prevention. J Fam Med Dis Prev 2018 Dec 31;4(4):1-8 [FREE Full text] [doi: 10.23937/2469-5793/1510094]
3. West C, Dyrbye L, Shanafelt T. Physician burnout: contributors, consequences and solutions. J Intern Med 2018 Jun;283(6):516-529 [FREE Full text] [doi: 10.1111/joim.12752] [Medline: 29505159]
4. Shin HS, Park H, Lee Y. The relationship between medical students' empathy and burnout levels by gender and study years. Patient Educ Couns 2022 Feb;105(2):432-439. [doi: 10.1016/j.pec.2021.05.036] [Medline: 34127334]
5. Kelsey E, West C, Cipriano P, Peterson C, Satele D, Shanafelt T, et al. Original research: suicidal ideation and attitudes toward help seeking in U.S. nurses relative to the general working population. Am J Nurs 2021 Nov 01;121(11):24-36. [doi: 10.1097/01.NAJ.0000798056.73563.fa] [Medline: 34629376]

6. Galanis P, Vraka I, Fragkou D, Bilali A, Kaitelidou D. Nurses' burnout and associated risk factors during the COVID-19 pandemic: A systematic review and meta-analysis. J Adv Nurs 2021 Aug;77(8):3286-3302 [FREE Full text] [doi: 10.1111/jan.14839] [Medline: 33764561]
7. Sharifi M, Asadi-Pooya A, Mousavi-Roknabadi R. Burnout among healthcare providers of COVID-19; a systematic review of epidemiology and recommendations. Arch Acad Emerg Med 2021;9(1):e7 [FREE Full text] [doi: 10.22037/aaem.v9i1.1004] [Medline: 33490964]
8. Chor W, Ng W, Cheng L, Situ W, Chong J, Ng L, et al. Burnout amongst emergency healthcare workers during the COVID-19 pandemic: A multi-center study. Am J Emerg Med 2021 Aug;46:700-702 [FREE Full text] [doi: 10.1016/j.ajem.2020.10.040] [Medline: 33129643]
9. Holmes MR, Rentrope CR, Korsch-Williams A, King JA. Impact of COVID-19 pandemic on posttraumatic stress, grief, burnout, and secondary trauma of social workers in the United States. Clin Soc Work J 2021;49(4):495-504 [FREE Full text] [doi: 10.1007/s10615-021-00795-y] [Medline: 33678920]
10. Stone KW, Kintziger KW, Jagger MA, Horney JA. Public health workforce burnout in the COVID-19 response in the U.S. Int J Environ Res Public Health 2021 Apr 20;18(8):4369 [FREE Full text] [doi: 10.3390/ijerph18084369] [Medline: 33924084]
11. Prasad K, McLoughlin C, Stillman M, Poplau S, Goelz E, Taylor S, et al. Prevalence and correlates of stress and burnout among U.S. healthcare workers during the COVID-19 pandemic: A national cross-sectional survey study. EClinicalMedicine 2021 May;35:100879 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100879] [Medline: 34041456]
12. Hall H. The effect of the COVID-19 pandemic on healthcare workers' mental health. JAAPA 2020 Jul;33(7):45-48. [doi: 10.1097/01.JAA.0000669772.78848.8c] [Medline: 32590533]
13. Pearman A, Hughes ML, Smith EL, Neupert SD. Mental health challenges of United States healthcare professionals during COVID-19. Front Psychol 2020;11:2065 [FREE Full text] [doi: 10.3389/fpsyg.2020.02065] [Medline: 32903586]
14. Walton M, Murray E, Christian M. Mental health care for medical staff and affiliated healthcare workers during the COVID-19 pandemic. Eur Heart J Acute Cardiovasc Care 2020 Apr;9(3):241-247 [FREE Full text] [doi: 10.1177/2048872620922795] [Medline: 32342698]
15. Ranney ML, Griffeth V, Jha AK. Critical supply shortages - the need for ventilators and personal protective equipment during the Covid-19 pandemic. N Engl J Med 2020 Apr 30;382(18):e41. [doi: 10.1056/NEJMp2006141] [Medline: 32212516]
16. Cao J, Wei J, Zhu H, Duan Y, Geng W, Hong X, et al. A study of basic needs and psychological wellbeing of medical workers in the fever clinic of a tertiary general hospital in Beijing during the COVID-19 outbreak. Psychother Psychosom 2020;89(4):252-254 [FREE Full text] [doi: 10.1159/000507453] [Medline: 32224612]
17. Sumner R, Kinsella E. "It's like a kick in the teeth": the emergence of novel predictors of burnout in frontline workers during Covid-19. Front Psychol 2021;12:645504 [FREE Full text] [doi: 10.3389/fpsyg.2021.645504] [Medline: 34113288]
18. Bender AE, Berg KA, Miller EK, Evans KE, Holmes MR. "Making sure we are all okay": healthcare workers' strategies for emotional connectedness during the COVID-19 pandemic. Clin Soc Work J 2021;49(4):445-455 [FREE Full text] [doi: 10.1007/s10615-020-00781-w] [Medline: 33456094]
19. Chirico F, Ferrari G, Nucera G, Szarpak L, Crescenzo P, Ilesanmi O. Prevalence of anxiety, depression, burnout syndrome, and mental health disorders among healthcare workers during the COVID-19 pandemic: a rapid umbrella review of systematic reviews. J Health Soc Sci 2021;6(2):209-220 [FREE Full text] [doi: 10.19204/2021/prvl7]
20. Forrest C, Xu H, Thomas L, Webb L, Cohen L, Carey T, HERO Registry Research Group. Impact of the early phase of the COVID-19 pandemic on US healthcare workers: results from the HERO Registry. J Gen Intern Med 2021 May;36(5):1319-1326 [FREE Full text] [doi: 10.1007/s11606-020-06529-z] [Medline: 33694071]
21. Chin E, Huynh B, Lo N, Hastie T, Basu S. Projected geographic disparities in healthcare worker absenteeism from COVID-19 school closures and the economic feasibility of child care subsidies: a simulation study. BMC Med 2020 Jul 15;18(1):218 [FREE Full text] [doi: 10.1186/s12916-020-01692-w] [Medline: 32664927]
22. Andraka-Christou B, Bouskill K, Haffajee RL, Randall-Kosich O, Golan M, Totaram R, et al. Common themes in early state policy responses to substance use disorder treatment during COVID-19. Am J Drug Alcohol Abuse 2021 Jul 04;47(4):486-496 [FREE Full text] [doi: 10.1080/00952990.2021.1903023] [Medline: 33909518]
23. Baccini L, Brodeur A. Explaining governors' response to the COVID-19 pandemic in the United States. American Politics Research 2020 Dec 01;49(2):215-220 [FREE Full text] [doi: 10.1177/1532673x20973453]
24. Kushner Gadarian S, Goodman S, Pepinsky T. Partisanship, Health Behavior, and Policy Attitudes in the Early Stages of the COVID-19 Pandemic. SSRN Journal 2020:1-23 [FREE Full text] [doi: 10.2139/ssrn.3562796]
25. 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]
26. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: A systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
27. Chandrasekaran R, Mehta V, Valkunde T, Moustakas E. Topics, trends, and sentiments of tweets about the COVID-19 pandemic: temporal infoveillance study. J Med Internet Res 2020 Oct 23;22(10):e22624 [FREE Full text] [doi: 10.2196/22624] [Medline: 33006937]

28. Valdez D, Ten Thij M, Bathina K, Rutter L, Bollen J. Social media insights into US mental health during the COVID-19: longitudinal analysis of Twitter data. J Med Internet Res 2020 Dec 14;22(12):e21418 [FREE Full text] [doi: 10.2196/21418] [Medline: 33284783]
29. Hemphill L, Russell A, Schöpke‐Gonzalez AM. What drives U.S. congressional members' policy attention on Twitter? Policy Internet 2020 Jun 28;13(2):233-256. [doi: 10.1002/poi3.245]
30. 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]
31. Casas A, Payson J, Nagler J, Bonneau R, Tucker JA. Social Media Data Reveal Patterns of Policy Engagement in State Legislatures. SSRN Journal 2020:1-37. [doi: 10.2139/ssrn.3698990]
32. Milligan M, Hoyt D, Gold A, Hiserodt M, Otto M. COVID-19 vaccine acceptance: influential roles of political party and religiosity. Psychol Health Med 2022 Oct;27(9):1907-1917. [doi: 10.1080/13548506.2021.1969026] [Medline: 34407721]
33. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. J Public Health Policy 2020 Dec;41(4):410-420 [FREE Full text] [doi: 10.1057/s41271-020-00247-w] [Medline: 32826935]
34. Grossman G, Kim S, Rexer JM, Thirumurthy H. Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. Proc Natl Acad Sci U S A 2020 Sep 29;117(39):24144-24153 [FREE Full text] [doi: 10.1073/pnas.2007835117] [Medline: 32934147]
35. Cook JM. Are American politicians as partisan online as they are offline? Twitter networks in the U.S. Senate and Maine State Legislature. Policy Internet 2016;8(1):55-71. [doi: 10.1002/poi3.109]
36. Guntuku S, Purtle J, Meisel Z, Merchant R, Agarwal A. Partisan differences in Twitter language among US legislators during the COVID-19 pandemic: cross-sectional study. J Med Internet Res 2021 Jun 03;23(6):e27300 [FREE Full text] [doi: 10.2196/27300] [Medline: 33939620]
37. Quorum. URL: https://www.quorum.us/ [accessed 2023-01-25]
38. List of United States state legislatures. Wikipedia. URL: https://en.wikipedia.org/w/index.php?title=List_of_United_States_state_legislatures&oldid=1135333667 [accessed 2023-01-26]
39. Members of the U.S. Congress. U.S. Congress. URL: http://www.congress.gov/ [accessed 2023-01-26]
40. Stokes D, Purtle J, Meisel Z, Agarwal A. State legislators' divergent social media response to the opioid epidemic from 2014 to 2019: longitudinal topic modeling analysis. J Gen Intern Med 2021 Nov;36(11):3373-3382 [FREE Full text] [doi: 10.1007/s11606-021-06678-9] [Medline: 33782896]
41. Engel-Rebitzer E, Stokes DC, Buttenheim A, Purtle J, Meisel ZF. Changes in legislator vaccine-engagement on Twitter before and after the arrival of the COVID-19 pandemic. Hum Vaccin Immunother 2021 Sep 02;17(9):2868-2872 [FREE Full text] [doi: 10.1080/21645515.2021.1911216] [Medline: 33970786]
42. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text] [doi: 10.5555/944919.944937]
43. Julia S, David R. Text Mining with R - A Tidy Approach. Tidytextmining. URL: https://www.tidytextmining.com/ [accessed 2023-02-15]
44. McCallum A. Mallet: A machine learning for language toolkit. University of Massachusetts. 2002. URL: https://mallet.cs.umass.edu/index.php/Main_Page [accessed 2023-02-15]
45. Agarwal AK, Wong V, Pelullo AM, Guntuku S, Polsky D, Asch DA, et al. Online reviews of specialized drug treatment facilities-identifying potential drivers of high and low patient satisfaction. J Gen Intern Med 2020 Jun;35(6):1647-1653 [FREE Full text] [doi: 10.1007/s11606-019-05548-9] [Medline: 31755009]
46. Ranard BL, Werner RM, Antanavicius T, Schwartz HA, Smith RJ, Meisel ZF, et al. Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. Health Aff (Millwood) 2016 Apr;35(4):697-705 [FREE Full text] [doi: 10.1377/hlthaff.2015.1030] [Medline: 27044971]
47. Hutto C, Gilbert E. VADER: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media. 2014 May 16 Presented at: Eighth International AAAI Conference on Weblogs and Social Media (ICWSM-14); June 1-4, 2014; Ann Arbor, MI p. 216-225. [doi: 10.1609/icwsm.v8i1.14550]
48. Kingdon J, Stano E. Agendas, Alternatives, and Public Policies. Boston, MA: Little, Brown; 1984.
49. Gold K, Schwenk T, Sen A. Physician suicide in the United States: updated estimates from the National Violent Death Reporting System. Psychol Health Med 2022 Aug;27(7):1563-1575. [doi: 10.1080/13548506.2021.1903053] [Medline: 33861665]
50. Ye GY, Davidson JE, Kim K, Zisook S. Physician death by suicide in the United States: 2012-2016. J Psychiatr Res 2021 Feb;134:158-165. [doi: 10.1016/j.jpsychires.2020.12.064] [Medline: 33385634]
51. Kalmoe M, Chapman M, Gold J, Giedinghagen A. Physician suicide: a call to action. Mo Med 2019;116(3):211-216 [FREE Full text] [Medline: 31527944]

Abbreviations

LDA: latent Dirichlet allocation
VADER: Valence Aware Dictionary and Sentiment Reasoner

Edited by A Mavragani; submitted 12.04.22; peer-reviewed by Z Butt, D Lambert; comments to author 18.01.23; revised version received 26.01.23; accepted 10.02.23; published 24.02.23.

Please cite as:
Abrams MP, Pelullo AP, Meisel ZF, Merchant RM, Purtle J, Agarwal AK
State and Federal Legislators' Responses on Social Media to the Mental Health and Burnout of Health Care Workers Throughout the COVID-19 Pandemic: Natural Language Processing and Sentiment Analysis
JMIR Infodemiology 2023;3:e38676
URL: https://infodemiology.jmir.org/2023/1/e38676
doi:10.2196/38676
PMID:37013000

Original Paper

Analyzing Discussions Around Rural Health on Twitter During the COVID-19 Pandemic: Social Network Analysis of Twitter Data

Wasim Ahmed[1*], BA, MSc, PhD; Josep Vidal-Alaball[2,3,4*], MD, MPH, PhD; Josep Maria Vilaseca Llobet[4,5*], MD, PhD

[1]Stirling University Management School, University of Stirling, Stirling, United Kingdom

[2]Unitat de Suport a la Recerca de la Catalunya Central, Fundació Institut Universitari per a la Recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Sant Fruitós de Bages, Spain

[3]Health Promotion in Rural Areas Research Group, Gerència Territorial de la Catalunya Central, Institut Català de la Salut, Sant Fruitós de Bages, Spain

[4]Faculty of Medicine, University of Vic – Central University of Catalonia, Vic, Spain

[5]Primary Care Service, Althaia Xarxa Assistencial Universitària de Manresa, Manresa, Spain

[*]all authors contributed equally

Corresponding Author:
Wasim Ahmed, BA, MSc, PhD
Stirling University Management School
University of Stirling
Airthrey Road
Stirling, FK9 4LA
United Kingdom
Phone: 44 01786 467333
Email: Wasim.Ahmed@stir.ac.uk

Abstract

Background: Individuals from rural areas are increasingly using social media as a means of communication, receiving information, or actively complaining of inequalities and injustices.

Objective: The aim of our study is to analyze conversations about rural health taking place on Twitter during a particular phase of the COVID-19 pandemic.

Methods: This study captured 57 days' worth of Twitter data related to rural health from June to August 2021, using English-language keywords. The study used social network analysis and natural language processing to analyze the data.

Results: It was found that Twitter served as a fruitful platform to raise awareness of problems faced by users living in rural areas. Overall, Twitter was used in rural areas to express complaints, debate, and share information.

Conclusions: Twitter could be leveraged as a powerful social listening tool for individuals and organizations that want to gain insight into popular narratives around rural health.

(JMIR Infodemiology 2023;3:e39209) doi:10.2196/39209

KEYWORDS

rural health; Twitter messaging; social media; COVID-19; SARS-CoV-2; coronavirus; social network analysis

Introduction

Globalization and proliferation of the world wide web and social media have increased the amount of information available internationally. Access to information can be crucial in rural areas as it can help break the traditional isolation that those living in rural areas experience. In the past, it was believed that rural communities were isolated, with poor access to web-based information and being excluded from social media. This is partially true [1]. However, in recent times, in both high-income and transitional countries, a remarkable number of individuals from rural areas are using social media to communicate to receive up-to-date information and access quality health support and services [2].

It is well known that several societal and health issues are unique to rural areas when compared to those in urban areas. This includes high poverty rates, less access to health care, a higher percentage of adults with health problems, and health issues related to exposure to chemicals used in farming.

Twitter is a popular form of social media, and its use by health care professionals has been studied extensively [3]. Some examples include the use of Twitter as a means of health promotion by large urban hospitals and clinics in the United States [4]. Moreover, Twitter has also been used as a new source of data to study depression and its wider determinants in deprived populations in India and Brazil and for predictive analytics and sentiment analysis [5].

A recent study analyzing the implications of Twitter in health-related research identified a wide variety of themes ranging from professional education in health care to big data, social marketing and substance use, physical and emotional well-being of young adults, and public health and health communication [6]. The analysis of social media provides a useful tool for public health specialists and government decision-makers to gain insight into population reactions and feelings [7], especially in times of uncertainty such as the one we are facing with the present pandemic [8].

Misinformation has been a problem on social media platforms such as Twitter. A systematic review of the prevalence of health misinformation on social media before the COVID-19 pandemic found that 2 of the 6 principal categories were vaccines (32%) and pandemics (10%). The prevalence of health misinformation was the highest on Twitter [9]. Another paper published after the onset of the COVID-19 pandemic suggests understudied research areas that need to be addressed to improve policy and practice in response to health misinformation; those research areas include (1) spatial, temporal, network, and cross-platform dynamics of misinformation sharing and (2) the focus on vulnerable populations [10]. A recent bibliometric study of the scientific literature on medical and health-related misinformation on social media found that the most popularly investigated social media platform is Twitter and that COVID-19 is a common topic investigated across all platforms [11].

A study by Cuomo et al [12] analyzed the geospatial distribution of Tweets related to COVID-19 to try to illustrate the full scope of the pandemic. The authors found that rural areas in the United States engaged in COVID-19–related social media conversations at later stages of the outbreak than urban areas [12]. A person's birthplace has been regarded as an important determinant of health [13]. The availability of resources in rural areas differs from that in urban areas, and this has an impact on population health [14,15]. Another problem in rural areas is the shortage of health professionals willing to work in these areas [16]. Some initiatives are being developed to promote interest in rural health in this context. One such initiative uses social media for this objective. This is the case of the Rural Family Medicine Café, which, since 2015, has been organizing regular meetings using social networks to put in contact health professionals who work or have an interest in rural health [17,18].

There are few studies investigating the use of Twitter in relation to rural health issues that analyze popular topics covered in these areas. This is particularly interesting at the time of the COVID-19 pandemic, as contrary to the initial beliefs that lower population density could protect against the virus, COVID-19 did not spare these areas [19].

The main overall aim of our study is to analyze the conversations related to rural health taking place on Twitter during the COVID-19 pandemic to better understand popular narratives being communicated. Twitter is a popular social networking platform, and our study aims to shed light on the content hosted on the platform related to rural health.

More specifically, the objectives of this study are to study a particular time frame to (1) develop an understanding of the content and debates being shared on Twitter related to rural health, (2) to identify influential users around rural health on Twitter, and (3) uncover the key hashtags and websites being shared.

By fulfilling these objectives, the study will gain an understanding of rural health conversations taking place on Twitter during a specific phase of the COVID-19 pandemic between June and August 2021. The results are likely to be of interest to other scholars working in these areas as well as public health organizations and activists.

Methods

Sampling Tweets

This study made use of the Twitter Archiving Google Sheets (TAGS) tool to retrieve 15,586 tweets matching the keyword "rural health." TAGS draws upon the Twitter Search application programming interface to retrieve tweets. Although an English keyword is used, other languages may also exist, for example, if they reply to a tweet using a different language but the original tweet was posted in English or if they quote or reply to a tweet. Tweets were retrieved from June 10 to August 6, 2021, encompassing 57 days during the COVID-19 pandemic. No particular geographical location was selected from which to retrieve tweets, and tweets could be sent and received from anywhere in the world where Twitter is available. This is not a limitation of this study per se, as Twitter does not provide accurate location-based data and many studies are conducted using keywords. It is important to mention that there are numerous definitions of "rural" [20]. These definitions differ in the cutoff point. To avoid disputes, we use the simple principle that if one thinks one is rural, one probably is.

Although it can be argued that tweets are in the public domain, the project was careful not to draw attention to individual users acting in a personal capacity (preventing unwanted exposure). However, the users and key tweets reproduced in this study derive either from accounts and users in the public domain, social media influencers, health organizations, politicians, and academic journals.

Data Analysis

In order to identify influential users, the metric of betweenness centrality (the influence a user exerts on other users by his/her tweets) was applied, which is derived from the network theory and has been used in this study to find Twitter users that have an influence in our data set. This methodology has been used in previous research [21-23]. This metric was used in this study as it can identify users located in strategic locations within the network and who are gatekeepers of information propagation. It is commonly used in social media research to find important users in a network. The betweenness centrality scores are unique to each network and can be used to benchmark each of the users and rank them from most to least influential by betweenness centrality. Providing a detailed overview of network visualizations is beyond the scope of this study. Those new to network visualizations may wish to examine research in this area, which outlines common network patterns and how to interpret them [24].

Social Network Analysis

The software NodeXL (Social Media Research Foundation) was used to conduct a social network analysis of the data [17]. The network graph was laid out using the Clauset-Newman-Moore layout algorithm that is integrated into NodeXL. Social network analysis is the process of investigating social structures using networks and graph theory. This entails identifying and analyzing relations among entities and features in a social system. In our case, we analyzed the relationship among users by examining interaction patterns (retweets, replies, mentions, etc). We examined the whole network without prefiltering.

Time-Series Analysis

Time-series analysis is the study of data over time (Multimedia Appendix 1); it is the process of examining a time series to understand it and make predictions about future trends based on past data. It has applications in many areas, including economics, biology and medicine, engineering, environmental science, and meteorology. In this study, we made use of time-series analysis to gain an understanding of the volume of all tweet types across time.

Content Analysis and Natural Language Processing

NodeXL was also used to identify co-occurring word pairs, which is a type of natural language processing modality. In a word pair analysis, we looked at all the words in our data set and compared them to each other using their co-occurrence statistics (Multimedia Appendix 2). We then used these statistics to find all the word pairs that are likely to occur together more often than expected by chance.

Ethics Approval

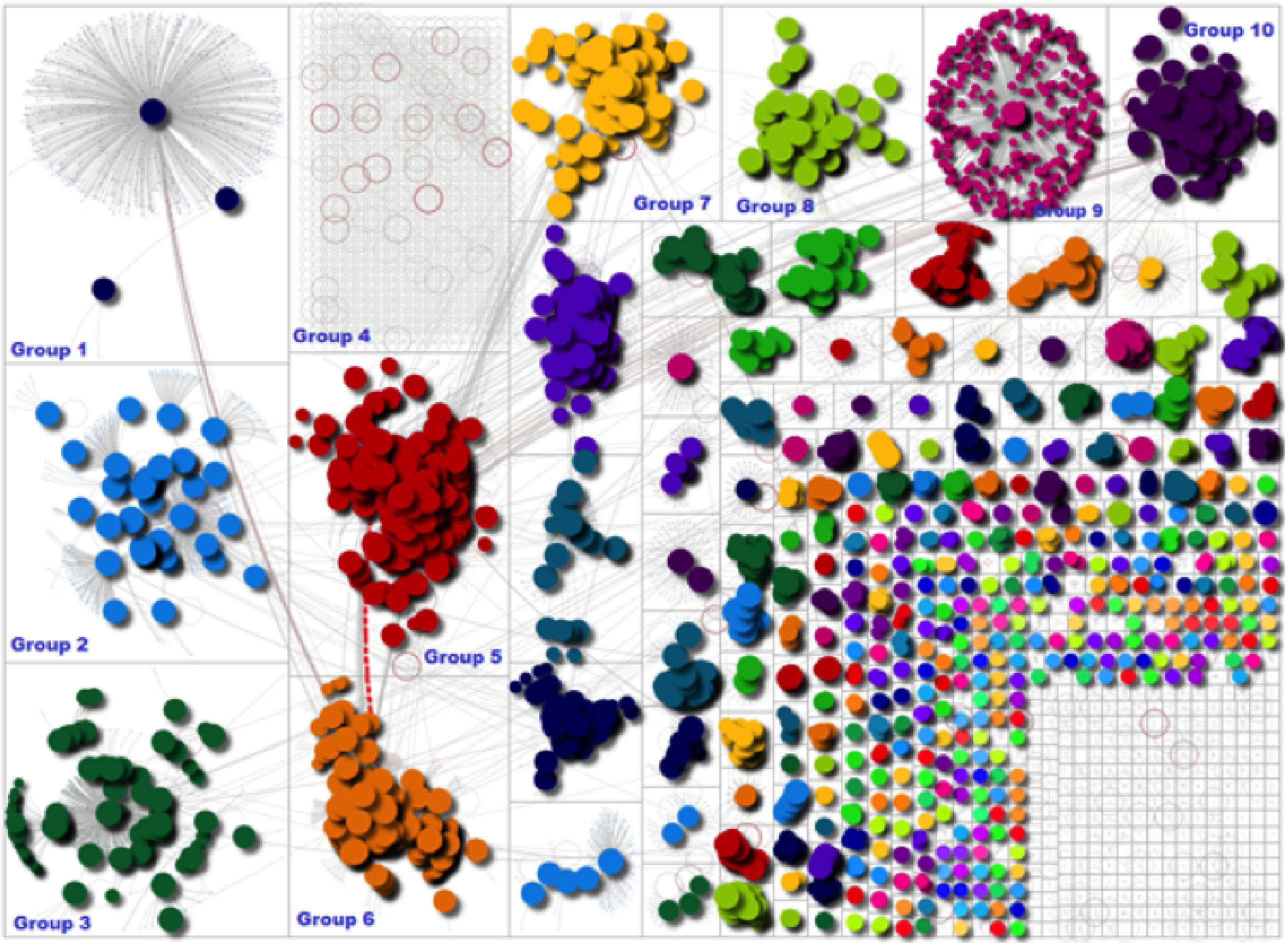
The study received ethical approval from Newcastle University (26055/2022).

Results

Results of Social Network Analysis

Figure 1 provides a visual representation of Twitter activity based on the data that were captured. The circles within the network represent individual Twitter users who were tweeting using the words "rural health," and the lines between them represent connections such as mentions and replies. Different colors are used to distinguish each of the groups, and they are listed from left to right, ranked by size, where group 1 is the largest cluster in the network, followed by group 2, and so forth.

Figure 1. Social network visualization and discussions across groups.



The figure highlights how there were several different groups of users who were conversing about different topics related to rural health. The largest group in the network (group 1) is that of a broadcast network where a user's tweet is retweeted with high frequency.

There are also several other smaller groups and broadcast networks providing the overall network a community shape. Other popular groups, such as groups 3, 5, 6, 7, and 10, indicate that there were different communities forming on Twitter, which were conversing about different topics related to rural health. There is little engagement (retweets and mentions) among users within some groups, but cross-group interaction was seen to occur between groups 5 and 9. Groups 1 and 9 appeared to be broadcast as network shapes. Multimedia Appendix 2 contains a full list of keywords associated with each of the clusters, providing insight into the types of topics that were being discussed.

More specifically, to provide more context on the content of the groups and communities represented within the visual, a range of news articles and reports were being amplified. For example, one article shared in group one was entitled, "India's healthcare workers are busting misinformation on WhatsApp." This was the most dominant news story being amplified in group 1. Aside from "rural, health" itself, the most popular word pair in group 1 was that of "fighting, covid" (n=873). Other interesting keywords identified within this group included "busting, myths" (n=873), indicating the combatting of misinformation, which was also linked to the aforementioned news article. If we cross-reference to the top 10 retweets and examine the tweet ranked as the third-most popular, it can be seen that this tweet uses the keyword "busting myths." In group 2, interesting word combinations (provided within the appendix) included "health, systems" (n=373), "expanding, medicaid" (n=256), "taxpayers, money" (n=254), and "affordable, health" (n=212). These keywords provide insight into the commonly used words and are helpful in understanding some of the topics that users were discussing. Multimedia Appendix 2 provides insight into groups 3 to 10 and an insight into some of the topics discussed.

Results of the Time-Series Analysis

Multimedia Appendix 1 provides an overview of the data set's unique edges (ie, tweets, retweets, mentions, etc). There appears to be a constant stream of Twitter activity, with 2 large peaks observed on June 18 and July 16, 2021, respectively. Overall, there appears to be much more activity taking place during June 2021. Upon investigating the peaks that were occurring within the data, it was found that these peaks tended to relate to spikes in retweets due to the tweets contained among the top 10.

Results Related to Key Users, Websites, Hashtags, and Retweets

Table 1 provides insight into the key users. The first key user is the account of Akhilesh Yadav, a socialist leader of India. This is followed by the Twitter account of the World Health Organization and the Rural and Remote Health Journal Twitter, an open-access academic international journal. In fourth place, we found the account of Senator Reverend Raphael Warnock, a US senator from Georgia, and in fifth place, the account of the National Rural Health Association, a US nonprofit organization with the mission to provide leadership on rural health issues through advocacy, communications, education, and research.

Table 2 provides information about the top websites used in tweets. The top website used in tweets and by far (877 occurrences) is from The Verge, an American technology news website. It features an article on how health care workers are combating misinformation about COVID-19 in rural India. The second-most used website in tweets is also related to India and is based on an article from The New York Times, which describes how the bodies floating at the river Ganges were buried at their shores and showed that the authorities were not telling the truth about the full extent of the death toll caused by COVID-19. The third-most used website was from a suspended account and no longer accessible. Finally, the fourth-most used website was from Gary Votour, who was running for the post of governor of South Carolina, and the fifth-most used website was an article from IndiaSpend, an Indian web-based journal, which discussed how Indian rural health centers were struggling with staff shortages, especially pharmacists and doctors.

Table 3 provides insight into the top hashtags used in the tweets. The most used hashtag is #appoint_pharmacist_for_rural_health, a hashtag used in a campaign to advocate for the appointment of pharmacists in rural India. The second-most used hashtag was #33yearsofpmk, a hashtag commemorating 33 years of the Paattali Makkal Katchi (working people's party), abbreviated as "PMK"—a political party in Tamil Nadu, India. The third-most used hashtag is directly related to rural health (#ruralhealth), and the fourth- and fifth-most used hashtags are 2 related hashtags, one in English and the other in Korean, to celebrate the birthday of Sunoo (birth name: Kim Sun-oo), a member of the Korean band ENHYPEN. This appears within the data because as result of the birthday packed lunches were delivered to the front-liners of the Los Banos, Laguna Rural Health Unit. The hashtags related to COVID-19 come in the 6th and 10th positions.

We also examined the top 10 retweets. It was found that the first, second, and fourth-ranking retweets were addressed to specific individuals. The most popular retweet was an appeal to the prime minister of India to appoint more rural doctors, and the second- and fourth-ranking retweets were related to a campaign to uncover water corruption in rural areas. The third-ranking retweet is a recognition of rural health activists who combat misinformation about COVID-19 in rural India. The other popular retweets had several purposes related to rural and public health: to report corruption related to rural health problems and the deplorable conditions of rural health care facilities, to congratulate a doctor by providing some key indicators of a rural health program milestone, to announce the building of health care facilities, and to report the shortage of health workforce and encourage professionals to work in rural areas.

Table 1. Key users by betweenness centrality.

User handle	Bio	Betweenness centrality
yadavakhilesh	Socialist Leader of India. Chief Minister of UP (2012 - 2017)	5185471
who	We are the #UnitedNations' health agency - #HealthForAll. Always check our latest tweets on #COVID19 for updated advice/information.	5150502
rrh_journal	Open-access, peer-reviewed journal providing an international evidence base to inform improvement in rural and remote health (free-to-read, no page charges)	4270264
senatorwarnock	United States Senator from Georgia. Pastor of Ebenezer Baptist Church.	3855435
ruralhealth	National Rural Health Association, 21k+ members nationwide, providing leadership and support at NRHA.	3000683
bprophetable	Only way to get good politicians is get rid of bad ones. I try to retweet facts and everyone's opinions including those I disagree with #FactsMatter	2476886
dainikbhaskar	India's Biggest Hindi Newspaper & News App. For Realtime News Updates, Local News for 2000 cities, Short Video News, Download our App: http://dainik-b.in/riOAhsOKg6	2403907
nytopinion	We amplify voices on the issues that matter to you. | Tell us what you think: letters@nytimes.com	2372269
timryan	Proud dad and husband, Ohio native, die-hard Browns fan. Running for U.S. Senate to fight like hell to cut workers in on the deal.	2349314
ruraldoctorsaus	Rural Doctors Association of Australia - promoting excellent medical care for rural and remote Australians.	2209906

Table 2. Top websites used in tweets.

Rank	Title	Tweets, n
1	India's Healthcare Workers Are Busting Misinformation On WhatsApp	877
2	The Ganges Is Returning the Dead. It Does Not Lie.	164
3	This tweet is from a suspended account	115
4	Official campaign website for Gary Votour for Governor of South Carolina	89
5	As Third Wave Looms, Rural Health Centres Struggle With Expired Drugs, Missing Doctors	69
6	Myth Vs Facts Government of India has been working towards effective COVID-19 management in rural India by sustained strengthening of the Rural health Infrastructure, and through focussed Public Health Measures in active collaboration with the States	61
7	Gary Votour for South Carolina Governor campaign	60
8	Chhattisgarh to privatise rural health infra; public health experts and activists demand roll back	55
9	Official Account Of Chhattisgarh Pradesh Congress Committee.	47
10	Barak Obama's twitter account, it reads: Today, the Supreme Court upheld the Affordable Care Act. Again. This ruling reaffirms what we have long known to be true: the Affordable Care Act is here to stay.	47

Table 3. Top hashtags used in tweets.

Rank	Top hashtags
1	appoint_pharmacist_for_rural_health
2	33yearsofpmk
3	ruralhealth
4	sunooourmiracleofjune
5	눈부신_선우의_열아홉번째_생일
6	covid19
7	pharmacistfederation
8	rural
9	medicaidsaveslivesact
10	covid

Results Related to Language and Geographical Locations

On examining the retweet count, the most widely used language was English, either in non-English native-speaking areas or in countries where other local languages are spoken. This is likely to be influenced by the keywords that were used to retrieve data as these were in English. In total, 4553 (80%) retweets are written only in English. Including tweets that mix English and other languages accounts for 94.5% (n=5350) of retweets. The second-most used language is Tagalog, mixed with English in the main body of the Tweet (n=407, 7 %). The third-most used language is Korean, but in this case used only as a hashtag; the retweets' main text is in English (n=390, 6.9%). The fourth-most used language is Hindi, only used in one of the top 10 retweets in our data set (n=311, 5.5%). Other languages can easily appear even if English keywords are used to retrieve the data because a tweet written in English can be quoted by a user writing in a different language, which would be included in our data set. The geographical locations of the debates are mainly India, Pakistan, Australia, the Philippines, and the United States. The most widely used language was English. Other languages used were Hindi, Korean, and Tagalog.

Discussion

Principal Findings

Although the role of social media in rural settings has been studied previously [2,18], to our knowledge, this is the first study on the specific use of Twitter in relation to rural health issues and has identified the common topics discussed in these settings at a specific point in time.

Our study also found that the key users related to this topic are individuals (mainly politicians) and organizations dealing with aspects related to rural health. The top websites used in the tweets specialized in neither health care nor public health. The tweets sometimes used wide audience sources, such as international newspapers (The New York Times) or local press. Key opinion leaders have a big influence on the spread of factual information [23], and health authorities could make more use of Twitter to publish news and articles related to rural health and COVID-19.

The most frequently used hashtags were able to uncover interesting and surprising connections to rural health. They included a celebration of the birthday of a top Korean boy band member and the anniversary of the foundation of an Indian political party. These occurred as packed lunches were donated to a rural health center on the front line due to the birthday, and in the case of the political party, it has strong relations with rural areas. The most used hashtag was related to a campaign requesting the appointment of a pharmacist in rural areas, indicating the shortage of pharmacists in these settings. The COVID-19 hashtag was also popular, being used in 2 different forms: "covid19" and "covid."

The top 10 retweets explicitly mention rural health, health care, or public health problems. The topics are generally of local interest, pointing at very specific issues. Even when rural health is part of a politician's campaign or a politician's comment, its interest is local or national. The main uses of Twitter identified in our study are complaints, debates, information sharing, acknowledgements, advertisements, and political campaigns. Regarding the geographical locations of the top tweets, the most influential tweets were derived from India. This is not surprising, given the size of India and the number of rural areas therein. The United States, the Philippines, and South Korea are also among the most frequent locations from where influential tweets were obtained.

The study has several limitations. A circumscribed 57-day time was examined, which may have excluded certain tweets falling outside this period. Another limitation is that the Twitter Search application programming interface can only retrieve data from public-facing Twitter accounts and not from private accounts; however, most accounts are set as public. Another limitation is that as our study retrieved data using a very specific keyword (rural health), our data may have excluded tweets from users who tweeted without using our target keyword. Furthermore, the study retrieved many tweets from other widely spoken languages, such as German or French, which may arise from the limited number of keywords used when retrieving data. Tweets from India occurred in higher frequency than those from other countries. This is potentially because of the huge rural population of the country; this is because India has the largest total rural population [25]. However, our aim was to examine content on Twitter, and content from India happened to be popular at the time we sampled data.

Assessing the needs for those living in rural communities has traditionally been challenging. Several circumstances have been a constraint: language as a barrier, isolation, lack of registries, difficulties to carry out interviews, location of the households, and expenditure to perform studies. Twitter could prove to be a solution for these problems and could be used as a social listening tool to identify the concerns and needs of rural communities. Our study shows that Twitter can be effectively used as a means of communication in rural areas and as a source of information on rural health. Moreover, the information existing on Twitter, when filtered by geographical locations, may be of interest to stakeholders, health care workers, politicians, patients, and communities in general.

Twitter could also be used strategically for those living in rural areas to communicate with one another, for sharing local updates, and to warn of disasters and areas to avoid. It could also be used to connect to share resources and supplies. This could be facilitated using domain-specific hashtags related to each area and widely advertised and popularized locally.

Conclusions

Twitter has been shown to be a powerful means of communicating about important issues around rural health. Twitter is a tool that can be used to raise awareness of the problems existing in rural health. When examining tweets in English, it was found that India has the most Twitter-related conversations on rural health. Twitter was used to discuss rural settings to express complaints, debate, share information, acknowledge somebody or something, and create advertisements or politician's campaigns. Twitter could be leveraged as a powerful source of information for individuals and organizations working on rural health and as a means to identify popular narratives and hot issues around this topic.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Time-series chart of Twitter activity.
[PNG File , 106 KB - infodemiology_v3i1e39209_app1.png]

Multimedia Appendix 2
Associations between word pairs and groups.
[DOCX File , 18 KB - infodemiology_v3i1e39209_app2.docx]

References

1. Greenberg AJ, Haney D, Blake KD, Moser RP, Hesse BW. Differences in access to and use of electronic personal health information between rural and urban residents in the United States. J Rural Health 2018 Feb 11;34 Suppl 1(Suppl 1):s30-s38 [FREE Full text] [doi: 10.1111/jrh.12228] [Medline: 28075508]
2. Mehmet M, Roberts R, Nayeem T. Using digital and social media for health promotion: a social marketing approach for addressing co-morbid physical and mental health. Aust J Rural Health 2020 Apr 22;28(2):149-158. [doi: 10.1111/ajr.12589] [Medline: 31970832]
3. Pershad Y, Hangge PT, Albadawi H, Oklu R. Social medicine: Twitter in healthcare. J Clin Med 2018 May 28;7(6) [FREE Full text] [doi: 10.3390/jcm7060121] [Medline: 29843360]
4. Kordzadeh N. Health promotion via Twitter: a case study of three medical centers in the USA. Health Promot Int 2022 Apr 29;37(2). [doi: 10.1093/heapro/daab126] [Medline: 34379755]
5. Thapa B, Torres I, Koya SF, Robbins G, Abdalla SM, Arah OA, et al. Use of data to understand the social determinants of depression in two middle-income countries: the 3-D Commission. J Urban Health 2021 Aug;98(Suppl 1):41-50 [FREE Full text] [doi: 10.1007/s11524-021-00559-6] [Medline: 34409557]
6. Yeung AWK, Kletecka-Pulker M, Eibensteiner F, Plunger P, Völkl-Kernstock S, Willschke H, et al. Implications of Twitter in health-related research: a landscape analysis of the scientific literature. Front Public Health 2021;9:654481 [FREE Full text] [doi: 10.3389/fpubh.2021.654481] [Medline: 34307273]
7. Vidal-Alaball J, Fernandez-Luque L, Marin-Gomez FX, Ahmed W. A new tool for public health opinion to give insight into telemedicine: Twitter poll analysis. JMIR Form Res 2019 May 28;3(2):e13870 [FREE Full text] [doi: 10.2196/13870] [Medline: 31140442]
8. De Rosis S, Lopreite M, Puliga M, Vainieri M. The early weeks of the Italian Covid-19 outbreak: sentiment insights from a Twitter analysis. Health Policy 2021 Aug;125(8):987-994 [FREE Full text] [doi: 10.1016/j.healthpol.2021.06.006] [Medline: 34176671]
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. Sylvia Chou W, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. Am J Public Health 2020 Oct;110(S3):S273-S275. [doi: 10.2105/ajph.2020.305905]
11. Yeung AWK, Tosevska A, Klager E, Eibensteiner F, Tsagkaris C, Parvanov ED, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. J Med Internet Res 2022 Jan 25;24(1):e28152 [FREE Full text] [doi: 10.2196/28152] [Medline: 34951864]
12. Cuomo RE, Purushothaman V, Li J, Cai M, Mackey TK. Sub-national longitudinal and geospatial analysis of COVID-19 tweets. PLoS One 2020 Oct 28;15(10):e0241330 [FREE Full text] [doi: 10.1371/journal.pone.0241330] [Medline: 33112922]
13. Cofiño Fernández R. [Your post code is more important for your health than your genetic code (1)]. Aten Primaria 2013 Mar;45(3):127-128 [FREE Full text] [doi: 10.1016/j.aprim.2013.02.001] [Medline: 23499154]
14. Lee H, Lin CC, Snyder JE. Rural-urban differences in health care access among women of reproductive age: a 10-year pooled analysis. Ann Intern Med 2020 Dec 01;173(11 Suppl):S55-S58 [FREE Full text] [doi: 10.7326/M19-3250] [Medline: 33253026]
15. Rahman M, White EM, Thomas KS, Jutkowitz E. Assessment of rural-urban differences in health care use and survival among Medicare beneficiaries with Alzheimer disease and related dementia. JAMA Netw Open 2020 Oct 01;3(10):e2022111 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.22111] [Medline: 33090226]
16. Lehmann U, Dieleman M, Martineau T. Staffing remote rural areas in middle- and low-income countries: a literature review of attraction and retention. BMC Health Serv Res 2008 Jan 23;8:19 [FREE Full text] [doi: 10.1186/1472-6963-8-19] [Medline: 18215313]
17. Floss M. How to Do a Rural Cafe: A Short Manual. 2018. URL: https://wonca.net/site/DefaultSite/filesystem/documents/Groups/Rural%20Practice/Manual%20Rural%20Cafe.pdf [accessed 2021-08-14]
18. Wheatley A, Floss M, Bakola M, Kampouraki M, Silveira B, Scott-Jones J. The Rural Family Medicine Café Project: a social media strategy to reduce occupational isolation and improve support for rural healthcare professionals. Front Public Health 2020;8:595255 [FREE Full text] [doi: 10.3389/fpubh.2020.595255] [Medline: 33330339]
19. Hart LG, Larson EH, Lishner DM. Rural definitions for health policy and research. Am J Public Health 2005 Jul;95(7):1149-1155. [doi: 10.2105/ajph.2004.042432]
20. Cuadros DF, Branscum AJ, Mukandavire Z, Miller FD, MacKinnon N. Dynamics of the COVID-19 epidemic in urban and rural areas in the United States. Ann Epidemiol 2021 Jul;59:16-20 [FREE Full text] [doi: 10.1016/j.annepidem.2021.04.007] [Medline: 33894385]
21. Ahmed W, Marin-Gomez X, Vidal-Alaball J. Contextualising the 2019 e-cigarette health scare: insights from Twitter. Int J Environ Res Public Health 2020 Mar 26;17(7):2236 [FREE Full text] [doi: 10.3390/ijerph17072236] [Medline: 32225020]
22. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 May 06;22(5):e19458 [FREE Full text] [doi: 10.2196/19458] [Medline: 32352383]
23. Fuster-Casanovas A, Das R, Vidal-Alaball J, Lopez Segui F, Ahmed W. The #VaccinesWork hashtag on Twitter in the context of the COVID-19 pandemic: network analysis. JMIR Public Health Surveill 2022 Oct 28;8(10):e38153 [FREE Full text] [doi: 10.2196/38153] [Medline: 36219832]
24. Smith M, Himelboim I, Rainie L, Shneiderman B. The Structures of Twitter Crowds and Conversations. In: Matei SA, Russell MG, Bertino E, editors. Transparency in Social Media: Tools, Methods and Algorithms for Mediating Online Interactions. Cham: Springer; 2015:67-108.
25. Rural population - Country Ranking. Index Mundi. 2019. URL: https://www.indexmundi.com/facts/indicators/SP.RUR.TOTL/rankings [accessed 2023-03-01]

Abbreviations

TAGS: Twitter Archiving Google Sheets

Edited by T Mackey; submitted 02.05.22; peer-reviewed by F Lopez Segui, C Pérez-Aldana; comments to author 15.08.22; revised version received 22.11.22; accepted 25.02.23; published 08.03.23.

Please cite as:
Ahmed W, Vidal-Alaball J, Vilaseca Llobet JM
Analyzing Discussions Around Rural Health on Twitter During the COVID-19 Pandemic: Social Network Analysis of Twitter Data
JMIR Infodemiology 2023;3:e39209
URL: https://infodemiology.jmir.org/2023/1/e39209
doi:10.2196/39209
PMID:36936067

Original Paper

Estimating Rare Disease Incidences With Large-scale Internet Search Data: Development and Evaluation of a Two-step Machine Learning Method

Jiayu Li[1], PhD; Zhiyu He[1], MSc; Min Zhang[1], PhD; Weizhi Ma[2], PhD; Ye Jin[3], MD; Lei Zhang[4], MD; Shuyang Zhang[5], MD; Yiqun Liu[1], PhD; Shaoping Ma[1], PhD

[1]Department of Computer Science and Technology, Tsinghua University, Beijing, China

[2]Institute for AI Industry Research, Tsinghua University, Beijing, China

[3]Department of Medical Research Center, Peking Union Medical College Hospital, Beijing, China

[4]Department of Nephrology, Peking Union Medical College Hospital, Beijing, China

[5]Department of Cardiology, Peking Union Medical College Hospital, Beijing, China

Corresponding Author:
Shuyang Zhang, MD
Department of Cardiology
Peking Union Medical College Hospital
No.1 Shuaifuyuan
Wangfujing, Dongcheng District
Beijing, 100730
China
Phone: 86 010 69151188
Email: shuyangzhang103@nrdrs.org

Abstract

Background: As rare diseases (RDs) receive increasing attention, obtaining accurate RD incidence estimates has become an essential concern in public health. Since RDs are difficult to diagnose, include diverse types, and have scarce cases, traditional epidemiological methods are costly in RD registries. With the development of the internet, users have become accustomed to searching for disease-related information through search engines before seeking medical treatment. Therefore, online search data provide a new source for estimating RD incidences.

Objective: The aim of this study was to estimate the incidences of multiple RDs in distinct regions of China with online search data.

Methods: Our research scale included 15 RDs in China from 2016 to 2019. The online search data were obtained from Sogou, one of the top 3 commercial search engines in China. By matching to multilevel keywords related to 15 RDs during the 4 years, we retrieved keyword-matched RD-related queries. The queries used before and after the keyword-matched queries formed the basis of the RD-related search sessions. A two-step method was developed to estimate RD incidences with users' intents conveyed by the sessions. In the first step, a combination of long short-term memory and multilayer perceptron algorithms was used to predict whether the intents of search sessions were RD-concerned, news-concerned, or others. The second step utilized a linear regression (LR) model to estimate the incidences of multiple RDs in distinct regions based on the RD- and news-concerned session numbers. For evaluation, the estimated incidences were compared with RD incidences collected from China's national multicenter clinical database of RDs. The root mean square error (RMSE) and relative error rate (RER) were used as the evaluation metrics.

Results: The RD-related online data included 2,749,257 queries and 1,769,986 sessions from 1,380,186 users from 2016 to 2019. The best LR model with sessions as the input estimated the RD incidences with an RMSE of 0.017 (95% CI 0.016-0.017) and an RER of 0.365 (95% CI 0.341-0.388). The best LR model with queries as input had an RMSE of 0.023 (95% CI 0.017-0.029) and an RER of 0.511 (95% CI 0.377-0.645). Compared with queries, using session intents achieved an error decrease of 28.57% in terms of the RER (P=.01). Analysis of different RDs and regions showed that session input was more suitable for estimating the incidences of most diseases (14 of 15 RDs). Moreover, examples focusing on two RDs showed that news-concerned session intents reflected news of an outbreak and helped correct the overestimation of incidences. Experiments on RD types further indicated that type had no significant influence on the RD estimation task.

Conclusions: This work sheds light on a novel method for rapid estimation of RD incidences in the internet era, and demonstrates that search session intents were especially helpful for the estimation. The proposed two-step estimation method could be a valuable supplement to the traditional registry for understanding RDs, planning policies, and allocating medical resources. The utilization of search sessions in disease detection and estimation could be transferred to infoveillance of large-scale epidemics or chronic diseases.

(JMIR Infodemiology 2023;3:e42721) doi:10.2196/42721

KEYWORDS

disease incidence estimation; rare disease; internet search engine; infoveillance; deep learning; public health

Introduction

Background

Rare diseases (RDs) refer to a group of diseases with very low prevalence (usually less than 0.05% of the population [1]). There are more than 7000 known RDs and more than 400 million people are affected by RDs worldwide [2]. Because of their diseases, patients with RDs often experience social discrimination and financial hardship [3]. Most RDs have a genetic or congenital cause, and over half of patients with RDs have varying degrees of disabilities [4]. The burden of disease management and income decrease due to the disorders have resulted in poverty being a common experience for families coping with RDs [5]. Therefore, RDs have become an essential concern in public health, attracting substantial research attention.

Disease surveillance (ie, detecting the incidences of diseases) is a common but crucial method for understanding RDs [6]. Traditional surveillance registries are based on consistent case reporting from workers in ubiquitous surveillance systems [7]. However, RDs incidence detection is challenging for traditional registry systems for several reasons: (1) the diagnosis of most RDs is extremely complicated, and it takes approximately 6-8 years to get an accurate diagnosis [2], resulting in complex registry records of RD patients; (2) different RDs belong to different clinical departments or systems, making it difficult to integrate data from various registry institutions; and (3) the cases of RDs are so scarce that maintaining timely reports will be a resource-intensive task.

Therefore, researchers have been seeking to detect or estimate the incidences of RDs with indirect information. For instance, various international and national platforms were constructed for collecting RD knowledge and incidences [6,8,9].

With the development of the internet, a tremendous amount of data was created online. Infoveillance (ie, using online information for syndromic surveillance [10]) has been successfully applied in many studies [11]. Diverse sources of online data greatly enrich the information for disease estimation, such as Wikipedia views [12], News views [13], medical forum blogs [14], and search engine data [15].

Nevertheless, to our knowledge, no study has yet explored the possibility of using infoveillance data in RD incidences estimation, and the existing research has not paid attention to the context information of disease-related data in the online environment, such as searching sessions in the search engines. However, comparing online search data to RD incidences and further estimating RD incidences is beneficial. Search engine data will locate the patients and families from the source, which is more convenient than a multiround clinical diagnosis and registry. In addition, search engines provide unlimited information, which can be used to break the barriers between RDs in different clinical departments. Hence, search engine data can make it possible to estimate multiple RDs in multiple locations simultaneously.

Prior Work

Because few studies have focused on estimating RD incidences with online information, we reviewed prior research about employing online data in detecting or estimating epidemic and chronic diseases, and evaluated their differences with respect to RD incidences estimation.

Since the spread of epidemic diseases will cause an increase of related online searches, several studies have focused on the detection and prediction of epidemic diseases using infoveillance methods [16]. The new approaches began with estimating trends of influenza [15,17]. Subsequently, the query volume of search engines has been widely used to detect flu [18,19], dengue [20], pandemic H1N1 [21], and other diseases. Beyond search data, Xu et al [22] further considered the influence of news, which was used to detect occurrences of hand-foot-and-mouth disease with related queries, news clicks, and page clicks, improving the disease detection performance. In recent years, geographical information has been considered for infoveillance. Researchers tried to predict flu trends in multiple locations simultaneously [19,23] or transferred a trained disease prediction model to new regions [24]. During the ongoing COVID-19 pandemic, web search data have also shown great utility in disease surveillance [25-27].

In addition to epidemics, infoveillance has also been utilized in chronic diseases and other disorders. Ram et al [28] tried to estimate the number of asthma patients at a specific hospital with data from Google Trends, Twitter, and nearby air quality. Correlation analysis between eye disease trends and related queries showed a significant interrelationship between disease cases and online data [29]. Tkachenko et al [30] revealed that Google Trends could detect early signs of diabetes by monitoring combinations of keywords in online search queries. Sleep disorders [31] and mental health problems [32] were also found to be related to search volumes.

These previous works on epidemics and chronic diseases showed great successes of infoveillance, which inspired us to apply search data for RDs incidence estimation. Nevertheless, existing methods cannot be used directly for RDs because RDs remarkably differ from epidemics or common chronic diseases.

In all previous studies based on search engine data, disease-related queries were extracted and the number (volume) of queries was used as the model input. However, RD-related search behaviors may be caused by cyberchondria (ie, an unfounded escalation of anxiety about common symptomatology), as search engines can potentially escalate medical concerns [33]. Our experiment also revealed that RD-related search behaviors are sparse, and only a minority of them are actually based on a concern about RDs. Therefore, besides query numbers, more information related to users' search process is needed for accurate RDs estimation.

Objective

The aim of this study was to estimate the incidences of multiple RDs in distinct regions using search engine data.

As RD-related search behaviors are sparse and complex, it is not suitable to utilize RD-related query numbers directly for RD incidence estimation. Therefore, we designed a two-step machine learning method to estimate RD incidences with the volume of search sessions that concern RDs. The RD-related *queries* were selected by matching the search logs with RD-specific keywords. The *search sessions* were constructed with the queries submitted in the period before and after the RD-related queries.

The two-step method is as follows. In the first step, the intents of search sessions are predicted. Users' search intents indicate their purpose when querying RD-related questions on the search engine. The intents vary when the users mention RD-related queries in the session, such as seeking medical resources for patients, learning about news, searching for answers to medical assignments, and out of curiosity. By identifying sessions specifically concerned with RDs, we could filter out the noise from the RD-related search data effectively. In the second step, the incidences of multiple RDs are estimated in multiple regions with the volume of different session intents. RD incidences could be estimated more accurately with the filtered session numbers. Following previous works on disease detection with search engine data [15,23,24,34], linear regression (LR) without autoregressive modeling of historical RD incidences was considered when estimating RD incidences from search session intents.

The novel aspects of this study are two-fold. First, to our best knowledge, this is the first study to utilize search engine data in the estimation of multiple RD incidences, paving a new direction for improved understanding of RDs. This study therefore provides a helpful supplement to traditional RD registry systems. Second, the proposed approach introduces search sessions, especially session intents, into search engine–based infoveillance. The experimental results showed significant improvement when session intents were considered. The search session information could also be applied for the infoveillance of other diseases.

Methods

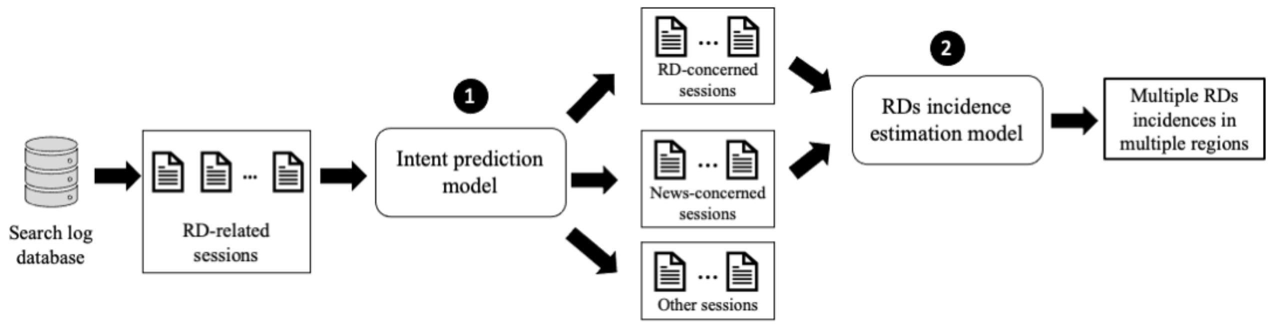
Overview and Framework

In this study, a two-step method was designed to estimate the incidences of RDs from search engine data. The first step was to distill RD-related search sessions and predict their intents into three categories: RD-concerned, news-concerned, and others. The second step was to estimate multiple RD incidences based on the volume of RD-concerned sessions and news-concerned sessions. Figure 1 shows an overview framework of the proposed two-step method.

The method was applied to search data of 15 RDs in 4 regions in China during 16 seasons from 2016 to 2019. To evaluate the results, we compared the estimated incidences with RD incidences collected from China's national multicenter clinical database of RDs [5].

Below, we describe the clinical RD incidences data (ie, the ground truth) and search data, followed by descriptions of the first and second steps in more detail, and the experimental settings.

Figure 1. Overview framework of the two-step rare disease (RD) incidences estimation method.



Data Collection

Ethical Approval

This study was approved by the Ethics Committee of Peking Union Medical College Hospital (S-k1790).

RD Types and Incidences

All data used in this study were anonymized statistics. A medical professional in the RD scenario helped us select RDs from the Compendium of China's First List of Rare Diseases (2018) [35]. A total of 15 RDs were selected, containing diseases from diverse departments and had stable long-term data in the registry database. Names and the types of the 15 RDs are listed in Multimedia Appendix 1. More details about the experiments evaluating the influence of RD types are provided in Multimedia Appendix 2.

We obtained the clinical RD incidences data from China's national multicenter clinical database of RDs [5]. The data set included anonymized confirmed RD cases from 2016 to 2019 reported by more than 300 hospitals across China. The cases were grouped by their diseases (1 of 15 RDs), confirmed time (16 seasons for 2016-2019), and permanent residence locations (one of the four regions in China's mainland [36]). The RD incidences were determined by dividing the case numbers by the regional population. Ultimately, we obtained incidences of the 15 RDs in 16 seasons (ie, 4 years) in four regions in China.

Online Search Data

We collected RD-related queries and their clicked documents from Sogou, one of the top-3 commercial search engines in China. The data were completely anonymized and no personalized information was collected. The side information included the search time and province located by IP address. No specific location was recorded.

First, we collected multisource medical knowledge to form keywords for each RD. Three levels of keywords, ranked by how closely they were associated with the RDs, were considered in our experiments: level 1 included RD-specific keywords, which helped to locate RD-related queries precisely from massive irrelevant queries; level 2 included RD-related nonspecific keywords to indicate how close the queries were related to an RD; and level 3 comprised general medical keywords, which helped determine whether the queries were likely to have medical-related concerns. Experts provided specific keywords about each RD, including disease names, specific genes, and specific treatments, which were defined as level 1 keywords. Based on China's Guide for the Diagnosis and Treatment of Rare Diseases (2019) [37], we extracted symptoms and pleiotropic treatments for each RD as level 2 keywords. An open medical lexicon [38] on general medical knowledge was treated as level 3 keywords. The lists of level 1 and level 2 keywords are provided in Multimedia Appendix 3, and the level 3 keywords are available from the open lexicon [38].

We matched and saved all queries that contained each level 1 keyword (corresponding to RD names, specific genes, or specific treatments) from all logs of the Sogou search database from 2016 to 2019. Search queries from all level 1 keywords were then merged to constitute the Query Set *Q*, including 2,749,257 queries related to 15 RDs. *Q* could be divided into three categories according to the matched keyword types: 2,615,272 name-related queries, 50,022 gene-related queries, and 83,963 treatment-related queries.

Finally, we introduced the *session* in users' search process, where a sequence of queries submitted by the same user within 30 minutes formed a session. To be specific, for each query q in Query Set Q for a user u, we backtracked u's query logs before query q until the interval between a certain query q_s and the previous query was greater than 30 minutes, and query q_s was then taken as the beginning of the session. We traced u's query logs after q until the interval between a certain query q_e and the next query was greater than 30 minutes, and query q_e was then taken as the end of the session. In this way, all sessions with at least one query in Q were distilled as the RD-related Session Set S, including 1,769,986 sessions. All queries in S were then marked with the highest-level keywords they contained. Queries containing level 1 keywords were selected as the *key queries* in the session. In this way, for each query in S, we collected the documents that the user clicked under the query. Due to privacy concerns, we only used the URL domains and positions (ie, the rank of the document in the list searched by the query) of the documents.

Session Intent Prediction

Session Intent

Session intent prediction is the first step of our two-step method, which serves to recognize the user intent behind each session in Session Set S, providing inputs for the second step. Users' search intents varied when using the search engine [39]. Although sessions in S all mentioned RD-related keywords, they might not come from RD patients or their family members who actually care about RDs. For instance, users might be searching for news, homework assignments, or just out of curiosity. Therefore, it is necessary to distinguish session intents (ie, users' intents when querying the sessions) in Session Set S. We grouped session intents into three categories: *RD-concerned*, *news-concerned*, and *others*. It was considered particularly important to distinguish the news-concerned sessions because breaking news would substantially increase the overall search volume, which would consequently influence the correlation between search volume and disease incidences [22].

Feature Extraction

Session-level features and sequences of query-level features were extracted for each session in S for predicting the session intent, considering both statistical features and semantic features.

The session-level and query-level statistical features are shown in Table 1. Among them, the Word_freq_change indicated whether a word appeared intensively in queries during a given period. This is a helpful feature to distinguish news-concerned sessions since breaking news will increase the frequency of some uncommon words. The word frequency change $C(w_i, t_k)$ of a word w_i in period (ie, season) t_k is defined as:

$$C(w_i, t_k)=[n(w_i, t_k)+\alpha]/[\sum^{K}_{j=1}n(w_i, t_j)/K+\alpha]$$

where $n(w_i, t_k)$ is the word frequency of w_i in period t_j, K is the number of periods, and α=1 for smoothing. At the query level, Word_freq_change is the mean value of all words in the query. At the session level, this feature represents the mean value of all queries in the session.

Both query and document semantic meanings were considered for the semantic features. The frequency of words and document URL domains were calculated separately for each of the three session intent classes. The words and URLs with a high frequency for one intent class and low frequencies for the other two classes were then selected as intent-specific words and URLs. The top 5 intent-specific words and URLs of each intent were selected, forming a set of 15 words and 15 URLs. A

30-dimension session-level vector was then used as a session feature to represent whether each word or URL appeared in a session. Moreover, whether level 1 keywords of each RD appeared in a query was represented with a multihot embedding vector of length 15 (ie, 15 RDs in the data set) as a query feature.

Finally, for a session S_i containing n_i queries, session-level features were concatenated as a vector, , including 8 dimensions for statistical features and 30 dimensions for semantic features, and query-level features formed a feature sequence , where is the feature vector of the mth query.

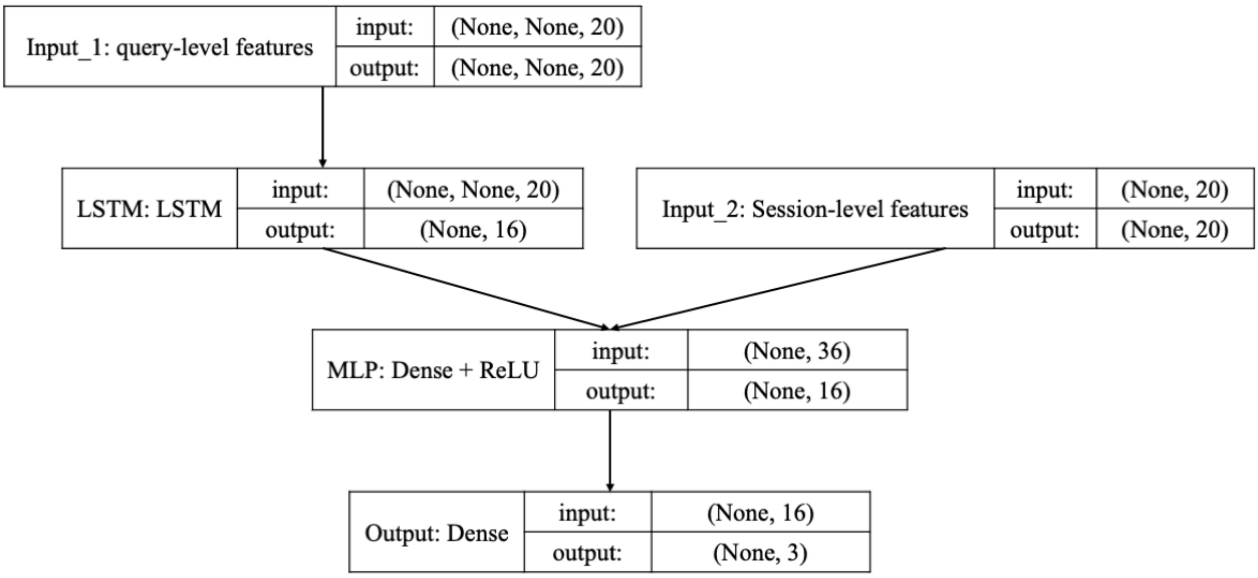
Table 1. Statistical features used for predicting session intents.

Feature name	Category	Description
Session_len	Session	Session length (ie, number of queries in a session)
Query_type	Query	Level of query (ie, the highest-level keywords a query contains)
Key_num	Session	Number of key (ie, level 1) queries in a session
Q2_num	Session	Number of level 2 queries in a session
Q3_num	Session	Number of level 3 queries in a session
Query_len	Query	Query length (ie, number of words in a query)
Click_num	Query	Number of clicked documents in a query
Sum_click_num	Session	Number of clicked documents in a session
Position_max	Query	Maximum position of clicked documents in the ranking list (set to 0 if no document is clicked)
All_position_max	Session	Maximum of Position_max of all queries in a session
Position_mean	Query	Average position of clicked documents in the ranking list
All_position_mean	Session	Average of Position_mean of all queries with clicked documents in a session
Word_freq_change	Query	Average word frequency change of all words in a query
All_word_freq_change	Session	Average of Word_freq_change of all queries in a session

Model Construction

After both sequential features and vector features were extracted for intent prediction, a combination of the long short-term memory (LSTM) and multilayer perceptron (MLP) algorithms was used to predict the session intents. The LSTM model is a recurrent neural network that is widely applied for modeling time-series data when the features are sequential [40]. In our work, an LSTM model was employed to transform the sequential features into a vector . Subsequently, and were concatenated and fed into a 1-layer MLP model with a rectified linear unit (ReLU) as an activation function to predict the session intents. The model structure is shown in Figure 2.

Figure 2. Model structure for session intent prediction. LSTM: long short-term memory; MLP: multilayer perceptron; ReLU: rectified linear unit.



Multiple RD Incidences Estimation

Input and Output Construction

To conduct the experiments on incidences estimation for 15 RDs in 16 seasons (ie, 4 years from 2016 to 2019) in 4 regions in China, we constructed the input and output of the second step for multiple RD incidences estimation as shown in Textbox 1.

For the ground truth labels, since the RDs incidence was very low (usually on the 1e–6 order of magnitude), the incidence was rescaled so that the maximum incidence was equal to 1.

Textbox 1. Input and output for multiple rare disease (RD) incidence estimation.

Input

- number of RD-concerned sessions x_{sd} (d_i, l_j, t_k)
- number of news-concerned sessions x_{sn} (d_i, l_j, t_k)

Output

- estimated incidence of RD $\hat{y}(d_i, l_j, t_k)$, where d_i, l_j, and t_k indicate the *i*th RD, *j*th region, and *k*th period, respectively

LR Model on Multiple RDs and Regions

Following previous research in infoveillance [15,25,34], we chose LR to estimate the incidences of multiple RDs. As the task was to estimate the incidences of multiple RDs in multiple regions, three variants of LR were constructed as incidence estimators.

The first LR model was a general LR, with all of the different RDs and regions estimated with the same set of parameters:

$$\hat{y}(d_i, l_j, t_k)=\alpha_d x_{sd}(d_i, l_i, t_k)+\alpha_n x_{sn}(d_i, l_j, t_k)+\beta,$$

where α_d, α_n, and β are learnable parameters.

The second LR model was an LR with specific parameters for disease (*LR Spec. D.* for short), where estimators of the same RD share parameters:

$$\hat{y}(d_i, l_j, t_k)=\alpha_d(d_i)x_{sd}(d_i, l_j, t_k)+\alpha_n(d_i)x_{sn}(d_i, l_j, t_k)+\beta(d_i)$$

and $\alpha_{n/d}(d_i)$ and $\beta(d_i)$ indicate the learnable parameters for the RD d_i.

The last LR model adopted specific parameters for both disease and regions (*LR Spec. D. L.* for short):

$$\hat{y}(d_i, l_j, t_k)=\alpha_d(d_i)\theta_d(l_j)x_{sd}(d_i, l_j, t_k)+\alpha_n(d_i)\theta_n(l_j)x_{sn}(d_i, l_j, t_k)+\beta(d_i)\Phi(l_j)$$

where $\alpha_{n/d}(d_i)$ and $\beta(d_i)$ are parameters for disease d_i, and $\theta_{n/d}(l_j)$ and $\Phi(l_j)$ are parameters for region l_j. All parameters are learnable in training the *LR Spec. D. L.* model.

Usefulness of News-Concerned Intents for RD Incidence Estimation

In RDs incidence estimation with session input, news-concerned intents were used as input for the LR models. We aimed to analyze the usefulness of the weights considering news about different diseases (d_i) and regions (l_j) in *LR Spec. D. L.* (ie, $\alpha_{sn}(d_i)\theta_{sn}(l_j)$) by displaying their values and distribution. Moreover, to explore how news-concerned sessions affect RDs incidence estimation dynamically, we tried to find RDs with outbreak news in 2018 and 2019, and display their session numbers, true incidences, and predicted incidences during the study period. In this way, we could illustrate how the news-specific parameters helped reduce the influence of a surge in query volume caused by outbreak news. In the experiment, we selected two diseases: Disease 1 (multiple sclerosis [MS]) and Disease 5 (amyotrophic lateral sclerosis [ALS]). MS represents a class of diseases that has received relatively less attention but gradually attracted public attention, which had related queries on around May 30 every year (ie, International MS Day). ALS is a relatively well-known RD with thousands of RD-concerned and news-concerned sessions, which attracted massive attention when Stephen William Hawking, a world-famous physicist who had ALS, died on March 14, 2018.

Evaluation Settings

Evaluation for Session Intent Prediction (Step 1)

Supervised training was employed to train the session intent prediction model in Figure 2. For the ground truth, a subset S_{anno} was selected from the session data set S to annotate manually. One hundred sessions were randomly selected from each month in 2016 and 2017, forming an S_{anno} data set of size 2400. Three annotators then labeled the sessions with one of the three intents: RD-concerned, news-concerned, and others. The final intent was voted on by the three annotators. The κ value [41] of the annotations was 0.719, indicating substantial consistency of annotating. Among the 2400 annotated sessions, 502 were RD-concerned, 143 were news-concerned, and 1755 belonged to the others category. Thus, a considerable percentage of sessions were not RD-concerned, indicating that it is necessary to distinguish the session intents. The 2400 sessions were randomly divided into a training set, validation set, and test set at an 8:1:1 ratio.

For model implementation, Python 3.6.13 was used for modeling and evaluation. Pytorch 1.7.1 was used as the framework for training the models. Macro-F1, accuracy, and F1 scores for each intent were used for performance evaluation.

Evaluation of Multiple RDs Incidence Estimation (Step 2)

For comparison, we also constructed query data as the input for RDs incidence estimation. The query input comprised the numbers of name-related, gene-related, and treatment-related queries of different RDs, regions, and periods. The structures of LR variants for the query input are the same as the equations presented in the previous subsection.

We compared different input types and LR models on the data set from 2016 to 2019, where data in 2016 and 2017 constituted the training set, data in 2018 served as the validation set, and data in 2019 served as the test set. The root mean square error (RMSE) and relative error rate (RER) were utilized for performance evaluation to obtain both the absolute error and relative error of the models:

where $y_{pred}(d_i, l_j, t_k)$ is the predicted output of LR models.

All experiments were conducted in the Python 3.6.13 environment and all methods were implemented with the Pytorch 1.7.1 library. Models were trained with the Adam optimizer until convergence on the validation set with a maximum of 1000 epochs.

Results

Summary Statistics of RDs Incidence and RD-Related Search Data

In general, the RDs incidence data set included more than 80,000 incidences from 2016 to 2019 in China (due to data privacy concerns, the specific number of incidences is not reported). The RD-related search data set included 2,749,257 RD-related queries and 1,769,986 sessions from 1,380,186 users. It is worth noting that repeated search was not a serious problem in our data set. On average, each user had 1.282 sessions, most users (n=1,193,362, 86.46%) had only one session, and 97.75% (n=1,349,105) of users contributed less than four sessions. This is mainly due to two reasons. First, the sessions grouped RD-related search queries that were submitted by a user over a short period of time; therefore, repeated sessions were less common for RD patients in our data set. Second, we distilled RD-related sessions by specific keywords for RDs (ie, level 1 keywords), and the provided results might be sufficiently clear that there was no need to repeat the search. Therefore, we adopted the intent prediction and incidence estimation tasks at the session level rather than the user level.

Furthermore, we considered four regions in our data set, which divided 31 provinces in China's mainland into four parts: East, West, Central, and Northeast. The populations of the four regions were 535.6 million, 378.1 million, 369.9 million, and 108.5 million, with gross domestic products of 7109 billion dollar, 2752 billion dollar, 2899 billion dollar, and 797 billion dollar, respectively (average of 4 years). In the RDs incidence data set, the sum of the incidences of 15 RDs was the highest in the West, followed by the East, Central, and Northeast regions. The incidence of different RDs varied among the four regions. For instance, MS and hemophilia had the largest incidences in the West, whereas ALS was the most frequently registered disease in the East. In the RD-related search data set, the average session and query numbers of the 4 years were 225,906.5 and 1,023,152.0 for the East; 91,357.5 and 413,361.3 for the West; 94,151.8 and 429,708.5 for the Central region; and 31,080.8 and 141,278.0 for the Northeast, respectively.

Generally, the East had the largest population, the most developed economy, and, accordingly, the highest number of queries and sessions. Overall, the session volume was proportional to the population. However, regional reported RD incidences and population did not always match, since the incidence of an RD in a given region might relate to whether it is a family genetic disease in the region, the diagnosis technique of the disease in that region, and other factors. Therefore, we considered the effect of region variables on the RD incidence estimation specifically.

Performance of Session Intent Prediction

The first-step session intent prediction was evaluated with the human-annotated test set of 240 sessions. In the three-category classification task, the model had a macro-F1 value of 0.452 and an accuracy of 0.682 on the test set. The F1 scores for RD-concerned sessions, news-concerned sessions, and other sessions were 0.397, 0.353, and 0.606, respectively. Some representative sessions with different intents are shown in Multimedia Appendix 4. All of the sessions were correctly predicted with the intent prediction model.

Finally, the model was applied to predict the intents of all 1,769,986 sessions in Session Set *S*, resulting in 426,031 RD-concerned sessions, 115,016 news-concerned sessions, and 1,228,939 other sessions. The RD-concerned and news-concerned sessions were grouped by their RDs, regions, and periods to form the session inputs $x_{sd}(d_i, l_j, t_k)$ and $x_{sn}(d_i, l_j, t_k)$ for comparing and estimating RDs incidence.

Performance of Incidence Estimation

Overall Performance

The incidence estimation results of different input types and LR models are shown in Table 2. Each experiment was repeated five times with different random seeds, and the average result and 95% CIs are reported. The null hypothesis was that there was no difference between the estimation results using query and session as the input. A two-sided *t*-test was performed on the results with different input types on the same model, and the *P* values are also reported in Table 2.

Session input had significantly better performance than query input on all models and metrics, which indicated the usefulness of considering search session intents in the RDs incidence estimation task. Comparing different models, *LR Spec. D. L.* exhibited the best performance, with RER=0.365 on session input and RER=0.511 on query input. However, the 95% CI was large. The instability was mainly due to the relatively large number of parameters in *LR Spec. D. L.* Further detailed comparison between session input and query input are shown in Multimedia Appendix 5.

Table 2. Relative error rate (RER) and root mean square error (RMSE) of rare disease incidence prediction with different linear regression (LR) models and input types.

Model	RER		RMSE	
	Average value (95% CI)	*P* value	Average value (95% CI)	*P* value
General LR				
Query input	0.998 (0.997-0.999)	<.001	0.042 (0.042-0.042)	<.001
Session input	0.864 (0.848-0.879)		0.039 (0.03-0.039)	
LR Spec. D.[a]				
Query input	0.887 (0.872-0.903)	<.001	0.037 (0.037-0.038)	<.001
Session input	0.720 (0.676-0.764)		0.030 (0.028-0.032)	
LR Spec. D. L.[b]				
Query input	0.511 (0.377-0.645)	.01	0.023 (0.017-0.029)	.008
Session input	0.365 (0.341-0.388)		0.017 (0.016-0.017)	

[a]Spec. D.: specific disease.

[b]Spec. D. L.: specific disease and location.

Usefulness of News-Concerned Intents for RDs Incidence Estimation

The weights considering news about different diseases d_i and regions l_j in the *LR Spec. D. L.* model (ie, $\alpha_{sn}(d_i)\theta_{sn}(l_j)$) are shown in Figure 3. The weights of news-concerned sessions were primarily negative, which confirmed our hypothesis that the effect of news should be deducted from the disease estimation, consistent with the findings of Xu et al [22]. The two outliers were Diseases 1 and 6, which had very small but positive parameters. There were too few news-concerned sessions (a few dozen) for these two diseases, and therefore they had little impact on the results. Moreover, since the volumes of search sessions and incidence were distinct, the magnitude of parameters varied among RDs.

To explore how news-concerned sessions affect RDs incidence estimation dynamically, we display two cases of RDs for Disease 1 (MS) and Disease 5 (ALS) in Figure 4. News-concerned session numbers, RD-concerned session numbers, and the true and predicted incidence (normalized to the range of 0 to 1) of RDs for each season during 2018 and 2019 are shown. Figure 4 demonstrates that outbreak news could be predicted with the intent prediction model, and the predicted incidence was corrected from the high query volume when the news-concerned sessions were considered. For MS, two peaks in news-concerned session numbers arose in the second seasons of 2018 and 2019 around May 30, International MS Day. By contrast, since the MS incidence was certainly not affected by MS Day, considering news-concerned sessions would reduce noise in session numbers for incidence estimation. News-concerned ALS sessions showed a noticeable peak in the 1st season in 2018, after Stephen William Hawking died on March 14, 2018. After considering the number of news-concerned sessions, the result was less affected by the outbreak news.

Figure 3. Weights of news-concerned session numbers in estimating the rare diseases incidence with the linear regression specific disease and location (LR Spec. D. L.) model.

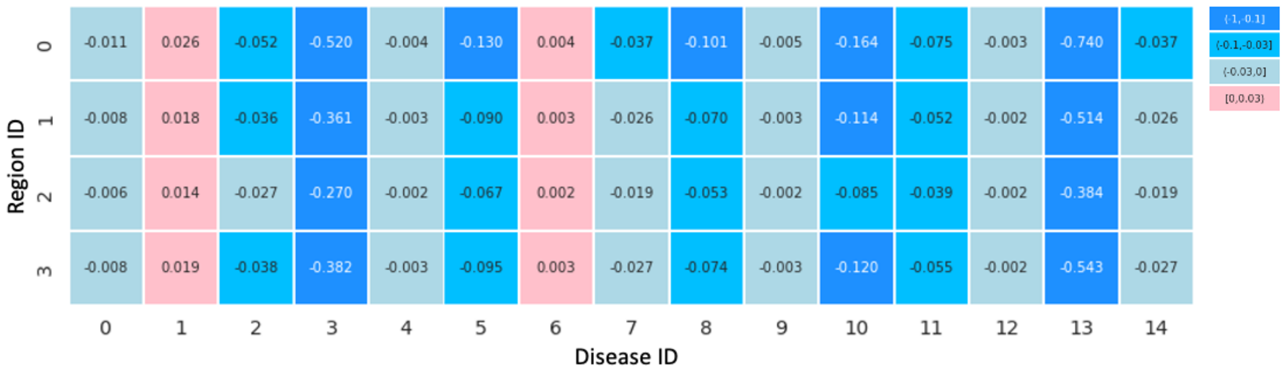
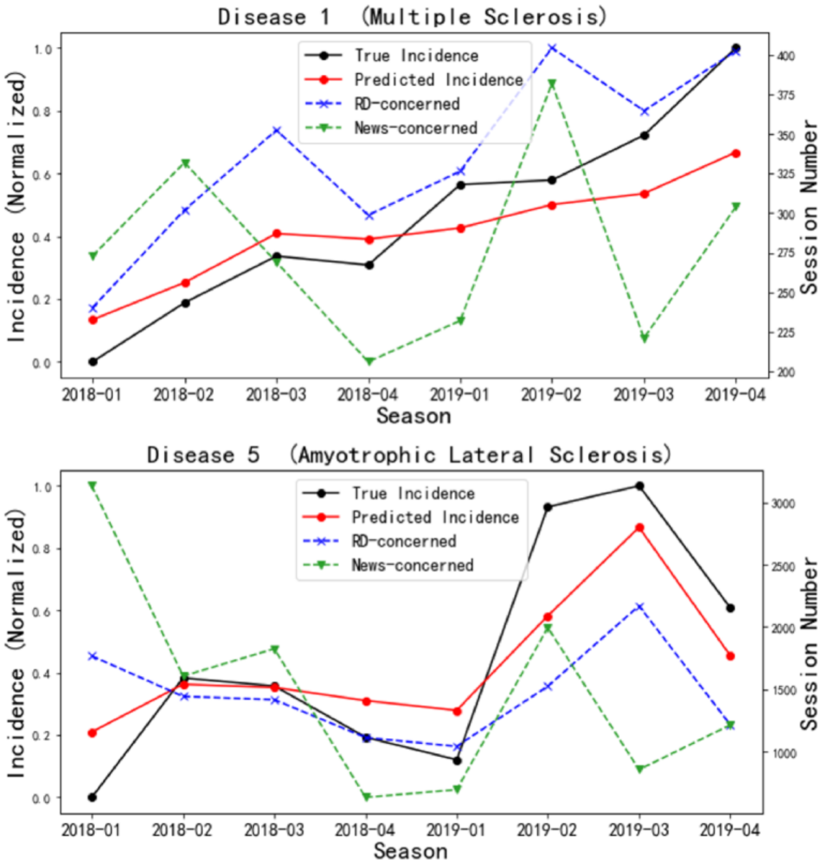


Figure 4. News-concerned session numbers, rare disease (RD)-concerned session numbers, and RDs true incidence and predicted incidence (normalized to the range of 0 to 1) of each season during 2018 and 2019 for Disease 1 (multiple sclerosis) and Disease 5 (amyotrophic lateral sclerosis).



Discussion

Principal Findings

The RD incidence estimation experiment on 15 RDs in 4 regions of China showed that RDs could be estimated with search engine logs, especially search session data. The RER of RDs incidence estimation was 0.365 for the session input and 0.511 for the query input. Considering the sparsity of RD cases, the RDs incidence estimation performance is encouraging.

The first step predicted session intents with a deep neural model. The prediction results indicated the necessity to distinguish the user intents in searching sessions. Among 1,769,986 RD-related sessions, only 426,031 (24.07%) were RD-concerned and 1,228,939 (69.43%) belonged to other intents. By identifying sessions concerned with RDs, irrelevant queries were effectively filtered from the data.

The second step, multiple RDs incidence estimation with LR, demonstrated that considering the volume of sessions rather than RD-related queries was significantly more helpful for disease estimation in most RDs and regions, as shown in Table 2 and Multimedia Appendix 5. Compared with queries, session intents helped estimate RDs incidence with an error decrease of 28.57% in terms of RER (*P*=.01). This illustrates the significant contribution of considering search sessions with more context for RD incidence detection. Moreover, as shown in Figure 3 and Figure 4, considering news-concerned session numbers in RDs incidence estimation was necessary and helpful.

When we considered the types of RDs (Multimedia Appendix 2), no significant differences were revealed between the similarity within each RD type and the similarity between different types. Adding RD type–specific parameters to the incidence estimation model also did not improve performance. Since the incidence and search query for RDs were both too sparse, their distributions might be less correlated with RD types. Moreover, RDs are often associated with genetics, and genetic variants vary among RDs of the same types, resulting in different distributions. The role of RD types is therefore considered to be relatively less important in RD-related infoveillance.

Comparison With Prior Work

To our knowledge, this study is the first to apply infoveillance in RDs incidence estimation, which provides a novel method to understand RDs. Compared with prior research on utilizing search engine data to estimate other diseases, a novel aspect of this study is that we considered the session context about disease-related queries and then utilized session intents to replace query volume for disease incidence estimation. Session inputs showed significant improvement on the RDs incidence estimation task. Although the sparsity of RD-related queries inspired the use of session information, the two-step method can be effectively transferred to other search engine–based disease detection and estimation tasks, as data noise pervasively exists online.

Limitations

This study has several limitations. First, the current data from the national multicenter clinical database of RDs were collected by retrospective reports. Due to the difficulty of RD diagnosis and the limited support of International Classification of Diseases 10th Revision codes for RDs, there might be delayed or unreported cases in the database. Therefore, the overestimations of incidence might reflect unreported cases, which was neglected in our analysis and discussions. In the future, it would be helpful to revisit patients in overestimated RDs and regions with privacy protection.

Second, 15 RDs with stable long-term data in the registry database were utilized for our experiments. These experiments could be applied to other RDs, whereas some RDs might not be estimated with our proposed methods, such as those with unclear symptoms, too low incidence, and low public awareness. Extending this method to more RDs and finding the boundary is promising future work.

Third, the level 1 keywords used for matching RD-related queries were provided by medical experts, which was time-consuming and might reflect knowledge bias. In the future, we will test automatic keyword discovery methods for RD-related keyword discovery.

Finally, a simple combination of LSTM and MLP was adopted for intent prediction in this study as the first attempt to integrate session intents in RDs incidence estimation. Since the numbers of RD-concerned and news-concerned sessions were much smaller than the numbers of sessions about other intents, the F1 scores of intent prediction about RD-concerned and news-concerned sessions were limited (0.397 and 0.353, respectively). Although challenging, accurate intent prediction is essential for capturing RD-concerned sessions precisely. Therefore, we aim to design neural predictors with more sophisticated network structures and more features about the sessions and queries to improve the session intent prediction accuracy, especially for RD-concerned and news-concerned sessions.

Conclusions

In this study, an experiment on multiple RDs in multiple regions showed that it is possible to estimate RDs incidence with online search engine data. The two-step estimation method illustrates promising performance improvement when session intents are considered in the RDs incidence estimation task. The use of session information can be transferred to infoveillance on other diseases.

This study did not aim to replace the clinical RD registry systems with search engine–based estimation. The two-step RDs incidence estimation model was designed as a supplement and prewarning method. For instance, if the model overestimates an RD in a region, this can remind experts of possible missing records from clinical registries or lack of medical support in the region. This method could help provide information for allocating medical resources and RD-related policy-making in the future. Moreover, with privacy protection, the method could offer advice to RD-concerned users of appropriate medical aids such as hospitals or institutes specialized in certain RDs. In conclusion, this study provides a promising method for understanding and locating RDs.

Acknowledgments

This work is supported by the Natural Science Foundation of China (grant number U21B2026), Tsinghua University Guoqiang Research Institute, and Tsinghua University-Peking Union Medical College Hospital Initiative Scientific Research Program (2019ZLH202).

Data Availability

The data sets generated and/or analyzed during the current study are not publicly available due to patients' privacy concerns, but are available from the corresponding author on reasonable request.

Authors' Contributions

All authors contributed thoughtful discussions of the work. JL conducted the models and experiments. ZH organized and analyzed the data. MZ, WM, and SZ guided the design of project. YJ and LZ provided the clinical rare disease incidences data and helped write the manuscript. YL and SM helped edit the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Rare disease (RD) names and types.
[DOCX File , 35 KB - infodemiology_v3i1e42721_app1.docx]

Multimedia Appendix 2
Influence of disease types on rare diseases incidence estimation.
[DOCX File , 238 KB - infodemiology_v3i1e42721_app2.docx]

Multimedia Appendix 3
Keyword lists.
[DOCX File , 63 KB - infodemiology_v3i1e42721_app3.docx]

Multimedia Appendix 4
Representative sessions with different intents.
[DOCX File , 35 KB - infodemiology_v3i1e42721_app4.docx]

Multimedia Appendix 5
Comparison between session input and query input for rare disease incidence estimation.
[DOCX File , 81 KB - infodemiology_v3i1e42721_app5.docx]

References

1. Griggs RC, Batshaw M, Dunkle M, Gopal-Srivastava R, Kaye E, Krischer J, Rare Diseases Clinical Research Network. Clinical research for rare disease: opportunities, challenges, and solutions. Mol Genet Metab 2009 Jan;96(1):20-26 [FREE Full text] [doi: 10.1016/j.ymgme.2008.10.003] [Medline: 19013090]
2. RARE Disease Facts. Global Genes. URL: https://globalgenes.org/rare-disease-facts/ [accessed 2022-08-29]
3. Dong D, Wang Y. Challenges of rare diseases in China. Lancet 2016 May 07;387(10031):1906. [doi: 10.1016/S0140-6736(16)30418-4] [Medline: 27203652]
4. Khoury MJ. Rare Disease Day 2022: The evolving impact of genomics and precision health. Centers for Disease Control and Prevention. 2022. URL: https://blogs.cdc.gov/genomics/2022/02/28/rare-disease-day-2022/ [accessed 2023-02-19]
5. Zhang S, Chen L, Zhang Z, Zhao Y. Orphan drug development in China: progress and challenges. Lancet 2019 Sep 28;394(10204):1127-1128. [doi: 10.1016/S0140-6736(19)32179-8] [Medline: 31571591]
6. Guo J, Liu P, Chen L, Lv H, Li J, Yu W, Administrative Group of National Rare Diseases Registry System of China. National Rare Diseases Registry System (NRDRS): China's first nation-wide rare diseases demographic analyses. Orphanet J Rare Dis 2021 Dec 18;16(1):515 [FREE Full text] [doi: 10.1186/s13023-021-02130-7] [Medline: 34922584]
7. Nsubuga P, White M, Thacker S, Anderson M, Blount S, Broome C, et al. Public health surveillance: a tool for targeting and monitoring interventions. In: Jamison DT, Breman JG, Measham AR, editors. The International Bank for Reconstruction and Development/The World Bank. New York: Oxford University Press; 2006.
8. Rubinstein YR, Groft SC, Bartek R, Brown K, Christensen RA, Collier E, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). Contemp Clin Trials 2010 Sep;31(5):394-404 [FREE Full text] [doi: 10.1016/j.cct.2010.06.007] [Medline: 20609392]
9. Ali S, Bryce J, Cools M, Korbonits M, Beun J, Taruscio D, et al. The current landscape of European registries for rare endocrine conditions. Eur J Endocrinol 2019 Jan 01;180(1):89-98 [FREE Full text] [doi: 10.1530/EJE-18-0861] [Medline: 30407922]
10. 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]
11. Barros JM, Duggan J, Rebholz-Schuhmann D. The application of internet-based sources for public health surveillance (infoveillance): systematic review. J Med Internet Res 2020 Mar 13;22(3):e13680 [FREE Full text] [doi: 10.2196/13680] [Medline: 32167477]
12. Generous N, Fairchild G, Deshpande A, Del Valle SY, Priedhorsky R. Global disease monitoring and forecasting with Wikipedia. PLoS Comput Biol 2014 Nov;10(11):e1003892 [FREE Full text] [doi: 10.1371/journal.pcbi.1003892] [Medline: 25392913]
13. Gozzi N, Tizzani M, Starnini M, Ciulla F, Paolotti D, Panisson A, et al. Collective response to media coverage of the COVID-19 pandemic on Reddit and Wikipedia: mixed-methods analysis. J Med Internet Res 2020 Oct 12;22(10):e21597 [FREE Full text] [doi: 10.2196/21597] [Medline: 32960775]
14. 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]
15. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009 Feb 19;457(7232):1012-1014. [doi: 10.1038/nature07634] [Medline: 19020500]
16. Mavragani A. Infodemiology and infoveillance: scoping review. J Med Internet Res 2020 Apr 28;22(4):e16206 [FREE Full text] [doi: 10.2196/16206] [Medline: 32310818]
17. Johnson H, Wagner M, Hogan W, Chapman W, Olszewski R, Dowling J, et al. Analysis of Web access logs for surveillance of influenza. Stud Health Technol Inform 2004;107(Pt 2):1202-1206. [Medline: 15361003]

18. Santillana M, Nguyen AT, Dredze M, Paul MJ, Nsoesie EO, Brownstein JS. Combining search, social media, and traditional data sources to improve influenza surveillance. PLoS Comput Biol 2015 Oct;11(10):e1004513 [FREE Full text] [doi: 10.1371/journal.pcbi.1004513] [Medline: 26513245]
19. Yang S, Ning S, Kou SC. Use internet search data to accurately track state level influenza epidemics. Sci Rep 2021 Feb 17;11(1):4023. [doi: 10.1038/s41598-021-83084-5] [Medline: 33597556]
20. Althouse BM, Ng YY, Cummings DAT. Prediction of dengue incidence using search query surveillance. PLoS Negl Trop Dis 2011 Aug 2;5(8):e1258 [FREE Full text] [doi: 10.1371/journal.pntd.0001258] [Medline: 21829744]
21. Malik MT, Gumel A, Thompson LH, Strome T, Mahmud SM. "Google flu trends" and emergency department triage data predicted the 2009 pandemic H1N1 waves in Manitoba. Can J Public Health 2011;102(4):294-297 [FREE Full text] [doi: 10.1007/BF03404053] [Medline: 21913587]
22. Xu D, Liu Y, Zhang M, Ma S, Cui A, Ru L. Predicting epidemic tendency through search behavior analysis. 2011 Presented at: Twenty-Second International Joint Conference on Artificial Intelligence; Catalonia, Spain; July 16-22, 2011. [doi: 10.5591/978-1-57735-516-8/IJCAI11-393]
23. Zou B, Lampos V, Cox I. Multi-task learning improves disease models from web search. 2018 Presented at: WWW '18: 2018 World Wide Web Conference; April 23-27, 2018; Lyon, France. [doi: 10.1145/3178876.3186050]
24. Zou B, Lampos V, Cox I. Transfer learning for unsupervised influenza-like illness models from online search data. 2019 Presented at: WWW '19: The Web Conference; May 13-17, 2019; San Fracisco, CA. [doi: 10.1145/3308558.3313477]
25. Cousins HC, Cousins CC, Harris A, Pasquale LR. Regional infoveillance of COVID-19 case rates: analysis of search-engine query patterns. J Med Internet Res 2020 Jul 30;22(7):e19483 [FREE Full text] [doi: 10.2196/19483] [Medline: 32692691]
26. Ye Z, Mao J, Liu Y, Zhang M, Ma S. Investigating COVID-19-related query logs of Chinese search engine users. Proc Assoc Inf Sci Technol 2020 Oct 22;57(1):e424 [FREE Full text] [doi: 10.1002/pra2.424] [Medline: 33173830]
27. Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, R Niakan Kalhori S. Predicting COVID-19 incidence through analysis of Google Trends data in Iran: data mining and deep learning pilot study. JMIR Public Health Surveill 2020 Apr 14;6(2):e18828 [FREE Full text] [doi: 10.2196/18828] [Medline: 32234709]
28. Ram S, Zhang W, Williams M, Pengetnze Y. Predicting asthma-related emergency department visits using big data. IEEE J Biomed Health Inform 2015 Jul;19(4):1216-1223. [doi: 10.1109/JBHI.2015.2404829] [Medline: 25706935]
29. Deiner MS, Lietman TM, McLeod SD, Chodosh J, Porco TC. Surveillance tools emerging from search engines and social media data for determining eye disease patterns. JAMA Ophthalmol 2016 Sep 01;134(9):1024-1030 [FREE Full text] [doi: 10.1001/jamaophthalmol.2016.2267] [Medline: 27416554]
30. Tkachenko N, Chotvijit S, Gupta N, Bradley E, Gilks C, Guo W, et al. Google Trends can improve surveillance of type 2 diabetes. Sci Rep 2017 Jul 10;7(1):4993. [doi: 10.1038/s41598-017-05091-9] [Medline: 28694479]
31. Cohen Zion M, Gescheit I, Levy N, Yom-Tov E. Identifying sleep disorders from search engine activity: combining user-generated data with a clinically validated questionnaire. J Med Internet Res 2022 Nov 23;24(11):e41288 [FREE Full text] [doi: 10.2196/41288] [Medline: 36416870]
32. Rochford B, Pendse S, Kumar N, De Choudhury M. Leveraging symptom search data to understand disparities in US mental health care: demographic analysis of search engine trace data. JMIR Ment Health 2023 Jan 30;10:e43253 [FREE Full text] [doi: 10.2196/43253] [Medline: 36716082]
33. White RW, Horvitz E. Cyberchondria. ACM Trans Inf Syst 2009 Nov 30;27(4):1-37. [doi: 10.1145/1629096.1629101]
34. Lampos V, Miller AC, Crossan S, Stefansen C. Advances in nowcasting influenza-like illness rates using search query logs. Sci Rep 2015 Aug 03;5:12760. [doi: 10.1038/srep12760] [Medline: 26234783]
35. Notice on the publication of the first catalogue of rare diseases. Chinese Government. URL: http://www.gov.cn/zhengce/zhengceku/2018-12/31/content_5435167.htm [accessed 2022-09-15]
36. Li M, He B, Guo R, Li Y, Chen Y, Fan Y. Study on population distribution pattern at the county level of China. Sustainability 2018 Oct 10;10(10):3598 [FREE Full text] [doi: 10.3390/su10103598]
37. Guidelines for the diagnosis and treatment of rare diseases (2019 edition). National Health Commission of the People's Republic of China. URL: http://www.nhc.gov.cn/yzygj/s7659/201902/61d06b4916c348e0810ce1fceb844333/files/e2113203d0bf45d181168d855426ca7c.pdf [accessed 2022-09-15]
38. Han S, Zhang Y, Ma Y, Tu C, Guo Z, Liu Z, et al. THUOCL: Tsinghua open Chinese lexicon. GitHub. URL: https://github.com/thunlp/THUOCL [accessed 2022-09-15]
39. Rose D, Levinson D. Understanding user goals in web search. 2004 Presented at: 13th international conference on World Wide Web; May 17-20, 2004; New York, NY. [doi: 10.1145/988672.988675]
40. Graves A. Long short-term memory. In: Supervised sequence labelling with recurrent neural networks. New York: Springer; 2012:37-45.
41. Viera A, Garrett J. Understanding interobserver agreement: the kappa statistic. Fam Med 2005 May;37(5):360-363 [FREE Full text] [Medline: 15883903]

Abbreviations

ALS: amyotrophic lateral sclerosis

LR: linear regression
LSTM: long short-term memory
MLP: multilayer perceptron
MS: multiple sclerosis
RD: rare disease
ReLU: rectified linear unit
RER: relative error rate
RMSE: root mean square error

Edited by T Mackey; submitted 15.09.22; peer-reviewed by S Shams, S Wakamiya, P Zhao; comments to author 11.01.23; revised version received 07.03.23; accepted 27.03.23; published 28.04.23.

Please cite as:
Li J, He Z, Zhang M, Ma W, Jin Y, Zhang L, Zhang S, Liu Y, Ma S
Estimating Rare Disease Incidences With Large-scale Internet Search Data: Development and Evaluation of a Two-step Machine Learning Method
JMIR Infodemiology 2023;3:e42721
URL: https://infodemiology.jmir.org/2023/1/e42721
doi:10.2196/42721
PMID:

Original Paper

Obesity-Related Discourse on Facebook and Instagram Throughout the COVID-19 Pandemic: Comparative Longitudinal Evaluation

Catherine Pollack[1,2], PhD; Diane Gilbert-Diamond[2,3,4], SCD; Tracy Onega[5], PhD; Soroush Vosoughi[6], PhD; A James O'Malley[1,7], PhD; Jennifer A Emond[1,3], PhD

[1]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, NH, United States
[2]Department of Epidemiology, Geisel School of Medicine at Dartmouth, Hanover, NH, United States
[3]Department of Pediatrics, Geisel School of Medicine at Dartmouth, Hanover, NH, United States
[4]Department of Medicine, Geisel School of Medicine at Dartmouth, Hanover, NH, United States
[5]Department of Population Health Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States
[6]Department of Computer Science, Dartmouth College, Hanover, NH, United States
[7]The Dartmouth Institute for Health Policy and Clinical Practice, Hanover, NH, United States

Corresponding Author:
Catherine Pollack, PhD
Department of Biomedical Data Science
Geisel School of Medicine at Dartmouth
1 Medical Center Drive
Lebanon, NH, 03756
United States
Phone: 1 5404973419
Email: Catherine.c.pollack.gr@dartmouth.edu

Abstract

Background: COVID-19 severity is amplified among individuals with obesity, which may have influenced mainstream media coverage of the disease by both improving understanding of the condition and increasing weight-related stigma.

Objective: We aimed to measure obesity-related conversations on Facebook and Instagram around key dates during the first year of the COVID-19 pandemic.

Methods: Public Facebook and Instagram posts were extracted for 29-day windows in 2020 around January 28 (the first US COVID-19 case), March 11 (when COVID-19 was declared a global pandemic), May 19 (when obesity and COVID-19 were linked in mainstream media), and October 2 (when former US president Trump contracted COVID-19 and obesity was mentioned most frequently in the mainstream media). Trends in daily posts and corresponding interactions were evaluated using interrupted time series. The 10 most frequent obesity-related topics on each platform were also examined.

Results: On Facebook, there was a temporary increase in 2020 in obesity-related posts and interactions on May 19 (posts +405, 95% CI 166 to 645; interactions +294,930, 95% CI 125,986 to 463,874) and October 2 (posts +639, 95% CI 359 to 883; interactions +182,814, 95% CI 160,524 to 205,105). On Instagram, there were temporary increases in 2020 only in interactions on May 19 (+226,017, 95% CI 107,323 to 344,708) and October 2 (+156,974, 95% CI 89,757 to 224,192). Similar trends were not observed in controls. Five of the most frequent topics overlapped (COVID-19, bariatric surgery, weight loss stories, pediatric obesity, and sleep); additional topics specific to each platform included diet fads, food groups, and clickbait.

Conclusions: Social media conversations surged in response to obesity-related public health news. Conversations contained both clinical and commercial content of possibly dubious accuracy. Our findings support the idea that major public health announcements may coincide with the spread of health-related content (truthful or otherwise) on social media.

(JMIR Infodemiology 2023;3:e40005) doi:10.2196/40005

KEYWORDS

obesity; Facebook; Instagram; COVID-19; social media; news; infodemiology; public health; online health information

Introduction

The SARS-CoV-2 virus and COVID-19 pandemic have fundamentally changed society. The first US case was reported on January 20, 2020, and the World Health Organization (WHO) declared COVID-19 a global pandemic on March 11. COVID-19 is an infectious respiratory disease associated with a range of symptoms, and severity may be amplified in individuals with chronic, preexisting conditions such as obesity. This link was first reported in the spring of 2020, and studies have estimated that obesity may increase the risk of hospitalization due to COVID-19 between 7% and 33% and death by 8% to 61% [1,2]. Patients with severe obesity are particularly susceptible; for example, studies have shown that these patients may have 7.36 higher odds of a need for invasive mechanical ventilation compared to normal-weight patients [3].

US mainstream media coverage of the association between COVID-19 severity and obesity peaked in October 2020, when then US president Trump contracted COVID-19 [4,5]. Given that the prevalence of obesity in the US now exceeds 40%, it is imperative to understand how discourse about the disease evolved both temporally and topically throughout the early stages of the pandemic [6]. Such knowledge can further elucidate how major events impact the public dialogue surrounding a chronic disease. On one hand, increased attention to obesity by the public may further understanding of obesity prevention and treatment; on the other hand, repeated negative portrayals of the disease, especially in the mainstream media, may amplify weight-related stigma [7].

Social media enables the documentation of heterogeneous opinions in near real time, making it an attractive avenue to assess shifts in opinion in response to major events. In particular, Facebook and Instagram are two popular social media platforms that were accessed by 70% and 59% of Americans on a daily basis in 2021, respectively [8]. The frequent use of these platforms by the public makes an evaluation of their content especially salient. While previous work on obesity discourse on these platforms during the pandemic has evaluated a small fraction of content, there has yet to be a comprehensive analysis of a large sample of public-facing content [9-11]. Reviewing a wide range of content on both Facebook and Instagram can further elucidate the interplay between mainstream and social media dialogue in the context of chronic diseases. Thus, the purpose of this study was to explore temporal and topical variations in obesity-related content on Facebook and Instagram throughout the first year of the pandemic to better contextualize the interplay between news media and public discourse as related to COVID-19 and obesity.

Methods

Overview

This study included temporal and topical analyses to characterize how obesity-related content evolved on Facebook and Instagram surrounding 4 major events related to COVID-19 in 2020. Two dates were selected given their relevance to the broader pandemic: January 20, the date of the first US case, and March 11, the date when the WHO declared COVID-19 a global pandemic. Two other dates were directly related to obesity: May 19, the approximate date that obesity and COVID-19 were linked, and October 2, the date when then US President Trump contracted COVID-19 and obesity was most discussed in the news media, according to data from Media Cloud, an open-source content aggregation tool [12]. Temporal analysis evaluated changes in the number of obesity-related posts and interactions on each platform around those dates. Topic modeling evaluated the 10 most frequent obesity-related themes on each platform, excluding content related to pet obesity.

Data Collection

Facebook and Instagram posts were collected from CrowdTangle, a public insights tool owned and operated by Facebook [13]. CrowdTangle's Facebook data encompassed public pages with over 25,000 likes or followers, public groups with over 95,000 members, US-based public groups with over 2000 members, and verified profiles (ie, user profiles that confirm the "authentic presence" of well-known public figures) [14]. CrowdTangle's Instagram data encompassed public profiles with over 50,000 followers and verified profiles [15]. All content in English between January 6, 2020, and October 16, 2020, that contained the words "obese" or "obesity" was extracted. Health ("headache" or "migraine") and nonhealth ("clarinet") control data were extracted for the same period. Keywords for controls were chosen based on their perceived independence from obesity and for posting frequency that was within a degree of magnitude of the obesity data. For all Facebook data, posts made on animal-specific pages were removed; this kind of information was not available for Instagram data. On Facebook, interactions were defined as any kind of post reaction (such as likes or wows), comments, and shares. On Instagram, interactions were defined as likes, comments, and shares. When available, data also included metadata for the page, group, or profile on which the post was made, such as the category of page or group (eg, new organization, hospital).

Temporal Modeling

Interrupted time series analysis was performed for each date using autoregressive integrated moving average (ARIMA) models. This method was chosen given its ability to control for highly cyclical and serially correlated data prior to each date and model complex postevent effects using one or a combination of transfer functions, including "pulse," "step," and "ramp" effects [16]. A pulse effect is characterized as an instantaneous increase on the day of the event followed by an immediate return to pre-event levels, a step effect is characterized as an instantaneous increase on the day of the event that is sustained after the event, and a ramp effect is a slope characterizing a differential rate of change in the outcome after the event [16]. All combinations of transfer functions were evaluated in separate models on the obesity data for a 29-day window around each date (ie, 14 days on either side of the event and the event itself). A 2-week postevent period was chosen to ensure that the impact of the event was captured. A shorter time window might not have captured the full extent of the event's effect, while a longer window might have increased the likelihood of including a confounding event that would have precluded the ability to establish an association between the event of interest and the change in behavior. The same length of time was chosen in the pre-event period for symmetry.

Each model was developed using the *auto.arima* function in the "forecast" package in R to identify *p*, *d*, and *q* parameters [17]. Here, *p* represents the number of autoregressive lags (ie, how many past values of the outcome are needed to predict the current value), *d* represents the degree of nonseasonal differences to reach stationarity (with "stationarity" defined as a mean and variance independent of time), and *q* represents the number of lagged errors required to predict the outcome (ie, the number of lags in the moving average component of the model). The transfer functions in the model with the lowest sample-corrected Akaike information criterion (AICc) were chosen for each date. Parameter selection for *p*, *d*, and *q* was repeated for control models using the best transfer function set from the obesity model. A sensitivity analysis was run on control models using the same *p* and *q* parameters as the obesity model—in all cases, AICc values were higher, so only results from the model with data-specific *p* and *q* values are presented (Multimedia Appendix 1 includes both).

Topic Modeling

Obesity-related posts were clustered into various topics using BERTopic, with a minimum topic size of 100 [18]. This minimum topic size was chosen to balance the size and similarity of the cluster. The BERTopic modeling process has demonstrated performance improvements over classical topic models, including latent Dirichlet allocation (LDA) and nonnegative matrix factorization when applied to both social media and public health data [19-23]. For example, a recent publication by de Groot and colleagues [23] showed that BERTopic generalizes well to short text domains (such as social media) and outperforms LDA in terms of coherence and diversity of topics. BERTopic first extracts document embeddings using bidirectional encoder representations from transformers (BERT), which generates numerical representations of textual data that preserve the context of the original text [24]. Embeddings then undergo dimensionality reduction before hierarchical clustering methods are applied to categorize them into topics. BERTopic was used independently on Facebook and Instagram data. Topic themes were assigned by a member of the research team with expertise in obesity medicine via qualitative examination of the top 3 exemplar points for each topic (ie, content located in the densest area of each cluster). Topics with exemplar posts that were focused on pet obesity were excluded. Temporal modeling, as described above, was conducted on the finalized set of top 10 topics.

For all analyses, a Bonferroni-adjusted critical value of *P*≤.003 was chosen by dividing the typical *P*≤.05 threshold by 16 (ie, 4 dates across 2 platforms with 2 types of content) to apply a conservative adjustment for multiple comparisons. Topic coherence was evaluated with c_v coherence, whereby values closer to 1 represent more intracluster similarity. Analyses were conducted in Python (version 3.8.8; Python Software Foundation) and R (version 4.1.12; R Foundation) using Visual Studio Code (version 1.63.2; Microsoft Corp) and RStudio (version 2021.09.0; Posit Software), respectively.

Ethical Considerations

Institutional review board approval was not required for this study given the public-facing nature of the social media data used [25].

Results

Aggregate Analysis

Facebook

Between January 6 and October 16, 2020, there were 175,242 posts across 66,497 public Facebook pages, groups, and pages in the CrowdTangle repository that contained the words "obesity" or "obese." There was no significant change in posting behavior in the 14 days after January 20 and March 11 compared to the 14 days prior (Table 1). There was a significant pulse ($_{p,2}$) effect on May 19 (ie, the approximate date when COVID-19 and obesity were linked), with a temporary increase of 405 posts (95% CI 166 to 645; *P*<.003). This was not observed in the health control data ($_{p,2}$=104, 95% CI –63.3 to 271; *P*=.224) or nonhealth control data ($_{p,2}$=87.4, 95% CI –23.7 to 198; *P*=.123). While the best model for this period also included a step parameter ($_{s,2}$) for a sustained effect in the 14 days after the event, this was not significant at the Bonferroni-adjusted threshold ($_{s,2}$=500, 95% CI 60.0 to 941; *P*=.026). The October 2 model included a pulse ($_{p,3}$) of 639 posts that was statistically significant (95% CI 359 to 883; *P*<.003). This effect was not observed in the health control data ($_{p,3}$=268, 95% CI 87.1 to 450; *P*=.004) or nonhealth control data ($_{p,3}$=196, 95% CI 27.4 to 364; *P*=.023) at the Bonferroni-adjusted threshold. The ramp parameter in the 14 days after the event in the obesity model was also not significant at the Bonferroni-adjusted threshold ($_{r,3}$=9.06, 95% CI 2.84 to 15.3; *P*=.004).

Table 1. Autoregressive integrated moving average models for Facebook posts per day. Values in italics denote statistical significance at the Bonferroni-adjusted threshold of $P≤.003$.

Category/date (parameters)	Estimate (95% CI)	*P* value[a]
Obesity		
January 20, 2020 (p=0, d=0, q=1, AICc[b]=341.24)		
Ramp ($_{r,0}$)[c]	–3.00 (–12.0 to 6.03)	.515
March 11, 2020 (p=5, d=2, q=0, AICc=367)		
Pulse ($_{p,1}$)[d]	–126 (–322 to 70.3)	.209
May 19, 2020 (*p=0, d=1, q=1, AICc=358.6*)		
Pulse ($_{p,2}$)	*405 (166 to 645)*	*<.003*
Step ($_{s,2}$)[e]	500 (60.0 to 941)	.026
October 2, 2020 (*p=3, d=0, q=0, AICc=382.94*)		
Pulse ($_{p,3}$)	*639 (395 to 883)*	*<.003*
Ramp ($_{r,3}$)	9.06 (2.84 to 15.3)	.004
Health control data		
January 20, 2020 (p=4, d=1, q=0, AICc=356.98)		
Ramp ($_{r,0}$)	–0.67 (–29.8 to 28.5)	.964
March 11, 2020 (p=0, d=1, q=0, AICc=360.43)		
Pulse ($_{p,1}$)[e]	–14.5 (–208 to 179)	.883
May 19, 2020 (*p=5, d=0, q=0, AICc=360.63*)		
Pulse ($_{p,2}$)	87.4 (–23.7 to 198)	.123
Step ($_{s,2}$)	*–71.8 (–94.8 to –48.8)*	*<.003*
October 2, 2020 (p=4, d=0, q=0, AICc=370.54)		
Pulse ($_{p,3}$)	196 (27.4 to 364)	.023
Ramp ($_{r,3}$)	–6.08 (–10.6 to –1.56)	.008
Nonhealth control data		
January 20, 2020 (p=0, d=0, q=0, AICc=287.6)		
Ramp ($_{r,0}$)	0.20 (–2.00 to 2.40)	.859
March 11, 2020 (p=1, d=0, q=0, AICc=295.2)		
Pulse ($_{p,1}$)	–29.2 (–93.6 to 35.2)	.374
May 19, 2020 (p=0, d=0, q=1, AICc=299.81)		
Pulse ($_{p,2}$)	2.88 (–58.3 to 64.1)	.927
Step ($_{s,2}$)	–17.0 (–54.0 to 20.0)	.368
October 2, 2020 (p=0, d=0, q=0, AICc=298.64)		
Pulse ($_{p,3}$)	–25.8 (–96.6 to 45.1)	.476

Category/date (parameters)	Estimate (95% CI)	*P* value[a]
Ramp ($_{r,3}$)	–0.95 (–3.51 to 1.60)	.464

[a]*P* values based on an autoregressive integrated moving average model for Facebook posts per day in the specified data set with a Bonferroni-adjusted significance threshold of *P*=.001.

[b]AICc: sample-corrected Akaike information criterion.

[c]Ramp (ω_r) functions are 0 before the intervention and (t–T+1) after the intervention (where *t* represents the current day and *T* represents the intervention date).

[d]Pulse (ω_p) functions are 1 if it is the day of the intervention and 0 all other days.

[e]Step (ω_s) functions are 0 before the intervention and 1 the day of and after the intervention.

Changes in interactions on obesity posts for the 14 days following January 20 and March 11 were not significant (Table 2). On May 19, there were significant pulse ($_{p,2}$=294,930, 95% CI 125,986 to 463,874; *P*<.003) and step ($_{s,2}$=473,247, 95% CI 235,680 to 711,814; *P*<.003) increases in interactions not significant in either control. The included ramp effect ($_{r,2}$) in the obesity model was also not significant at the Bonferroni-adjusted threshold ($_{r,2}$=–38,596, 95% CI –64,268 to –12.923; *P*=.003). The October 2 pulse effect ($_{p,3}$) was significant in both the obesity model and health control data, although there were approximately 5000 more interactions on average in the obesity data set ($_{p,3}$=182,814, 95% CI 160,524 to 205,105; *P*<.003) relative to the health control data ($_{p,3}$=177,855, 95% CI 96,952 to 258,758; *P*<.003). The best model for this date also included a step parameter ($_{s,3}$), but it was not significant at the Bonferroni-adjusted threshold ($_{p,3}$=5791, 95% CI 1449 to 10,134; *P*=.009).

Table 2. Autoregressive integrated moving average models for Facebook interactions. Values in italics denote statistical significance at the Bonferroni-adjusted threshold of *P*=.003.

Category/date (parameters)	Estimate (95% CI)	*P* value[a]
Obesity		
January 20, 2020 (p=2, d=1, q=0, AICc[b]=691.73)		
Step ($\omega_{s,0}$)[c]	–61,169 (–139,294 to 16,957)	.125
March 11, 2020 (p=3, d=1, q=1, AICc=681.75)		
Pulse ($\omega_{p,1}$)[d]	–3298 (–60,578 to 53,981)	.910
May 19, 2020 (*p=2, d=0, q=0, AICc=762.64*)		
Pulse ($\omega_{p,2}$)	*294,930 (125,986 to 463,874)*	*<.003*
Step ($\omega_{s,2}$)	*473,247 (235,680 to 711,814)*	*<.003*
Ramp ($\omega_{r,2}$)[e]	–38,596 (–64,268 to –12,923)	.003
October 2, 2020 (*p=3, d=0, q=0, AICc=661.31*)		
Pulse ($\omega_{p,3}$)	*182,814 (160,524 to 205,105)*	*<.003*
Step ($\omega_{s,3}$)	5791 (1449 to 10,134)	.009
Health control data		
January 20, 2020 (p=2, d=1, q=0, AICc=667.46)		
Step ($\omega_{s,0}$)	37,827 (–9152 to 84,807)	.115
March 11, 2020 (*p=0, d=1, q=1, AICc=686.64*)		
Pulse ($\omega_{p,1}$)	*176,502 (93,695 to 259,308)*	*<.003*
May 19, 2020 (p=0, d=0, q=0, AICc=699.9)		
Pulse ($\omega_{p,2}$)	–10,492 (–86,771 to 65,786)	.787
Step ($\omega_{s,2}$)	1826 (–43,247 to 46,890)	.937
Ramp ($\omega_{r,2}$)	–1139 (–5542 to 3265)	.612
October 2, 2020 (*p=0, d=0, q=0, AICc=706.39*)		
Pulse ($\omega_{p,3}$)	*177,855 (96,952 to 258,758)*	*<.003*
Step ($\omega_{s,3}$)	–49.6 (–29,591 to 29,492)	.997
Nonhealth control data		
January 20, 2020 (p=2, d=1, q=0, AICc=479.66)		
Step ($\omega_{s,0}$)	–107 (–1835 to 1621)	.903
March 11, 2020 (p=0, d=0, q=0, AICc=504.12)		
Pulse ($\omega_{p,1}$)	–2189 (–4738 to 359)	.092
May 19, 2020 (p=1, d=2, q=0, AICc=491.06)		
Pulse ($\omega_{p,2}$)	–331 (–3632 to 2969)	.844
Step ($\omega_{s,2}$)	377 (–5442 to 6196)	.899
Ramp ($\omega_{r,2}$)	–589 (–4759 to 3582)	.782
October 2, 2020 (p=0, d=0, q=0, AICc=484.48)		
Pulse ($\omega_{p,3}$)	–138 (–1901 to 1626)	.878
Step ($\omega_{s,3}$)	66.1 (–578 to 710)	.840

[a]*P* values based on an autoregressive integrated moving average model for Facebook interactions per day in the specified data set with a Bonferroni-adjusted significance threshold of *P*=.001.

[b]AICc: sample-corrected Akaike information criterion.

[c]Step (ω_s) functions are 0 before the intervention and 1 the day of and after the intervention.

[d]Pulse (ω_p) functions are 1 if it is the day of the intervention and 0 all other days.

[e]Ramp (ω_r) functions are 0 before the intervention and (t–T+1) after the intervention (where *t* represents the current day and *T* represents the intervention date).

Instagram

Between January 6 and October 16, 2020, there were 18,129 posts across 3170 unique usernames in the CrowdTangle repository containing "obese" or "obesity." Of the 4 dates, only a ramp effect after January 20 ($_{r,0}$) was significant ($_{r,0}$=–1.04, 95% CI –1.33 to –0.76; *P*<.003). There was a pulse effect ($_{p,2}$) on May 19 of approximately 61 ($_{p,2}$=61.1, 95% CI 18.5 to 104; *P*=.005) additional posts, although this was not significant at the Bonferroni-adjusted threshold. For both dates, no significant effect was observed in the health control or nonhealth control data (Multimedia Appendix 1). For interactions, there was a pulse effect ($_{p,2}$) on May 19 of an estimated 226,017 (95% CI 107,323 to 344,708; *P*<.003) additional interactions relative to the surrounding window. This was not observed in either the health control ($_{p,2}$=–13,005, 95% CI –218,774 to 192,764; *P*=.880) or nonhealth control ($_{p,2}$=2161, 95% CI –35,026 to 39,349; *P*=.909) data. Similarly, there was a pulse ($_{p,3}$) of 156,974 (95% CI 89,757 to 224,192; *P*<.003) additional interactions on obesity content on October 2 not observed at the Bonferroni-adjusted threshold in the health control ($_{p,3}$=–14,864; 95% CI –246,793 to 217,063; *P*=.900) or nonhealth control ($_{p,3}$=26,307; 95% CI 7774 to 44,840 *P*=.005) data.

Topic Analysis

Facebook—General Description

Of 175,242 obesity-related posts, 87,470 (49.9%) could not be classified into a topic; a random sample of these can be found in Multimedia Appendix 2. The remaining posts were clustered into 245 different topics with a c_v topic coherency of 0.76. Of the initial most frequent topics (Multimedia Appendix 2), 2 were removed because they were related to pet obesity, and 1 was removed because it consisted of 1 redundant post on a male supplement not directly related to obesity. The remaining 10 largest topics contained 19,485 posts, representing 22.2% (19,485/87,772) of classifiable posts and 11.1% (19,485/175,242) of all obesity-related posts. Themes included COVID-19 (n=3849), childhood obesity (n=2443), sugary drinks (n=2425), bariatric surgery (n=2413), weight loss stories (n=2090), "clickbait" (ie, catchy content designed to increase clicks; n=1494), cancer (n=1355), sleep (n=1166), yoga (n=1130), and heart disease (n=1120). Posts related to weight loss stories had the highest median interactions (20, IQR 3-94), while posts related to yoga had the fewest (2, IQR 0-9). Of posts from pages with a labeled category, the most frequent categories included general health sites (sugary drinks, weight loss stories, and sleep themes), hospitals (bariatric surgery, cancer, and heart disease themes), media and news companies (COVID-19 and clickbait themes), nonprofit organizations (childhood obesity theme), and yoga and Pilates (yoga theme).

Facebook—Temporal Modeling

The distribution of most frequent topics changed around each date (Figure 1). In the 29 days surrounding and including January 20, the least popular topics were yoga and COVID-19, while the most were childhood obesity and bariatric surgery. These topics remained the most popular around March 11, while the least popular were yoga and heart disease. The distribution changed surrounding May 19, whereby COVID-19 and clickbait were the most popular topics while sleep and heart disease were the least popular. This again changed around October 2, when COVID-19 and heart disease were the most popular topics and sleep and yoga were the least popular.

Each topic also had distinct daily posting behavior (Figure 2, Multimedia Appendix 3). While there was no significant difference for any topic around January 20, 5 topics showed a change in posting behavior around March 11. The COVID-19 topic experienced an average daily increase of approximately one additional post per day ($_{r,1}$=0.69, 95% CI 0.42 to 0.69; *P*<.003). A significant, gradual decline was observed for sugary drinks ($_{,1}$=–0.48, 95% CI –0.78 to –0.19; *P*<.003) and weight loss stories ($_{r,1}$=–0.32, 95% CI –0.52 to –0.12; *P*<.003), while an immediate, sustained decline was observed for childhood obesity ($_{s,1}$=–10.2, 95% CI –12.7 to –7.66; *P*<.003) and clickbait ($_{s,1}$=–3.70, 95% CI –5.49 to –1.92; *P*<.003). In contrast, clickbait experienced an immediate pulse of content ($_{p,2}$=25.2, 95% CI 14.1 to 36.4; *P*<.003) on May 19, coupled with a sustained increase ($_{s,1}$=22.1, 95% CI 14.8 to 29.4; *P*<.003) and gradual decrease ($_{r,2}$=–2.58, 95% CI –3.48 to –1.69; *P*<.003). The cancer topic also experienced a pulse increase ($_{p,2}$=6.34, 95% CI 2.80 to 9.88; *P*<.003), while weight loss stories experienced a step decrease ($_{s,2}$=–2.85, 95% CI –4.52 to –1.19; *P*<.003). No topics met the Bonferroni-adjusted statistical significance threshold for October 2, although clickbait experienced a sustained increase of approximately two additional posts per day ($_{p,3}$=1.95, 95% CI 0.17 to 3.73; *P*=.032).

Figure 1. Distribution of the top 10 most frequent topics for Facebook around each date of interest.

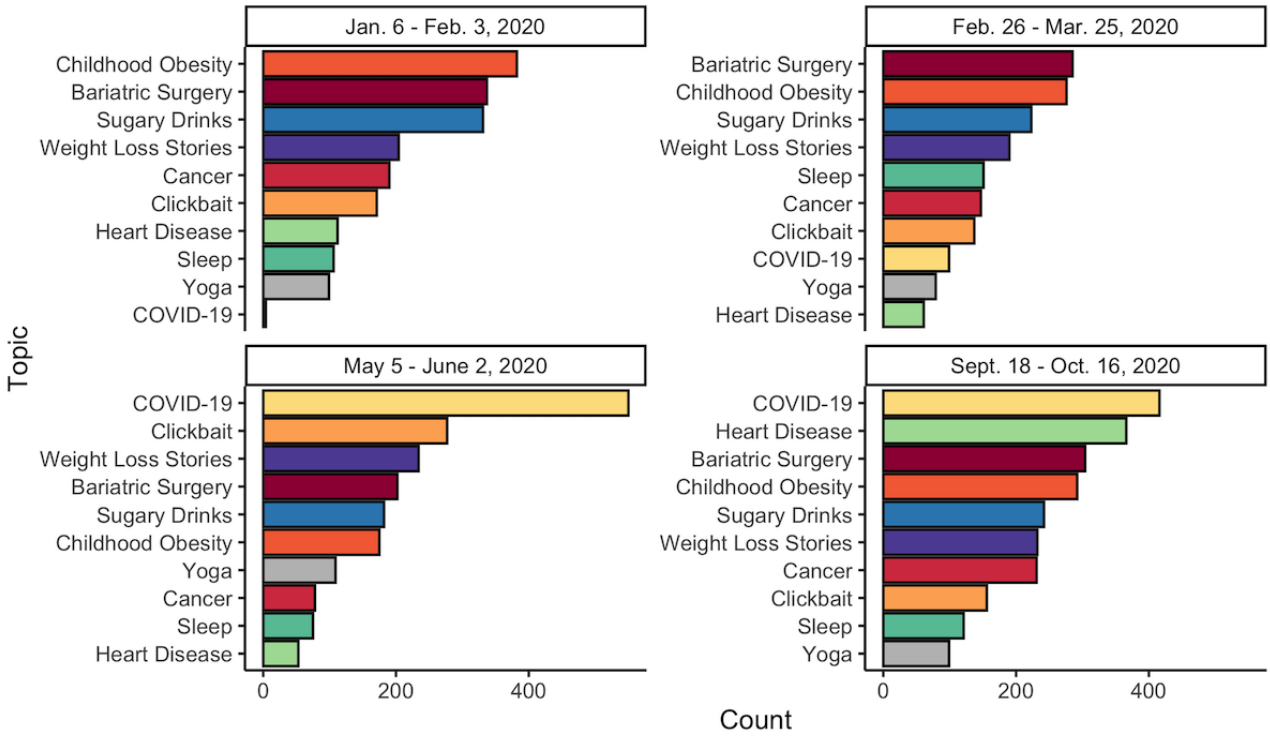
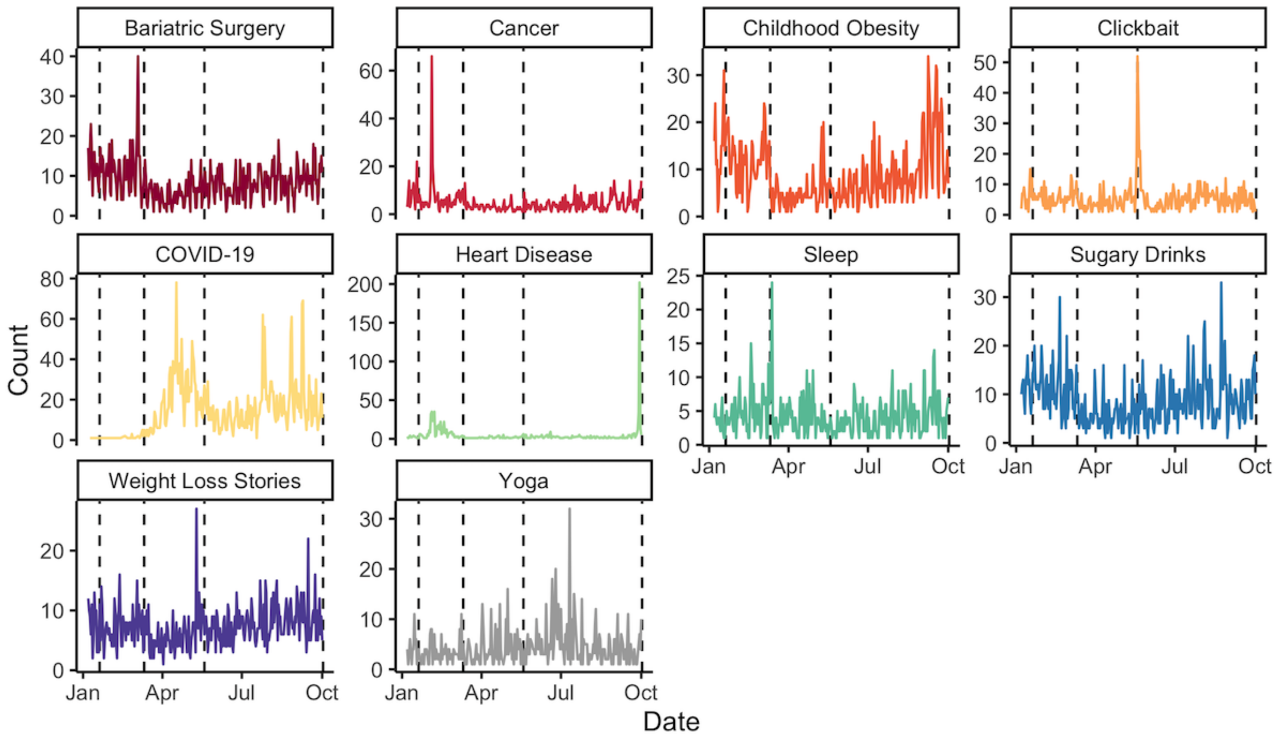


Figure 2. Longitudinal variations in top 10 most frequent topics about obesity on Facebook. The dashed lines indicate the 4 key dates of interest (January 20, March 11, May 19, and October 2, 2020).



Instagram—General Description

Of the 18,129 obesity-related Instagram posts, 6856 (37.8%) could not be classified into a topic; a random sample of these can be found in Multimedia Appendix 4. The remaining posts were clustered into 28 different topics with a c_v topic coherency of 0.57. Of the initial 10 largest topics, 1 (with n=769 posts) was removed for its pet-specific content. The remaining 10 largest topics represented 60.9% (6865/11,273) of classifiable posts and 37.9% (6865/18,129) of all posts (Multimedia Appendix 4). Some themes overlapped with Facebook, including weight loss stories (n=2718 posts), COVID-19 (n=1069 posts), bariatric surgery (n=363 posts), childhood obesity (n=331 posts), and sleep (n=312 posts). Additional topics included keto diet (n=588 posts), specific weight loss programs (n=415 posts), calories (n=391 posts), sugar (n=341 posts), and responses to a UK government policy (n=337 posts). Weight loss stories had the highest median overall interactions (544, IQR 154-1733), while the bariatric surgery topic had the fewest (51, IQR 17-114).

Instagram—Temporal Modeling

The ranking of topic frequency was consistent around each date (Figure 3). Weight loss stories were the most frequent in each of the 4 windows, and COVID-19 was the second most frequent in 3 of the 4. The only exception was the first window, in which COVID-19 was the least frequent and the keto diet was the second most frequent. The least frequent topic varied within the other 3 windows—the weight loss program was the least frequent around March 11, sleep was the least frequent around May 19, and responses to the UK government policy were the least frequent around October 2.

Two Instagram topics changed significantly surrounding the 4 dates (Figure 4, Multimedia Appendix 5). On January 20, there was a significant pulse increase (${}_{p,0}$) in keto posts (${}_{p,0}$=4.88, 95% CI 1.84 to 7.93; *P*<.003), and on October 2, there was a sustained increase (${}_{s,3}$) in posts about calories (${}_{s,3}$=1.36, 95% CI 0.51 to 2.21; *P*<.003). Other topics also showed a change in posting behavior that did not reach the Bonferroni-adjusted significance threshold. Posts on the childhood obesity topic experienced a sustained decrease of 1.17 posts (95% CI –2.22 to –0.11; *P*=.030). This also occurred on March 11, with an immediate, sustained decrease of 1.59 posts (95% CI –3.05 to –0.12; *P*=.034). Weight loss stories experienced a gradual ramp (${}_{r,1}$) decrease of 0.61 posts (95% CI –1.02 to –0.21; *P*=.003). In the 14 days following May 19, topics related to weight loss stories (${}_{r,2}$=–0.23, 95% CI –0.44 to –0.03; *P*=.026), COVID-19 (${}_{r,2}$=–0.20, 95% CI –0.38 to –0.02; *P*=.028), and calories (${}_{r,2}$=–0.14, 95% CI –0.25 to –0.02; *P*=.022) experienced a gradual ramp decline. Around October 2, topics related to sugar (${}_{p,3}$=1.50, 95% CI 0.34 to 2.66; *P*=.011) and childhood obesity (${}_{p,3}$=1.83, 95% CI 0.18 to 3.49; *P*=.030) experienced a pulse increase, weight loss stories experienced a step increase (${}_{s,3}$=2.43, 95% CI 0.55 to 4.30; *P*=.011), and COVID-19 topics experienced a ramp increase (${}_{r,3}$=0.18, 95% CI 0.04 to 0.32; *P*=.010).

Figure 3. Distribution of the top 10 most frequent topics for Instagram around each date of interest.

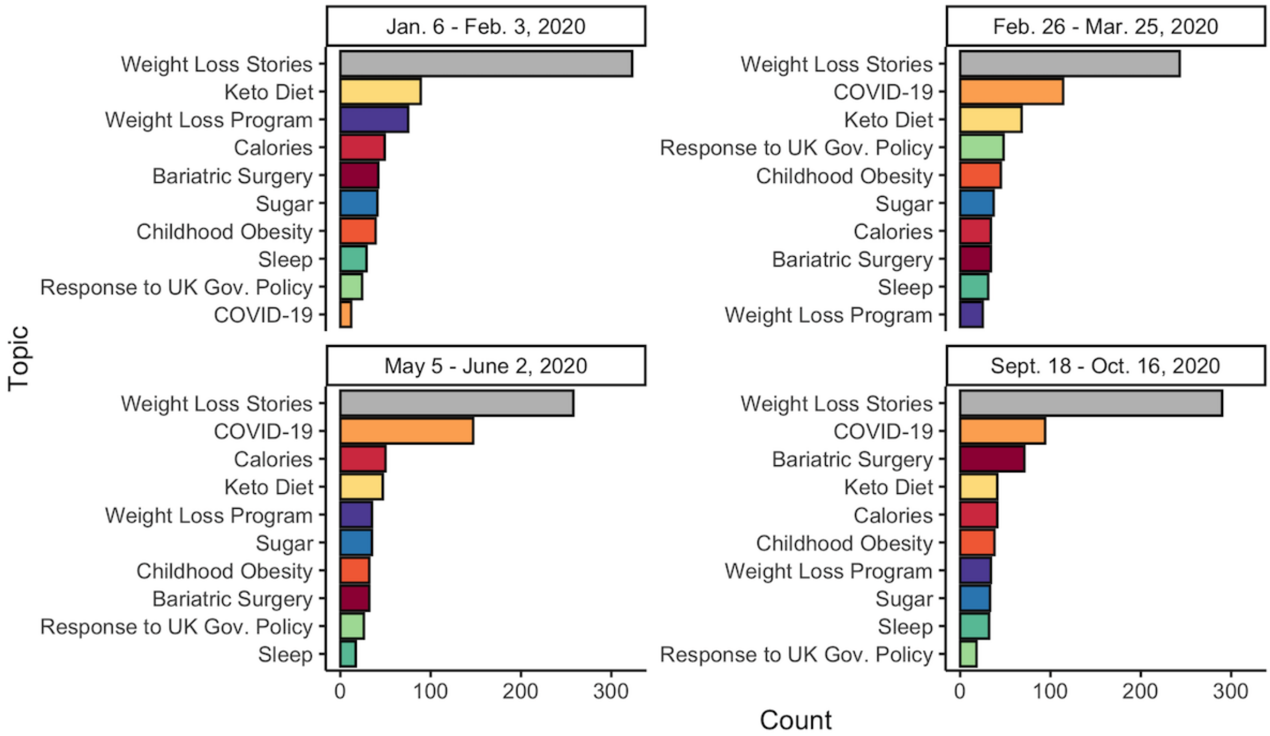
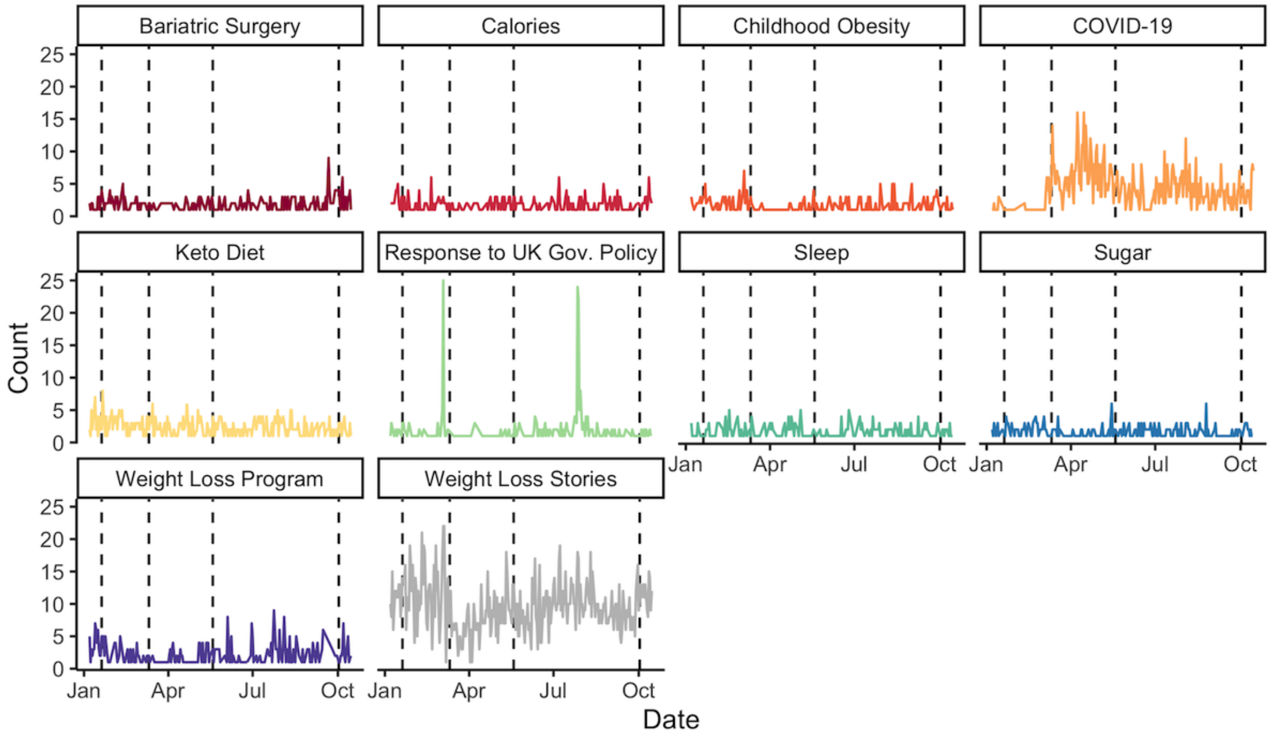


Figure 4. Longitudinal variations in top 10 most frequent topics about obesity on Instagram. The dashed lines indicate the 4 key dates of interest (January 20, March 11, May 19, and October 2, 2020).



Discussion

This study is the first to comprehensively evaluate obesity-related content throughout the pandemic on Facebook and Instagram. On Facebook, obesity-related content surged around the dates of 4 key news stories related to obesity or COVID-19. Posting behavior of obesity-related content on Instagram was not affected, although changes in interactions occurred. Frequent content on each platform had some overlapping themes (ie, COVID-19, bariatric surgery, childhood obesity, weight loss stories, and sleep), while other topics varied in popularity. These findings demonstrate how social media conversations regarding prevalent health conditions may be influenced by news media and global events.

On Facebook, there were immediate changes in posting and interaction behavior for obesity-related content on both May 19 and October 2 that were not sustained in the following 14 days. A similar effect was observed for interactions on Instagram. The lack of statistical significance in the control data for any of the May 19 outcomes provides evidence that change in online discussion about obesity was specific to the reported association between COVID-19 and obesity that was shared in the mainstream media on that day. For October 2, a significant pulse effect was observed for interactions on Facebook posts in the health control data, while the nonhealth control data remained insignificant. Since the keyword for the health control data (ie, “headache”) is also a symptom of COVID-19, this may suggest that the surge in interactions occurred on posts that discussed the same topic as the obesity data (ie, the prognosis of then US president Trump, who had contracted COVID-19).

When broken down by topic, 5 frequent topics overlapped. Three (ie, bariatric surgery, pediatric obesity, and sleep) were clinical in nature and received the fewest interactions from users, suggesting that social media may not be an ideal platform to communicate this kind of content. In contrast, weight loss stories were present on both platforms and received a high number of interactions. This consistency may suggest that individuals feel comfortable sharing personal stories on these platforms as a show of support for others, and frequent mentions of obesity online or in mainstream media may empower individuals to communicate their own experiences with obesity or weight loss [26]. However, anecdotal stories may spread commonly held falsehoods about weight loss or give viewers unrealistic expectations. This is especially problematic on Instagram, which has faced scrutiny over how it may detrimentally impact body image among its adolescent female user base [26,27].

Prior work has pointed to a limited amount of healthy dietary advice on Instagram among posts with the hashtags #weightloss or #quarantine15 [10,11]. The present work adds to that literature, as 3 of the top 10 most frequent themes on Instagram were focused on some kind of diet or food group (ie, keto diet, calories, and sugar). Exemplar posts for each of these categories often included compelling language that promised a lifechanging transformation (in the case of keto) or warned of imminent dangers if immediate changes were not made (in the case of calories and keto). This kind of catchy language was also dominant in the “clickbait” category on Facebook, which consisted primarily of short phrases that encouraged readers to click on either a linked article or shared post. Although the words “obesity” or “obese” were not present in the analyzed text itself, the fact that these posts were included in the data suggests that other information in the posts (such as the image text or link descriptions) included the keywords. Future work is needed to perform an in-depth analysis on this topic to understand the types of links shared and how frequently individuals interacted with them.

The practical implications and importance of these finding are 3-fold. Broadly, the ability to isolate the impact that media mentions of public health topics have on social media discussion contributes to the growing body of literature that demonstrates how social media can help gauge public opinions during times of crisis [27,28]. This study demonstrates that by comparing a public health topic of interest to multiple controls one can obtain quantitative estimates of the effect that major announcements or stories about the disease have on dialogue. While the present study focused on 2 obesity-related events covered by many media outlets, future work could identify stories covered by only a few outlets to try to estimate precise effects of those channels. Additionally, the swift and substantial response across social media platforms to obesity- and COVID-19–related stories in the media emphasizes the need for both researchers and media outlets to consider how premature public health announcements may contribute to the spread of online misinformation. Future work should study whether this strong relationship is present across other health topics, time periods, and platforms. Finally, this study adds to the growing body of literature that demonstrates the utility and power of BERTopic in analyzing both public health and social media data [19-23].

Strengths of this study include its expansion on prior work to understand obesity discourse more broadly, inclusion of multiple social media platforms, and evaluation of both temporal and topical patterns. However, there are several limitations that are important to note. First, there were no demographic data for users who created and viewed the content, which is a common challenge of epidemiologic research on social media. This study attempted to address this in part by using multiple platforms, which broadened the possible generalizability of the study. For instance, while approximately 71% of US adults aged between 18 and 29 years report ever using each platform, only 13% of US adults over the age of 65 years report using Instagram, compared to 50% for Facebook. Differences exist in other demographic groups as well, including race, income, and education [8]. Because each group may vary in how they perceive and discuss obesity (as well as in their underlying risk factors for the disease), future multi-platform studies are of the utmost importance to characterize perceptions of multiple groups. Second, while some content related exclusively to pet obesity was removed during topic analysis, future work could refine this process to ensure that all animal obesity content was removed. Third, only English-language content was evaluated; future work could be expanded to examine content in other languages. Fourth, only a finite number of topics was evaluated; future work could attempt to conduct a more holistic analysis, including exploration of outlier posts that could not be classified or adjustment of the minimum topic size in the BERTopic algorithm. Future work could also explore topics outside of the top 10, as these only represented about 22.2% (19,485/87,772) of classifiable posts about obesity on Facebook. Fifth, the use of a Bonferroni-adjusted threshold resulted in a conservative evaluation of the results, which may have biased findings toward the null (ie, fewer associations were made between the dates of interest and posting or interaction behavior). Finally, this study only looked at 4 dates of interest; future work could evaluate additional dates that occurred either before the pandemic or after 2020.

Overall, these findings suggest that the pandemic had distinct impacts on the frequency of and attention to obesity-related conversations on 2 popular social media platforms. Posts about obesity and corresponding interactions did not shift after two COVID-19–specific dates (ie, January 28 and March 11), suggesting that events of public health significance that do not relate to obesity do not dramatically alter conversations about the disease on Facebook and Instagram. In contrast, posts and interactions about obesity increased after 2 dates of importance to both COVID-19 and obesity (ie, May 19 and October 2). This pattern was not observed in health and nonhealth control data for the same time period, demonstrating how the relationship between COVID-19 and obesity amplified discussions about obesity. Clinical topics were similar between the platforms, as were weight loss stories. Dietary topics were more prevalent on Instagram, while “clickbait” was more prevalent on Facebook. Taken together, these results suggest that the impact of major public health events (including mainstream media attention and government campaigns) on social media discourse can be successfully isolated and monitored. Public health officials should consider leveraging social media campaigns to prevent the spread of misleading, deleterious content, such as misinformation that may spike around such events.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Models with matched p and q parameters.
[DOCX File , 24 KB - infodemiology_v3i1e40005_app1.docx]

Multimedia Appendix 2
Representative documents from Facebook topics.
[DOCX File , 100 KB - infodemiology_v3i1e40005_app2.docx]

Multimedia Appendix 3

Model selection and equations for Facebook topics.
[DOCX File , 16 KB - infodemiology_v3i1e40005_app3.docx]

Multimedia Appendix 4
Representative documents from Instagram topics.
[DOCX File , 98 KB - infodemiology_v3i1e40005_app4.docx]

Multimedia Appendix 5
Model selection and equations for Instagram topics.
[DOCX File , 16 KB - infodemiology_v3i1e40005_app5.docx]

References

1. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, the Northwell COVID-19 Research Consortium, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. JAMA 2020 May 26;323(20):2052-2059 [FREE Full text] [doi: 10.1001/jama.2020.6775] [Medline: 32320003]
2. Kompaniyets L, Goodman AB, Belay B, Freedman DS, Sucosky MS, Lange SJ, et al. Body mass index and risk for COVID-19-related hospitalization, intensive care unit admission, invasive mechanical ventilation, and death - United States, March-December 2020. MMWR Morb Mortal Wkly Rep 2021 Mar 12;70(10):355-361 [FREE Full text] [doi: 10.15585/mmwr.mm7010e4] [Medline: 33705371]
3. Simonnet A, Chetboun M, Poissy J, Raverdy V, Noulette J, Duhamel A, The LICORN and the Lille COVID-19 and Obesity study group. High prevalence of obesity in Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) requiring invasive mechanical ventilation. Obesity (Silver Spring) 2020 Jul;28(7):1195-1199 [FREE Full text] [doi: 10.1002/oby.22831] [Medline: 32271993]
4. Keith T, Rascoe A, Katkov M, Wise A, Ordoñez F. Trump Takes 'Precautionary' Treatment After He And First Lady Test Positive For Virus. National Public Radio. URL: https://www.npr.org/sections/latest-updates-trump-covid-19-results/2020/10/02/919385151/president-trump-and-first-lady-test-positive-for-covid-19 [accessed 2022-03-03]
5. Older, overweight and male: Trump's COVID risk factors make him vulnerable. Reuters. URL: https://www.reuters.com/article/health-coronavirus-usa-trump-risk-int-idUSKBN26N23I [accessed 2022-03-03]
6. Hales C, Carroll M, Fryar C, Ogden C. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/products/databriefs/db360.htm [accessed 2021-10-26]
7. Flint S. Stigmatizing media portrayal of obesity during the coronavirus (COVID-19) pandemic. Front Psychol 2020;11:2124 [FREE Full text] [doi: 10.3389/fpsyg.2020.02124] [Medline: 33013540]
8. Auxier B, 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 2021-10-26]
9. Busam B, Solomon-Moore E. Public understanding of childhood obesity: qualitative analysis of news articles and comments on Facebook. Health Commun 2023 May;38(5):967-980. [doi: 10.1080/10410236.2021.1985859] [Medline: 34605342]
10. Lucibello KM, Vani MF, Koulanova A, deJonge ML, Ashdown-Franks G, Sabiston CM. #quarantine15: A content analysis of Instagram posts during COVID-19. Body Image 2021 Sep;38:148-156 [FREE Full text] [doi: 10.1016/j.bodyim.2021.04.002] [Medline: 33892438]
11. Jebeile H, Partridge SR, Gow ML, Baur LA, Lister NB. Adolescent exposure to weight loss imagery on Instagram: A content analysis of "top" images. Child Obes 2021 Jun;17(4):241-248. [doi: 10.1089/chi.2020.0351] [Medline: 33691078]
12. Media Cloud. Media Cloud. URL: https://mediacloud.org [accessed 2022-02-01]
13. Content Discovery and Social Monitoring Made Easy. CrowdTangle. URL: https://www.crowdtangle.com/ [accessed 2023-05-02]
14. Verified Badges. Instagram Help Center. URL: https://help.instagram.com/854227311295302 [accessed 2022-03-17]
15. What data is CrowdTangle tracking? CrowdTangle. URL: http://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking [accessed 2022-01-25]
16. Schaffer AL, Dobbins TA, Pearson S. Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: a guide for evaluating large-scale health interventions. BMC Med Res Methodol 2021 Mar 22;21(1):58 [FREE Full text] [doi: 10.1186/s12874-021-01235-8] [Medline: 33752604]
17. Hyndman RJ, Khandakar Y. Automatic time series forecasting: The forecast package for R. J Stat Soft 2008;27(3):1-22 [FREE Full text] [doi: 10.18637/JSS.v027.i03]
18. Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. ArXiv Preprint posted online Mar 11, 2022. [FREE Full text]
19. Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Front Sociol 2022;7:886498 [FREE Full text] [doi: 10.3389/fsoc.2022.886498] [Medline: 35602001]

20. Baird A, Xia Y, Cheng Y. Consumer perceptions of telehealth for mental health or substance abuse: a Twitter-based topic modeling analysis. JAMIA Open 2022 Jul;5(2):ooac028 [FREE Full text] [doi: 10.1093/jamiaopen/ooac028] [Medline: 35495736]
21. Hendry D, Darari F, Nurfadillah R, Khanna G, Sun M, Condylis P, et al. Topic modeling for customer service chats. 2021 Presented at: 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS); October 23-25, 2021; Depok, Indonesia. [doi: 10.1109/icacsis53237.2021.9631322]
22. Scarpino I, Zucco C, Vallelunga R, Luzza F, Cannataro M. Investigating topic modeling techniques to extract meaningful insights in Italian long COVID narration. BioTech (Basel) 2022 Sep 03;11(3):41 [FREE Full text] [doi: 10.3390/biotech11030041] [Medline: 36134915]
23. de Groot G, Aliannejadi M, Haas M. Experiments on generalizability of BERTopic on multi-domain short text. 2022 Presented at: Widening Natural Language Processing 2022 Workshop; December 7, 2022; Abu Dhabi, United Arab Emirates.
24. Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. ArXiv Preprint posted online May 24, 2019. [FREE Full text]
25. Office for Human Research Protections (OHRP). 2018 Requirements (2018 Common Rule). Office for Human Research Protections (OHRP). 2017. URL: https://tinyurl.com/28dx8p2y [accessed 2021-05-19]
26. King G, Schneer B, White A. How the news media activate public expression and influence national agendas. Science 2017 Nov 10;358(6364):776-780. [doi: 10.1126/science.aao1100] [Medline: 29123065]
27. Starbird K, Maddock J, Orand M, Achterman P, Mason R. Rumors, false flags, and digital vigilantes: misinformation on Twitter after the 2013 Boston Marathon bombing. 2014 Presented at: iConference 2014; March 4-7, 2014; Berlin, Germany. [doi: 10.9776/14308]
28. Vidal-Alaball J, Fernandez-Luque L, Marin-Gomez FX, Ahmed W. A new tool for public health opinion to give insight into telemedicine: Twitter poll analysis. JMIR Form Res 2019 May 28;3(2):e13870 [FREE Full text] [doi: 10.2196/13870] [Medline: 31140442]

Abbreviations

AICc: sample-corrected Akaike information criterion
ARIMA: autoregressive integrated moving average
BERT: bidirectional encoder representations from transformers
LDA: latent Dirichlet allocation
WHO: World Health Organization

Edited by R Cuomo; submitted 31.05.22; peer-reviewed by A Wahbeh, PCI Pang; comments to author 05.11.22; revised version received 30.01.23; accepted 27.03.23; published 16.05.23.

Please cite as:
Pollack C, Gilbert-Diamond D, Onega T, Vosoughi S, O'Malley AJ, Emond JA
Obesity-Related Discourse on Facebook and Instagram Throughout the COVID-19 Pandemic: Comparative Longitudinal Evaluation
JMIR Infodemiology 2023;3:e40005
URL: https://infodemiology.jmir.org/2023/1/e40005
doi:10.2196/40005
PMID:37191990

Original Paper

Exploring Chronic Pain and Pain Management Perspectives: Qualitative Pilot Analysis of Web-Based Health Community Posts

Claire Harter[1], MA; Marina Ness[1], MPH; Aleah Goldin[1], BA; Christine Lee[2], PharmD, PhD; Christine Merenda[2], RN, MPH; Anne Riberdy[2], MA; Anindita Saha[2], BSc; Richardae Araojo[2], MSc, PharmD; Michelle Tarver[2], MD, PhD

[1]Inspire, Arlington, VA, United States

[2]US Food and Drug Administration, Silver Spring, MD, United States

Corresponding Author:
Christine Lee, PharmD, PhD
US Food and Drug Administration
WO32 RM2306
10903 New Hampshire Ave
Silver Spring, MD, 20993
United States
Phone: 1 240 402 4228
Email: ChristineS.lee@fda.hhs.gov

Abstract

Background: Patient perspectives are central to the US Food and Drug Administration's benefit-risk decision-making process in the evaluation of medical products. Traditional channels of communication may not be feasible for all patients and consumers. Social media websites have increasingly been recognized by researchers as a means to gain insights into patients' views about treatment and diagnostic options, the health care system, and their experiences living with their conditions. Consideration of multiple patient perspective data sources offers the Food and Drug Administration the opportunity to capture diverse patient voices and experiences with chronic pain.

Objective: This pilot study explores posts from a web-based patient platform to gain insights into the key challenges and barriers to treatment faced by patients with chronic pain and their caregivers.

Methods: This research compiles and analyzes unstructured patient data to draw out the key themes. To extract relevant posts for this study, predefined keywords were identified. Harvested posts were published between January 1, 2017, and October 22, 2019, and had to include #ChronicPain and at least one other relevant disease tag, a relevant chronic pain management tag, or a chronic pain management tag for a treatment or activity specific to chronic pain.

Results: The most common topics discussed among persons living with chronic pain were related to disease burden, the need for support, advocacy, and proper diagnosis. Patients' discussions focused on the negative impact chronic pain had on their emotions, playing sports, or exercising, work and school, sleep, social life, and other activities of daily life. The 2 most frequently discussed treatments were opioids or narcotics and devices such as transcutaneous electrical nerve stimulation machines and spinal cord stimulators.

Conclusions: Social listening data may provide valuable insights into patients' and caregivers' perspectives, preferences, and unmet needs, especially when conditions may be highly stigmatized.

(JMIR Infodemiology 2023;3:e41672) doi:10.2196/41672

KEYWORDS

chronic pain; pain management; online health community

Introduction

Patients are at the heart of what the US Food and Drug Administration (FDA) does and are vital to the agency's work of protecting public health by ensuring the safety and efficacy of drugs, biological products, and medical devices [1]. Understanding patient perspectives can aid the agency in numerous ways; review staff can better understand patient experience, consider symptom management and side effects, impact of treatment on quality of life, and risk-benefit profiles.

Traditionally, the patient voice is heard through channels such as participation at formal meetings, letters to the agency, docket comments, or survey responses. It is important to recognize that not all patients are familiar or comfortable with using traditional ways of communicating with organizations such as the FDA. Confounded with this barrier are the unique challenges that racial and ethnic minorities and underserved populations encounter, such as mistrust [2]. Medical mistrust can hinder communication and sharing of information such as medical history and patient experiences [3]. This reticence to share experiences is often further amplified in discussions of stigmatized disease conditions such as chronic pain [4].

Over the past few decades, advances in technology have enabled researchers and health care providers to gain insights into patients' perspectives in ways that have not been previously possible. Social media websites have been increasingly recognized as a platform for patients to gather information, explore options, and share their experiences [5]. With over 80% of Americans seeking and sharing health information online through blogs, microblogging (eg, Twitter), social networking (eg, Facebook), and video and file-sharing sites (eg, YouTube), social media cannot be ignored [6,7]. Social listening is one potential avenue that can be leveraged to gain insights into the patient experience.

Incorporation of the patient voice is an important aspect of regulatory decision-making, supported by the 21st Century Cures Act (Cures Act). The Cures Act builds on the FDA's ongoing work to incorporate patients' perspectives into the development of regulated products and regulatory decision-making process [8].

The Center for Devices and Radiological Health aims to ensure patients are at the center of its regulatory decision-making process. It does this through encouraging patient engagement, the incorporation of clinical outcome assessments in medical device clinical investigations, and the collection of patient preference information [9]. The FDA Office of Minority Health and Health Equity (OMHHE) also supports efforts to amplify equity of voices through its Enhance Equity Initiative by (1) expanding OMHHE's diverse stakeholder community, (2) supporting research to leverage novel and big data sources to understand diverse patient perspectives, preferences, and unmet needs, and (3) supporting expansion of culturally and linguistically tailored health education [10].

Many internet users seek health information through online health communities and other social media, including sharing information and well as finding value in peer-generated health information [11-13]. This study supports the priority, Empowering Patients and Consumers, within the report Advancing Regulatory Science at FDA: Focus Areas of Regulatory Science. This priority advances understanding of ways to engage patients and consumers to better understand the US patients' and public's perspectives and preferences related to outcomes [14].

This pilot study explores the potential for social listening data to expand our understanding of its use for gathering patients' and caregivers' perspectives. The goal of this project was to gain insights into the key challenges and barriers faced by persons living with chronic pain (PLWCP), how they mitigate or treat chronic pain, perspectives and experience with medication dependence and addiction, and how they describe their pain and measures of success. Utilizing multiple patient perspective data sources offers the FDA the opportunity to capture diverse patient voices and experiences on chronic pain.

Methods

The Inspire research team regularly compiles and analyzes unstructured patient data to draw out key themes. Over 1,700,000 members have joined Inspire through its website [15], to share their patient journey, ask and answer questions, and engage with other members who know what they are going through by writing posts and responding to others' posts. These members belong to one or more of over 240 communities focused on specific conditions or disease areas.

To extract relevant posts for this study, predefined keywords and TextRazor tags [16] were identified and used to extract Inspire posts. Harvested posts were published between January 1, 2017, and October 22, 2019, the latter being the date the posts were extracted. The first data set comprised all chronic pain posts that contained #ChronicPain and at least one other relevant disease tag such as #Migraine or #NervePain. The second data set contained #ChronicPain plus a relevant chronic pain management tag or a chronic pain management tag for a treatment or activity specific to chronic pain (eg, #SpinalCordStimulator). Table 1 shows the full set of keywords and TextRazor tags used for harvesting posts. All keywords and tags accounted for misspellings and variations in spelling, and TextRazor tags additionally accounted for synonyms.

Table 1. Full set of keywords and TextRazor tags used for harvesting web-based posts.

Category, keywords	TextRazor tags
Chronic pain[a]	
Chronic pain	ChronicPain
Common diagnoses related to chronic pain	
Back pain	BackPain, LowBackPain
Carpal tunnel syndrome	CarpalTunnel, CarpalTunnelSyndrome
Complex regional pain syndrome	ComplexRegionalPainSyndrome
Failed back surgery syndrome	FailedBackSyndrome
Fibromyalgia	Fibromyalgia
Migraine	Migraine
Multiple sclerosis	MultipleSclerosis
Muscle spasms, muscle pain	MusclePain, Spasm
Nerve pain	NervePain
Neuropathy	NeuropathicPain, PeripheralNeuropathy
Peripheral vascular disease	PeripheralArteryDisease
Phantom limb pain	PhantomPain
Sickle cell disease	SickleCellDisease
Spasticity	Spasticity
Spinal cord injury	SpinalCordInjury
Cerebral palsy	CerebralPalsy
Medication-assisted treatment	
Buprenorphine, Butrans	Buprenorphine, BuprenorphineNaloxone
Dilaudid	Dilaudid
Evzio	Evzio
Suboxone	Suboxone
Fentanyl	Fentanyl
Hydrocodone	Hydrocodone
Hydromorphone	Hydromorphone
Methadone	Methadone
Morphine	Morphine
Naloxone, Norco	Naloxone, Narcan
Naltrexone	Naltrexone, LowDoseNaltrexone
Norco	Norco
Oxy, oxycodone, oxycontin	Oxycodone
Percocet	Percocet
Stadol	Butorphanol
Vicodin	AcetaminophenHydrocodone
Demerol, meperidine	Demerol
Prescription pain medication	PainMedication
Chronic pain medication	PainMedication
Opioid	Opioid, Opiate, OpioidEpidemic
Devices	

Category, keywords	TextRazor tags
TENS[b]	TranscutaneousElectricalNerveStimulation, Neurostimulation, Transcutaneous electrical nerve stimulationunit
Pain pump~[c]	PainPump
Drug pump	DrugPump
Implantable pump	Implantable pump
Opioid pump	Opioid pump
Patient-controlled analgesia pump	Patient-controlled analgesia pump
Spinal pump~	Spinal pump~
PNS[d]	peripheral nerve stimulator, PercutaneousTibialNerveStimulation
Spinal cord stimulator~	SpinalCordStimulator
Pain management	
Pain management	PainManagement
Addiction, addicted	Addiction
Health care provider, physician, pain specialist	Physician

[a]This tag was included in all posts other than those focused on chronic pain–specific treatment.

[b]TENS: transcutaneous electrical nerve stimulation.

[c]Tilde (~) indicates posts focused on a chronic pain–specific treatment.

[d]PNS: peripheral nerve stimulator.

The data pull yielded 3156 posts with the following information recorded for each: post title, post content, unique user token of author, time stamp of when the post was published, geographic location of where the post was published, and gender and age per self-report from initial registration or user profile. Posts were subsequently excluded if they did not contain any text (eg, only contained images or videos) or were duplicates of other posts. For the data set about chronic pain management, posts were read to ensure the inclusion of content about management strategies for chronic pain and not only comorbidities.

In-depth analyses were performed on approximately a third of all posts (ie, 920 posts after duplicates and image and video-only posts were removed). Within these posts, approximately half were about chronic pain (494 posts) and the remaining half (426 posts) were about chronic pain management. The 2 data sets—the chronic pain data set and the chronic pain management data set—were examined individually and had different codebooks. The codebooks were developed following a 4-level hierarchy of decision-making: during open coding, text was carefully analyzed from each post to identify preliminary themes (level 1), and then preliminary codes were discussed among the coders (level 2). After reviewing the data, codes were finalized (level 3) and then for better characterization further divided into subcodes (level 4), thereby ensuring a robust model of consensus-based analysis, which means that the final tags did not stem from 1 analyst but 2. In this case, both coders discussed and reached consensus on what the codes and subcodes should be, and then the posts were tagged accordingly. Themes and subthemes were developed using a data-driven approach, relying on a constant comparative method that closely followed that of Osadchiy et al's [17] social listening study. Inspire's research team first created a data coding tracker in the targeted Inspire data pull, which identified the overarching topics by which analysts would organize the analysis. Next, analysts created a data codebook, which identified the terms and topics that could be coded under each tracker column for each post. Using this codebook, researchers manually read, analyzed, and tagged each post for key trends and topics. All disagreements were resolved by discussion with team members talking through their coding logic and coming to a consensus.

As seen in an overview of the codebooks (Table 2), the analyses consisted of 3 main parts: (1) lexical analysis, which investigated rhetorical strategies within posts about chronic pain, (2) identification of treatment types and sources for posts about chronic pain management, and (3) content analysis about key challenges and measures of success for both data sets. In order to establish the themes and subthemes for classification, a random sampling of posts was read, characterized, and discussed. Once the categories for coding were agreed upon, posts were reread and all posts subsequently coded.

Table 2. Overview of the codebooks.

Codebook, theme	Subtheme
Disease codebook	
Chronic pain lexicon	• metaphor/imagery: "severe" • scale/level: "flare," "worsening," "extreme," "constant," "exhausting," "aching," "horrible," "debilitating" • other lexicon
Symptoms	• fatigue • depression, anxiety • irritable bowel syndrome symptoms • insomnia • nausea, vomiting • confusion, brain fog • dizziness, vertigo • neuropathy • other symptoms
Key challenges	• quality of life impact • poor disease management • lack of diagnosis or misdiagnosis • stigma and social impact • bad health care provider • emotional impact • comorbidities • lack of support • finding health care provider • flares • limited health literacy • impact on loved ones • loss of independence/autonomy
Measures of success	• good disease management • finding support • successful diagnosis • improving quality of life • finding good health care provider • health literacy • decreased stigma • remission • maintaining autonomy
Treatment codebook	
Treatment type	• opioid or narcotic • device • alternative: item • alternative: activity • anticonvulsant • surgery or procedure • sedative or anesthetic • nonsteroid anti-inflammatory drugs • steroid • muscle relaxant • antidepressant • other treatment types
Specific treatments	• oxycodone/oxycontin • marijuana/cannabis • spinal cord stimulator • gabapentin • physical therapy • diet • Lyrica • tramadol • exercise • TENS[a] • other specific treatments

Codebook, theme	Subtheme
Mode of administration	• oral • subcutaneous • transcutaneous • topical • sublingual • other mode of administration
Treatment source	• health care provider prescription • over the counter or web-based store • illegal source • friends or family
Treatment emotions	• negative (angry, desperate, afraid…) • positive (hopeful, satisfied, grateful…) • neutral (cautious, curious, confused…)
Key challenges and barriers	• tolerability, side effects • lack of efficacy • access: health care provider • stigma • addiction, dependence • access: legal • quality of life impact • lack of health literacy • low dosage • difficult administration • other challenges
Measures of success	• improving quality of life • efficacy • access • tolerability • reducing medications • lack of stigma • lack of addiction • other measures of success

[a]TENS: transcutaneous electrical nerve stimulation.

Ethical Approval

The New England Independent Review Board and the FDA both approved this study, finding it minimal risk and met the requirements for a waiver of consent (New England IRB# 120190469; the FDA #: 2023-OC-060). The informed consent process was waived for this study because this was secondary data analysis.

Results

A total of 920 posts by 360 authors who resided in the United States were manually analyzed. When posts contained direct references to a "self" (and the type of self could be determined on the basis of analysis of content), the authors were classified as either patients or caregivers. In posts identifying the author (865/920, 94%), the majority were patients (813/865, 93.9%) followed by caregivers (52/865, 6%). If an author mentioned being both a patient and caregiver, then the author was only classified as the former for the purposes of this research. Per registration and profile data, self-reported gender was collected for 310 (86.1%) of the 360 authors: 89% (276/310) identified as female and 10.9% (34/310) identified as male. During the time of the post extraction, there was no option for nonbinary gender selection on Inspire. Age was also self-reported for 84.4% (304/360) of the authors, with the majority in 40-69 years of age (see Table 3). Overall information on race or ethnicity could not be discerned, as most user profiles lacked such information. This information was not collected on Inspire at the time of the post extraction.

Table 3. Age data of the authors (n=304).

Age range (years)	Values, n (%)
18-29	13 (4.3)
30-39	39 (12.8)
40-49	58 (19.1)
50-59	73 (24)
60-69	93 (30.6)
>70	28 (9.2)

The specific diseases and conditions mentioned most frequently in association with chronic pain were fibromyalgia (43/360, 11.9%), Ehlers-Danlos syndrome (33/360, 9.2%), complex regional pain syndrome (19/360, 5.3%), cancer (18/360, 5%), and chronic migraine (18/360, 5%). More than 65 other chronic pain conditions were mentioned less than 5% of the time, including back injury, scleroderma, and rheumatoid arthritis. Nearly half of the authors of posts (162/360, 45%) who mentioned a specific comorbidity also wrote about experiencing multiple comorbidities, with an average of 2.5 conditions mentioned per author on average.

Within the first data set (ie, specific to chronic pain and not its management), the Inspire research team identified 5 rhetorical themes among the posts that contextualized personal experiences of living with chronic pain. The team categorized the 5 themes in this study as subjective scales, examples of quality of life impact, frequency and length of pain descriptors, illustrative characterizations of pain, and self-validating language based on the content (see Table 4). Often a single post contained 2 or more of these themes, and all of them were used to impart information about pain intensity or quality. Moreover, rarely (in <2% of posts) did posts contain mitigating language such as mild, minimal, moderate, tolerable, or stable to describe the chronic pain. When such adjectives or adverbs were used, they were wielded to reflect how authors perceived others such as health care providers' perspectives of chronic pain.

> *…Well unfortunately in my area there are no temporomandibular joint dysfunction (TMJ) dysfunction support groups as TMJ is viewed as a mild condition not worthy of even having a support group.* [Person with fibromyalgia]

Table 4. Five rhetorical themes among the posts on personal experiences of living with chronic pain.

Theme	Function	Examples
Subjective scales	Convey how patients feel relative to their baseline levels of chronic pain	• "new level of pain" • "very severe" • "immense" • "worsening" • "manage my pain level"
Examples of quality of life impact	Show concrete examples of how chronic pain impacts various aspects of life	• "disruptive" • "disabling" • "daily struggle" • "barely tolerable"
Frequency and length of pain descriptors	Demonstrate the regularity of chronic pain	• "daily" • "intermittent" • "unceasing • "progressive"
Illustrative characterizations of pain	Pinpoints differences in quality of chronic pain experienced	• "burning" • "throbbing" • "radiating" • "sharp"
Self-validating language	Emphasizes the authenticity and weight of lived chronic pain experiences	• "legitimate" • "actual" • "real" • "serious"

Nearly all published posts about chronic pain contained content about the key challenges (437/494, 88.4%) with impact on quality of life the most frequent challenge mentioned (73/437, 16.7%), with quality of life defined as performing daily activities such as cooking and bathing as well as interacting with others. The full complement of key challenges can be found in Table 5.

Table 5. Full complement of key challenges (n=437).

Key challenge	Values, n (%)
Impact on quality of life	73 (16.7)
Managing disease	45 (10.3)
Proper diagnosis	44 (10.1)
Stigma and social impact	39 (8.9)
Relationship with health care providers	33 (7.5)
Emotional impact of disease	31 (7.1)
Navigating comorbidities	27 (6.2)
Lack of support from loved ones	27 (6.2)
Flare-ups	16 (3.7)
Limited health literacy	14 (3.2)
Loss of autonomy	3 (0.7)
Other	85 (19.4)

Approximately 37.8% (187/494) of the posts discussed measures of success for living with chronic pain. The top measures of success within these posts were having good disease management (53/187, 28.3%), maintaining social support (49/187, 26.2%), getting a proper diagnosis (48/187, 25.7%), improving quality of life (47/187, 25.1%), and working with health care providers by willing to listen and advocate for them (47/187, 25.1%). Other measures of success included developing greater health literacy (18/187, 9.6%), noticing less stigma around chronic pain (9/187, 4.8%), being in remission (8/187, 4.3%), and feeling increased autonomy (5/187, 2.7%).

Of the 426 coded posts about chronic pain management, 96.2% (410/426) mentioned a category of chronic pain relief. Opioids or narcotics were mentioned most often (105/410, 25.6%) with oxycodone discussed most frequently (44/105, 41.9%), followed by tramadol (13/105, 12.3%). Only a minority of posts mentioned anticonvulsants (29/410, 7%) such as pregabalin (14/29, 48.3%) or gabapentin (19/29, 65.5%). Few posts mentioned surgery or procedures (24/410, 5.9%) or sedatives or anesthetics (23/410, 5.6%), with lidocaine and acetaminophen equally represented (9/23, 39.1% each). A full accounting of the pain management types can be found in Table 6.

Table 6. Pain management type (n=410).

Pain management type	Values, n (%)
Opioids/narcotics	105 (25.6)
Device (eg, spinal cord stimulators, TENS[a])	56 (13.7)
Alternative substances (eg, cannabis)	48 (11.7)
Alternative interventions (eg, physical therapy, diets, exercise)	44 (10.7)
Anticonvulsants	29 (7.1)
Surgery	24 (5.9)
Sedatives/anesthetics	23 (5.6)
Nonsteroidal anti-inflammatory drugs	20 (4.9)
Steroids	17 (4.1)
Muscle relaxants	14 (3.4)
Antidepressants	13 (3.2)
Other	17 (4.1)

[a]TENS: transcutaneous electrical nerve stimulation.

Treatment sources were reported in 30.7% (142/462) of the posts, with the majority of these indicating that the treatment under discussion was prescribed by a health care provider (103/142, 72.2%). Rarely did posts refer to over-the-counter or web-based vendors (28/142, 19.7%). An even smaller subset of posts mentioned procurement through the street or from friends or family (11/142, 7.7%). When discussing treatment sources, particularly for opioids, posts often made a point to mention having at least at one point a legitimate script from a health care provider.

> *…When the pharmacy refused to fill a legitimate script, I was left in a very bad way. My husband*

> *couldn't stand to watch me suffer; we are law-abiding people, but he was going to buy me something off the street - those ninnys in office don't see they are driving patients with chronic pain to despair.* [Person with Ehlers-Danlos syndrome]

Challenges about chronic pain management were mentioned in 38.5% (164/426) of the posts. The 2 most frequent challenges discussed in posts were tolerability (55/164, 33.5%) and lack of efficacy (54/164, 32.9%). Nearly a third of the posts mentioned difficulty accessing treatments from health care providers (49/164, 29.9%); a smaller number of posts mentioned stigma around their condition (27/164, 16.5%) and addiction (28/164, 17.1%). Some posts featured challenges such as legal access (20/164, 12.2%), low health literacy (18/164, 10.9%), quality of life impact (16/164, 9.8%), low dosage (15/164, 9.1%), difficulty with administration (14/164, 8.5%), and other challenges (15/164, 9.1%), including cost and time. When broaching these challenges, many of the posts were contextualized within the opioid crisis. Overall, authors seemed conflicted, recognizing that long-term opioid usage leads to dependence but also feeling exasperated by not being able to find other treatments with similar levels of pain relief.

> *…Until recently, I took more medication, and was able to function better. However, I can see the pressure my rheumatologist is under to limit the prescription of narcotics, and I do not want to cause problems for her. This makes me sad, because I am in constant pain and my mobility and quality of life are severely affected. A couple of days ago, I broke down in tears because I was in so much pain, yet so conflicted about taking more medication; sometimes, I feel quite downhearted about it.* [Person with Ehlers-Danlos syndrome]

Measures of success for chronic pain management were featured in approximately a third of posts (137/426, 32.2%). Within posts, measures of success included improved quality of life (87/137, 63.5%) and efficacy (86/137, 62.8%). Posts that mentioned personal experience with opioids often stated improved quality of life as the primary reason they preferred or were grateful for opioids.

> *…I also had 10 opioid pills. The second day home a pain came suddenly to my gut and chest area. It was late on a Friday night. A stabbing pain, more intense than any I have ever had. Thank god for the pain medication. [...] the difference between screaming in pain and resting somewhat comfortably is not something I would want to live without.* [Person with Ehlers-Danlos syndrome and cancer]

Other measures of success were access to chronic pain management (27/137, 19.7%) and tolerability of the management (26/137, 18.9%). Less commonly mentioned were reduction of medications or dosages (12/137, 8.8%), decreased stigma (11/137, 8%), not being dependent or addicted (8/137, 5.8%), and other measures (11/137, 8%) such as ease of administration, health literacy, and compliance.

Discussion

This study explores the potential for utilizing social listening data to expand our understanding of its use for gathering patients' and caregivers' perspectives of chronic pain. It is important to understand user-generated content about chronic pain and chronic pain management from social media and web-based peer-to-peer health networks. In addition, key challenges and barriers faced by PLWCP as well as how they mitigate or treat chronic pain were identified from these platforms. For example, there were some key differences in chronic pain discussions between general social media and peer-to-peer health networks. In general, research has documented that social media sites (eg, Reddit, Instagram, Tumblr, Pinterest, Twitter) act as venues for patients seeking others' advice and stages from which to legitimize their experiences and build empathy [18,19]. In this way, digital conversations and narratives help make invisible chronic pain visible and combat the culture of disbelief, that is, the failure to accept an individual's account of his or her pain as true [20-22]. On networks such as Inspire, the audience within the venue is more targeted and includes only other patients, caregivers, and the occasional health care provider. Yet, even in this relatively safe environment, we found that authors of chronic pain habitually felt the need to use rhetorical appeals to ground and situationalize their questions and advice. This may, in part, reflect the extent to which the culture of disbelief is internalized by patients and caregivers and impacts their chronic pain experiences.

It is in this context that posts about relief for chronic pain also exist. Studies within health care spaces have revealed that patients felt disrespected and suspected of drug-seeking when seeking chronic pain management even before the height of the opioid crisis [23]. Part of the issue may be differences in patients' and health care providers' relative priorities for pain management. Patients' top priorities are generally reduction of pain intensity, followed by diagnosing the cause of the pain, whereas health care providers' top priorities are generally improving function, followed by reducing medication side effects [24]. Approximately 24.6% (105/426) of the chronic pain management posts from Inspire mentioned opioids or narcotics. Although there is awareness within these posts that the long-term regular usage of opioids can lead to dependence and that misuse of opioids is common, it is important to note that many PLWCP either (1) do not consider themselves at risk for addiction or (2) consider this risk less important than immediate relief from pain. This matches what other studies have found, with the reasoning there being that patients tended to regard themselves as exceptions since they were genuinely in pain and were not engaging in aberrant behaviors such as asking for early refills or taking more medications than prescribed [25].

Patients are keenly aware of the stigma surrounding opioids or narcotics and crave other efficacious management strategies, which can be seen in the language they use within their posts. PLWCP who mention using opioids in Inspire posts frequently assert that they take the "lowest possible dose" or that this is the "only treatment which has been successful" or that they "only take the medication as needed." When compared to opioids or narcotics, other chronic pain management strategies tended to be positioned as ineffective. For instance, when marijuana or gabapentin was mentioned in posts, these treatments were portrayed as unsuccessful as compared to the immediate and long-lasting relief of opioids. Even so, some PLWCP reported moderate success with anticonvulsants and sedatives, although both anticonvulsants and sedatives were mentioned less than 30 times each, and these results should be taken with caution. Similarly, there appears to be increasing awareness that medical devices such as spinal cord stimulators and TENS (transcutaneous electrical nerve stimulation) machines may help alleviate chronic pain. Patients without personal exposure to such devices expressed hope and curiosity about them, actively seeking out personal anecdotes of PLWCP.

To adequately address chronic pain, we need to have a greater awareness of the multifaceted discussions that PLWCP are having online, particularly on digital peer-to-peer health networks. As seen in the key challenges mentioned in Inspire posts, many PLWCP felt as though they exposed themselves to social and institutional barriers that have made them feel even more vulnerable and isolated than before when attempting to reduce pain intensity. Nearly a third of the posts about chronic pain management mentioned difficulty accessing treatments from health care providers (49/164, 29.9%), followed by stigma (27/164, 20.7%). Even those who did not mention chronic pain management in their posts reported stigma and social impact (39/437, 8.9%) and having poor relations with health care providers (33/437, 7.6%). As other studies have documented, the health care system has not always been structured to reflect a continuum of care for pain, resulting in barriers that can impede persons with chronic pain from receiving timely access to care [26]. Analysis of web-based conversations, especially those directed to and for other patients and caregivers, should inform how we attempt to address chronic pain barriers and measures of success. Particularly important is better understanding patient and caregiver perceptions of the available treatment options and what approaches might encourage them to try management strategies that have a low risk of dependency.

The findings in our report are subject to several limitations. First, because of the digital divide, those who post on web-based peer-to-peer health networks are not representative of the general population. Although this is beginning to change in the age of mobile-friendly websites, this still means that those who are unable to afford a mobile device or have easy access to Wi-Fi are limited in their ability to participate in these networks. Second, this study had a relatively small sample of posts mentioning anticonvulsants, sedatives, and treatment devices for chronic pain. Future studies should further investigate patient perspectives of these chronic pain management strategies, as this literature is still in its infancy. The 5 themes in our study did not have any theoretical framework to support the rhetoric or related research fields, which is a limitation. Researchers have become increasingly interested in the social context of chronic pain conditions, including pain severity, physical disability, pain behaviors, and psychological distress, and have developed theoretical models [27]. In the future, theoretical models should be incorporated to support analysis of constructs. Another limitation was that only 1 source of data was used for the analysis, which was Inspire-only data. Future studies should expand data sources to include additional social media platforms. Finally, while anonymity is a valuable benefit to participating in a web-based peer-to-peer health network, it also creates difficulties when systematically analyzing user-generated content. Key demographics in this study such as gender and age could not be determined unless patients chose to self-identify upon registration or later via their profile pages. Further, demographic information about race and ethnicity was not collected originally at the time of platform registration, thereby severely limiting the analysis of these characteristics. Recognition of this limitation spurred Inspire to collect race and ethnicity data from new members, thereby improving opportunities for health equity research across their platforms. Additionally, it is important to consider that although the use of social media by patients for health-related reasons is growing rapidly, not all social media platforms are ideal or may appeal to all patients. This study only examined 1 condition on 1 online health community platform, that is, Inspire. Future studies should incorporate other diseases and web-based platforms to gather a more comprehensive understanding [11,28,29]. Lastly, studies should include other potential stakeholders such as family members and health professionals to understand their perspectives on chronic pain management.

This study underscores the role of user-generated content in web-based peer-to-peer health networks to help the health care community better understand the treatment and management experience of some patients with chronic pain. Our results suggest that these conversations could help inform our conceptualization of chronic pain challenges and measures of success, which is especially crucial to capture, considering the culture of disbelief. The rhetorical strategies used in posts on Inspire indicate the extent to which this culture impacts even content written to others with akin experiences. PLWCP are aware of the stigma surrounding certain chronic pain treatments options and crave efficacious management strategies; yet, authors of posts perceived strategies other than opioids to be less effective for substantial long-term relief. Even so, some PLWCP reported moderate success with anticonvulsants and sedatives, and some PLWCP appear to be aware that medical devices such as spinal cord stimulators and TENS machines may help alleviate chronic pain. More analysis is needed of the multifaceted discussions that PLWCP are having with each other online. Particularly important is better understanding patient and caregiver perceptions of relief with available chronic pain methods and what may encourage patients to try strategies that can be safely used to manage chronic pain over long periods of time.

Acknowledgments

The authors would like to thank Julie Hsieh, PhD, for providing valuable feedback and for her editorial assistance on this manuscript. This project and publication were supported by the Food and Drug Administration (FDA) of the US Department of Health and Human Services (HHS) as part of a financial assistance award (U01FD005978) totaling US $125,000, with 100% funded by FDA/HHS.

Disclaimer

The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by Food and Drug Administration/Health and Human Services or the US government.

Conflicts of Interest

None declared.

References

1. Learn about patient engagement at the FDA. USFDA. URL: https://www.fda.gov/patients/learn-about-fda-patient-engagement [accessed 2023-01-20]
2. University of California - San Francisco. One in 5 patients report discrimination in health care: Blacks most likely to name race; Whites and Hispanics age, weight or income. ScienceDaily. 2017 Dec 17. URL: https://www.sciencedaily.com/releases/2017/12/171214140840.htm [accessed 2023-01-20]
3. Bazargan M, Cobb S, Assari S. Discrimination and Medical Mistrust in a Racially and Ethnically Diverse Sample of California Adults. Ann Fam Med 2021;19(1):4-15 [FREE Full text] [doi: 10.1370/afm.2632] [Medline: 33431385]
4. Waugh OC, Byrne DG, Nicholas MK. Internalized stigma in people living with chronic pain. J Pain 2014 May;15(5):550.e1-550.10. [doi: 10.1016/j.jpain.2014.02.001] [Medline: 24548852]
5. Househ M, Borycki E, Kushniruk A. Empowering patients through social media: the benefits and challenges. Health Informatics J 2014 Mar;20(1):50-58 [FREE Full text] [doi: 10.1177/1460458213476969] [Medline: 24550564]
6. Decamp M. Physicians, social media, and conflict of interest. J Gen Intern Med 2013 Feb;28(2):299-303 [FREE Full text] [doi: 10.1007/s11606-012-2251-x] [Medline: 23129160]
7. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
8. 21st century cures act. US FDA. URL: https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act [accessed 2023-01-20]
9. Benz HL, Saha A, Tarver ME. Integrating the Voice of the Patient Into the Medical Device Regulatory Process Using Patient Preference Information. Value Health 2020 Mar;23(3):294-297 [FREE Full text] [doi: 10.1016/j.jval.2019.12.005] [Medline: 32197723]
10. Enhance EQUITY initiative. US FDA. URL: https://www.fda.gov/consumers/minority-health-and-health-equity/enhance-equity-initiative [accessed 2023-01-20]
11. Rupert DJ, Moultrie RR, Read JG, Amoozegar JB, Bornkessel AS, O'Donoghue AC, et al. Perceived healthcare provider reactions to patient and caregiver use of online health communities. Patient Educ Couns 2014 Sep;96(3):320-326 [FREE Full text] [doi: 10.1016/j.pec.2014.05.015] [Medline: 24923652]
12. Hodgkin P, Horsley L, Metz B. The emerging world of online health communities. Stanford Social Innovation Review 2018:e. [doi: 10.48558/9jgh-h283]
13. Horrell LN, Lazard AJ, Bhowmick A, Hayes S, Mees S, Valle CG. Attracting Users to Online Health Communities: Analysis of LungCancer.net's Facebook Advertisement Campaign Data. J Med Internet Res 2019 Nov 04;21(11):e14421 [FREE Full text] [doi: 10.2196/14421] [Medline: 31682589]
14. Focus areas of regulatory science report. US FDA. URL: https://www.fda.gov/science-research/advancing-regulatory-science/focus-areas-regulatory-science-report [accessed 2023-01-20]
15. Health and wellness support groups and communities: patients engagement platform. Inspire. URL: https://www.inspire.com/ [accessed 2023-01-20]
16. The natural language processing API. TextRazor Ltd. London, UK URL: https://www.textrazor.com [accessed 2023-01-20]
17. Osadchiy V, Mills JN, Eleswarapu SV. Understanding Patient Anxieties in the Social Media Era: Qualitative Analysis and Natural Language Processing of an Online Male Infertility Community. J Med Internet Res 2020 Mar 10;22(3):e16728 [FREE Full text] [doi: 10.2196/16728] [Medline: 32154785]
18. Merolli M, Gray K, Martin-Sanchez F. Therapeutic affordances of social media: emergent themes from a global online survey of people with chronic pain. J Med Internet Res 2014 Dec 22;16(12):e284 [FREE Full text] [doi: 10.2196/jmir.3494] [Medline: 25533453]
19. Guidry JPD, Benotsch EG. Pinning to Cope: Using Pinterest for Chronic Pain Management. Health Educ Behav 2019 Aug;46(4):700-709. [doi: 10.1177/1090198118824399] [Medline: 30767569]

20. Nicola M, Correia H, Ditchburn G, Drummond P. Invalidation of chronic pain: a thematic analysis of pain narratives. Disabil Rehabil 2021;43(6):861-869. [doi: 10.1080/09638288.2019.1636888] [Medline: 31290347]
21. Sendra A, Farré J. Communicating the experience of chronic pain through social media: patients' narrative practices on Instagram. Journal of Communication in Healthcare 2020 Apr 21;13(1):46-54. [doi: 10.1080/17538068.2020.1752982]
22. Newton BJ, Southall JL, Raphael JH, Ashford RL, LeMarchand K. A narrative review of the impact of disbelief in chronic pain. Pain Manag Nurs 2013 Sep;14(3):161-171. [doi: 10.1016/j.pmn.2010.09.001] [Medline: 23972867]
23. Upshur CC, Bacigalupe G, Luckmann R. "They don't want anything to do with you": patient views of primary care management of chronic pain. Pain Med 2010 Dec;11(12):1791-1798. [doi: 10.1111/j.1526-4637.2010.00960.x] [Medline: 21029353]
24. Henry SG, Bell RA, Fenton JJ, Kravitz RL. Goals of chronic pain management: do patients and primary care physicians agree and does it matter? Clin J Pain 2017;33(11):955-961. [doi: 10.1097/ajp.0000000000000488]
25. Matthias MS, Johnson NL, Shields CG, Bair MJ, MacKie P, Huffman M, et al. "I'm Not Gonna Pull the Rug out From Under You": Patient-Provider Communication About Opioid Tapering. J Pain 2017 Nov;18(11):1365-1373 [FREE Full text] [doi: 10.1016/j.jpain.2017.06.008] [Medline: 28690000]
26. Glajchen M. Chronic pain: treatment barriers and strategies for clinical practice. J Am Board Fam Pract 2001;14(3):211-218 [FREE Full text] [Medline: 11355054]
27. Leonard MT, Cano A, Johansen AB. Chronic pain in a couples context: a review and integration of theoretical models and empirical evidence. J Pain 2006 Jun;7(6):377-390 [FREE Full text] [doi: 10.1016/j.jpain.2006.01.442] [Medline: 16750794]
28. Rupert DJ, Moultrie RR, Read JG, Amoozegar JB, Bornkessel AS, O'Donoghue AC, et al. Perceived healthcare provider reactions to patient and caregiver use of online health communities. Patient Educ Couns 2014 Sep;96(3):320-326 [FREE Full text] [doi: 10.1016/j.pec.2014.05.015] [Medline: 24923652]
29. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: A systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16:442 [FREE Full text] [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]

Abbreviations

FDA: Food and Drug Administration
OMHHE: Office of Minority Health and Health Equity
PLWCP: persons living with chronic pain
TENS: transcutaneous electrical nerve stimulation

Edited by T Mackey; submitted 06.09.22; peer-reviewed by N Hu, A Pal; comments to author 02.11.22; revised version received 11.04.23; accepted 24.04.23; published 30.05.23.

Please cite as:
Harter C, Ness M, Goldin A, Lee C, Merenda C, Riberdy A, Saha A, Araojo R, Tarver M
Exploring Chronic Pain and Pain Management Perspectives: Qualitative Pilot Analysis of Web-Based Health Community Posts
JMIR Infodemiology 2023;3:e41672
URL: https://infodemiology.jmir.org/2023/1/e41672
doi:10.2196/41672
PMID:37252767

Original Paper

Expert Credibility and Sentiment in Infodemiology of Hydroxychloroquine's Efficacy on Cable News Programs: Empirical Analysis

Dobin Yim[1], PhD; Jiban Khuntia[2], PhD; Elliot King[1], PhD; Matthew Treskon[3], MS, MLS; Panagis Galiatsatos[4], MHS, MD

[1]Loyola University, Baltimore, MD, United States

[2]Health Administration Research Consortium, University of Colorado Denver, Denver, CO, United States

[3]Johns Hopkins University, Baltimore, MD, United States

[4]Johns Hopkins School of Medicine, Baltimore, MD, United States

Corresponding Author:
Jiban Khuntia, PhD
Health Administration Research Consortium
University of Colorado Denver
1475 Lawrence St
Denver, CO, 80202
United States
Phone: 1 3038548024
Email: jiban.khuntia@ucdenver.edu

Abstract

Background: Infodemic exacerbates public health concerns by disseminating unreliable and false scientific facts to a population. During the COVID-19 pandemic, the efficacy of hydroxychloroquine as a therapeutic solution emerged as a challenge to public health communication. Internet and social media spread information about hydroxychloroquine, whereas cable television was a vital source. To exemplify, experts discussed in cable television broadcasts about hydroxychloroquine for treating COVID-19. However, how the experts' comments influenced airtime allocation on cable television to help in public health communication, either during COVID-10 or at other times, is not understood.

Objective: This study aimed to examine how 3 factors, that is, the credibility of experts as doctors (DOCTOREXPERT), the credibility of government representatives (GOVTEXPERT), and the sentiments (SENTIMENT) expressed in discussions and comments, influence the allocation of airtime (AIRTIME) in cable television broadcasts. SENTIMENT pertains to the information credibility conveyed through the tone and language of experts' comments during cable television broadcasts, in contrast to the individual credibility of the doctor or government representatives because of the degree or affiliations.

Methods: We collected transcriptions of relevant hydroxychloroquine-related broadcasts on cable television between March 2020 and October 2020. We coded the experts as DOCTOREXPERT or GOVTEXPERT using publicly available data. To determine the sentiments expressed in the broadcasts, we used a machine learning algorithm to code them as POSITIVE, NEGATIVE, NEUTRAL, or MIXED sentiments.

Results: The analysis revealed a counterintuitive association between the expertise of doctors (DOCTOREXPERT) and the allocation of airtime, with doctor experts receiving less airtime ($P<.001$) than the nonexperts in a base model. A more nuanced interaction model suggested that government experts with a doctorate degree received even less airtime ($P=.03$) compared with nonexperts. Sentiments expressed during the broadcasts played a significant role in airtime allocation, particularly for their direct effects on airtime allocation, more so for NEGATIVE ($P<.001$), NEUTRAL ($P<.001$), and MIXED ($P=.03$) sentiments. Only government experts expressing POSITIVE sentiments during the broadcast received a more extended airtime ($P<.001$) than nonexperts. Furthermore, NEGATIVE sentiments in the broadcasts were associated with less airtime both for DOCTOREXPERT ($P<.001$) and GOVTEXPERT ($P<.001$).

Conclusions: Source credibility plays a crucial role in infodemics by ensuring the accuracy and trustworthiness of the information communicated to audiences. However, cable television media may prioritize likeability over credibility, potentially hindering this goal. Surprisingly, the findings of our study suggest that doctors did not get good airtime on hydroxychloroquine-related discussions on cable television. In contrast, government experts as sources received more airtime on hydroxychloroquine-related discussions. Doctors presenting facts with negative sentiments may not help them gain airtime. Conversely, government experts expressing positive sentiments during broadcasts may have better airtime than nonexperts. These findings have implications on the role of source credibility in public health communications.

(JMIR Infodemiology 2023;3:e45392) doi:10.2196/45392

KEYWORDS

source credibility; infodemic; infoveillance; broadcasting; cable television; COVID-19

Introduction

Background

An infodemic is an expression that blends the words information and epidemic. An infodemic occurs when accurate or inaccurate information rapidly spreads everywhere; the overabundance makes it difficult for people to find trustworthy sources and reliable guidance when needed [1,2]. The dispersion of facts and rumors often bleed into each other in an infodemic, as the information spreads concerns and fears among the public [2,3]. Subsequently, it becomes challenging to learn the correct and essential information.

Prior studies on health information retrieval, spread, and dissemination in flu contexts have asserted infodemiology as a vital area of research needing more attention to explore deeper nuanced mechanisms of health communications [3,4]. Combating infodemics involves awareness, literacy, fact-checking, monitoring (infoveillance), and the nondistortion of facts [2]. More studies would help design and monitor accurate health communication strategies that can disseminate scientific facts to inform public health and policy [5-7].

Early work in media and information management has suggested that people are more likely to be persuaded when a source presents itself as credible while disseminating information [8-10]. A relevant concept of medium credibility would evaluate the medium through which the message is delivered and the characteristics of the message source, such as how social media or newspapers influence persuasion [11]. News and media channels must identify areas where there is a knowledge translation gap between best evidence (what some experts know) and practice (what most people do or believe), as well as markers for "high-quality" information to curb the spread of misinformation [3].

Research must inform how scientific credibility in communication helps manage the spread of information. In this context, source credibility is a concept that focuses on the origin of the fact, message, or information. The source may refer to the government, a nonprofit agency, or a corporation. News and media agencies ratify information through experts from scientific institutions, agencies, or academia to provide credibility [12]. The audience may consider these experts as primary sources. Thus, it is crucial to understand how the source credibility affects the expert-ratified information dissemination during infodemics, which is the objective of this study.

Infodemic During COVID-19

The issue of infodemic was quite apparent during the COVID-19 pandemic, with several pieces of information spreading swiftly; the accuracy of the fact-checking was questionable [13]. In February 2020, as the gravity of the threat posed by COVID-19 came to be recognized internationally, the Director-General of the World Health Organization declared that the world must fight not only the epidemic but also an infodemic [14].

The rapid spread of COVID-19 raised many difficult questions, including what the origin of the virus was, how transmissible it was, how lethal it would be, what mitigation measures might be required to minimize its impact, and how effective the potential treatments and therapeutic drugs were. Given the array of questions to which there were no known answers, the number of COVID-19 cases skyrocketed, and therefore the consumption of information about the pandemic soared [15]. Several studies revealed that the COVID-19–related content found on many social media platforms was inconsistent and unreliable [16-18], leading to infodemic challenges during this period of uncertainty.

Overview of Hydroxychloroquine in Public Discourse

The spread of information about hydroxychloroquine in public discourse during COVID-19 is an exemplary infodemic. The idea that hydroxychloroquine could be an effective therapeutic for COVID-19 began circulating in China in January 2020. Subsequently, it spread through social media in Nigeria; Vietnam; France; and ultimately in the United States in early March when Paul Sperry, a conservative author, tweeted it on March 9. On March 13, investor James Toldano tweeted a link to a Google Document he had coauthored with Gregory Rigano, a lawyer, touting the benefits of hydroxychloroquine.

In March 2020, the idea that hydroxychloroquine could be effective against COVID-19 was first raised publicly with subsequent infodemics [19]. On March 16, Lara Ingraham discussed the drug with Dr Anthony Fauci on her show, and on March 18, Rigano was interviewed on The Tucker Carlson Show; both the shows were broadcasted on FOX News. On the same day, a reporter asked about the potential of hydroxychloroquine as a therapeutic for COVID-19 at a White House briefing. On March 19, at another White House briefing, President Trump touted the drug as a "potential game changer." On March 28, the Food and Drug Administration issued an emergency use authorization, empowering doctors to prescribe hydroxychloroquine to fight COVID-19. Approximately 1 month later, on April 24, the Food and Drug Administration cautioned against using hydroxychloroquine as a treatment for COVID-19, and on June 15, it rescinded the emergency use authorization. In addition, subsequent clinical trials established hydroxychloroquine as an ineffective treatment for COVID-19. The National Institutes of Health stated that hydroxychloroquine was ineffective for COVID-19 in November 2020.

The public interest sparked by the media coverage was evident. The number of prescriptions for using hydroxychloroquine increased from approximately 30,000 in February 2020 to >220,000 in March. However, the number of prescriptions reduced to approximately 100,000 in April and 35,000 in May [20]. There is some evidence that the publicity given to hydroxychloroquine as a therapeutic for COVID-19 led to shortages of the drug for patients who need it for other reasons [21]. On November 9, 2020, the National Institutes of Health issued a press release based on a study that appeared the same day in the Journal of the American Medical Association, stating that hydroxychloroquine does not provide a clinical benefit to adults hospitalized with COVID-19 [22].

Research Gap and Questions

Prior research points to the role of social and other media in scientific credibility and health communication contexts [5-7]. The role of source credibility as a persuasive element remains relatively unexplored [8-10]. More specifically, given the consequential nature of the context of broadcasting information about hydroxychloroquine in public discourse during COVID-19 [23,24], misleading information spread [25,26] points to the need to conduct research exploring source credibility as an element in health communications.

Existing literature that explores hydroxychloroquine in public discourse by using social and other media is sparse. A prior study has identified and characterized scientific authority–related discussions about hydroxychloroquine, alluding to medical experts' credibility aspect of sources [27]. Other studies have explored how emotional-moral words correlate with a higher likelihood of being retweeted, how emotions are essential in making content contagious on social media [27-30], and how moral emotions shaped information spread on Twitter and other media about hydroxychloroquine as a solution to COVID-19 [29,30].

News broadcasts played a substantial role in disseminating information about hydroxychloroquine. Broadcasts used experts from institutions, agencies, or universities, who may have been perceived as the primary source by the audience [12]. It is crucial to understand whether this expert-ratified information was helpful. However, no study provides insights into how expert opinions during the broadcast provided credibility. To address this research gap in the context of hydroxychloroquine in public discourse during COVID-19, we asked, (1) Do credible information sources influence the infodemic process? If so, how? and (2) Which attributes of the source credibility influence the dynamics of information spread?

Study Road Map

This study examines how 3 factors—the credibility of experts as doctors (DOCTOREXPERT), the credibility of government representatives (GOVTEXPERT), and the sentiments (SENTIMENT) expressed in cable television discussion broadcasts—influence the allocation of airtime (AIRTIME) for hydroxychloroquine in public discourse during COVID-19. The data were collected from transcripts of cable television broadcasts and coded using machine learning algorithms. We used Tobit regression models to estimate the effect of experts' credibility and sentiment on airtime. The implications of the findings of our analysis are discussed.

Methods

Sampling Period and Strategy

The study period spans from March 1, 2020, to November 30, 2020. The first mention of hydroxychloroquine on cable news was on March 1, 2020. We noted that the first mention of hydroxychloroquine as a potential treatment for COVID-19 symptoms in a tweet by Elon Musk with a link to a Google Document occurred on March 16, 2020. However, discussions regarding the potential use were present on social media earlier. The National Institutes of Health declared the drug ineffective against COVID-19 on November 30, 2020. The data collection and coding process for this study followed several steps: the identification of the days in which hydroxychloroquine was most discussed on 3 primary cable news networks from March 2020 to November 2020; collection of the broadcast videos; identification of the experts and collection of information about them, calculating the amount of airtime the medical experts on each network received during the discussion of hydroxychloroquine; and an assessment of sentiments expressed in their remarks.

Data Collection Process

The study's data set comes from Stanford Cable TV News Analyzer [31], which collects data from the Internet Archive for television data set that consists of >300,000 video recordings. A vital feature of the Stanford Cable TV News Analyzer augments the Internet Archive data set with a trend dashboard that helps to create a curated database of video segments from cable news, enabling us to conduct focused searches. One key feature of the analyzer is its keyword search query tool, which allows us to identify video segments where specific words are spoken by participants by using the transcript of the video as a reference. This functionality provides valuable insights into experts' sentiments as expressed in the cable television broadcasts.

According to the Stanford Cable TV News Analyzer, a video segment is defined as an approximately 3-minute interval from a cable news show in which at least 1 panel expert mentions the keyword (eg, hydroxychloroquine) in the news transcript. The daily totals indicate the interest cable news networks had in hydroxychloroquine. The search query was performed at the "daily" level; thus, the daily aggregation unit generated a time trend chart to identify the peak periods of hydroxychloroquine-related discussions on the 3 US cable news networks. We defined peak periods as days in which the search results of hydroxychloroquine returned at least ≥20 video segments. We removed dates during which the total daily number of video segments aired was <20 to focus on the high-interest level periods, resulting in 565 unique video segments. Table 1 provides information on the broadcasts on key dates.

Table 1. Information about the broadcasts.

Dates	April 6	April 22	April 23	April 24	May 6	May 14	May 18	May 19	May 20	May 21	May 22	May 23	July 28	July 29	July 30	July 21	August 3	October 2
Total videos, n	29	43	37	43	22	33	25	56	38	29	32	29	28	51	24	24	27	24
Seconds per episode	6.62	5.58	5.35	4.05	1.99	4	11.52	11.04	5.84	2.28	6	4.34	6	6.35	3.5	3.75	4.44	2.05
Minutes per day	3.2	4	3.3	2.9	0.73	2.2	4.8	10.3	3.7	1.1	3.2	2.1	2.8	5.4	1.4	1.5	2	0.82

Once we identified the dates when cable news prominently featured discussions about hydroxychloroquine with panels of experts, we used the query tool by entering 2 variations of hydroxychloroquine, "Hydroxychloroquine" and "Hydroxy," as the keywords. We added the names of 3 main cable news channels: "FOX, CNN, and MSNBC." We also limited the search by adding the term "aired between March 1, 2020, and November 30, 2021." The search was performed using a publicly available Python package on open-source GitHub Archives [32] to query the television archive database. We modified an original Python script (*get_news_identifiers.py*) to implement the search strategy for the videos that matched the key dates. We found 1147 videos, of which 425 (37.05%) were from Cable News Network (CNN), 357 (31.12%) were from FOX News, and 365 (31.82%) were from MSNBC cable networks. We then retrieved the full-text captioning of the videos using another script (ie, *scrape_archive_org.py*), returning HTML files as output, with captions demarcated to the minute. The script identifies and parses the text segment based on the start and end of the time stamps identified from the previous data-coding process. Then, we filtered a subset of these videos whose full text included the word "hydroxychloroquine" or "hydroxy." Filtering for hydroxychloroquine yielded 585 videos (CNN: n=273, 46.7%; FOX News: n=117, 20%; and MSNBC: n=195, 33.3%). Upon final review, we removed 10 videos because they were duplicates, resulting in 575 videos.

Experts' Information in the Broadcasts

The sampled videos were then shared with coders that marked the expert speaker, comment start time, and comment end time. For each video segment identified during the peak period dates, we obtained the names and affiliations of the experts and measured the amount of airtime they received by marking the time stamps of their first and last appearances within the segment. A custom Python script extracted the text of the expert speaker to the nearest minute. Because the time marker of the transcript is at a 1-minute interval, the parsing procedures may include extraneous text, such as the host's introduction of the expert in the output text. Although the added text by the host may introduce potential errors in extracting the expert's core message, the nature of the content is related and relevant; thus, the validity of the analysis would remain intact.

The coding process involved a team of 3 researchers and 3 graduate students who analyzed each person featured in the video segment the show hosts interviewed. Typically, the identifying information about a person, such as their name, credentials, and affiliation, appeared at the bottom of the screen. The coders categorized a person as an expert if their credentials listed a terminal doctorate in medicine or a relevant scientific discipline such as microbiology or epidemiology. Otherwise, the person was coded as a nonexpert. If the video segment did not provide complete credentials and affiliations, the coders searched Google and LinkedIn to verify their expert status. Individuals whose incomplete information could not be verified were excluded from the data set. The coders deliberated on individuals who sounded knowledgeable to include or exclude in the experts' categories, with the inclusion criteria that evaluating or providing expert inputs on hydroxychloroquine's effectiveness as therapeutic for COVID-19-related symptoms requires a scientific or clinical understanding of its applicability as a new treatment alternative. We excluded politicians, lobbyists, lawyers, news contributors, correspondents, hosts, and political appointees holding administrative positions in organizations who did not have academic credentials or prior professional experience in the medical-related field.

We measured the amount of airtime received by experts by recording the start and end time stamps of the conversations between the news host and the experts. The duration of the host's introduction was subtracted from the calculation. If the conversation involved multiple exchanges between the host and the expert, the total duration of the expert's appearance was recorded. In cases where multiple experts were featured in the show, each expert's contribution was captured separately. We addressed syntax and duration calculation errors in the samples and removed samples with missing data. In total, we identified 354 unique experts.

Sentiment Analysis of the Samples Broadcasts

The entire corpus was processed using latent Dirichlet allocation-based topic modeling and an automated sentiment analysis program using Amazon Web Service Comprehend (AWSC). This cloud-based automated service uses machine learning to process the videos' full text for sentiment analysis. This process involves training a classifier on a labeled data set to predict sentiment polarity, including positive, negative, and neutral categories. AWSC is similar to other commercial software applications such as Linguistic Inquiry and Word Count (LIWC) or NVivo and open-source programming languages such as Python and R, which provide sentiment classifier packages such as Natural Language Toolkit, Gensim, and topic modeling. These packages enable automatic tabulation and numerical calculation of sentiment scores for sentences, paragraphs, and documents. Typically, sentiment scores range from 0 (lowest) to 1 (highest) for discrete sentiment polarity or from −1 (negative) to 1 (positive) for a combined sentiment polarity scale. However, AWSC was preferred because of the ease and appropriateness of analyzing extensive text data, scalability, and the latest features while accurately identifying positive and negative sentiments in text.

Sentiment analysis approaches have been used in prior research to understand mediated and health-related content; for example, a study analyzed positive, negative, neutral, and ambiguous tones of tweets on e-cigarettes [33]. Studies have used computer-aided sentiment analysis to annotate sentiments' directionality to better understand sentiments experienced following alcohol-induced blackouts [34] and on breast cancer social networks [35].

The sentiment analysis process follows a lexicon-based classification that categorizes each word in each text as positive, negative, or neutral based on a predefined dictionary. For instance, words such as "joy," "happy," and "excited" are classified as positive sentiment words, whereas "angry," "scared," and "sad" fall into the negative sentiment category. The number of words identified in each sentiment category or polarity can then be compared across different documents. The count of sentiment polarity occurrences in each document can be normalized on a standard scale or combined into a single scale, such as −1 to 1, where negative values indicate negative polarity, positive values denote positive sentiment, and 0 represents neutral sentiment. However, inferring document sentiment solely based on the relative occurrence of sentiment-embedded words can be challenging, especially in cases where such words are sparse, such as in academic manuscripts. A machine learning approach trains a model on a prelabeled data set of documents and their corresponding sentiment outcomes as discrete categories or numerical scores to derive a predictive sentiment classifier, overcoming the sparse-word challenge. This model can then predict the sentiment for a focal set of documents. To better understand what might drive the sentiment scores, we isolated the comments made by the experts and ranked them according to their sentiment scores. Examples of expert comments with high positive or negative sentiment scores are shown in Textbox 1.

Textbox 1. Examples of positive and negative sentiments in experts' comments.

Examples of positive sentiment broadcast comments

- "We continue to study the effectiveness of Hydroxychloroquine and other therapies in the treatment and prevention of the virus, and we will keep the American people fully informed of our fighting. Hydroxychloroquine is looking like it's having some good results. i hope that would be a phenomenal thing but we have it right now." [Mehmet Oz, FOX News on April 4, 2020; score: 0.982]
- "...Hydroxychloroquine that the doctor was talking about in test tubes and seems to be more effective against the virus, and this is the one that has been used more or less around the world. this is the one that the French looked at and had a pretty profound response...I'm very happy about the University of Minnesota is testing and studying this drug. The University of Washington is giving six patients, and what it looks like it's coming out about this drug is it works better if it is used early in the process before the coronavirus covid-19 really takes on steam. so that's what I am looking at." [Marc Siegel, FOX News, on March 24, 2020; score: 0.983]

Examples of negative sentiment broadcast comments

- "Some compounds in a test tube appear to have an anti-viral capacity and are worthless in humans. A recent example of a compound like that is Hydroxychloroquine, which in vitro appeared to have antiviral capabilities but, tested in human beings, is worthless...you hear proponents of this people say I have seen it with my own eyes have incredible power, which you know is all well and good. it sounds great, and maybe that person actually believes it, but that is not actually how science works." [Jonathan Reiner, Cable News Network (CNN), on August 17, 2020; score: 0.996]
- "I can't prescribe Hydroxychloroquine for my lupus patients because so many other people have gotten prescriptions who don't need them. you can see how the misinformation actually leads to pretty bad consequences for patients...it is pretty bad." [Kavita Patel, MSNBC, on April 5, 2020; score: 0.975]
- "The American corporations...are globalist and they want to push a global agenda and make sure that when the time comes for China to be open to that they aren't on the wrong side of China's propaganda arm the Chinese government. That's why they are allowing it. if people are telling people, Hydroxychloroquine doesn't work. Saying that they will die if they take it. they are being allowed to get." [Harmeet Dhillon, FOX News, on April 1, 2020; score: 0.973]

Ethical Considerations

The data collected for this study were obtained from publicly available sources. The study did not involve any interaction with users. Therefore, ethical approval was not required for this study.

Sample Statistics

Table 2 shows the descriptive statistics and pairwise correlations among the key variables used in this study. On average, experts were featured for 264.72 seconds per cable news show in which they appeared. Their statements expressed an average score of 0.16 for positive sentiments, 0.29 for negative sentiments, 0.33 for neutral sentiments, and 0.21 for mixed sentiments. Of the 565 video segments analyzed, 354 (62.7%) featured experts and 64 (11.3%) featured government affiliates. Of the 565 video segments analyzed, the largest number of monthly totals was aired in April, with 171 (30.3%) video segments, followed by 150 (26.6%) segments in May, 100 (17.7%) segments in March, 87 (15.4%) segments in August, 45 (8%) segments in July, and 12 (2.1%) segments in October.

Table 2. Summary statistics and pairwise correlations among key variables (number of observations=565).

Variables	Value, mean (SD)	Value, range	ln(AIRTIME)	DOCTOREXPERT	GOVTEXPERT	POSITIVE	NEGATIVE	NEUTRAL	MIXED	March	April	May	July	August	October
ln(AIRTIME)	5.13 (1.00)	1.39-8.15	1.00	—[a]	—	—	—	—	—	—	—	—	—	—	—
DOCTOREXPERT	0.63 (0.48)	0-1	−0.14	1.00	—	—	—	—	—	—	—	—	—	—	—
GOVTEXPERT	0.11 (0.32)	0-1	0.16	−0.14	1.00	—	—	—	—	—	—	—	—	—	—
POSITIVE	0.16 (0.18)	0-0.94	0.21	0.02	0.25	1.00	—	—	—	—	—	—	—	—	—
NEGATIVE	0.29 (0.23)	0-0.97	−0.07	−0.06	−0.16	−0.54	1.00	—	—	—	—	—	—	—	—
NEUTRAL	0.33 (0.21)	0-0.98	−0.16	−0.09	0.03	−0.05	−0.37	1.00	—	—	—	—	—	—	—
MIXED	0.21 (0.21)	0-0.93	0.05	0.13	−0.07	−0.22	−0.25	−0.55	1.00	—	—	—	—	—	—
March	0.18 (0.38)	0-1	0.16	−0.01	0.24	0.27	−0.30	0.08	0.01	−0.53	1.00	—	—	—	—
April	0.30 (0.46)	0-1	−0.01	0.12	−0.02	0.03	−0.04	0.03	−0.01	−0.39	−0.31	1.00	—	—	—
May	0.27 (0.44)	0-1	−0.11	−0.11	−0.18	−0.13	0.17	0.00	−0.07	−0.02	−0.28	−0.40	1.00	—	—
July	0.08 (0.27)	0-1	−0.20	0.02	0.08	−0.02	0.03	−0.06	0.05	0.32	−0.14	−0.19	−0.18	1.00	—
August	0.15 (0.36)	0-1	0.14	−0.01	−0.06	−0.18	0.19	−0.12	0.07	0.69	−0.20	−0.28	−0.26	−0.13	1.00
October	0.02 (0.14)	0-1	0.02	−0.06	−0.05	0.03	−0.06	0.08	−0.05	0.40	−0.07	−0.10	−0.09	−0.04	−0.06

[a]Not applicable.

Study Variables

The unit of analysis is the expert's appearance on a cable news show per video segment. The dependent variable in this study is AIRTIME. AIRTIME is measured by calculating the difference between the start and end time of a guest's appearance on the cable news network's show in seconds. The values were log transformed to mitigate the skewed distribution of airtime. Table 2 displays that, on average, AIRTIME is 5.13 or 264.72 seconds.

A total of 6 independent variables are of interest in this study. The first 2 are DOCTOREXPERT and GOVTEXPERT. The second set includes the 4 types of sentiments expressed in the broadcasts: POSITIVE, NEGATIVE, NEUTRAL, and MIXED. The independent variable, DOCTOREXPERT, identified a featured guest's expertise on the subject matter because of the advanced doctorate degree and subsequent clinical practice involvements. If the featured guest had a degree in medicine or an advanced degree in a relevant scientific discipline such as microbiology or epidemiology, the variable was coded as 1 and otherwise as 0. The study sample featured an approximately equal distribution of experts (354/565, 62.7%) and nonexperts.

The second independent variable, GOVTEXPERT, identified a featured guest's affiliation with a government organization. An affiliation variable with other organizations, such as academic institutions, health organizations, news organizations, or private practice, was also considered. Only one affiliation type was associated with each featured guest, and the variable was coded as 1 for the affiliation and otherwise as 0. Of the various affiliations, only the government affiliation (64/565, 11.3%) was considered for this study, as other affiliations did not show any statistical significance to explain airtime. Together, these variables comprise a featured guest's credibility in their expertise to report facts or opinions about hydroxychloroquine as a legitimate therapeutic for COVID-19.

The machine learning algorithm measured the 4 variables associated with sentiments expressed in featured guests' statements. Each measurement scale ranged from 0 to 1, with 1 representing the highest level of sentiment expressed. In general, featured guests showed more sentiments in their statements, with combined sentiment scores of 0.67: POSITIVE, NEGATIVE, MIXED, and NEUTRAL sentiments scored 0.16, 0.29, 0.21, and 0.33, respectively.

Statistical Analyses

The empirical model examined the relationship between experts' credibility, experts' sentiments expressed during broadcasts, and the airtime they received. The models included controls for months and days adjust for variations in the opportunities and interests of experts, and variation in the number of cable news appearances over time appearing on cable news networks to discuss hydroxychloroquine. Tobit regression was used that accounted for extreme airtime values at the upper and lower bounds, where some experts who should have received airtime did not appear on the show. The specified and estimated interaction models build on a base model that specify credibility variables' direct effects on airtime. We then each one of the 3 highly correlated sentiment-dummy variables separately in regressions to avoid multicollinearity, improve the accuracy and comprehensiveness of the analysis, and draw comparable insights about each variable. We added dummy variables reflecting cable news channels to cluster the error variances that may arise from the repeated measures of cable news shows. Including time dummy variables and cable news clustering variables minimizes the bias associated with the model specification. Textbox 2 shows the interaction model specifications that were estimated, in which *i* denotes one broadcast as the unit of analysis:

Textbox 2. Interaction model with DOCTOREXPERT and GOVTEXPERT.

- *1.1: ln(Airtime)i = β0 + β1DOCTOREXPERTi + β2GOVTEXPERTi + β3DOCTOREXPERTi × GOVTEXPERTi + Controlsi + εi*
- *1.2: ln(Airtime)i = β0 + β1DOCTOREXPERTi + β2GOVTEXPERTi + β3POSITIVEi + β4DOCTOREXPERTi × GOVTEXPERTi + β5DOCTOREXPERTi × POSITIVEi + β6GOVTEXPERTi × POSITIVEi + Controlsi + εi*
- *1.3: ln(Airtime)i = β0 + β1DOCTOREXPERTi + β2GOVTEXPERTi + β3NEGATIVEi + β4DOCTOREXPERTi × GOVTEXPERTi + β5DOCTOREXPERTi × NEGATIVEi + β6GOVTEXPERTi × NEGATIVEi + Controlsi + εi*
- *1.4: ln(Airtime)i = β0 + β1DOCTOREXPERTi + β2GOVTEXPERTi + β3NEUTRALi + β4DOCTOREXPERTi × GOVTEXPERTi + β5DOCTOREXPERTi × NEUTRALi + β6GOVTEXPERTi × NEUTRALi + Controlsi + εi*
- *1.5: ln(Airtime)i = β0 + β1DOCTOREXPERTi + β2GOVTEXPERTi + β3MIXEDi + β4DOCTOREXPERTi × GOVTEXPERTi + β5DOCTOREXPERTi × MIXEDi + β6GOVTEXPERTi × MIXEDi + Controlsi + εi*

Results

Overview

Cable television broadcasts used in this sample for hydroxychloroquine span approximately 5 to 265 seconds, with high participation of academic doctor experts (354/565, 62.7%) but fewer government experts (64/565, 11.3%). The broadcasts were equally positive, negative, or mixed, but with a higher neutral sentiment coefficient score. Doctors received less airtime (correlation of −0.14 with AIRTIME) as compared with nonexperts, but government experts received more airtime (correlation of 0.16 with AIRTIME). In general, featured guests showed more sentiments in their statements, with combined sentiment scores of 0.67: positive, negative, mixed, and neutral sentiments scored 0.16, 0.29, 0.21, and 0.33, respectively.

The results of the Tobit regression model estimation are shown in Table 3. There are 5 sets of columns, with the first column displaying the coefficient estimates of each variable and the second column displaying the *P* values. First, with respect to the DOCTOREXPERT variable, the coefficient estimate is negative and statistically significant (−0.181; *P*=.01); however, although its valence is primarily negative, its statistical significance is inconsistent across specifications, suggesting that other factors likely moderate the effect of DOCTOREXPERT on airtime.

Table 3. Full interaction model of Tobit regression results.[a,b]

Variables	DV[c], airtime (seconds; log transformed)									
	1.1		1.2		1.3		1.4		1.5	
	All (SE)	*P* value	All (SE)	*P* value	All (SE)	*P* value	All (SE)	*P* value	All (SE)	*P* value
DOCTOREXPERT	−0.181 (0.070)	.01	−0.117 (0.085)	.17	0.147 (0.104)	.16	−0.197 (0.044)	<.001	−0.035 (0.191)	.86
GOVTEXPERT	0.674 (0.084)	<.001	0.281 (0.133)	.04	0.798 (0.054)	<.001	0.940 (0.561)	.09	0.612 (0.011)	<.001
POSITIVE	N/A[d]	N/A	0.267 (0.211)	.21	N/A	N/A	N/A	N/A	N/A	N/A
NEGATIVE	N/A	N/A	N/A	N/A	0.247 (0.057)	<.001	N/A	N/A	N/A	N/A
NEUTRAL	N/A	N/A	N/A	N/A	N/A	N/A	−0.723 (0.235)	<.001	N/A	N/A
MIXED	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.566 (0.266)	.03
DOCTOREXPERT × GOVTEXPERT	−0.631 (0.281)	.03	−0.667 (0.288)	.02	−0.873 (0.230)	<.001	−0.693 (0.353)	.05	−0.819 (0.356)	.02
DOCTOREXPERT × POSITIVE	N/A	N/A	0.286 (0.424)	.50	N/A	N/A	N/A	N/A	N/A	N/A
DOCTOREXPERT × NEGATIVE	N/A[d]	N/A	N/A	N/A	−0.675 (0.148)	<.001	N/A	N/A	N/A	N/A
DOCTOREXPERT × NEUTRAL	N/A	N/A	N/A	N/A	N/A	N/A	0.344 (0.419)	.41	N/A	N/A
DOCTOREXPERT × MIXED	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	−0.225 (0.474)	.64
GOVTEXPERT × POSITIVE	N/A	N/A	0.938 (0.084)	<.001	N/A	N/A	N/A	N/A	N/A	N/A
GOVTEXPERT × NEGATIVE	N/A	N/A	N/A	N/A	−0.706 (0.356)	.05	N/A	N/A	N/A	N/A
GOVTEXPERT × NEUTRAL	N/A	N/A	N/A	N/A	N/A	N/A	−1.080 (1.162)	.35	N/A	N/A
GOVTEXPERT × MIXED	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.194 (0.831)	.82

[a]A set of control variables, including dummy variables for months and days, are included in the model.

[b]1.1: number of observations=564, log pseudolikelihood=−728.06, Akaike information criterion=1460.11; 1.2: number of observations=437, log pseudolikelihood=−536.41, Akaike information criterion=1076.82; 1.3: number of observations=437, log pseudolikelihood=−539.31, Akaike information criterion=1082.62; 1.4: number of observations=437, log pseudolikelihood=−533.23, Akaike information criterion=1070.47; 1.5: number of observations=437, log pseudolikelihood=−538.87, Akaike information criterion=1081.73.

[c]DV: dependent variable.

[d]N/A: not applicable.

Second, the GOVTEXPERT variable shows a positive coefficient and moderate to solid statistical significance across specifications, indicating that experts affiliated with the government received more airtime. Third, sentiments generally show positive coefficients compared with neutral sentiments. The positive (0.267; *P*=.21), negative (0.247; *P*<.001), and mixed (0.566; *P*=.03) sentiments are *positively associated* with airtime. However, neutral sentiment (−0.723; *P*<.001) is negatively associated with airtime. However, the coefficient for positive sentiment is not statistically significant, suggesting that a positive opinion may depend on other contextual factors.

The interaction term between DOCTOREXPERT and GOVTEXPERT is negative and statistically significant (−0.631; *P*=.03 for base model specification) across specifications, indicating that the 2 operationalized credibility variables amplify one another. Specifically, government-affiliated experts with a doctorate received less airtime compared with nonexperts.

More interestingly, we found that extreme valence sentiments, such as positive and negative sentiments, interact with the credibility variables for DOCTOREXPERT and GOVTEXPERT affiliation differently. For positive sentiments, there was a statistically significant interaction with GOVTEXPERT (0.938; *P*<.001) but not with DOCTOREXPERT (0.286; *P*=.50). For negative sentiments, we observed a significant interaction with both DOCTOREXPERT (−0.675; P<.001) and GOVTEXPERT (−0.706; P=.05), indicating that the relationship between expert affiliation and sentiment influences airtime differently depending on the valence of the sentiment.

These findings indicate that when experts express clear sentiments, it can directly impact the airtime they receive. Specifically, positive sentiments positively moderate the credibility of experts in gaining more airtime, whereas negative sentiments negatively moderate the credibility of experts in receiving less airtime than nonexperts. This suggests that the audience may be more interested in hearing positive news from authoritative sources and less interested in hearing negative news.

However, we found no statistical significance for neutral and mixed sentiments. This may suggest that regardless of credibility, neutral sentiments do not directly impact the amount of airtime received. One plausible explanation is that neutral sentiments may be perceived as uninteresting, and mixed sentiments may be perceived as confusing, resulting in less airtime dedicated to these sentiments.

Robustness Checks

We checked the robustness of the Tobit regression results across the base models and with different interactions between the sentiment and experts' relevant variables. The results remained relatively similar, with minor changes to the values of the coefficients.

We checked whether the results were affected by the software or procedure for coding the sentiment values. We acknowledge that our choice of AWSC to conduct sentiment analyses is based on a specific set of assumptions around the model. Nevertheless, we checked with Empath, VADER (Valence Aware Dictionary and Sentiment Reasoner), LIWC techniques, and AWSC tools for coding sentiment values. Empath and VADER are available as Python packages that rely on a lexicon-based approach to sentiment analysis using predefined dictionaries of words and phrases with assigned sentiment scores. VADER can handle negations and context-dependent sentiment classification and detect the intensity of emotions and sentiments; however, its accuracy may be lower than that of other methods. Our regression results, with the coded variables from Empath, VADER, and LIWC, remain similar, with some variations in the statistical significance. Broadly, we can say that LIWC and Empath are inconsistent because their sentiment methodology counts, but it does not adjust for the context, whereas results from AWSC and VADER both show consistent results.

We conducted a regression analysis using 3 sentiment polarity scores, whose values were predicted on a scale of 0 to 1. Unfortunately, the variational inflation factor on a simplified specification model consisting of all 3 sentiment polarities shows a variance inflation factor score above 2.5, a general index threshold for indicating multicollinearity, thereby limiting our ability to use the 3 sentiment dummies in the same models. We then merged the 3 categories into one variable, in which case the results came to be positive and showed significant interaction with DOCTOREXPERT and not significant with GOVTEXPERT variables. However, this does not indicate the positive or negative sentiment effects expressed in the comments.

Additional Analyses

We conducted additional analyses to further delve into the details related to using expert sources in discussions about hydroxychloroquine as a therapeutic for COVID-19 on cable news networks and to understand any potential differences in using expert sources and messages among the networks. We found that nonexpert sources were used more frequently than expert sources to discuss the therapeutic validity of hydroxychloroquine and that the amount of airtime allocated to expert sources decreased over time (Figure 1).

We also found that a small number of experts accounted for a significant proportion of the total airtime allotted to experts on each network. The top 5 voices represented >40% of airtime on CNN and MSNBC and slightly >50% on FOX News (Table 4).

Figure 1. Comparison of experts versus nonexperts airtime across cable networks.

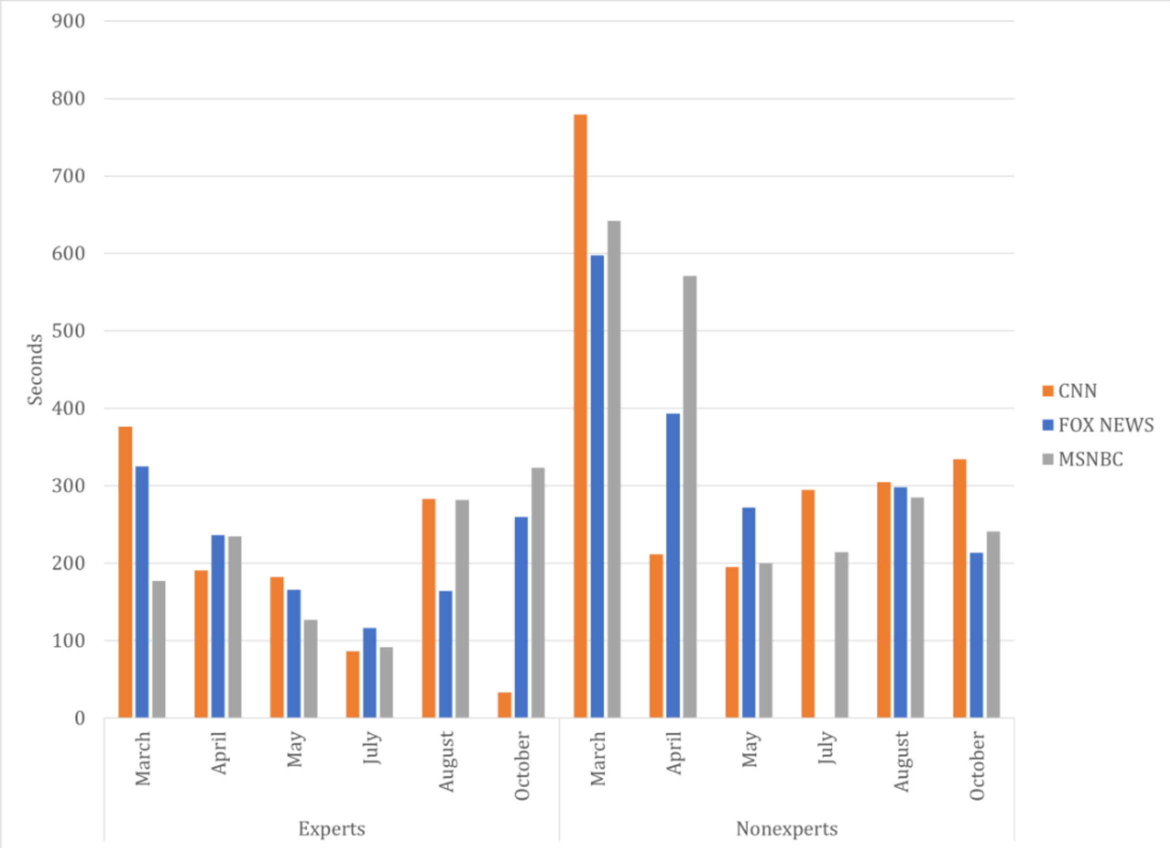


Table 4. Share of total airtime by featured experts across cable networks.

Name	Share (%)
CNN[a]	
Sanjay Gupta	18.49
Anthony Fauci	9.57
Peter Hotez	4.89
Celine Gounder	4.73
Jonathan Reiner	4.51
FOX News	
Mehmet Oz	23.02
Deborah Birx	11.26
Marc Siegel	6.89
Nicole Saphier	5.51
Stephen Hahn	5.2
MSNBC	
Kavita Patel	15.78
Amesh Adalja	8.43
Natalie Azar	7.52
Vin Gupta	6.64
Ezekiel Emanuel	6

[a]CNN: Cable News Network.

The top 5 experts represented >40% of the airtime allotted to experts on CNN and MSNBC. On FOX News, the leading 5 experts accounted for slightly >50% of the airtime. An analysis of the top 3 experts shows some variability among the networks. The CNN medical correspondent Dr Sanjay Gupta received the most airtime, followed by Dr Fauci and Dr Peter Hotez, an expert in infectious diseases and vaccine development and dean of the National School of Tropical Medicine at Baylor College of Medicine.

On FOX News, Dr Mehmet Oz, a celebrity, received the most airtime, followed by Dr Deborah Birx and Dr Marc Siegel, the FOX News medical correspondent. MSNBC does not have a dedicated medical correspondent. On MSNBC, Dr Kavita Patel, a former Federal Administration Official associated with the Center for Health Policy at the Brookings Institution, received the most airtime, followed by Dr Amesh Adalja, a senior scholar at the Johns Hopkins Center for Health Security, and Dr Natalie Azar, a National Broadcasting Company (NBC) News Medical Contributor and a professor at New York University Langone School of Medicine. Furthermore, our analysis revealed that the sentiment in the broadcast toward hydroxychloroquine was marked by a heated exchange of opinions and charged sentiments in contrast to a measured and thoughtful discussion. Both experts and nonexperts exhibited a range of sentiments, with positive, negative, and mixed sentiments occurring more frequently than neutral sentiments (Figures 2 and 3). Although some experts expressed negative views on the effectiveness of hydroxychloroquine and the negative consequences surrounding its use, others expressed positive sentiments and highlighted the ongoing studies on its potential use in the treatment and prevention of COVID-19. However, the experts emphasized the importance of studying the drug and informing the public.

Figure 2. Comparison of sentiments across cable networks and top experts in sampled broadcasts. CNN: Cable News Network.

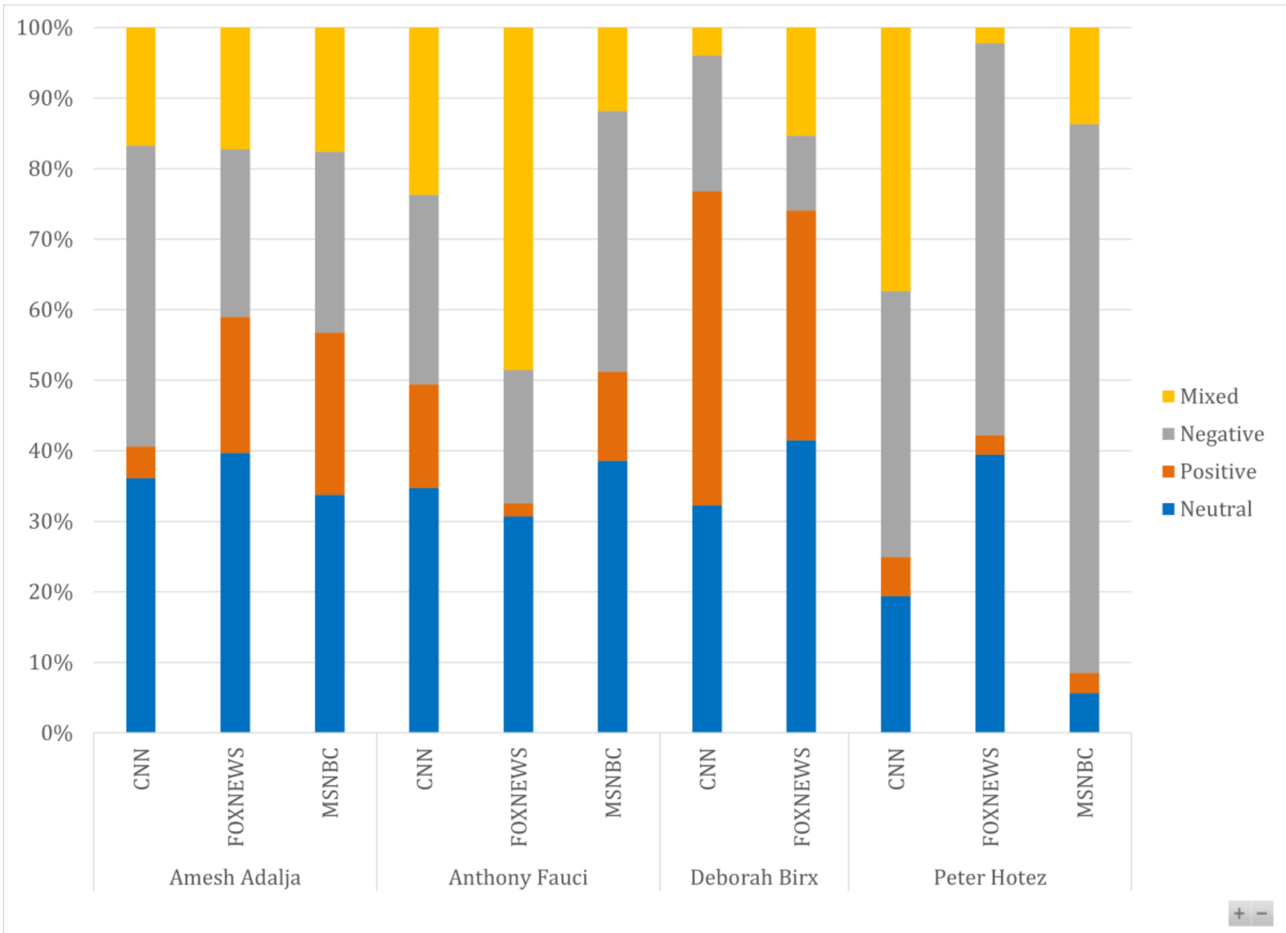
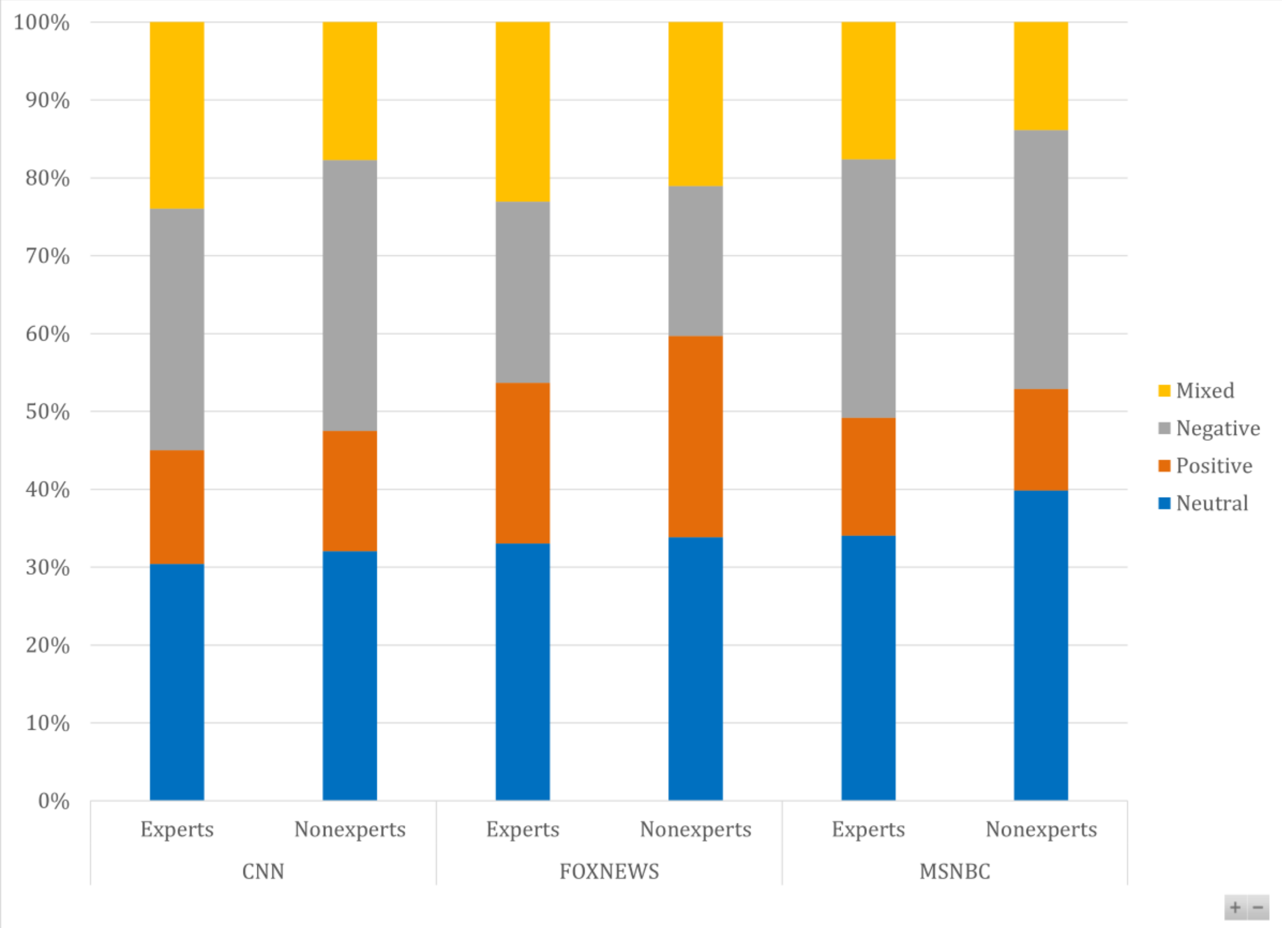


Figure 3. Comparison of sentiments across cable networks in sampled broadcasts. CNN: Cable News Network.



Discussion

Explanation of Key Findings

Before discussing the implications of the findings of this study, we highlight the key findings. This study informs a substantial issue regarding how information is generated and disseminated to the public through cable television networks. Experts with advanced degrees, such as MDs and PhDs, are often seen as highly credible sources of information on public health issues. However, the findings suggest that they would receive less airtime on cable television. Academic expertise credibility may not be sufficient to be perceived as a persuasive dimension by the audience. Alternatively, these experts may not need much airtime except for what is taken to ratify the credibility. Government officials receive more time than nonexperts, even accounting for the positive or negative sentiments expressed during the broadcast. This could be due to their perceived authority from their positional power, confidence on camera, or experience with media appearances.

It is important to note that positive sentiments are generally associated with more airtime, whereas neutral or "boring" sentiments are negatively associated with airtime. Interestingly, displaying negative sentiments alone does not necessarily lead to less airtime. Instead, when doctors display negative sentiment, it impacts their airtime negatively. Mixed sentiments, in contrast, seem to be positively associated with airtime.

Thus, the findings raise concerns about accurate and comprehensive health information disseminated in cable television broadcasts. Given that academic experts do not get much airtime compared with government officials, does the public get a complete perspective? Should the media consider a balanced representation of credible sources to ensure the public gets accurate and comprehensive health information?

Implications

The findings of this study have several practice and policy implications. First, this study draws insights into the influence of cable television on health communications and specifically highlights that broadcasters must be careful about the information they disseminate. Government officials get more airtime than academic experts, which may be because of their optimistic or biased statements. Academic experts who can provide more scientific facts do not get much airtime. Experts' choices and preparations must be made carefully to instill credibility [12], a lack of which polarizes and politicizes health communications, which was evident during the pandemic [23,24].

This study sheds light on the specific context of information spread in cable broadcasting and its comparison with prior research on the spread of health information in social media [27]. Emotional-moral words correlate with a higher likelihood of being retweeted, and emotions are essential in making content contagious on social media [27-30]. Studies have shown how moral emotions shaped information spread on Twitter and other media about hydroxychloroquine as a solution to COVID-19 [29,30].

The findings have implications to highlight some elements of cable television discussions around hydroxychloroquine that differ from social media in 2 ways. First, a prior study has also identified and characterized scientific authority–related discussions about hydroxychloroquine in Twitter, alluding to medical experts' credibility aspect of sources [27]. The findings of this study contrast with the earlier claim highlighting that there is less airtime for experts or authority figures in cable television broadcasting compared with the prominence of authority figures used on Twitter—a meaningful comparison as both authority figures and moral emotions shaped information spread on Twitter and other media about hydroxychloroquine as a solution to COVID-19 [29,30]. The contrast in the use of the expert's credibility suggests that the medium of the discourse (cable vs Twitter) influences what type of content is spread or prominent; specifically, in the context of embedding the content with sentiments, both media have very different orientations for dissemination. These findings add further insights into how source credibility and health communications across different mediums differ in their shape, context, and ways of propagation.

Content broadcasts for scientific topics differ qualitatively from the information diffused through social media. The content broadcast on television is the product of a collaborative activity that includes scientists, journalists, editors, experts, and the public. The centerpiece of the collaboration is the interaction between journalists and their sources, often subject matter experts in their domain. The information provided by these experts helps shape and illuminate the story [36]. In the routine practice of science and medical journalism, journalists generally rely on material published in well-respected peer-reviewed journals, administrators of respected institutions, researchers, and sources that have previously spoken to the press [37]. The findings of this study raise a substantial challenge to public health communicators and specialists who are frequently advised to build working relationships with journalists [38]. However, different cable news networks may develop their relationships with different sources, and a few sources dominated the discussion about hydroxychloroquine.

Moreover, despite the available scientific data, or the lack of data in the initial stages, the expert sources on the different networks expressed different sentiments regarding its efficacy. Because viewers generally do not watch all 3 cable news networks, the information they received was dictated by the network they watch. For instance, FOX News viewers' perspectives on the appropriateness of taking hydroxychloroquine differed sharply from CNN and MSNBC viewers. Particularly in the peak periods, when hydroxychloroquine was most frequently mentioned on cable news networks, the focus of the stories was not specifically on the merits or demerits of the drug. The expert opinions were expressed within the context of a broader newsworthy event. The experts also shared insights and discussed with nonexperts on the same broadcasts. The opinions of journalists, politicians, and others were often as prevalent as those of medical and scientific experts.

The mistaken suggestion that hydroxychloroquine could be used to treat COVID-19 had a real-world impact. Prescriptions written for the drug soared, resulting in thousands of people taking ineffective and potentially harmful treatment, which put pressure on the drug supply for those who needed it. The dynamics of the discussion about hydroxychloroquine are evidence of the development of filter bubbles and the polarization of critical public health information on cable news networks. They have an impact on decision-making and health outcomes. The divergences in outlook are not quickly addressed by typical health communication bromides, and public health officials should deliver consistent information in an appropriate format through channels of communication to which people attend. It requires different strategies to mitigate conflicting sentiments on complex public health issues.

Given the value of academic expertise to ensure that the public can access accurate and comprehensive health information, source credibility needs to be shown to the public in a way that they can assess and trust. This is a "trust-in-media" issue that goes beyond only viewership to be responsible for informing the public on significant health issues. We recommend that the channels indicate experts' credibility during broadcasts. This will help the audience to reflect on the comments appropriately.

Limitations and Directions for Further Research

This study has a few limitations that future studies may be able to address. First, the study's data set focuses on the US viewership of cable news, focusing on the 3 major news networks. Therefore, our findings may not be generalizable to viewers elsewhere. Focusing on other issues around broadcasting may provide more nuanced and enriched explanations for the effect on the credibility-airtime associations. We did not capture everything in our models, and future studies may explore many such factors. This study was contextualized to the hydroxychloroquine-related discussions during the COVID-19 pandemic. The generalizability of other contexts remains a limitation that can only be explained after similar models have been applied to varied contexts in future studies. Another limitation of this study is that we used the cross-sectional data set to examine the relationships between variables. We believe that with multiple years of data from the same or similar contexts, future research will be able to provide causal inferences.

It could be argued that specific television programs, such as morning news or current affairs roundups, have prior agreements to allocate fixed interval times of airtime to featured guests to adhere to scripted formats. However, it becomes difficult to script and allocate a set time for individual experts regarding controversial topics such as hydroxychloroquine for COVID-19 as a therapeutic. In such cases, it is unlikely that experts are given predetermined amounts of airtime in a live, real-time show. Instead, the allocated airtime may depend more on their accessibility and the quality of the individual experts' opinions [39]. For instance, the ability of experts to explain complicated information may increase allocated airtime. Alternatively, the depth or relevance of experts' opinions may increase allocated airtimes. However, we acknowledge that the relationship between expert credibility and airtime allocation in this context does not necessarily indicate causality. A more robust analysis with other explanatory variables, experiments, or a panel data–oriented study design may be needed to establish causality. Thus, this study is exploratory and focuses on several controversial discussions held on major cable television channels in the United States regarding the role of hydroxychloroquine in the treatment of COVID-19. We used this unique and significant context to explore how the credibility of experts and the credibility of information influence the allocation of airtime in cable television.

Conclusions

This study focused on the message that the credibility of broadcast sources is essential. These findings call for responsible behavior from broadcasters. The perceived credibility of the origin of the information is a critical determinant in guiding viewers' evaluation of whether the information is true or false and consequently, the viewers' opinion on the issue under discussion [40]. As was evident during the pandemic, discussions on the efficacy of hydroxychloroquine as a therapeutic for COVID-19 could potentially mislead the public into believing that there was a cure for COVID-19 that did not exist [25]. A cacophony of voices clamors for attention to any given topic, including politicians, journalists, and government officials. In the case of pandemics and other medical issues, doctors, scientific experts, and public discussions about hydroxychloroquine were no different [26]. Television channels need to be careful about health communications from experts.

Acknowledgments

The authors acknowledge and thank the Internet Archive television News for the publicly available data used in this study. The authors thank the Loyola University Maryland undergraduate work-study program student coders for helping with data preparation. The authors appreciate the funding provided to the coders through the Undergraduate Work-Study program at Loyola University to offer avenues for undergraduate students an opportunity to contribute to research studies. JK expressly acknowledges the Health Administration Research Consortium at the Business School of the University of Colorado Denver for providing a platform for the stimulating discussion insights provided on this topic.

Conflicts of Interest

None declared.

References

1. Understanding the infodemic and misinformation in the fight against COVID-19. Pan American Health Organization. 2020. URL: https://www.paho.org/en/documents/understanding-infodemic-and-misinformation-fight-against-covid-19 [accessed 2023-05-29]
2. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res 2020 Jun 29;22(6):e21820 [FREE Full text] [doi: 10.2196/21820] [Medline: 32589589]
3. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med 2002 Dec 15;113(9):763-765. [doi: 10.1016/s0002-9343(02)01473-0] [Medline: 12517369]
4. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
5. 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]
6. 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]
7. Mavragani A. Infodemiology and infoveillance: scoping review. J Med Internet Res 2020 Apr 28;22(4):e16206 [FREE Full text] [doi: 10.2196/16206] [Medline: 32310818]
8. Hovland CI, Janis IL, Kelley HK. Communication and Persuasion: Psychological Studies of Opinion Change. New Haven, CT, USA: Yale University Press; 1953.
9. Berlo DK, Lemert JB, Mertz RJ. Dimensions for evaluating the acceptability of message sources. Public Opin Q 1969 Jan;33(4):563-576 [FREE Full text]
10. Greenberg BS, Miller GR. The effects of low‐credible sources on message acceptance. Speech Monogr 1966 Jun;33(2):127-136 [FREE Full text] [doi: 10.1080/03637756609375488]
11. Golan GJ. New perspectives on media credibility research. Am Behav Sci 2010 Sep;54(1):3-7 [FREE Full text] [doi: 10.1177/0002764210376307]
12. Boyce T. Journalism and expertise. Journal Stud 2006;7(6):889-906 [FREE Full text] [doi: 10.1080/14616700600980652]
13. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 09;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
14. Calleja N, AbdAllah A, Abad N, Ahmed N, Albarracin D, Altieri E, et al. A public health research agenda for managing infodemics: methods and results of the first WHO infodemiology conference. JMIR Infodemiology 2021 Sep 15;1(1):e30979 [FREE Full text] [doi: 10.2196/30979] [Medline: 34604708]

15. 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]
16. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
17. Banerjee D, Meena KS. COVID-19 as an "infodemic" in public health: critical role of the social media. Front Public Health 2021 Mar 18;9:610623 [FREE Full text] [doi: 10.3389/fpubh.2021.610623] [Medline: 33816415]
18. Chandrasekaran R, Mehta V, Valkunde T, Moustakas E. Topics, trends, and sentiments of tweets about the COVID-19 pandemic: temporal infoveillance study. J Med Internet Res 2020 Oct 23;22(10):e22624 [FREE Full text] [doi: 10.2196/22624] [Medline: 33006937]
19. Samuels E, Kelly M. How false hope spread about hydroxychloroquine to treat COVID-19-and the consequences that followed. The Washington Post. 2020 Apr 13. URL: https://www.washingtonpost.com/politics/2020/04/13/how-false-hope-spread-about-hydroxychloroquine-its-consequences/ [accessed 2023-05-30]
20. Bull-Otterson L, Gray EB, Budnitz DS, Strosnider HM, Schieber LZ, Courtney J, et al. Hydroxychloroquine and chloroquine prescribing patterns by provider specialty following initial reports of potential benefit for COVID-19 treatment - United States, January-June 2020. MMWR Morb Mortal Wkly Rep 2020 Sep 04;69(35):1210-1215 [FREE Full text] [doi: 10.15585/mmwr.mm6935a4] [Medline: 32881845]
21. Catillon M, Majumder MS, Manzi SF, Mandl KD. News coverage and drug shortages during the COVID-19 pandemic. medRxiv. Preprint posted on online on October 14, 2020. [doi: 10.1101/2020.10.12.20211656]
22. Hydroxychloroquine does not benefit adults hospitalized with COVID-19. National Institute of Health. 2020 Nov 09. URL: https://www.nih.gov/news-events/news-releases/hydroxychloroquine-does-not-benefit-adults-hospitalized-covid-19 [accessed 2023-05-30]
23. Hart PS, Chinn S, Soroka S. Politicization and polarization in COVID-19 news coverage. Sci Commun 2020 Oct;42(5):679-697 [FREE Full text] [doi: 10.1177/1075547020950735]
24. Jurkowitz M, Mitchell A. Cable tv and COVID-19: how Americans perceive the outbreak and view media coverage differ by main news source. Pew Research Center. 2020 Apr 01. URL: https://www.pewresearch.org/journalism/2020/04/01/cable-tv-and-covid-19-how-americans-perceive-the-outbreak-and-view-media-coverage-differ-by-main-news-source/ [accessed 2023-05-30]
25. McAweeney E. Who benefits from health misinformation? Medium. 2020 Mar 30. URL: https://points.datasociety.net/who-benefits-from-health-misinformation-8d094804058d [accessed 2023-05-30]
26. Blevins JL, Edgerton E, Jason DP, Lee JJ. Shouting into the wind: medical science versus “B.S.” in the Twitter maelstrom of politics and misinformation about hydroxychloroquine. Soc Media Soc 2021 Apr;7(2):1-14 [FREE Full text] [doi: 10.1177/20563051211024977]
27. 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 Jan;8(1):1-15 [FREE Full text] [doi: 10.1177/20539517211013843]
28. Brady WJ, Wills JA, Jost JT, Tucker JA, Van Bavel JJ. Emotion shapes the diffusion of moralized content in social networks. Proc Natl Acad Sci U S A 2017 Jul 11;114(28):7313-7318 [FREE Full text] [doi: 10.1073/pnas.1618923114] [Medline: 28652356]
29. 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]
30. Solovev K, Pröllochs N. Moral emotions shape the virality of COVID-19 misinformation on social media. In: Proceedings of the ACM Web Conference 2022. 2022 Presented at: WWW '22; April 25-29, 2022; Virtual Event p. 3706-3717 URL: https://dl.acm.org/doi/10.1145/3485447.3512266 [doi: 10.1145/3485447.3512266]
31. The Stanford Cable TV News Analyzer. URL: https://tvnews.stanford.edu/ [accessed 2023-05-04]
32. GitHub. URL: https://github.com [accessed 2023-05-10]
33. Martinez LS, Hughes S, Walsh-Buhi ER, Tsou MH. "Okay, we get it. You vape": an analysis of geocoded content, context, and sentiment regarding e-cigarettes on Twitter. J Health Commun 2018;23(6):550-562. [doi: 10.1080/10810730.2018.1493057] [Medline: 29979920]
34. Merrill JE, Ward RM, Riordan BC. Posting post-blackout: a qualitative examination of the positive and negative valence of tweets posted after "blackout" drinking. J Health Commun 2020 Feb 01;25(2):150-158 [FREE Full text] [doi: 10.1080/10810730.2020.1719242] [Medline: 31986999]
35. Kim E, Hou J, Han JY, Himelboim I. Predicting retweeting behavior on breast cancer social networks: network and content characteristics. J Health Commun 2016;21(4):479-486. [doi: 10.1080/10810730.2015.1103326] [Medline: 27007166]
36. Conrad P. Uses of expertise: sources, quotes, and voice in the reporting of genetics in the news. Public Underst Sci 1999 Oct;8(4):285-302 [FREE Full text] [doi: 10.1088/0963-6625/8/4/302]
37. Nelkin D. Selling Science: How the Press Covers Science and Technology. 2nd edition. New York, NY, USA: W. H. Freeman & Company; 1995.

38. Donovan J. Concrete recommendations for cutting through misinformation during the COVID-19 pandemic. Am J Public Health 2020 Oct;110(S3):S286-S287. [doi: 10.2105/AJPH.2020.305922] [Medline: 33001730]
39. Hertzum M. How do journalists seek information from sources? A systematic review. Inf Process Manag 2022 Nov;59(6):103087 [FREE Full text] [doi: 10.1016/j.ipm.2022.103087]
40. Swire B, Berinsky AJ, Lewandowsky S, Ecker UK. Processing political misinformation: comprehending the Trump phenomenon. R Soc Open Sci 2017 Mar;4(3):160802 [FREE Full text] [doi: 10.1098/rsos.160802] [Medline: 28405366]

Abbreviations

AWSC: Amazon Web Service Comprehend
CNN: Cable News Network
LIWC: Linguistic Inquiry and Word Count
VADER: Valence Aware Dictionary and Sentiment Reasoner

Edited by T Mackey; submitted 28.12.22; peer-reviewed by SH Kim, M Haupt; comments to author 01.02.23; revised version received 21.03.23; accepted 15.05.23; published 27.06.23.

Please cite as:
Yim D, Khuntia J, King E, Treskon M, Galiatsatos P
Expert Credibility and Sentiment in Infodemiology of Hydroxychloroquine's Efficacy on Cable News Programs: Empirical Analysis
JMIR Infodemiology 2023;3:e45392
URL: https://infodemiology.jmir.org/2023/1/e45392
doi:10.2196/45392
PMID:37204334

Original Paper

YouTube Videos on Nutrition and Dental Caries: Content Analysis

Memphis Long[1], BSc, RD; Laura E Forbes[2], PhD, RD; Petros Papagerakis[3], PhD; Jessica R L Lieffers[1], PhD, RD

[1]College of Pharmacy and Nutrition, University of Saskatchewan, Saskatoon, SK, Canada

[2]Department of Family Relations & Applied Nutrition, University of Guelph, Guelph, ON, Canada

[3]Faculty of Dental Medicine, Laval University, Quebec City, QC, Canada

Corresponding Author:
Jessica R L Lieffers, PhD, RD
College of Pharmacy and Nutrition
University of Saskatchewan
105 Wiggins Road
Saskatoon, SK, S7N 5E5
Canada
Phone: 1 3069668594
Email: jessica.lieffers@usask.ca

Abstract

Background: Dental caries is the most common health condition worldwide, and nutrition and dental caries have a strong interconnected relationship. Foods and eating behaviors can be both harmful (eg, sugar) and healthful (eg, meal spacing) for dental caries. YouTube is a popular source for the public to access information. To date, there is no information available on the nutrition and dental caries content of easily accessible YouTube videos.

Objective: This study aimed to analyze the content of YouTube videos on nutrition and dental caries.

Methods: In total, 6 YouTube searches were conducted using keywords related to nutrition and dental caries. The first 20 videos were selected from each search. Video content was scored (17 possible points; higher scores were associated with more topics covered) by 2 individuals based on the inclusion of information regarding various foods and eating behaviors that impact dental caries risk. For each video, information on video characteristics (ie, view count, length, number of likes, number of dislikes, and video age) was captured. Videos were divided into 2 groups by view rate (views/day); differences in scores and types of nutrition messages between groups were determined using nonparametric statistics.

Results: In total, 42 videos were included. Most videos were posted by or featured oral health professionals (24/42, 57%). The mean score was 4.9 (SD 3.4) out of 17 points. Videos with >30 views/day (high view rate; 20/42, 48% videos) had a trend toward a lower score (mean 4.0, SD 3.7) than videos with ≤30 views/day (low view rate; 22/42, 52%; mean 5.8, SD 3.0; *P*=.06), but this result was not statistically significant. Sugar was the most consistently mentioned topic in the videos (31/42, 74%). No other topics were mentioned in more than 50% of videos. Low–view rate videos were more likely to mention messaging on acidic foods and beverages (*P*=.04), water (*P*=.09), and frequency of sugar intake (*P*=.047) than high–view rate videos.

Conclusions: Overall, the analyzed videos had low scores for nutritional and dental caries content. This study provides insights into the messaging available on nutrition and dental caries for the public and guidance on how to make improvements in this area.

(JMIR Infodemiol 2023;3:e40003) doi:10.2196/40003

KEYWORDS

dental caries; diet; nutrition; YouTube; internet; consumer health information

Introduction

Dental caries (or tooth decay) occurs when dietary fermentable carbohydrates (eg, simple sugars) are metabolized by bacteria in the mouth (eg, *Streptococcus mutans*) to produce a highly acidic environment that can degrade tooth structures (eg, enamel) [1]. Dental caries is the most common disease worldwide. According to the 2019 Global Burden of Disease Study, 2.0 billion people worldwide had untreated dental caries in permanent teeth, and 0.5 billion children aged 0 to 14 years had untreated caries in their deciduous teeth [2]. Untreated dental caries is more common than cardiovascular diseases, diabetes, cancers, mental disorders, and chronic respiratory diseases worldwide [3]. The World Health Organization recommends that oral health care become part of universal health care and that there is an increased emphasis on the prevention of oral diseases [4].

Although dental caries can be attributed to numerous interrelated factors described elsewhere [1,5], many foods and dietary habits have been identified as important risk factors. Sugar, which is a fermentable carbohydrate, is a major driving force for the development of dental caries [6,7]. The World Health Organization strongly recommends that children and adults consume <10% of their calories as free sugars because of the association between this dietary component and dental caries. They also conditionally recommend reducing free sugar intake to <5% of the total energy consumed because of an additional protective effect of lower intakes on dental caries risk [8]. Sugary drinks (eg, soft drinks and juice) have also been linked to dental caries, and limiting the consumption of these drinks has been recommended [1,9-14]. Furthermore, foods that are both sugary and starchy (eg, cakes and donuts) are thought to be more cariogenic than foods containing sugar alone; this outcome is likely owing to the sugar being retained on teeth for longer periods due to the stickiness of the starch [15,16]. In addition, more frequent consumption of sugar (including consumption between meals) is associated with an increased risk of dental caries than less frequent consumption of sugar [1,17]. Poor-quality diets can also cause nutrient deficiencies (eg, vitamin D and calcium) that can cause issues with tooth formation and mineralization, making them susceptible to caries development [7,18].

Dietary factors can also prevent the development of dental caries. Foods that are thought to benefit teeth are whole grains, fruits, vegetables, high-quality proteins, and dairy products such as cheese and milk; spacing out meals is also thought to be beneficial [1,14,19,20]. Furthermore, xylitol is thought to be beneficial for dental caries prevention through different mechanisms, including replacing sugar intake in the diet; stimulation of saliva; and inhibition of the growth of cariogenic bacteria [21]. Although diet is an important determinant of dental caries, many studies have reported that health professionals experience substantial barriers in providing diet counseling for this issue, and often, this service may not be provided [22].

Previous research has found that it is common for the public to access web-based sources (eg, internet and social media) to obtain health information [23-25]. For example, Shahab et al [23] found that in the United States, 78.2% of individuals who had ever used the internet had used this source in the previous year to access health information. The Tracking Nutrition Trends survey conducted by the Canadian Foundation for Dietetic Research in 2015 found that 49% of survey respondents from Canada used the internet, social media, and blogs to obtain information on food and nutrition. They also found that this activity was more common among the younger respondents [26]. There are numerous reasons why members of the public may seek health information from web-based sources, including to obtain more knowledge on a health condition, supplement information obtained from health providers, explore embarrassing topics, and obtain support from others [27,28]. However, despite the popularity of web-based information, there are concerns with accessing these sources, including the presence of misinformation and potential harms of making decisions based on unsubstantiated claims [27,29].

YouTube is one source of web-based information that deserves attention. It is a video sharing platform founded in 2005 and is the second most highly trafficked website globally, with 34.6 billion visits each month [30]. In 2020, there were 2.3 billion users of YouTube globally, and this has steadily increased over the last several years [31]. The content uploaded to YouTube is extensive. For example, for every minute as of February 2020, a total of 500 hours of video content was uploaded [32]. There are many reasons why people use YouTube, including to learn new things, problem-solving, entertainment, self-care (eg, destress and relaxation), and to improve skills [33]. In a 2019 report, approximately 70% of YouTube users reported that this platform is the first website they go to when trying to learn [34]. YouTube can also be easily accessed through different devices, including computers, tablets, and mobile phones. Substantial interest has been generated around the use of YouTube for health-related purposes. To date, a few studies have shown that YouTube videos can be beneficial for improving health-related knowledge, attitudes, and behaviors [28]. However, despite the popularity of this platform and the interest in its use for health-related purposes, the content of YouTube videos is not reviewed to ensure accuracy and comprehensiveness.

To date, several studies have been conducted on the content of health information available on YouTube. These studies have been summarized in different review articles [28,35-38]. These articles have reported that, in general, videos do not comprehensively cover various health topics and that the content quality of videos varies widely, with many studies reporting a high prevalence of poor-quality videos or nonuseful videos and a low prevalence of good-quality videos. However, some high-quality videos are available in some topic areas [37,38]. In addition, many studies have found either no relationship between video quality and engagement (eg, views and likes) or a negative relationship (ie, as quality decreases, engagement increases) [38]. These articles have also found that videos tend to be of higher quality when they feature health professionals (eg, physicians) or health organizations [37]. Although numerous studies have assessed the content of various types of health-related YouTube videos, to our knowledge, no studies have examined the content of YouTube videos related to nutrition and dental caries. Owing to the high prevalence of dental caries worldwide, the strong relationship that diet has with dental caries, the popularity of YouTube, and the barriers experienced by health professionals providing support on this issue, information on this topic is needed.

The purpose of this study was to analyze the content of YouTube videos regarding dental caries and nutrition that are easily accessible using default search settings. We were also interested in examining nutrition messaging according to creator type and engagement.

Methods

Ethical Considerations

This study was exempt from ethical review from the University of Saskatchewan Behavioural Ethics Office as per Article 2.2 of the Tri-Council Policy Statement (TCPS): Ethical Conduct for Research Involving Humans—TCPS 2 (2018) [39].

Video Selection

Our strategy was to search for YouTube videos that would be most accessible to the public searching for educational content regarding nutrition and dental caries. Google Keyword Planner [40] was used to select 2 dental caries–related keywords and 3 nutrition-related keywords. For dental caries, the 2 top keywords associated with this concern were *tooth decay* and *dental cavities*. The top 3 keywords for nutrition were *nutrition*, *diet*, and *food*. This resulted in a total of 6 searches: *tooth decay and nutrition*, *tooth decay and diet*, *tooth decay and food*, *dental cavities and nutrition*, *dental cavities and diet*, and *dental cavities and food*. Videos were eligible for inclusion if they were in English, <20 minutes in duration, and included information about nutrition and dental caries. We chose <20 minutes in duration as the inclusion criteria because similar time frames have been used in previous related work [41,42]. A 2018 study also found that 90% of respondents preferred instructional and informational videos to be <20 minutes [43].

The YouTube searches were conducted on May 17, 2021, using the default settings on YouTube to best replicate the search strategy used by the public. The searches were conducted by ML, a female undergraduate nutrition student, using Google Chrome's Incognito mode to prevent bias when conducting the searches. ML opened a new incognito window to complete each search. Each of the 6 YouTube searches were completed in a sequence, and the first 20 videos were recorded from each search. The first 20 videos were chosen because similar numbers have been used in other related studies [44-46]. We also chose the first 20 videos because previous work has found that most people who use the internet do not look past the first search results page [47]. For each video, the title, publisher, country of origin, total number of views, date posted, URL, length in minutes, whether the video was an animation, and the number of likes and dislikes were recorded by ML in Microsoft Excel 365. Transcripts for each video were also downloaded from the YouTube website.

Video Scoring System

Owing to the numerous dietary factors that can affect the risk of dental caries, a scoring system was developed to be used for this study. This type of approach (scoring system or presence or absence of content in videos) has been used in other related YouTube content analysis studies [48-52]. The scoring system had 17 possible points, with higher scores indicating that more topic areas were covered. Table 1 lists each of the topic areas. The inclusion of misinformation in videos was not considered in the scoring tool. This approach has also been used elsewhere [51].

Table 1. Scoring system to assess messaging in the included YouTube videos (total possible score: 17 points).

Message assessed in each video	Score, n
Dental caries mechanism	1
Factors that increase the risk of dental caries (or poor oral health)	
Acidic foods and beverages	1
Any mention of sugar	1
Sugary drinks (eg, soda, fruit juices, energy drinks, and sweetened coffee and sweetened tea)	1
Sticky foods (eg, dried fruit)	1
Frequency of sugar intake (eg, frequent and prolonged intake of simple sugars or limiting snacking or eating sugary foods with meals or eating sticky foods alone)	1
Candy (either in general or specific types of candy)	1
Snack foods high in sugar and starch (eg, cookies, cakes, and pastries)	1
Factors that reduce the risk of dental caries (or promote good oral health)	
Chewing sugar-free gum or eating sugar-free candy or xylitol	1
Vegetables and fruit (including specific vegetables and fruits)	1
Protein from high-quality sources (eg, meats, nuts, seeds, and legumes)	1
Whole grains	1
Water	1
Dairy products (both in general or mentioning specific products)	1
Drink beverages with a straw	1
Brush teeth after meals or brush teeth at least 2 times/day	1
Mention food guide or food label reading	1

This scoring system was developed partially based on the Academy of Nutrition and Dietetics' most recently published position paper on Oral Health and Nutrition [1], which identifies dietary patterns and eating behaviors associated with an increased and decreased risk of dental caries. In this position paper, dietary patterns and behaviors that were identified as causing an increased risk of dental caries included sugar intake, sugary beverage intake, candy intake, starchy and sugary food intake, sticky food intake, and frequency of consuming sugary foods and beverages. For dietary patterns and behaviors associated with decreased risk, the position paper included sugar-free gum and candies, vegetables and fruit, high-quality protein foods, and whole-grain foods.

Dietary factors were also added to the scoring tool based on content from other high-quality evidence-based sources related to nutrition and dental caries, including the National Health Service Health Scotland Oral Health and Nutrition Guidance for Professionals June 2012 [14], the 2015 Joint Position Statement on Oral Health and Nutrition from the Dietitians Association of Australia and Dental Health Services Victoria [20], and a Chairside Dietary Assessment tool developed by a dietitian published by the *Journal of the American Dental Association* [53]. These additional protective factors included dairy products, water, and drinking with a straw. Acidic foods and beverages were also included because they have been found to cause dental erosion [54,55]. Caution surrounding acidic foods and beverages is mentioned by both the Scottish and Australian guidelines listed earlier [14,20]. We also examined videos for mention of the food guide or food label reading, as these are common recommendations for general healthy eating and were mentioned in the Scottish guidelines [14], and for information about toothbrushing [14,20]. The recommendation to drink with a straw was found in the chairside assessment tool [53]. In addition, the mechanism of dental caries was also included in the scoring system (ie, including information about how bacteria in the mouth convert sugar into acid and damage tooth structures).

Data Analysis

Videos were scored using the 17-point scoring system independently by 2 individuals (ML and JRLL) using information presented in either text listed in the video or what was said verbally. Discrepancies were discussed until a consensus was reached.

Information on video characteristics (ie, view count, length, number of likes, number of dislikes, video age, viewing rate [views/day; calculated by taking the number of views and dividing by number of days since the video was uploaded] [48], like rate [likes/view; calculated by taking the number of likes and dividing by the number of views], and dislike rate [dislikes/view; calculated by taking the number of dislikes and dividing by the number of views]) were summarized using descriptive statistics (mean, SD, median, and range) determined using Microsoft Excel 365 and SPSS (version 28; IBM Corp).

Each video was categorized into 1 of 4 groups based on the author or presenter featured in the video. The four groups were as follows: (1) oral health professionals (OHPs; eg, dentists, dental hygienists, dental practice groups, dental offices, or commercial content reviewed by OHPs), (2) health professionals who are not OHPs including complementary and alternative medicine providers (eg, microbiologists, chiropractors, and naturopaths), (3) government (videos posted by government sources that could feature any type of health professional), and (4) no health professional credentials or unknown credentials (eg, social media influencers with no credentials and bloggers). Videos were categorized into 2 roughly equal-sized groups based on view rate to examine differences between the most viewed videos compared with less commonly viewed videos.

Inferential statistics were determined using SPSS Statistics (version 28). Fisher exact test was used to determine whether there were significant differences between categorical variables, and the Mann-Whitney *U* test and Kruskal-Wallis test were used to determine whether there were significant differences between continuous variables. The Bonferroni correction was used to correct for multiple comparisons. Spearman correlations were used to examine the relationships between 2 continuous variables. *P* values of <.05 were considered significant.

Results

Search Results

In total, 120 videos from the 6 searches were considered for inclusion; 78 (65%) videos were removed from the analysis because they (1) were duplicate videos (n=65, 54.2%) or (2) did not meet inclusion criteria (n=13, 10.8%; ie, video did not mention anything related to diet and dental caries, n=9, 7.5%; video was >20 min, n=3, 2.5%; and video was not in English, n=1, 0.8%). After these videos were removed, 42 videos were eligible for analysis.

Video Characteristics

Characteristics of the included videos are provided in Table 2. Most videos were posted by or featured OHPs (24/42, 57%), followed by those with no health professional credentials or unknown credentials (10/42, 24%), health professionals who were not OHPs including complementary and alternative medicine providers (6/42, 14%), and the government (2/42, 5%). Notably, 17% (7/42) of the videos were presented as cartoons.

Most videos originated from the United States (25/42, 60%), followed by the United Kingdom (4/42, 10%), India (4/42, 10%), Canada (3/42, 7%), Australia (2/42, 5%), Indonesia (1/42, 2%), and Italy (1/42, 2%). For 5% (2/42) of the videos, we were unable to identify the country of origin. Included videos were on average 4 minutes and 40 seconds in length (SD 3 min and 9 s; range 47 s to 16 min and 35 s) and had been posted for a median of 926.5 (range 164-3917) days.

Table 2. Characteristics of the YouTube videos on nutrition and dental caries included for the analysis (n=42).

	All videos (n=42)	OHPs[a] (n=24)	Health professionals who are not OHPs including complementary and alternative medicine providers (n=6)	Government (n=2)	No health professional credentials or unknown credentials (n=10)
Video age (days)					
Median (range)	926.5 (164-3917)	927 (164-3917)	1694 (663-2160)	580 (356-804)	778 (192-3255)
Mean (SD)	1202 (952)	1257 (1055)	1609 (569)	580 (317)	949 (906)
Length					
Median (range)	3 min and 56 s (47 s to 16 min and 35 s)	3 min and 45 s (1 min and 2 s to 16 min and 35 s)	4 min and 54 s (2 min and 29 s to 9 min and 14 s)	3 min and 4 s (1 min and 51 s to 4 min and 16 s)	3 min and 41 s (47 s to 9 min and 18 s)
Mean (SD)	4 min and 40 s (3 min and 9 s)	4 min and 58 s (3 min and 36 s)	5 min and 8 s (2 min and 26 s)	3 min and 4 s (1 min and 43 s)	4 min and 0 s (2 min and 37 s)
View count					
Median (range)	21,533 (394-3,768,733)	17,741 (394-1,512,464)	718,780 (3080-3,768,733)	11,526 (4564-18,488)	52,383 (1485-1,854,382)
Mean (SD)	292,689 (706,004)	119,821 (320,627)	1,150,249 (1,403,640)	11,526 (9846)	249,270 (569,009)
Viewing rate (views/number of days since posting)					
Median (range)	29.6 (0.3-4863.2)	15.6 (0.3-4863.2)	490.9 (1.4-2252.7)	28.8 (5.7-51.9)	44.8 (6.0-2499.2)
Mean (SD)	375.1 (945.6)	343.8 (1085.5)	687.2 (813.1)	28.8 (32.7)	321.5 (773.2)
Number of likes					
Median (range)	422 (0-33,000)	257 (0-8100)	20,500 (43-30,000)	12 (0-24)	609 (28-33,000)
Mean (SD)	4130 (8832)	1145 (2194)	17,491 (12,905)	12 (17.0)	4100 (10,196)
Like rate					
Median (range)	0.015 (0-0.047)	0.015 (0-0.034)	0.020 (0.0080-0.041)	0.0026 (0-0.0053)	0.018 (0.005-0.047)
Mean (SD)	0.017 (0.012)	0.016 (0.010)	0.024 (0.014)	0.0026 (0.0037)	0.020 (0.013)
Number of dislikes					
Median (range)	14 (0-1200)	12 (0-633)	429 (0-1200)	0 (0-0)	18 (0-1100)
Mean (SD)	159 (325)	80 (180)	553 (532)	0 (0)	146 (339)
Dislike rate					
Median (range)	0.00057 (0-0.0046)	0.00061 (0-0.0046)	0.00060 (0-0.00075)	0 (0-0)	0.00055 (0-0.0012)
Mean (SD)	0.00063 (0.00074)	0.00075 (0.00092)	0.00049 (0.00028)	0 (0)	0.00054 (0.00038)

[a]OHP: oral health professional.

Overall, the 42 included videos had 12,292,954 total views recorded. Videos published by health professionals who are not OHPs (including complementary and alternative medicine providers) had the most views (median 718,780, range 3080-3,768,733 views/video; total views: 6,901,491), followed by videos published by or featuring those with no health professional credentials or unknown credentials (median 52,383, range 1485-1,854,382; total views: 2,492,704), videos that were published by or featured OHPs (median 17,741, range 394-1,512,464; total views: 2,875,707), and videos from the government (median 11,526, range 4564-18,488; total views: 23,052). The mean viewing rate (views/day) was similar between videos posted by those with no health professional credentials or unknown credentials (321.5, SD 773.2; range 6.0-2499.2) and OHPs (343.8, SD 1085.5; range 0.3-4863.2); however, videos by health professionals who are not OHPs (including complementary and alternative medicine providers) had a higher mean view rate (mean 687.2, SD 813.1; range 1.4-2252.7).

Nutrition Messaging

The mean video score for all included videos (42/42, 100%) was 4.9 (SD 3.4; of a maximum possible total of 17), with scores varying from 0 to 13. Table 3 provides a breakdown of the information on scoring by creator type. Videos published by the government and OHPs had a higher mean score (government: 6.5, SD 0.7; OHPs: 5.7, SD 3.8) compared with the scores of videos published by other health professionals (including complementary and alternative medicine providers) or those with no health professional credentials or unknown credentials (other health professionals: 4.0, SD 1.3; no health professional credentials or unknown credentials: 3.4, SD 3.3). However, there was no statistically significant difference in the video scores between the creator type (*P*=.29). Of note, 14% (6/42) of the videos had a score of 0; these videos were published by individuals with no health professional credentials or unknown credentials (n=3, 50%) and OHPs (n=3, 50%).

We investigated the correlation between total video scores and public engagement with videos. No significant Spearman correlations were found between the total video score and total views (−0.114; *P*=.47), view rate (−0.196; *P*=.21), total likes (−0.200; *P*=.20), like rate (−0.202; *P*=.20), total dislikes (−0.156; *P*=.32), and dislike rate (−0.199; *P*=.21).

To further examine nutrition messaging and video engagement, we divided all videos (42/42, 100%) into 2 similar-sized groups based on view rate. The high–view rate category (>30 views/day; 20/42, 48% videos) consisted of 7 videos by OHPs, 7 videos from the no health professional credentials or unknown credentials category, 5 videos by health professionals who are not OHPs (including complementary and alternative medicine providers), and 1 video by the government. The low–view rate category (≤30 views/day; 22/42, 52% videos) consisted of 17 videos by OHPs, 3 videos in the no health professional credentials or unknown credentials category, 1 video by health professionals who are not OHPs (including complementary and alternative medicine providers), and 1 video by the government. Videos with >30 views/day (20/42, 48% videos) had a mean score of 4.0 (SD 3.7) compared with videos with ≤30 views/day (22/42, 52%), which had a mean score of 5.8 (SD 3.0); there was a trend toward the scores being different between groups (*P*=.06; Table 4), but this result was not statistically significant.

Table 5 provides an in-depth breakdown of the different nutrition messages for all videos (42/42, 100% videos). In addition, information on the breakdown of messaging in low–view rate videos (≤30 views/day) versus high–view rate videos (>30 views/day) is also presented.

Table 3. Scores of YouTube videos on nutrition and dental caries by type of creator (n=42).

	All videos (n=42)	OHPs[a] (n=24)	Health professionals who are not OHPs including complementary and alternative medicine providers (n=6)	Government (n=2)	No health professional credentials or unknown credentials (n=10)
Total score (out of 17)					
Median (range)	4.5 (0-13)	5.5 (0-13)	3.5 (3-6)	6.5 (6-7)	3.5 (0-10)
Mean (SD)	4.9 (3.4)	5.7 (3.8)	4.0 (1.3)	6.5 (0.7)	3.4 (3.3)

[a]OHP: oral health professional.

Table 4. Scores of YouTube videos on nutrition and dental caries by view rate (n=42).

	All videos	Low video view rate: ≤30 views/day (range 0.3-29.8; n=22)	High video view rate: >30 views/day (range 35.5-4863.2; n=20)
Total score (out of 17)[a]			
Median (range)	4.5 (0-13)	6 (0-11)	3.5 (0-13)
Mean (SD)	4.9 (3.4)	5.8 (3.0)	4.0 (3.7)

[a]*P*=.06 for the difference between high–view rate videos and low–view rate videos (Mann-Whitney *U* test).

Table 5. Nutrition and dental caries messaging included in the analyzed YouTube videos by view rate (n=42).

	All videos (n=42), n (%)	Low video view rate: ≤30 views/day (range 0.3-29.8; n=22), n (%)	High video view rate: >30 views/day (range 35.5-4863.2; n=20), n (%)	*P* value
Inclusion of specific type of information				
Dental caries mechanism	20 (48)	12 (55)	8 (40)	.37
Any mention of sugar	31 (74)	17 (77)	14 (70)	.73
Sugary drinks	20 (48)	11 (50)	9 (45)	.77
Snack foods high in sugar and starch	17 (40)	10 (45)	7 (35)	.54
Candy	13 (31)	9 (41)	4 (20)	.19
Frequency of sugar intake[a]	13 (31)	10 (45)	3 (15)	.047
Acidic foods and beverages[a]	11 (26)	9 (41)	2 (10)	.04
Sticky foods	10 (24)	6 (27)	4 (20)	.72
Vegetables and fruit	15 (36)	8 (36)	7 (35)	.99
Brush teeth after eating or brush teeth at least 2 times/day	12 (29)	7 (32)	5 (25)	.74
Drink water[b]	12 (29)	9 (41)	3 (15)	.09
Protein from high-quality sources	11 (26)	6 (27)	5 (25)	.99
Dairy products	9 (21)	5 (23)	4 (20)	.99
Sugar-free gum or sugar-free candy or xylitol	6 (14)	5 (23)	1 (5)	.19
Whole grains	5 (12)	3 (14)	2 (10)	.99
Drink beverages with a straw	2 (5)	0 (0)	2 (10)	.22
Mention food guide or food label reading	0 (0)	0 (0)	0 (0)	

[a]*P*<.05 for difference between high–view rate videos and low–view rate videos (Fisher exact test).

[b]*P*<.10 for difference between high–view rate videos and low–view rate videos (Fisher exact test).

In total, 48% (20/42) of the videos contained information on how dental caries are formed. There were no significant differences in the percentage of low–view rate and high–view rate videos that provided this message (*P*=.37).

Overall, 74% (31/42) of the videos contained information about sugar being a cause of dental caries. Of note, guidance on the specific amounts of sugar to consume was not mentioned in any video. Almost half (20/42, 48%) of the videos mentioned that sugary drinks (either in general or specific beverages) were a cause of dental caries. Snack foods high in sugar and starch were mentioned as a risk factor for dental caries in 40% (17/42) of the videos. Candy and sticky foods were mentioned as factors that increase the risk of dental caries in 31% (13/42) and 24% (10/42) of videos, respectively. There were no significant differences in the proportion of low–view rate and high–view rate videos that provided each of the abovementioned messages related to sugary foods and drinks (sugar: *P*=.73; sugary drinks: *P*=.77; snack foods high in sugar and starch: *P*=.54; candy: *P*=.19; and sticky foods: *P*=.72).

Messaging on the frequency of sugar intake was present in 31% (13/42) of the videos. A higher percentage of low–view rate videos contained this message compared with high–view rate videos (10/22, 45% vs 3/20, 15%; *P*=.047).

Acidic foods and beverages being harmful toward oral health were mentioned in 26% (11/42) of videos. This message was more often present in low–view rate videos than in high–view rate videos (9/22, 41% vs 2/20, 10%; *P*=.04).

Compared with harmful foods and behaviors, those that are healthful were mentioned less often. Eating more vegetables and fruit (either in general or specific vegetables or fruits) was the most common healthful behavior mentioned; this message was mentioned in 36% (15/42) of the videos. Eating high-quality protein sources (eg, legumes, pulses, nuts, meat, fish, and seafood) was mentioned in just 26% (11/42) of the videos. In addition, 21% (9/42) of the videos mentioned that dairy products (in general or specific products such as cheese, yogurt, and milk) were beneficial. Whole grains were recommended in 12% (5/42) of the videos. No statistically significant differences were found in the proportion of low–view rate videos versus high–view rate videos that contained each of the healthful food messages listed earlier (vegetables and fruit: *P*=.99; high-quality protein: *P*=.99; dairy products: *P*=.99; and whole grains: *P*=.99).

Drinking water was mentioned as being protective toward dental caries in 29% (12/42) of the videos. Only 2 videos specifically spoke about the consumption of fluoridated water. Drinking water was more often mentioned in low–view rate videos than in high–view rate videos (9/22, 41% vs 3/20, 15% of videos; *P*=.09), but this result was not statistically significant. Brushing teeth at least twice a day or after eating was mentioned in 29% (12/42) of videos. Sugar-free gum or sugar-free candy or xylitol was also discussed in only a few videos (6/42, 14%). Videos rarely recommended drinking beverages with a straw (2/42, 5% of videos). The food guide or food label reading was not discussed in any of the videos. Notably, there were no statistically significant differences in the proportion of low–view rate and high–view rate videos that contained messages for brushing teeth (*P*=.74), sugar-free gum or sugar-free candy or xylitol (*P*=.19), and drinking beverages with a straw (*P*=.22).

Discussion

Principal Findings

To our knowledge, this is the first study that has focused on investigating nutrition and dental caries content on YouTube. These results are important because nutrition is strongly related to dental caries risk, dental caries is common, and YouTube is a popular web-based platform for the public to access information. These results provide insights into future directions for YouTube content in this area of public health importance.

Overall, we found that the 42 included videos had a low mean score (4.9, SD 3.4 out of 17 points), indicating that few relevant topics on nutrition and dental caries were covered in the videos. This finding is similar to the findings of other studies that have examined health-related content on YouTube. For example, in a study on oral cancer YouTube videos, Hassona et al [48] found that included videos provided "inadequate descriptions" of oral cancer risk factors. Similarly, in a study on oral hygiene instruction in YouTube videos, Smyth et al [56] found that none of the included videos addressed all topics of interest, and the authors had concerns about the messages presented in some videos. In addition, a recent review article found that the comprehensiveness of YouTube videos on various health topics was low [38]. Similar concerns have also been reported in pediatric oral health education leaflets. Arora et al [57] found that nutrition messaging in these types of leaflets was incomplete. Our results suggest that members of the public accessing YouTube for information on nutrition and dental caries may not get the complete information on this topic needed to fully optimize diets to prevent this issue.

We found that sugar was the most consistent topic mentioned in the included videos (mentioned in 31/42, 74% of videos). No other topic we assessed was mentioned in more than half of the videos. In a content analysis of nutrition information in oral health education leaflets from the United Kingdom, Morgan et al [58] also found that sugar was the most common topic covered and that there was variability in the number of topics covered. We also found that fewer YouTube videos covered foods and beverages to consume to decrease the risk of dental caries (eg, vegetables and fruit). This finding contrasts with the findings of previous studies on oral health leaflets, which showed a high prevalence of messages regarding what foods to consume. For example, Morgan et al [58] found that 73% and 70% of assessed oral health leaflets recommended vegetables and fruit for snacks and drinking only milk and water, respectively. In addition, Arora et al [57] found that 81% of leaflets recommended water and 53% recommended consuming milk. In addition, 44% of the leaflets recommended drinking fluoridated or tap water. Individuals accessing YouTube videos for information on nutrition and dental caries have a high chance of receiving messaging regarding sugar but are less likely to obtain evidence-based messaging about what foods to eat to prevent dental caries. Limited messaging about what foods to eat provides the public with incomplete information, which may affect their ability to make meaningful changes.

As we were conducting our analyses, we noticed that some videos mentioned that concepts surrounding nutrition and dental caries (and specifically sugar) were common knowledge to the public using statements such as "everyone knows," "most people know," and "we all know." These statements contradict studies that have shown that the level of nutritional knowledge related to oral health in different populations may not be ideal [59-61]. When designing future YouTube videos on this topic, it is important to address the amount and frequency of sugar consumption, and it is also important to acknowledge that there are many other foods and eating behaviors that can influence the risk of dental caries, that it is a complex relationship, and that the information may be new to viewers.

Videos created by the government and OHPs had higher mean scores than those produced by health professionals who were not OHPs (including complementary and alternative medicine providers) and individuals with no health professional credentials or unknown credentials. However, these score differences were not statistically significant. Other studies on YouTube video health content generally find that videos produced by health professionals and professional associations are of better quality than those that are not produced by health professionals and professional associations (eg, advertisements) [35,37]. We generally found this to be the case in our study but not always. For example, a couple of videos in our study featuring OHPs had a score of 0. One possible reason could be that the nutrition content in nondietetic health profession programs (including dental programs) is often limited, and there are many barriers toward providing this training; therefore, OHPs may not have in-depth training in this area [62].

Our analysis revealed that there were small nonsignificant negative correlations between various engagement measures (eg, total views, view rate, total likes, like rate, total dislikes, and dislike rate) and video score (out of 17 points; range −0.202 to −0.114). In general, these results align with other content analysis studies on health-related YouTube videos. In a recent review article, Osman et al [38] found that 84% and 74% of the included studies that assessed correlations between engagement and video quality found no correlations or negative correlations for video quality versus number of views and video quality versus number of likes, respectively. However, when we divided our included videos into low–view rate and high–view rate videos, we found that low–view rate videos had a trend toward higher overall score compared with high–view rate videos, but this result was not statistically significant. We also found that low–view rate videos were more likely to have messaging related to the frequency of sugar intake (*P*=.047), acidic foods and beverages (*P*=.04), and water (*P*=.09) compared with high–view rate videos. Messaging regarding the frequency of sugar intake is especially important because the frequency of sugar intake is thought to be possibly more important than the amount of sugar in terms of dental caries risk [63]. In addition, it is expected that individuals will eat sugar; therefore, information on how to best consume this dietary component to prevent dental caries is an important message. Warren et al [37] have previously mentioned that higher engagement with poor-quality videos could suggest that the public may have difficulty determining quality health-related YouTube content. Health professionals have an important role to provide more education to the public about how to select quality videos related to nutrition and dental caries. In addition, oral health, nutrition, and other professionals play important roles in producing evidence-based videos that are engaging and can be easily found by the public. Haslam et al [28] provided a list of strategies that can be used by creators to help make their videos more accessible.

As we watched and scored the videos, we observed that there was some contradictory diet advice related to some evidence-based items included in our 17-item scoring tool, both between videos and within videos, that was worthy of discussion. However, these contradictory messages are not evidence-based and could cause confusion for the public. We will discuss a few examples below, including sugary foods and beverages, whole grains, and milk products.

First, there was some contradictory advice about sugar-rich foods and beverages, where evidence-based guidelines suggest avoidance for dental caries prevention. Juice, which is a sugary beverage, was sometimes recommended or recommended over other sugary drinks. For example, some contradictory advice included recommending calcium-fortified juice, mentioning that unsweetened juice was beneficial for teeth because of vitamin C, and suggesting that juice was not as harmful as other sugary beverages. In addition, dried fruit, a sticky food that is highlighted as a sugary food as part of evidence-based guidelines [1,14,20], was mentioned as healthful in a couple of videos because of the presence of phytochemicals. Although a review article from 2016 has suggested that evidence regarding dried fruit and dental caries is limited [64], these foods are high in sugar. Finally, honey (including manuka honey) was identified as a better sugar choice in a couple of instances. Although a few studies have suggested that honey has antibacterial and antibiofilm properties, clinical evidence of the impact of honey on dental caries is not conclusive, and more studies are needed [65]. In summary, these products are all high in sugar; therefore, caution is needed surrounding these foods regarding dental caries risk.

Second, 2 videos in this data set advised limiting or avoiding whole grains because of concerns surrounding phytic acid causing dental caries, which contradicts evidence-based recommendations to consume whole grains. These videos recommended the consumption of grain products, where phytic acid has been reduced. Phytic acid is an antinutrient found in nuts, seeds, grains, and legumes and is known to bind some trace elements (eg, calcium, iron, and zinc), which can make them unavailable for absorption [66,67]. However, phytic acid should not be a concern when eaten as part of a mixed diet [67], and the benefits of consuming whole grains in Western countries outweigh the potential risks of phytic acid [68]. Currently, there is no strong evidence that phytic acid causes dental caries.

Third, there were a few videos that mentioned consuming dairy products is a risk factor for dental caries (eg, coffee creamers and yogurt owing to the carbohydrate content). Although the main sugar in milk products (lactose) is cariogenic, it is not as cariogenic as other sugars, and milk products contain many other beneficial components for dental caries prevention (eg, casein, calcium, and phosphorus). To date, evidence points toward milk being low cariogenic and possibly anticariogenic [69]. In addition, there was a recommendation to consume raw dairy products. This finding is concerning because raw milk is illegal to sell in many jurisdictions (eg, Canada), and milk pasteurization is often mandatory to avoid severe illnesses [70,71]. Although some cheeses made with raw milk that meet certain criteria can be sold in jurisdictions where pasteurization is mandatory, certain groups (eg, children, older adults, women who are pregnant, and individuals with weakened immune systems) are at risk of harmful effects from consuming these products [71]. Currently, there is no strong evidence suggesting that raw milk is beneficial for preventing dental caries or promoting better oral health.

These contradictory messages may cause confusion to viewers about whether the abovementioned foods or beverages are harmful or healthful regarding oral health and dental caries. These findings are consistent with the findings from the study by Morgan et al [58]. The authors of this UK study found that there were also inconsistencies in the nutrition and oral health information in different leaflets, including confusing information [58]. Arora et al [57] also identified confusing messaging regarding nutrition and oral health in pediatric oral health education leaflets in Australia, including confusing messaging around milk. This observation is important because when the public is exposed to contradictory advice, it has the potential to confuse them. The identification of areas of contradictory information is useful for health professionals looking to develop resources on this topic in the future.

We also found that there were some included videos that mentioned complementary and alternative medicine approaches to optimize oral health that were not included as part of our 17-item scoring tool. Some examples included consumption of probiotic supplements or probiotic-rich foods (mentioned in 4 videos), oil pulling (mentioned in 3 videos, with 1 video stating that this process was unpleasant and not recommended), and various recipes of home remedies for mixtures applied directly to the teeth or mouthwashes with various ingredients such as coconut oil, garlic, mustard oil, turmeric, clove oil, and salt (4 videos). In addition, 8 videos promoted or highlighted the consumption of vitamin K (usually K2) often in conjunction with vitamins A and D. In these videos, foods or supplements for these nutrients were recommended. Vitamin and mineral supplements (eg, vitamin D, calcium, magnesium, and vitamin K2) were also highlighted in 4 videos. Although some of these approaches (eg, probiotics and vitamin D) have generated substantial interest in the research and clinical communities regarding oral health and have shown promise for positive outcomes related to dental caries (eg, probiotics [72-79] and vitamin D [80-82]), for many of these approaches, there is a lack of evidence, and they are not recommended by professional associations (eg, oil pulling is not recommended by the American Dental Association [83], and probiotics are currently not recommended for dental caries prevention by the Canadian Pediatric Society [84]). It is important that health professionals are aware of these types of recommendations being made on the internet and are prepared to answer questions and generate evidence-based content related to these topics to help the public make informed decisions.

Limitations

A limitation of our study was that although we attempted to imitate search strategies used by the public to capture readily accessed YouTube videos, this might not be a completely accurate representation of the actual approaches used. However, we used Google Keyword Planner to plan searches and selected videos that appeared first in the results. Furthermore, our small sample size might be a limitation, but it is consistent with other studies assessing the health content of YouTube videos [35]. In the future, a study using a larger sample size of videos to evaluate content on this topic may be beneficial. We also excluded videos lasting for >20 minutes. In addition, misinformation was not considered as part of our scoring system. In the future, a study incorporating misinformation into a scoring approach in this topic and including longer videos would be worthwhile.

Conclusions

Our study found that most YouTube videos regarding nutrition and dental caries feature OHPs, and many videos cover a limited selection of topics. With the high prevalence of dental caries in the general population, the strong link between nutrition and dental caries, and the popularity of YouTube, there is a strong need for quality content containing evidence-based recommendations and information regarding this topic on this platform.

Acknowledgments

This research was funded by a Saskatchewan Health Research Foundation Establishment Grant (4503).

Conflicts of Interest

None declared.

References

1. Touger-Decker R, Mobley C, Academy of Nutrition and Dietetics. Position of the academy of nutrition and dietetics: oral health and nutrition. J Acad Nutr Diet 2013 May;113(5):693-701. [doi: 10.1016/j.jand.2013.03.001] [Medline: 23601893]
2. Wen PY, Chen MX, Zhong YJ, Dong QQ, Wong HM. Global burden and inequality of dental caries, 1990 to 2019. J Dent Res 2022 Apr;101(4):392-399. [doi: 10.1177/00220345211056247] [Medline: 34852668]
3. Global oral health status report: towards universal health coverage for oral health by 2030. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240061484 [accessed 2023-03-19]
4. Oral health. World Health Organization. 2023 Mar 14. URL: https://www.who.int/news-room/fact-sheets/detail/oral-health [accessed 2023-03-19]
5. Selwitz RH, Ismail AI, Pitts NB. Dental caries. Lancet 2007 Jan 06;369(9555):51-59. [doi: 10.1016/S0140-6736(07)60031-2] [Medline: 17208642]
6. Touger-Decker R, van Loveren C. Sugars and dental caries. Am J Clin Nutr 2003 Oct;78(4):881S-892S. [doi: 10.1093/ajcn/78.4.881S] [Medline: 14522753]
7. Hujoel PP, Lingström P. Nutrition, dental caries and periodontal disease: a narrative review. J Clin Periodontol 2017 Mar;44 Suppl 18:S79-S84. [doi: 10.1111/jcpe.12672] [Medline: 28266117]
8. Guideline: sugars intake for adults and children. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241549028 [accessed 2023-03-29]
9. Bleich SN, Vercammen KA. The negative impact of sugar-sweetened beverages on children's health: an update of the literature. BMC Obes 2018 Feb 20;5:6 [FREE Full text] [doi: 10.1186/s40608-017-0178-9] [Medline: 29484192]
10. Wilder JR, Kaste LM, Handler A, Chapple-McGruder T, Rankin KM. The association between sugar-sweetened beverages and dental caries among third-grade students in Georgia. J Public Health Dent 2016;76(1):76-84. [doi: 10.1111/jphd.12116] [Medline: 26339945]
11. Moss ME, Luo H, Rosinger AY, Jacobs MM, Kaur R. High sugar intake from sugar-sweetened beverages is associated with prevalence of untreated decay in US adults: NHANES 2013-2016. Community Dent Oral Epidemiol 2022 Dec;50(6):579-588 [FREE Full text] [doi: 10.1111/cdoe.12725] [Medline: 34939664]
12. Evans EW, Hayes C, Palmer CA, Bermudez OI, Cohen SA, Must A. Dietary intake and severe early childhood caries in low-income, young children. J Acad Nutr Diet 2013 Aug;113(8):1057-1061 [FREE Full text] [doi: 10.1016/j.jand.2013.03.014] [Medline: 23706351]

13. Policy on dietary recommendations for infants, children, and adolescents. American Academy of Pediatric Dentistry. 2021. URL: https://www.aapd.org/globalassets/media/policies_guidelines/p_recdietary.pdf [accessed 2023-01-22]
14. Oral health and nutrition guidance for professionals. National Health Service, Health Scotland. 2012 Jun. URL: https://www.scottishdental.org/wp-content/uploads/2014/10/OralHealthAndNutritionGuidance.pdf [accessed 2022-01-23]
15. Firestone AR, Schmid R, Mühlemann HR. Cariogenic effects of cooked wheat starch alone or with sucrose and frequency-controlled feedings in rats. Arch Oral Biol 1982;27(9):759-763. [doi: 10.1016/0003-9969(82)90026-7] [Medline: 6959584]
16. Moynihan PJ. The role of diet and nutrition in the etiology and prevention of oral diseases. Bull World Health Organ 2005 Sep;83(9):694-699 [FREE Full text] [Medline: 16211161]
17. Tinanoff N, Palmer CA. Dietary determinants of dental caries and dietary recommendations for preschool children. J Public Health Dent 2000;60(3):197-206. [doi: 10.1111/j.1752-7325.2000.tb03328.x] [Medline: 11109219]
18. Tungare S, Paranjpe AG. Diet and nutrition to prevent dental problems. StatPearls. 2021 Aug 11. URL: https://www.ncbi.nlm.nih.gov/books/NBK534248/ [accessed 2023-03-19]
19. Moynihan P. Foods and factors that protect against dental caries. Nutr Bull 2001 Dec 25;25(4):281-286 [FREE Full text] [doi: 10.1046/j.1467-3010.2000.00033.x]
20. Joint position statement on interdisciplinary collaboration between accredited practising dietitians, nutrition and oral health professionals for oral health and nutrition. Dietitians Association of Australia, Dental Health Services Victoria. 2021 Nov. URL: https://dietitiansaustralia.org.au/sites/default/files/2022-06/FINAL%20VERSION%20JPS%20OH_Nutrition%20Collab_20062022.pdf [accessed 2023-07-19]
21. Riley P, Moore D, Ahmed F, Sharif MO, Worthington HV. Xylitol-containing products for preventing dental caries in children and adults. Cochrane Database Syst Rev 2015 Mar 26;2015(3):CD010743 [FREE Full text] [doi: 10.1002/14651858.CD010743.pub2] [Medline: 25809586]
22. Lieffers JR, Vanzan AG, Rover de Mello J, Cammer A. Nutrition care practices of dietitians and oral health professionals for oral health conditions: a scoping review. Nutrients 2021 Oct 13;13(10):3588 [FREE Full text] [doi: 10.3390/nu13103588] [Medline: 34684589]
23. Shahab L, Brown J, Gardner B, Smith SG. Seeking health information and support online: does it differ as a function of engagement in risky health behaviors? Evidence from the health information national trends survey. J Med Internet Res 2014 Nov 06;16(11):e253 [FREE Full text] [doi: 10.2196/jmir.3368] [Medline: 25380308]
24. Prestin A, Vieux SN, Chou WS. Is online health activity alive and well or flatlining? Findings from 10 years of the health information national trends survey. J Health Commun 2015;20(7):790-798. [doi: 10.1080/10810730.2015.1018590] [Medline: 26042588]
25. Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, Allen SV, Moser RP, Hesse BW. Online health information seeking among US adults: measuring progress toward a healthy people 2020 objective. Public Health Rep 2019 Nov;134(6):617-625 [FREE Full text] [doi: 10.1177/0033354919874074] [Medline: 31513756]
26. Tracking nutrition trends 2015. Canadian Foundation for Dietetic Research. 2015. URL: https://www.cfdr.ca/Downloads/Tracking-Nutrition-Trends/Tracking-Nutrition-Trends-IX-2015.aspx [accessed 2022-01-22]
27. 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]
28. Haslam K, Doucette H, Hachey S, MacCallum T, Zwicker D, Smith-Brilliant M, et al. YouTube videos as health decision aids for the public: an integrative review. Can J Dent Hyg 2019 Mar 01;53(1):53-66 [FREE Full text] [Medline: 33240342]
29. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 02;41:433-451 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
30. Neufeld D. The 50 most visited websites in the world. Visual Capitalist. 2021. URL: https://www.visualcapitalist.com/the-50-most-visited-websites-in-the-world/ [accessed 2022-01-22]
31. Iqbal M. YouTube revenue and usage statistics. Business of Apps. 2022. URL: https://www.businessofapps.com/data/youtube-statistics/ [accessed 2022-01-22]
32. Ceci L. Hours of video uploaded to YouTube every minute as of February 2020. Statista. 2021. URL: https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/ [accessed 2022-01-22]
33. Why people turn to YouTube. Google. URL: https://www.thinkwithgoogle.com/consumer-insights/consumer-trends/why-people-turn-to-youtube/ [accessed 2022-01-22]
34. Berkowitz IS, Davis CH, Smith H. Watchtime Canada: how YouTube connects creators & consumers, Ryerson university faculty of communication and design. Ryerson University Faculty of Communication & Design. 2019. URL: https://sites.google.com/view/watchtime-2019 [accessed 2022-01-23]
35. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. Health Informatics J 2015 Sep;21(3):173-194 [FREE Full text] [doi: 10.1177/1460458213512220] [Medline: 24670899]
36. Drozd B, Couvillon E, Suarez A. Medical YouTube videos and methods of evaluation: literature review. JMIR Med Educ 2018 Feb 12;4(1):e3 [FREE Full text] [doi: 10.2196/mededu.8527] [Medline: 29434018]

37. Warren CJ, Sawhney R, Shah T, Behbahani S, Sadeghi-Nejad H. YouTube and men's health: a review of the current literature. Sex Med Rev 2021 Apr;9(2):280-288. [doi: 10.1016/j.sxmr.2020.09.002] [Medline: 33610492]
38. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. BMC Med Educ 2022 May 19;22(1):382 [FREE Full text] [doi: 10.1186/s12909-022-03446-z] [Medline: 35590410]
39. TCPS 2 (2018) – chapter 2: scope and approach. Government of Canada Panel on Research Ethics. 2018. URL: https://ethics.gc.ca/eng/tcps2-eptc2_2018_chapter2-chapitre2.html [accessed 2023-03-19]
40. Keyword planner. Google Ads. URL: https://ads.google.com/intl/en_ca/home/tools/keyword-planner/ [accessed 2022-01-23]
41. Nason K, Donnelly A, Duncan HF. YouTube as a patient-information source for root canal treatment. Int Endod J 2016 Dec;49(12):1194-1200. [doi: 10.1111/iej.12575] [Medline: 26551481]
42. Brachtenbach T, Cardenas R, Pate H, Solis L, Chinnery L, Bowers A, et al. YouTube: searching for answers about breast cancer. Breast Dis 2020;39(2):85-90. [doi: 10.3233/BD-200445] [Medline: 32250286]
43. Video length: how long should instructional videos be? TechSmith. URL: https://www.techsmith.com/blog/video-length/ [accessed 2023-03-19]
44. Jia JS, Mehran N, Purgert R, Zhang QE, Lee D, Myers JS, et al. Marijuana and glaucoma: a social media content analysis. Ophthalmol Glaucoma 2021 Jul;4(4):400-404. [doi: 10.1016/j.ogla.2020.11.004] [Medline: 33242684]
45. Luo C, Zheng X, Zeng DD, Leischow S. Portrayal of electronic cigarettes on YouTube. BMC Public Health 2014 Oct 03;14:1028 [FREE Full text] [doi: 10.1186/1471-2458-14-1028] [Medline: 25277872]
46. Bromberg JE, Augustson EM, Backinger CL. Portrayal of smokeless tobacco in YouTube videos. Nicotine Tob Res 2012 Apr;14(4):455-462 [FREE Full text] [doi: 10.1093/ntr/ntr235] [Medline: 22080585]
47. How to optimize your SEO results through content creation. Forbes. 2017. URL: https://www.forbes.com/sites/ajagrawal/2017/08/30/how-to-optimize-your-seo-results-through-content-creation/?sh=737aa6b72aa3 [accessed 2023-03-19]
48. Hassona Y, Taimeh D, Marahleh A, Scully C. YouTube as a source of information on mouth (oral) cancer. Oral Dis 2016 Apr;22(3):202-208. [doi: 10.1111/odi.12434] [Medline: 26718020]
49. ElKarmi R, Hassona Y, Taimeh D, Scully C. YouTube as a source for parents' education on early childhood caries. Int J Paediatr Dent 2017 Nov;27(6):437-443. [doi: 10.1111/ipd.12277] [Medline: 27882621]
50. Basch CH, Blankenship EB, Goff ME, Yin J, Basch CE, DeLeon AJ, et al. Fluoride-related YouTube videos: a cross-sectional study of video contents by upload sources. J Dent Hyg 2018 Dec;92(6):47-53. [Medline: 30643004]
51. Castillo J, Wassef C, Wassef A, Stormes K, Berry AE. YouTube as a source of patient information for prenatal repair of myelomeningocele. Am J Perinatol 2021 Jan;38(2):140-144. [doi: 10.1055/s-0039-1694786] [Medline: 31430813]
52. Camm CF, Russell E, Ji Xu A, Rajappan K. Does YouTube provide high-quality resources for patient education on atrial fibrillation ablation? Int J Cardiol 2018 Dec 01;272:189-193. [doi: 10.1016/j.ijcard.2018.08.066] [Medline: 30177231]
53. Marshall TA. Chairside diet assessment of caries risk. J Am Dent Assoc 2009 Jun;140(6):670-674. [doi: 10.14219/jada.archive.2009.0252] [Medline: 19491162]
54. Salas MM, Nascimento GG, Vargas-Ferreira F, Tarquinio SB, Huysmans MC, Demarco FF. Diet influenced tooth erosion prevalence in children and adolescents: results of a meta-analysis and meta-regression. J Dent 2015 Aug;43(8):865-875. [doi: 10.1016/j.jdent.2015.05.012] [Medline: 26057086]
55. Bartlett DW, Fares J, Shirodaria S, Chiu K, Ahmad N, Sherriff M. The association of tooth wear, diet and dietary habits in adults aged 18-30 years old. J Dent 2011 Dec;39(12):811-816 [FREE Full text] [doi: 10.1016/j.jdent.2011.08.014] [Medline: 21911033]
56. Smyth RS, Amlani M, Fulton A, Sharif MO. The availability and characteristics of patient-focused YouTube videos related to oral hygiene instruction. Br Dent J 2020 May;228(10):773-781. [doi: 10.1038/s41415-020-1527-5] [Medline: 32444751]
57. Arora A, Doan J, Martinez J, Phan C, Kolt GS, Bhole S, et al. Content analysis of nutritional information in paediatric oral health education leaflets. BMC Pediatr 2017 Feb 20;17(1):58 [FREE Full text] [doi: 10.1186/s12887-017-0814-z] [Medline: 28219353]
58. Morgan MZ, McFarlane E, Stewart KF, Hunter ML, Fairchild RM. An assessment of nutritional information in oral health education leaflets. Community Dent Health 2010 Jun;27(2):81-88 [FREE Full text] [Medline: 20648884]
59. Badrasawi MM, Hijjeh NH, Amer RS, Allan RM, Altamimi M. Nutrition awareness and oral health among dental patients in Palestine: a cross-sectional study. Int J Dent 2020 Feb 26;2020:3472753 [FREE Full text] [doi: 10.1155/2020/3472753] [Medline: 32184828]
60. Sachdev PK, Freeland-Graves J, Babaei M. Development and validation of the dental nutrition knowledge competency scale for low-income women. Public Health Nutr 2020 Mar;23(4):691-700 [FREE Full text] [doi: 10.1017/S1368980019002714] [Medline: 31588884]
61. Sachdev PK, Freeland-Graves J, Ranjit N, Babaei M. Role of dental nutrition knowledge and socioecological factors in dental caries in low-income women. Health Educ Behav (Forthcoming) 2021 Apr 08:10901981211004149. [doi: 10.1177/10901981211004149] [Medline: 33829899]
62. DiMaria-Ghalili RA, Mirtallo JM, Tobin BW, Hark L, Van Horn L, Palmer CA. Challenges and opportunities for nutrition education and training in the health care professions: intraprofessional and interprofessional call to action. Am J Clin Nutr 2014 May;99(5 Suppl):1184S-1193S [FREE Full text] [doi: 10.3945/ajcn.113.073536] [Medline: 24646823]

63. van Loveren C. Sugar restriction for caries prevention: amount and frequency. Which is more important? Caries Res 2019;53(2):168-175 [FREE Full text] [doi: 10.1159/000489571] [Medline: 30089285]
64. Sadler MJ. Dried fruit and dental health. Int J Food Sci Nutr 2016 Dec;67(8):944-959. [doi: 10.1080/09637486.2016.1207061] [Medline: 27415591]
65. Deglovic J, Majtanova N, Majtan J. Antibacterial and antibiofilm effect of honey in the prevention of dental caries: a recent perspective. Foods 2022 Sep 02;11(17):2670 [FREE Full text] [doi: 10.3390/foods11172670] [Medline: 36076855]
66. Schlemmer U, Frølich W, Prieto RM, Grases F. Phytate in foods and significance for humans: food sources, intake, processing, bioavailability, protective role and analysis. Mol Nutr Food Res 2009 Sep;53 Suppl 2:S330-S375. [doi: 10.1002/mnfr.200900099] [Medline: 19774556]
67. Petroski W, Minich DM. Is there such a thing as "anti-nutrients"? A narrative review of perceived problematic plant compounds. Nutrients 2020 Sep 24;12(10):2929 [FREE Full text] [doi: 10.3390/nu12102929] [Medline: 32987890]
68. Brouns F. Phytic acid and whole grains for health controversy. Nutrients 2021 Dec 22;14(1):25 [FREE Full text] [doi: 10.3390/nu14010025] [Medline: 35010899]
69. Woodward M, Rugg-Gunn AJ. Chapter 8: milk, yoghurts and dental caries. Monogr Oral Sci 2020;28:77-90. [doi: 10.1159/000455374] [Medline: 31940625]
70. Food safety: the facts about raw milk. Ministry of Health and Ministry of Long Term Care, Government of Ontario. 2015. URL: https://www.health.gov.on.ca/en/public/programs/publichealth/foodsafety/milk_facts.aspx [accessed 2022-02-01]
71. Raw or unpasteurized milk. Government of Canada. 2013. URL: https://www.canada.ca/en/health-canada/services/milk-infant-formula/raw-or-unpasteurized-milk.html [accessed 2022-02-01]
72. Gruner D, Paris S, Schwendicke F. Probiotics for managing caries and periodontitis: systematic review and meta-analysis. J Dent 2016 May;48:16-25. [doi: 10.1016/j.jdent.2016.03.002] [Medline: 26965080]
73. Jørgensen MR, Castiblanco G, Twetman S, Keller MK. Prevention of caries with probiotic bacteria during early childhood. Promising but inconsistent findings. Am J Dent 2016 Jun;29(3):127-131. [Medline: 27505986]
74. Sivamaruthi BS, Kesika P, Chaiyasut C. A review of the role of probiotic supplementation in dental caries. Probiotics Antimicrob Proteins 2020 Dec;12(4):1300-1309. [doi: 10.1007/s12602-020-09652-9] [Medline: 32307660]
75. Lin TH, Lin CH, Pan TM. The implication of probiotics in the prevention of dental caries. Appl Microbiol Biotechnol 2018 Jan;102(2):577-586. [doi: 10.1007/s00253-017-8664-z] [Medline: 29192351]
76. Cagetti MG, Mastroberardino S, Milia E, Cocco F, Lingström P, Campus G. The use of probiotic strains in caries prevention: a systematic review. Nutrients 2013 Jul 05;5(7):2530-2550 [FREE Full text] [doi: 10.3390/nu5072530] [Medline: 23857225]
77. Bonifait L, Chandad F, Grenier D. Probiotics for oral health: myth or reality? J Can Dent Assoc 2009 Oct;75(8):585-590 [FREE Full text] [Medline: 19840501]
78. Shi J, Wang Q, Ruan G, Chen Y, Zhao M, Shi D, et al. Efficacy of probiotics against dental caries in children: a systematic review and meta-analysis. Crit Rev Food Sci Nutr 2022 May 24:1-18. [doi: 10.1080/10408398.2022.2077693] [Medline: 35607893]
79. Saïz P, Taveira N, Alves R. Probiotics in oral health and disease: a systematic review. Appl Sci 2021;111(7):8070 [FREE Full text] [doi: 10.3390/app11178070]
80. Hujoel PP. Vitamin D and dental caries in controlled clinical trials: systematic review and meta-analysis. Nutr Rev 2013 Feb;71(2):88-97. [doi: 10.1111/j.1753-4887.2012.00544.x] [Medline: 23356636]
81. Botelho J, Machado V, Proença L, Delgado AS, Mendes JJ. Vitamin D deficiency and oral health: a comprehensive review. Nutrients 2020 May 19;12(5):1471 [FREE Full text] [doi: 10.3390/nu12051471] [Medline: 32438644]
82. Carvalho Silva C, Mendes R, Manso MD, Gavinha S, Melo P. Prenatal or childhood serum levels of vitamin D and dental caries in paediatric patients: a systematic review. Oral Health Prev Dent 2020 Sep 04;18(1):653-667. [doi: 10.3290/j.ohpd.a45089] [Medline: 32895648]
83. Oil pulling. American Dental Association. URL: https://www.mouthhealthy.org/all-topics-a-z/oil-pulling/ [accessed 2023-04-13]
84. Schneider R, Sant'Anna A. Using probiotics in paediatric populations. Paediatr Child Health 2022 Dec 27;27(8):482-502. [doi: 10.1093/pch/pxac087] [Medline: 36583073]

Abbreviations

OHP: oral health professional
TCPS: Tri-Council Policy Statement

Edited by T Mackey; submitted 02.06.22; peer-reviewed by N Bardach, H Cheng; comments to author 22.10.22; revised version received 26.04.23; accepted 14.06.23; published 10.08.23.

Please cite as:
Long M, Forbes LE, Papagerakis P, Lieffers JRL
YouTube Videos on Nutrition and Dental Caries: Content Analysis
JMIR Infodemiology 2023;3:e40003
URL: https://infodemiology.jmir.org/2023/1/e40003
doi:10.2196/40003
PMID:37561564

Original Paper

Assessment of the Early Detection of Anosmia and Ageusia Symptoms in COVID-19 on Twitter: Retrospective Study

Carole Faviez[1,2*], PhD; Manissa Talmatkadi[3*], MSc; Pierre Foulquié[3], MSc; Adel Mebarki[3], MSc; Stéphane Schück[3], MD; Anita Burgun[1,2,4], MD, PhD; Xiaoyi Chen[1,2,5], PhD

[1]Centre de Recherche des Cordeliers, Université Paris Cité, Sorbonne Université, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR 1138, Paris, France

[2]Health Data- and Model- Driven Knowledge Acquisition (HeKA), Inria Paris, Paris, France

[3]Kap Code, Paris, France

[4]Department of Medical Informatics, Hôpital Necker-Enfant Malades, Assistance Publique - Hôpitaux de Paris (AP-HP), Paris, France

[5]Data Science Platform, Imagine Institute, Université Paris Cité, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR 1163, Paris, France

[*]these authors contributed equally

Corresponding Author:
Carole Faviez, PhD
Centre de Recherche des Cordeliers, Université Paris Cité
Sorbonne Université
Institut National de la Santé et de la Recherche Médicale (INSERM) UMR 1138
24 Boulevard du Montparnasse
Paris, 75015
France
Phone: 33 142754200
Email: carole.faviez@inserm.fr

Abstract

Background: During the unprecedented COVID-19 pandemic, social media has been extensively used to amplify the spread of information and to express personal health-related experiences regarding symptoms, including anosmia and ageusia, 2 symptoms that have been reported later than other symptoms.

Objective: Our objective is to investigate to what extent Twitter users reported anosmia and ageusia symptoms in their tweets and if they connected them to COVID-19, to evaluate whether these symptoms could have been identified as COVID-19 symptoms earlier using Twitter rather than the official notice.

Methods: We collected French tweets posted between January 1, 2020, and March 31, 2020, containing anosmia- or ageusia-related keywords. Symptoms were detected using fuzzy matching. The analysis consisted of 3 parts. First, we compared the coverage of anosmia and ageusia symptoms in Twitter and in traditional media to determine if the association between COVID-19 and anosmia or ageusia could have been identified earlier through Twitter. Second, we conducted a manual analysis of anosmia- and ageusia-related tweets to obtain quantitative and qualitative insights regarding their nature and to assess when the first associations between COVID-19 and these symptoms were established. We randomly annotated tweets from 2 periods: the early stage and the rapid spread stage of the epidemic. For each tweet, each symptom was annotated regarding 3 modalities: symptom (yes or no), associated with COVID-19 (yes, no, or unknown), and whether it was experienced by someone (yes, no, or unknown). Third, to evaluate if there was a global increase of tweets mentioning anosmia or ageusia in early 2020, corresponding to the beginning of the COVID-19 epidemic, we compared the tweets reporting experienced anosmia or ageusia between the first periods of 2019 and 2020.

Results: In total, 832 (respectively 12,544) tweets containing anosmia (respectively ageusia) related keywords were extracted over the analysis period in 2020. The comparison to traditional media showed a strong correlation without any lag, which suggests an important reactivity of Twitter but no earlier detection on Twitter. The annotation of tweets from 2020 showed that tweets correlating anosmia or ageusia with COVID-19 could be found a few days before the official announcement. However, no association could be found during the first stage of the pandemic. Information about the temporality of symptoms and the psychological impact of these symptoms could be found in the tweets. The comparison between early 2020 and early 2019 showed no difference regarding the volumes of tweets.

Conclusions: Based on our analysis of French tweets, associations between COVID-19 and anosmia or ageusia by web users could have been found on Twitter just a few days before the official announcement but not during the early stage of the pandemic. Patients share qualitative information on Twitter regarding anosmia or ageusia symptoms that could be of interest for future analyses.

(JMIR Infodemiology 2023;3:e41863) doi:10.2196/41863

KEYWORDS

social media; COVID-19; anosmia; ageusia; infodemiology; symptom; Twitter; psychological; tweets; pandemic; rapid stage; epidemic; information; knowledge; online health; tweets; misinformation; education; online education; ehealth; qualitative

Introduction

In recent years, social media, with its widespread usage and large user base, have gained significant attention as potential sources of information for public health surveillance. The systematic collection, analysis, and interpretation of health-related information from social media allow for various applications, ranging from the spread of infectious diseases (eg, HIV, SARS, and influenza), to vaccination uptake, antibiotics consumption, and alcohol consumption. The emergence of the COVID-19 pandemic has significantly accelerated research efforts in infodemiology and infoveillance using social media data. As an example, a search on PubMed using "social media" and "symptoms" as keywords revealed 579 publications from 2012 to 2019, and after the outbreak of COVID-19, the number of publications retrieved using the same keywords increased to 423 in 2020 alone, then 611 and 622 in 2021 and 2022, respectively, around 40% of which were COVID-19–related (163/423 in 2020, 288/611 in 2021, and 279/622 in 2022), reflecting the exponential growth of studies investigating the role of social media in symptom surveillance. Among social media platforms, Twitter stands out as an event-reactive tool with high posting frequency. For example, there have been attempts to use Twitter mining to monitor vaccine adverse events, showing that, for example, sore to touch, fatigue, and headache were the most common adverse effects in the United States [1], and symptoms related with appetite, allergy, injection site, and clots in the United Kingdom [2]. During the pandemic, it has been used for discussing various dimensions of the pandemic, including epidemiology, economy, as well as clinical and emotional aspects [3,4]. While social media provides an opportunity to directly communicate health information to the public, health related testimonies posted on the internet may also be used for early detection of symptoms and diseases. As an example, Lopreite et al [5] analyzed the data from Twitter to uncover early warning signals of COVID-19 outbreaks in Europe in the winter season 2019-2020, and showed that unexpected levels of pneumonia-related tweets were raised across a number of European countries in early 2020 prior to official announcement. For example, they identified an increase in the number of tweets mentioning dry cough during the weeks leading to the peak in February 2020.

In this study, we aim to assess the early detection of anosmia and ageusia symptoms associated with COVID-19 using Twitter data, contributing to the growing field of digital epidemiology and infodemiology. We focused on these 2 symptoms because their relation with COVID-19 was unknown at the beginning of the pandemic. After the initial outbreak of COVID-19, reported findings varied across countries and time. For example, the list of symptoms at the onset of illness reported by Wuhan clinicians in early 2020 [6] only includes fever, cough, and myalgia or fatigue, sputum production, headache, hemoptysis, and diarrhea. Later on, symptoms such as anosmia, dysgeusia, headache, and muscle pain have been noted, along with more reports of central and peripheral nervous system involvement [7,8]. Anosmia, that is, the loss of the sense of smell, and ageusia, that is, the loss of the sense of taste, have been associated with SARS-CoV-2 test positivity, regardless of the population, and illness duration or complexity [9]. In March 2020, national and international institutions (eg, [10]) proposed to add these symptoms to the list of screening tools for possible COVID-19 infection. They were followed by all health authorities worldwide, including the Center for Disease Control and Prevention and the World Health Organization, who added "new loss of taste or smell" to the list of COVID-19 symptoms. Anosmia, in particular, has been seen in patients ultimately testing positive for the coronavirus with no other symptoms. Moreover, the estimated prevalence varies a lot across studies. A multicenter European study looking at patients with mild to moderate COVID-19 disease found a prevalence of 85.6% and 88% of olfactory and gustatory dysfunctions, respectively [11]. Contrasting with this high prevalence in European COVID-19 cohorts, a prevalence of only 5.1%-5.6% impairment of smell or taste (3% in severe, 6%-7% in nonsevere) was observed in China during January and February 2020 [8]. In between, a cross-sectional survey in Italy found that 33.9% of the patients admitted before March 2020 had olfactory or taste disorders [12]. Finally, a cohort study including patients with clinically diagnosed or laboratory-confirmed COVID-19 from 28 centers, representing 13 countries and 4 continents concluded that anosmia or ageusia was the second mostly self-reported neurological symptom after headache by patients with COVID-19, which led to a global prevalence of 26% [13]. This study was performed by the Global Consortium Study of Neurological Dysfunction in COVID-19 and the European Academy of Neurology Neuro-COVID Registry after anosmia and ageusia were added to the list of symptoms, more precisely from March to September or October 2020.

In this context, our objective is to perform a retrospective analysis of Twitter data to investigate to what extent anosmia and ageusia were reported and associated with COVID-19 in the tweets before their official recognition. As these symptoms are uncommon outside of COVID-19, their unique representation and potential as early indicators of the disease can help differentiate COVID-19 from other respiratory diseases such as the flu. Having detected that anosmia and ageusia were COVID-19 symptoms earlier could have helped diagnose and isolate some patients. Therefore, the findings of this retrospective study contribute to the broader field of infoveillance, providing insights into the utility of Twitter data for early detection and monitoring of symptoms of a new disease.

Methods

Tweet Extraction

We considered all the tweets posted between January 1, 2020, and March 31, 2020, which corresponds to the early stages of the pandemic in France. The Detec't (Kap Code) tool [14] was used to extract all nonretweet tweets in French containing at least one of the keywords related to the sense of smell (eg, anosmia, smell, and olfactive) or taste (eg, ageusia, taste, and gustative) and posted during the first 3 months of 2020. The established list of French keywords and their English translation is given in Table S1 in Multimedia Appendix 1.

An additional data set of tweets posted in early 2019 was also extracted to enable the comparison between 2019 and 2020. The objective of this comparison was to assess if the volume of declarations of anosmia or ageusia with unknown causes in early 2019 was lower than the one in early 2020, which could imply that tweets in early 2020 were related to COVID-19.

Symptom Detection

Anosmia and ageusia specific data sets were created based on symptom detection in tweets. A fuzzy matching method was applied to the set of smell- or taste-related tweets. More precisely, we reused the dictionary of symptoms based on MedDRA (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) and enhanced by colloquial terms, which includes the preferred terms (PTs), the associated low level terms, and manual enrichment [15]. Terms associated with the 2 PTs "anosmia" and "ageusia" were used for detection, based on the method that we previously developed and tuned [15], that is, exact matching for short terms containing less than 6 characters and fuzzy matching for other terms. This method was proven to be effective in improving the detection of symptoms with spelling mistakes or in the plural form. Tweets containing these terms were used to populate the anosmia and ageusia specific data sets.

Media Coverage of Symptoms

To determine whether there was a temporal gap between the declaration of anosmia and ageusia by patients on the internet and their coverage by media, we compared the coverage of these symptoms in social media and in traditional media. Anosmia- and ageusia-related news articles were extracted using Brandwatch Consumer Research (Brandwatch) tool [16]. Content classified as "News" was counted to describe traditional media coverage over time. Cross correlation [17] between volumes of tweets and news were computed for each symptom to measure the similarity between the series. Dissimilarities would suggest the identification of a temporal gap.

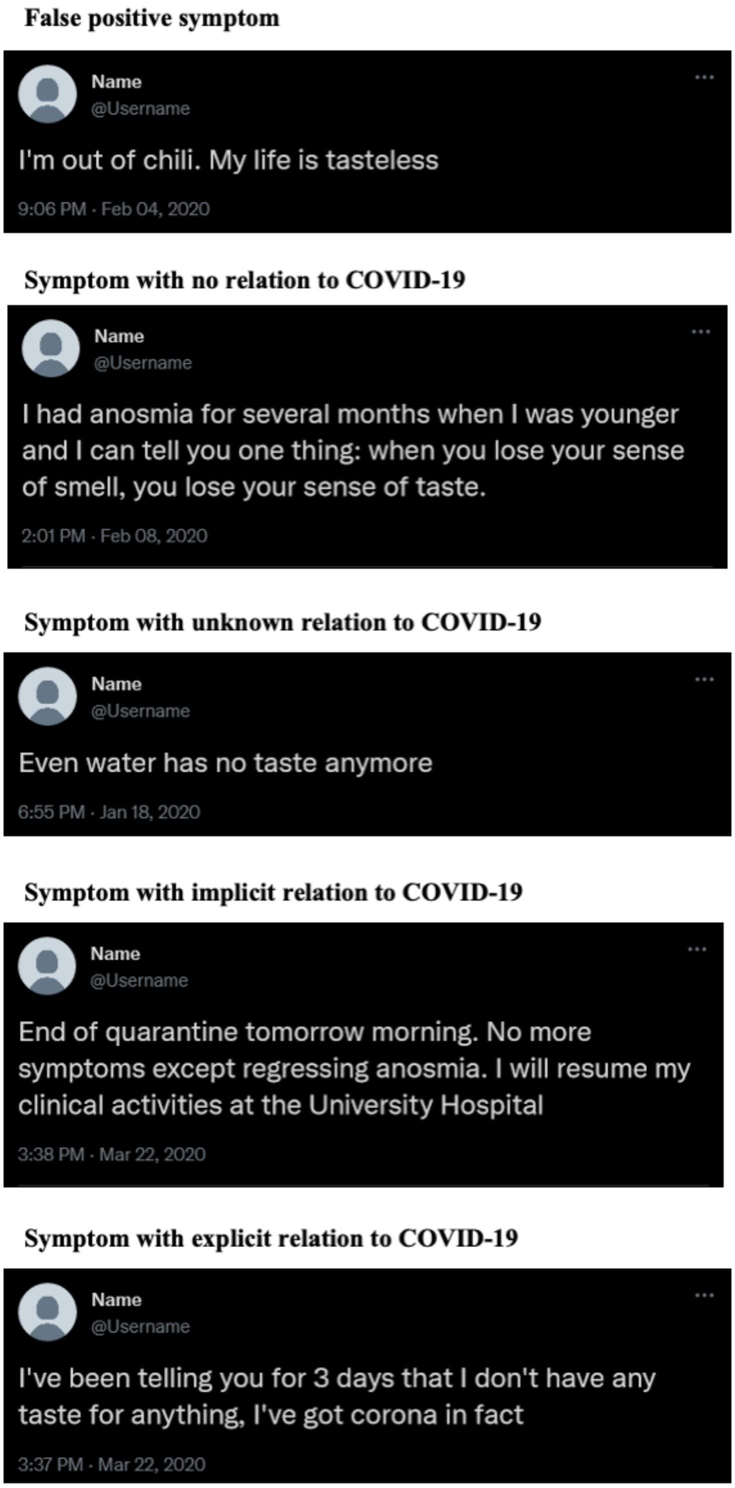
Tweet Annotation

We considered 2 periods of time to perform random sampling for the tweet annotation: before and after March 10, 2020. The choice of this date was based on the number of cases in France, and the official alert of anosmia and ageusia as COVID-19 symptoms. The first period (January 1-March 10) is an early stage of the epidemic (daily new cases<400, cumulative cases<1500), and the second period (March 10-31) corresponds to the rapid spread of the epidemic, as 1 week after the beginning of the second period, the lockdown started (March 17, 2020), and 3 days later anosmia or ageusia was confirmed as symptoms of COVID-19 by the French health authorities (March 20, 2020). Manual annotation was performed by one of the authors (CF) for a set of randomly selected tweets of the same sample size for period #1 and #2 of 2020 and the corresponding period #1 of 2019, for both symptoms. This resulted in 6 annotated data sets. The goal of this annotation task was to identify mentions of symptoms that could be attributable to COVID-19.

In total, 3 types of annotations were performed. For each detected term, we first annotated if it corresponded to a genuine symptom, that is, reference to a lack of smell or to a lack of taste, or to another meaning. Indeed, a tweet can contain a term from MedDRA associated with the PT "anosmia" and "ageusia" without being a symptom, for example in case of polysemy. For example, terms like "pas de gout" (no taste) may denote ageusia (yes=true positive), but may be also used in other contexts like fashion, thus corresponding to a false positive. Then for each genuine symptom, we further annotated if it was COVID-19–related (yes, no, or unknown), and if, according to the author of the tweet, it was a symptom experienced by an individual (yes, no, or unknown), either the author of the tweet, a relative, or another patient. Regarding the relation to COVID-19, we considered that a tweet was COVID-19–related when the relation was explicitly mentioned by the author or strongly suspected (implicit relation) based on contextual information and the interpretation by the annotator. Figure 1 displays 5 examples of annotated ageusia-anosmia related tweets translated from French.

The fourth example in Figure 1 does not explicitly mention COVID-19 but mentions quarantine which is highly suggestive of an infection by COVID-19. The fifth example expresses ageusia and links it with COVID-19 ("corona").

Figure 1. Examples of ageusia- and anosmia-related tweets translated from French to English.



Analysis and Comparison Between Periods and to the Previous Year

Temporal evolution of symptom declarations was analyzed in 2 ways. First, we assessed the temporal evolution of volumes of tweets in each annotation category (symptom yes or no, related to COVID-19 yes, no, or unknown, and experienced yes, no, or unknown). A 3-day cumulative time series were displayed for the 2 periods of 2020 to smooth potential missing data due to the sampling. Second, we focused on the volumes of tweets regarding experienced symptoms with unexplicit causes, and compared them in period #1 of 2020 and 2019. A significant increase in early 2020 compared with early 2019 could suggest that tweets in early 2020 were related to COVID-19 and consequently that cases of COVID-19–related anosmia or ageusia could have been detected earlier in Twitter.

Ethical Considerations

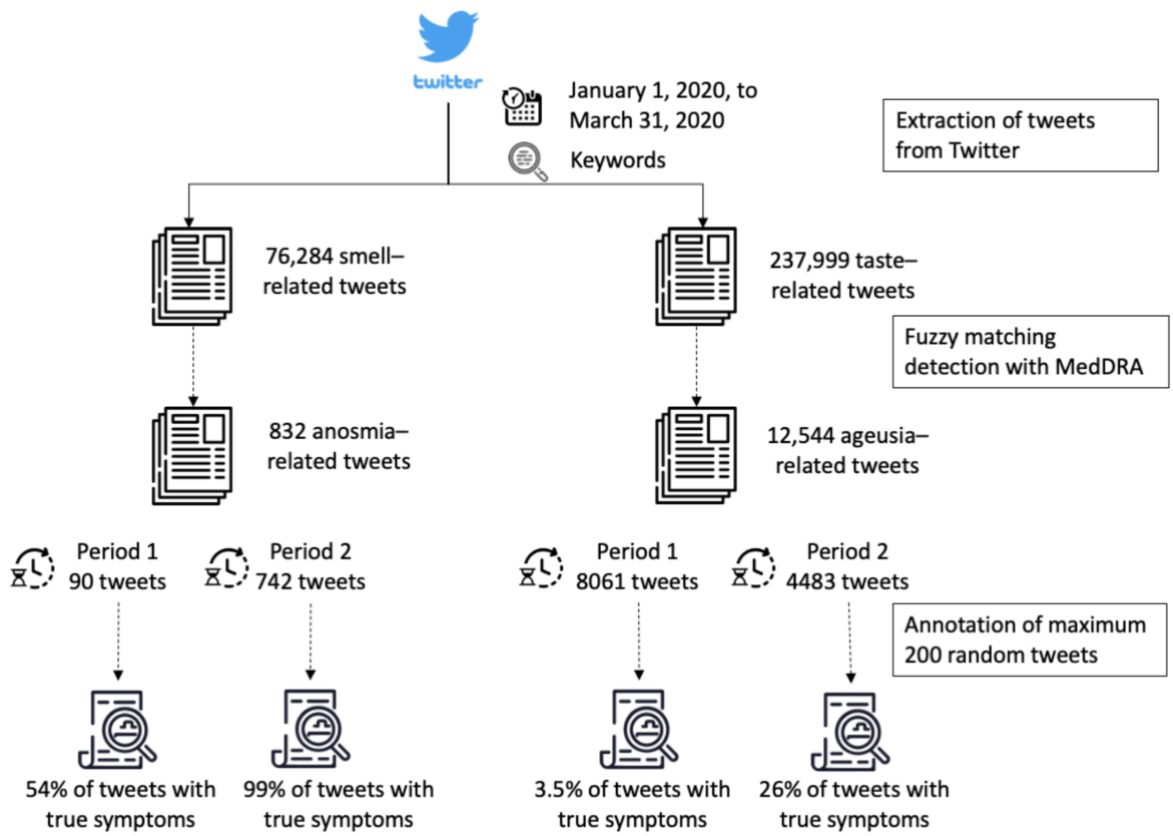
This retrospective and observational study is not a human subject research as it is based on the secondary analysis of publicly available data for public health research. Therefore, institutional review board approval is not required. Despite that, ethical issues were still considered: usernames were deidentified by removing account identity and other personal information such as email address and tweets cited and discussed were paraphrased (translated from French to English) thus nonreidentifiable.

Results

Extracted Tweets With Symptoms of Interest

A set of 307,290 French tweets posted between January 1, 2020, and March 31, 2020, was extracted. Further, 76,284 tweets were obtained with smell-related terms, and 237,999 tweets were obtained with taste-related terms. Symptom fuzzy matching identified 832 tweets containing at least 1 term potentially related to anosmia (90 tweets for period #1 and 742 for period #2) and 12,544 tweets containing at least 1 term potentially related to ageusia (8061 tweets for period #1 and 4483 for period #2). Similarly, 94 tweets potentially related to anosmia and 4721 tweets potentially related to ageusia were extracted for period #1 in 2019. For each period and each symptom of interest, 200 randomly selected tweets were manually annotated. If less than 200 tweets fit the criteria, all the tweets were manually reviewed and annotated, for example, 90 tweets for anosmia in period #1 in 2020. Figure 2 displays a flowchart of the extraction process and the associated volumes of tweets at each step and for each period of 2020.

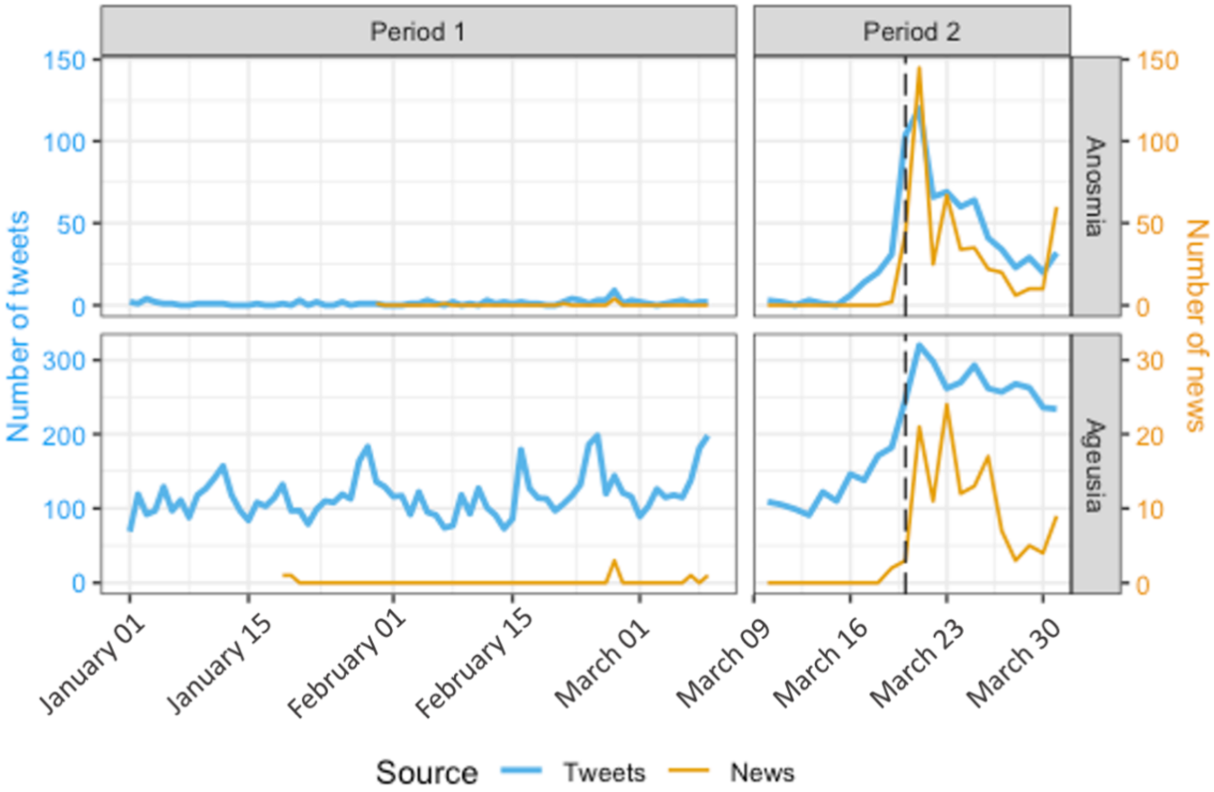
Figure 2. Flowchart of the analysis process.



Media Coverage

In total, 549 unique news articles related to anosmia, ageusia, or both were extracted. Only 11 (2%) of them were published in period #1. We compared the temporal evolution of the number of new articles containing anosmia and ageusia to that of tweets (Figure 3). A slight volume peak can be identified for anosmia on the 27th of February, which is Anosmia Awareness Day. A steep increase is observable for both sources on March 20 when these symptoms were officially linked to COVID-19 (denoted as a vertical perforated line). We can see that the increase in the number of publications happens a few days earlier for the tweets than for the news. However, computed cross-correlations show that series are correlated at 87% (anosmia) and 77% (ageusia) without lag, which suggests that there is no significant temporal gap between the declaration of anosmia and ageusia by patients on the internet and their coverage by media. Nevertheless, it shows a strong reactivity of Twitter to events.

Figure 3. Comparison of the evolution of anosmia- and ageusia-related tweets and news articles between January, 2020, and March 2020.



Detailed Analysis of Anosmia-Related Tweets

Regarding anosmia, among the 290 manually annotated anosmia-related tweets, 247 tweets contained genuine anosmia symptoms (49/90 and 198/200 tweets during periods 1 and 2, respectively), 36 were false positives (34/90 and 2/200 tweets during periods 1 and 2, respectively), and 7 tweets from period 1 were too short or unexplicit (eg, presence of irony) to conclude.

Among the 247 tweets containing mentions of the symptom anosmia:

- 165 tweets contained general information or discussions regarding anosmia, without any description of personal experiences (34 in period #1 and 131 in period #2). In total, 95/165 of these tweets were COVID-19–related. The remaining tweets (70/165) addressed topics such as the Anosmia Awareness Day, which is observed on February 27th every year as well as general information regarding anosmia and irony tweets.
- 67 tweets reported on experienced anosmia (15 in period #1 and 52 in period #2). Further, 32/67 tweets were COVID-19-related. Two-thirds of them provided information regarding co-occurring symptoms and about half of them provided the chronology of symptoms. The non–COVID-19–related tweets included mentions of long-term anosmia, and anosmia due to stuffy nose. Interestingly, there were 30 cases of anosmia without any explicit cause (11 in period #1 and 19 in period #2), that might be due to COVID-19 (eg, "I have not been able to smell anything for three days without having a stuffy nose").
- 15 tweets from period #2 could not be classified.

Regarding the relation to COVID-19, anosmia related to COVID-19 was only found during period #2 (136 tweets). A quarter of these tweets reported on experienced anosmia (32/136).

To summarize the results for anosmia: (1) the performance of the detection method was high, with a precision of 85% (247/290); (2) before March 10, 2020, only 90 tweets containing a term related with anosmia were identified, leading to manually reviewing a total of 290 tweets; (3) among these 290 tweets, a total of 32 tweets (11%) reported experiencing anosmia associated with COVID-19. All these tweets were found during period #2; (4) the World Anosmia Day, on February 27, generated more tweets on anosmia, unrelated to COVID-19; and (5) some qualitative information such as co-occurring symptoms and the temporality of symptoms could be found in a large proportion of tweets from patients with COVID-19.

Detailed Analysis of Ageusia-Related Tweets

As with anosmia, the first step consisted in solving polysemy issues, that is, differentiating between terms that are related to the symptom ageusia and other polysemic usages of the terms in other contexts (eg, no taste in fashion or music and really insipid food). The initial corpus of tweets with terms potentially related to ageusia was very noisy: among the 400 manually annotated ageusia-related tweets, only 59 tweets (14.8%) contained genuine ageusia symptoms (7 and 52 tweets during periods 1 and 2, respectively), whereas 330 were false positives (186 and 144 tweets during periods 1 and 2, respectively), and 11 tweets were too short or unexplicit (eg, ironic messages) to conclude (7 and 4 during periods 1 and 2, respectively).

Among the 59 tweets containing mentions of the symptom ageusia, (1) 12 tweets contained general information or discussions regarding ageusia, without any description of personal experiences (1 in period #1 and 11 in period #2); (2) 43 tweets reported on experienced loss or absence of taste (6 in period #1 and 37 in period #2). Eighteen tweets out of 43 (41.9%) were COVID-19–related; and (3) 4 tweets could not be classified (all of them after March 10).

Regarding the relation to COVID-19, a large proportion of tweets mentioning the symptom ageusia did not mention the cause (33/59). Ageusia related to COVID-19 was only found during period #2 in 42.4% of the tweets (25/59). Further, 72% of these tweets reported experienced ageusia (18/25), including 7 tweets on its negative psychological impact (eg, "(...) since friday, complete ageusia and anosmia. Life has become really sad, I mean really really sad" [March 24, 2020].

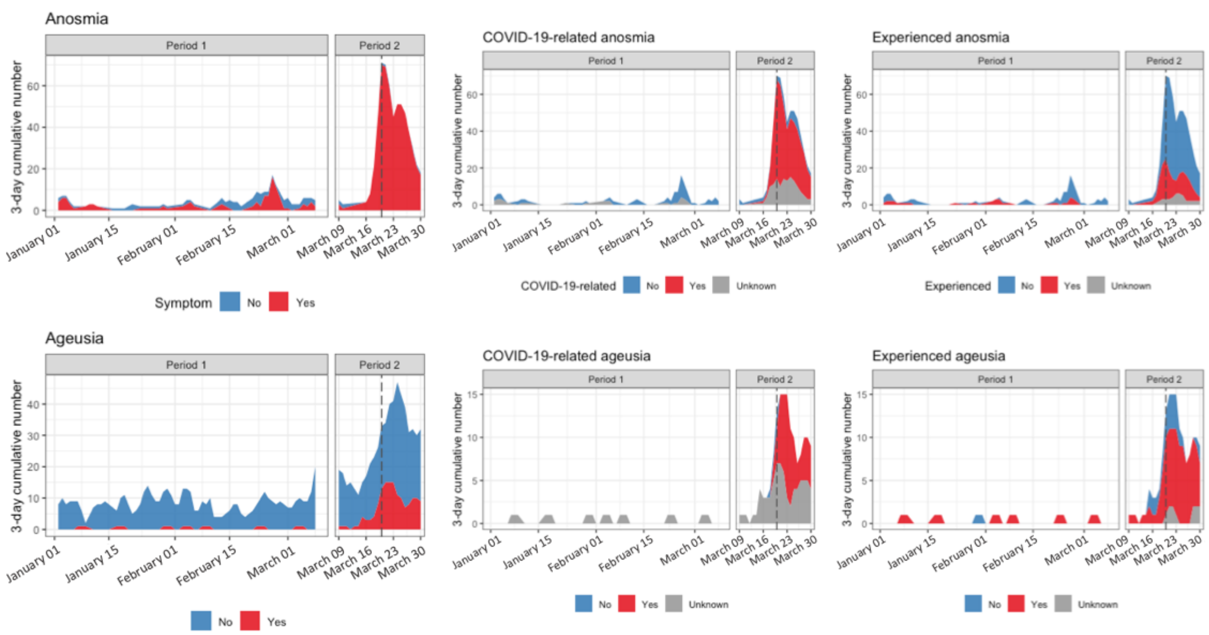
To summarize the results for ageusia: (1) the overall precision of the symptom detection method was very low, with less than 15% (59/400) of the tweets containing genuine ageusia symptoms. This was due to the presence of highly polysemic terms (eg, "goût"); (2) before March 10, there was no post reporting on ageusia experienced by a patient with COVID-19 while after March 10, we found 18 tweets out of 200 (9%) reporting on it; and (3) patients share qualitative information, such as the psychological negative impact of ageusia.

Time Evolution

In this section, the temporal evolution of declared symptoms and their links to COVID-19 are discussed. We provide results based on the tweets that have been manually reviewed for each symptom and each period. Evolution over time of symptoms declaration, link to COVID-19 and experience of the symptom by the author is studied and plotted.

The 3-day rolling total number of tweets in each annotated category was calculated for each symptom. The evolution of proportion for all annotation groups is shown in Figure 4. The proportion of tweets related to the symptom ageusia among ageusia-related tweets (Figure 4, left) increased from the beginning to the end of the second period but remained low (from 3.5%, 7/200 before March 10 to 26%, 52/200 after, see Figure 2). As for anosmia, the proportion of tweets with genuine symptoms was 54% (49/90) before March 10 contrasting with 99% (198/200) after (Figure 2 and Figure 4, left).

Figure 4. A 3-day cumulative time evolution of manually reviewed anosmia- and ageusia-related tweets. Occurrences of genuine anosmia and ageusia in all the tweets that were manually reviewed (left). True positive symptoms are broken down in proportions of the tweets expressing COVID-19–related symptoms or not (middle) and experience of symptoms (right).



The number of tweets relating experienced symptoms during period #1 was very low for both ageusia and anosmia (Figure 4, right). Over period #2, despite a global increase in the number of tweets, the proportion of tweets reporting on experienced symptoms did not drastically evolve, and remained very low for anosmia and important for ageusia. Regarding COVID-19–related mentions of symptoms (Figure 4, middle), tweets associated with COVID-19 were identified only from March 17 for anosmia and March 20 for ageusia. The proportion of tweets relating COVID-19 symptoms among tweets relating genuine symptoms reached 62% (21/34) during the last 10 days of period #2 for ageusia, contrasting with 22% (4/18) during the beginning of period #2. For anosmia, the proportion remained high and quite constant over the whole period #2, around 69% (136/198). These results show the important impact on Twitter of the media coverage of these symptoms, which confirms the high reactivity of Twitter to news, and the fact that

Twitter can be a source of interest to find qualitative information regarding an epidemic.

As these symptoms were confirmed as related to COVID-19 by Health authorities on Friday, March 20th, we performed a deeper analysis to assess if experienced COVID-19 related anosmia or ageusia (experienced=yes, COVID-19–related=yes) were detected on social media before this date, at the very beginning of period #2. The evolution of the 3-day cumulative number of tweets related to symptoms (all, COVID-19–related, experienced) showed a peak on March 20 and March 23 for anosmia and ageusia, respectively, but some tweets related to COVID-19 could be found a few days earlier.

Regarding anosmia, 6 tweets were found before March 20 (from March 17 to March 19) mentioning (explicitly or implicitly) experienced COVID-19 symptoms including anosmia. Further, 5 out of 6 tweets were written by health professionals about their patients (eg, "I started to get it the third time, generally I see one patient for anosmia per month, this time it is 3 in 48 hours" [March 17, 2020]). Further, 1 tweet was written by a web user who had experienced COVID-19 with anosmia ("I share this diagnosis because it is exactly what I have had for 2 days (...) some of my relative also have the same symptoms #anosmia # Covid" [March 19, 2020]). Some tweets posted before mentioned experienced anosmia but the relation to COVID-19 cannot be verified (eg, "has someone already had anosmia? #anosmia #help" [February 14, 2020], "I wear perfume but I don't smell it, I brush my teeth but I don't even taste the tooth path... It is really hard for me because I usually pay a lot of attention to smell" [January 23, 2020]).

Regarding ageusia, no tweet could be found before March 20 mentioning experienced COVID-19 symptoms including ageusia. Experienced ageusia with unexplicit cause could be found earlier, with unknown (eg, "even my favorite meals don't taste anything anymore" [January 8, 2020], "even water doesn't taste anything anymore" [January 18, 2020]) or high probability (eg, "Yes... I smelt something for the first time after 10 days, I still don't taste anything anymore" [January 19, 2020]) to be associated with COVID-19.

Comparison With 2019

To assess whether a signal about anosmia and ageusia could have been detected from Twitter, we compared the number of tweets mentioning experienced ageusia or anosmia with unexplicit causes (experienced=yes, COVID-19–related=unknown) posted during the same periods in 2019 and 2020. Using the same methodology as with 2020, we extracted 105 and 4701 tweets for anosmia and ageusia from January 1, 2019, to March 9, 2019, respectively. All anosmia-related tweets and 200 randomly sampled ageusia-related tweets were manually annotated.

Regarding anosmia, 75 tweets (respectively 49) mentioning the anosmia symptom were identified in 2019 (respectively 2020). Among them, 15 tweets (respectively 11) mentioning experienced anosmia without any cause were found.

Regarding ageusia, 8 tweets (respectively 7) mentioning the ageusia symptom were identified in 2019 (respectively 2020), and among them 7 tweets (respectively 6) mentioning experienced ageusia without any cause were found.

We could not identify any notable quantitative or qualitative difference between 2019 and 2020 that could have been early signals for COVID-19–related anosmia and ageusia in 2020.

Discussion

Principal Results

In this paper, our objective was to assess if the association between anosmia or ageusia and COVID-19 could have been detected earlier by mining social media, that is, before the relation was identified and published in journals. We tested such a hypothesis on a set of French tweets posted between January, 2020, and March 2020. We focused on these two symptoms for three reasons: (1) they were associated with COVID-19 later than other symptoms, these terms were even little known by the population before this pandemic, (2) they were present in a large proportion of the population with moderate symptoms in Europe but less in Asia earlier, and (3) they are quite uncommon outside of COVID-19, so their potential as early indicators of the disease could have helped differentiate COVID-19 from other diseases. Regarding the comparison with news articles, computed cross-correlations showed a strong correlation between the temporal evolution of the number of new articles containing anosmia and ageusia to that of tweets without lag, which suggests that there was no significant temporal gap between the declaration of anosmia and ageusia by patients on the internet and their coverage by media. Nevertheless, these results show a strong reactivity of Twitter to news. To obtain more insights, given the significant noise in symptom extraction, we performed a manual review of the anosmia- and ageusia-related tweets. We showed that only 6 tweets associating anosmia to COVID-19 could be identified among our set of manually annotated tweets before the news coverage, and these tweets were posted only 3 days before. Further, 5 out of 6 of these early warning messages (83%) were written by health professionals. No association could be found earlier. These results tend to demonstrate that Twitter could not have helped to identify the relation between COVID-19 and anosmia or ageusia during the early stage of the pandemic. Eventually, to assess if anosmia- and ageusia-related tweets with unexplicit causes from early 2020 could have been due to COVID-19, we compared their volume to the number of tweets mentioning experienced ageusia or anosmia with unexplicit causes from the same period from 2019. However, we could not identify any notable quantitative or qualitative difference between 2019 and 2020.

Most of the tweets associating anosmia or ageusia and COVID-19 were posted later, that is, after March 20th. These tweets contained qualitative information, such as psychological impact for ageusia and chronology of associated symptoms for anosmia, which could be of interest for further analyses. For example, it could be interesting to compare such insights with information from more traditional sources, especially as, according to Sarker et al [18], these symptoms were not reported in clinical studies at this time. We think the reactivity of Twitter can make it a useful tool to provide qualitative information that could complement what is already known in case of any health alert.

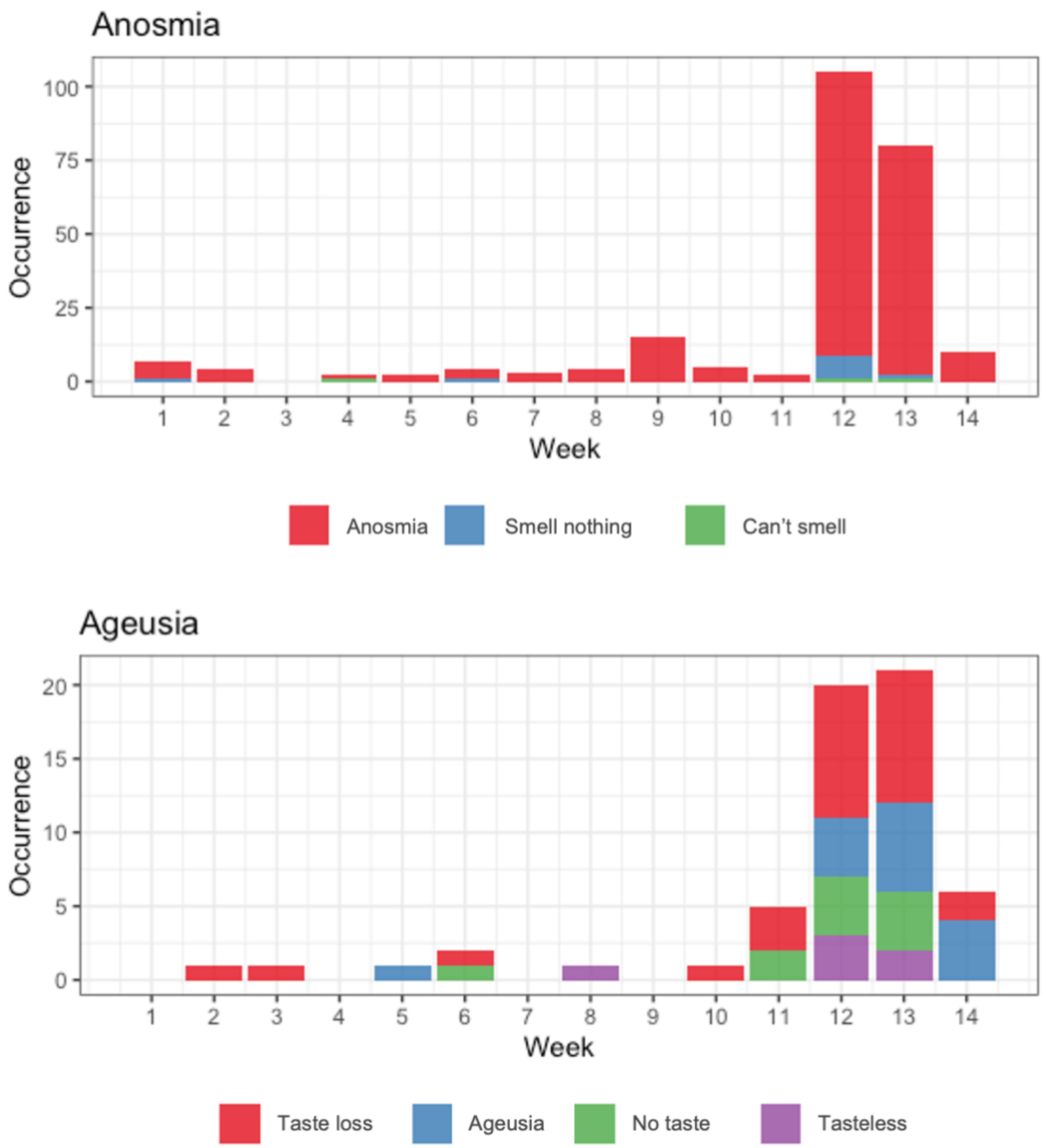
Comparison With Prior Work

A PubMed search using “social media” and “covid” as keywords retrieved 6764 publications, while adding the keyword “symptoms” restricted the results to 750 publications. Indeed, many prior studies explored the topic and sentiment expressed in social media (eg, [3,4,19,20]) misinformation and fake news spread in social media during COVID-19 (eg, [21-23]), and mental health (eg, [24,25]). As for symptoms, some efforts were made to improve the extraction of COVID-19–related symptoms from social media tweets for further investigation without a specific infoveillance objective (eg, [15,26]). Shen et al [27] performed an observational infoveillance study with a focus on the early stage of COVID-19 outbreak in China (before March 31, 2020) by assessing the relationship between reports of symptoms in Chinese social media platform Weibo and the daily confirmed infected cases. Ding et al [28] performed an infodemiological study to assess how discussion of symptoms changed over time and to identify correlations between frequency of the top 5 commonly mentioned symptoms in post and daily COVID-19 statistics (new cases and new deaths) in the United States. However, anosmia and ageusia were not mentioned in these 2 studies. Indeed, Shen et al [27] established a broad list of symptoms that did not include anosmia and ageusia probably due to the low prevalence of these symptoms observed in China during the first 3 months of 2020 [8], and Ding et al [28] focused on the top 5 commonly and most mentioned symptoms. Regarding anosmia and ageusia, Sarker et al [18] found that these 2 symptoms were frequently reported in English tweets, but not in clinical studies, and consequently that COVID-19 symptoms identified from Twitter may complement those identified in clinical settings. However, they performed their analysis in May 2020 and did not analyze the temporal characteristics of the distributions on Twitter, so it is unclear if reported anosmia and ageusia symptoms were found before the media coverage. Sudre et al [29] measured the association between different symptoms and COVID-19, based on data collected through 6 surveillance platforms and demonstrated that anosmia or ageusia was the strongest, most consistent symptom of COVID-19. However, they used the data collected between April, 2020, and June 2020, which is later after the official announcement. In this study, we conducted a comprehensive analysis of French tweets between January, 2020, and March 2020, focusing on the early detection and association of anosmia and ageusia with COVID-19. The temporal evolution of declared symptoms and their association with COVID-19 was assessed, allowing us to gain insights into the temporal trends and patterns of anosmia and ageusia discussions on Twitter. The comparison with media coverage and the comparison with the anosmia and ageusia related tweets in 2019 strengthened our conclusion.

Limitations

The exact terms “ageusia” and “anosmia” were quite unknown by the lay public before the overmediatization of these symptoms. Figure 5 shows the different words used to express COVID-19–related symptoms over our analysis period. The word “agueusie” per se is rarely used to denote this symptom before March 20 (week 12). After this date, its use takes up a growing proportion of the vocabulary used. This suggests an effect of mediatization on how patients express their symptoms, which shows a positive evolution in patients' health literacy through media outreach, also thanks to Twitter reactivity. In contrast, the medical term “anosmie” is the dominant term to express loss of sense of smell over the whole period. This can be explained by the tweets about the Anosmia Awareness Day on February 27, an international event that brings attention to the patients without the sense of smell, whatever the cause.

Figure 5. Anosmia- and ageusia-related terms for symptom detection.



The absence of use of precise medical terms in the tweets made it more difficult to extract these symptoms, especially before March 20, 2020. We considered synonymous terms for symptom detection, but these synonyms are much more polysemous: for example, “pas de goût” (no taste) can have different meanings, including taste, flavor, liking, fondness, and preference; consequently, on the evaluation subset, only 59 out of 400 (14.7%) ageusia-related tweets were actually dealing with the symptom ageusia.

Conclusions

Based on our analysis of French tweets, associations between COVID-19 and anosmia or ageusia by web users could have been found on Twitter only a few days before the official announcement but not during the early stage of the pandemic. The comparison between early 2020 and early 2019 showed no difference regarding the volumes of tweets, and no significant temporal gap between the declaration of anosmia and ageusia by patients on the internet and their coverage by media could be found.

Although we could not find early signals linking anosmia or ageusia with COVID-19 on Twitter, these tweets can potentially provide valuable information. First, we have demonstrated that Twitter was highly reactive to the news, with a significant increase after March 20. Second, our analysis showed that tweets contained qualitative information, such as psychological impact for ageusia and chronology of associated symptoms for anosmia. We think the reactivity of Twitter can make it a useful tool to provide qualitative information that could complement what is already known in case of any health alert.

Acknowledgments

The authors received no specific funding for this study. The first two authors, CF and MT contributed equally to this paper. CF, AB, and XC conceived this study. MT, PF, and SS contributed to the data curation. CF performed the annotation. CF, MT, PF, and XC contributed to the formal analysis, the methodology, and the visualization. CF, MT, PF, AB, and XC drafted this paper. All authors contributed to the data interpretation and critical revision of this paper.

Data Availability

The data sets generated and analyzed during this study are not publicly accessible in accordance with European Union General Data Protection Regulation requirements. However, some anonymized examples can be obtained from the corresponding author upon reasonable request.

Conflicts of Interest

MT, PF, AM, and SS are members of Kap Code, a start-up that owns Detec't.

Multimedia Appendix 1
List of anosmia and ageusia related keywords for the extraction.
[XLSX File (Microsoft Excel File), 6 KB - infodemiology_v3i1e41863_app1.xlsx]

References

1. Lian AT, Du J, Tang L. Using a machine learning approach to monitor COVID-19 Vaccine Adverse Events (VAE) from Twitter data. Vaccines (Basel) 2022;10(1):103 [FREE Full text] [doi: 10.3390/vaccines10010103] [Medline: 35062764]
2. 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]
3. 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;22(4):e19016 [FREE Full text] [doi: 10.2196/19016] [Medline: 32287039]
4. Schück S, Foulquié P, Mebarki A, Faviez C, Khadhar M, Texier N, et al. Concerns discussed on Chinese and French social media during the COVID-19 lockdown: comparative infodemiology study based on topic modeling. JMIR Form Res 2021;5(4):e23593 [FREE Full text] [doi: 10.2196/23593] [Medline: 33750736]
5. Lopreite M, Panzarasa P, Puliga M, Riccaboni M. Early warnings of COVID-19 outbreaks across Europe from social media. Sci Rep 2021;11(1):2147 [FREE Full text] [doi: 10.1038/s41598-021-81333-1] [Medline: 33495534]
6. Huang C, Xu X, Cai Y, Ge Q, Zeng G, Li X, et al. Mining the characteristics of COVID-19 patients in China: analysis of social media posts. J Med Internet Res 2020;22(5):e19087 [FREE Full text] [doi: 10.2196/19087] [Medline: 32401210]
7. Xydakis MS, Dehgani-Mobaraki P, Holbrook EH, Geisthoff UW, Bauer C, Hautefort C, et al. Smell and taste dysfunction in patients with COVID-19. Lancet Infect Dis 2020;20(9):1015-1016 [FREE Full text] [doi: 10.1016/S1473-3099(20)30293-0] [Medline: 32304629]
8. Mao L, Jin H, Wang M, Hu Y, Chen S, He Q, et al. Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. JAMA Neurol 2020;77(6):683-690 [FREE Full text] [doi: 10.1001/jamaneurol.2020.1127] [Medline: 32275288]
9. Vaira LA, Salzano G, Deiana G, De Riu G. Anosmia and ageusia: common findings in COVID-19 patients. Laryngoscope 2020;130(7):1787 [FREE Full text] [doi: 10.1002/lary.28692] [Medline: 32237238]
10. Anosmia. American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS). URL: https://www.entnet.org/covid-19/anosmia/ [accessed 2023-06-26]
11. Lechien JR, Chiesa-Estomba CM, De Siati DR, Horoi M, Le Bon SD, Rodriguez A, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. Eur Arch Otorhinolaryngol 2020;277(8):2251-2261 [FREE Full text] [doi: 10.1007/s00405-020-05965-1] [Medline: 32253535]
12. Giacomelli A, Pezzati L, Conti F, Bernacchia D, Siano M, Oreni L, et al. Self-reported olfactory and taste disorders in patients with severe acute respiratory coronavirus 2 infection: a cross-sectional study. Clin Infect Dis 2020;71(15):889-890 [FREE Full text] [doi: 10.1093/cid/ciaa330] [Medline: 32215618]
13. Chou SH, Beghi E, Helbok R, Moro E, Sampson J, Altamirano V, GCS-NeuroCOVID Consortium and ENERGY Consortium. Global incidence of neurological manifestations among patients hospitalized with COVID-19-A report for the GCS-NeuroCOVID consortium and the ENERGY consortium. JAMA Netw Open 2021;4(5):e2112131 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.12131] [Medline: 33974053]

14. Abdellaoui R, Schück S, Texier N, Burgun A. Filtering entities to optimize identification of adverse drug reaction from social media: how can the number of words between entities in the messages help? JMIR Public Health Surveill 2017;3(2):e36 [FREE Full text] [doi: 10.2196/publichealth.6577] [Medline: 28642212]
15. Faviez C, Foulquié P, Chen X, Mebarki A, Quennelle S, Texier N, et al. Fuzzy matching for symptom detection in tweets: application to COVID-19 during the first wave of the pandemic in France. Stud Health Technol Inform 2021;281:896-900. [doi: 10.3233/SHTI210308] [Medline: 34042803]
16. Brandwatch consumer research. Brandwatch. URL: https://www.brandwatch.com/products/consumer-research/features/ [accessed 2022-11-17]
17. Dalla V, Liudas G, Phillips PCB. Robust tests for white noise and cross-correlation. Econometric Theory 2020;38:1-29 [FREE Full text] [doi: 10.1017/S0266466620000341]
18. Sarker A, Lakamana S, Hogg-Bremer W, Xie A, Al-Garadi MA, Yang YC. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. J Am Med Inform Assoc 2020;27(8):1310-1315 [FREE Full text] [doi: 10.1093/jamia/ocaa116] [Medline: 32620975]
19. Chandrasekaran R, Mehta V, Valkunde T, Moustakas E. Topics, trends, and sentiments of tweets about the COVID-19 pandemic: temporal infoveillance study. J Med Internet Res 2020;22(10):e22624 [FREE Full text] [doi: 10.2196/22624] [Medline: 33006937]
20. De Santis E, Martino A, Rizzi A. An infoveillance system for detecting and tracking relevant topics from Italian tweets during the COVID-19 event. IEEE Access 2020;8:132527-132538 [FREE Full text] [doi: 10.1109/ACCESS.2020.3010033] [Medline: 34786279]
21. Broniatowski DA, Kerchner D, Farooq F, Huang X, Jamison AM, Dredze M, et al. Twitter and Facebook posts about COVID-19 are less likely to spread misinformation compared to other health topics. PLoS One 2022;17(1):e0261768 [FREE Full text] [doi: 10.1371/journal.pone.0261768] [Medline: 35020727]
22. 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]
23. Pierri F, DeVerna MR, Yang KC, Axelrod D, Bryden J, Menczer F. One year of COVID-19 vaccine misinformation on twitter: longitudinal study. J Med Internet Res 2023;25:e42227 [FREE Full text] [doi: 10.2196/42227] [Medline: 36735835]
24. Haddad JM, Macenski C, Mosier-Mills A, Hibara A, Kester K, Schneider M, et al. The impact of social media on college mental health during the COVID-19 pandemic: a multinational review of the existing literature. Curr Psychiatry Rep 2021;23(11):70 [FREE Full text] [doi: 10.1007/s11920-021-01288-y] [Medline: 34613542]
25. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
26. Sarker A. LexExp: a system for automatically expanding concept lexicons for noisy biomedical texts. Bioinformatics 2021;37(16):2499-2501 [FREE Full text] [doi: 10.1093/bioinformatics/btaa995] [Medline: 33244602]
27. Shen C, Chen A, Luo C, Zhang J, Feng B, Liao W. Using reports of symptoms and diagnoses on social media to predict COVID-19 case counts in Mainland China: observational infoveillance study. J Med Internet Res 2020;22(5):e19421 [FREE Full text] [doi: 10.2196/19421] [Medline: 32452804]
28. Ding Q, Massey D, Huang C, Grady CB, Lu Y, Cohen A, et al. Tracking self-reported symptoms and medical conditions on social media during the COVID-19 pandemic: infodemiological study. JMIR Public Health Surveill 2021;7(9):e29413 [FREE Full text] [doi: 10.2196/29413] [Medline: 34517338]
29. Sudre CH, Keshet A, Graham MS, Joshi AD, Shilo S, Rossman H, et al. Anosmia, ageusia, and other COVID-19-like symptoms in association with a positive SARS-CoV-2 test, across six national digital surveillance platforms: an observational study. Lancet Digit Health 2021;3(9):e577-e586 [FREE Full text] [doi: 10.1016/S2589-7500(21)00115-1] [Medline: 34305035]

Abbreviations

PT: preferred term

Edited by A Mavragani; submitted 12.08.22; peer-reviewed by A Chen, D Alexis, F Medina; comments to author 14.06.23; revised version received 29.06.23; accepted 01.08.23; published 25.09.23.

Please cite as:
Faviez C, Talmatkadi M, Foulquié P, Mebarki A, Schück S, Burgun A, Chen X
Assessment of the Early Detection of Anosmia and Ageusia Symptoms in COVID-19 on Twitter: Retrospective Study
JMIR Infodemiology 2023;3:e41863
URL: https://infodemiology.jmir.org/2023/1/e41863
doi:10.2196/41863
PMID:37643302

Original Paper

Corpus-Based Discourse Analysis of a Reddit Community of Users of Crystal Methamphetamine: Mixed Methods Study

Andrew Lustig[1*], MSc, MD; Gavin Brookes[2*], PhD

[1]Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[2]Department of Linguistics and English Language, Lancaster University, Lancaster, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Gavin Brookes, PhD
Department of Linguistics and English Language
Lancaster University
County South, Bailrigg
Lancaster, LA1 4YL
United Kingdom
Phone: 44 1524 593045
Email: g.brookes@lancaster.ac.uk

Abstract

Background: Methamphetamine is a highly addictive stimulant that affects the central nervous system. Crystal methamphetamine is a form of the drug resembling glass fragments or shiny bluish-white rocks that can be taken through smoking, swallowing, snorting, or injecting the powder once it has been dissolved in water or alcohol.

Objective: The objective of this study is to examine how identities are socially (discursively) constructed by people who use methamphetamine within a subreddit for people who regularly use crystal meth.

Methods: Using a mixed methods approach, we analyzed 1000 threads (318,422 words) from a subreddit for regular crystal meth users. The qualitative component of the analysis used concordancing and corpus-based discourse analysis to identify discursive themes informed by assemblage theory. The quantitative portion of the analysis used corpus linguistic techniques including keyword analysis to identify words occurring with statistically marked frequency in the corpus and collocation analysis to analyze their discursive context.

Results: Our findings reveal that the subreddit contributors use a rich and varied lexicon to describe crystal meth and other substances, ranging from a neuroscientific register (eg, methamphetamine and dopamine) to informal vernacular (eg, meth, dope, and fent) and commercial appellations (eg, Adderall and Seroquel). They also use linguistic resources to construct symbolic boundaries between different types of methamphetamine users, differentiating between the esteemed category of "functional addicts" and relegating others to the stigmatized category of "tweakers." In addition, contributors contest the dominant view that methamphetamine use inevitably leads to psychosis, arguing instead for a more nuanced understanding that considers the interplay of factors such as sleep deprivation, poor nutrition, and neglected hygiene.

Conclusions: The subreddit contributors' discourse offers a "set and setting" perspective, which provides a fresh viewpoint on drug-induced psychosis and can guide future harm reduction strategies and research. In contrast to this view, many previous studies overlook the real-world complexities of methamphetamine use, perhaps due to the use of controlled experimental settings. Actual drug use, intoxication, and addiction are complex, multifaceted, and elusive phenomena that defy straightforward characterization.

(JMIR Infodemiology 2023;3:e48189) doi:10.2196/48189

KEYWORDS

methamphetamine; social media; substance-related disorders; discourse analysis; mental health; mixed methods; corpus analysis; web-based health

Introduction

Methamphetamine is a highly addictive stimulant that affects the central nervous system. Crystal methamphetamine (or "crystal meth") is a form of the drug resembling glass fragments or shiny bluish-white rocks that can be taken through smoking, swallowing (pill), snorting, or injecting the powder once it has been dissolved in water or alcohol [1]. This study examines the social construction of identity among people who use methamphetamine. Using a corpus-based approach to discourse analysis (introduced in the Methods section), we interrogate a data set (or "corpus") comprising 500 threads from a subreddit for people who regularly use crystal meth. More specifically, our analysis takes an inductive approach that seeks to identify discursive themes within the threads, using exploratory methods within the corpus linguistics toolkit (introduced in more detail in the Methods section). Our interpretation of observed themes is informed by assemblage theory (introduced in the next section).

Methamphetamine has received extensive media, policy, and research attention that has emphasized its perceived association with violence, physical deterioration, and mental health problems [2]. Government agencies, scientists, and others have characterized methamphetamine as dangerous [3]. Chronic methamphetamine use has been implicated in memory loss [4], psychosis [5], cognitive deficits [6], mood disturbances [7], and dental problems [8]. Both the scholarly literature and the popular press [9] have positioned methamphetamine use and its effects as a public health crisis. Because its use is highly stigmatized in some parts of society [10], methamphetamine use is often carried out in secret or only with sympathetic fellow users to avoid the judgment of those who do not endorse such practices [2].

Social media affords an opportunity to openly and honestly discuss a variety of habits, practices, and health conditions, which may be difficult to address offline [11] due to the theorized "online disinhibition effect" [12]. Some social media platforms promote more radically honest disclosure by granting participants anonymity or pseudonymity [13,14], which may facilitate more frank disclosure about practices and attitudes toward drug use. The social media platform Reddit in particular permits the use of characteristic "throwaway" accounts which can be discarded after a single use. This allows users to abandon an account after posting information that they do not want to be associated with their other web-based activity or their offline identity [15]. Barak et al [16] have described how participation in digital support groups allows people with particular experiences to share information, provide and accept emotional support, socialize, and form relationships [16].

In contrast to the treatment of opioid use disorders, where there is a robust evidence base supporting the use of agonist therapy such as methadone and buprenorphine in treatment, a similar approach has not been found to be effective in the treatment of amphetamine-type stimulant use disorders. Recent meta-analyses have demonstrated that agonist therapy—that is, prescribing stimulants for the treatment of amphetamine-type stimulant use disorder—has no use in promoting sustained abstinence in people with methamphetamine use disorder [17]. Such efforts at treatment are based on the presumption that addiction exists independently of efforts to describe it. This medically informed model of methamphetamine addiction views it as a disease that is the result of disordered brain processes. Neurochemical [18], neurophysiological [19], and genetic [20] explanations are invoked in the literature to explain the effects of methamphetamine use, as well as the syndromes of addiction and dependence. These explanations embrace the disease model of addiction but ultimately fail to provide a coherent view of the complex nature of addiction [21].

An assemblage approach to amphetamine use offers some insight into conceptualizing substance use in its full complexity. This approach does not dispute the biological basis of addiction but adds to it by describing the manifold other factors that bear upon addictive behaviors. It conceptualizes addiction as a collection of heterogeneous bodies, objects, and forces that form a complex of addiction. Examples of such factors include the people one associates with while using drugs, the spaces where drugs are consumed, cultural mores regarding drug use, and economic considerations determining the price of drugs at the site of consumption. It conceptualizes drug use and drug addiction as a dynamic collection of heterogeneous elements including bodies, materials, spaces, forces, and signs [22]. Assemblage thinking serves to emphasize the real conditions in which drug problems emerge by way of the entire cast of human and nonhuman, distal, and proximate forces at work in such problems [22].

In this study, we provide what is, to our knowledge, the first examination of how methamphetamine use as a habit and as an addiction is discursively produced, reproduced, and contested on a subreddit specifically for people using methamphetamine. Because of its unstructured nature, this site permits contributors to discuss drug use in its full complexity without biases that may be introduced by structured or semistructured interview guides. In contrast to studies that recruit diagnostically homogenous groups, this subreddit is available to anyone with an internet connection. This admits the possibility of a heterogeneous, messy collection of contributors who relate to methamphetamine use in different ways.

Although much public health research about methamphetamine has focused on its harms, a more recent strand of research advocates a more comprehensive and inclusive approach to scholarship around drug use and addiction, contemplating both its harms and pleasures [23]. This study has the advantage of collecting data from an open forum (ie, a subreddit) where contributors can write about any topic pertaining to methamphetamine use of their choosing, in the process discursively constructing, theoretically, any kind of relationship between themselves and methamphetamine use. In this way, such interactional contexts provide promising sites for adopting a more comprehensive and inclusive approach to understanding this topic and how users relate to this in terms of their "normal life."

This study represents, to our knowledge, the first digital ethnography of a web-based community of methamphetamine users and provides novel insights into the discursive practices through which this community, and indeed the practice of methamphetamine use are socially constructed. In particular, the analysis reveals a different perspective on drug-induced psychosis compared to conventional scholarly literature, with methamphetamine presented as not necessarily inducing psychosis, when paired with regular sleep, good nutrition, and meaningful activities. We will argue that this perspective challenges the traditional view that directly links drug use to psychosis, highlighting the significance of context, setting, and other behaviors in these experiences. This research thus sheds new light on the discursive dynamics surrounding real-world drug use, hinting at its complexity and multifaceted nature. Following the reporting of the results, we consider how the insights produced through our analysis may guide harm reduction efforts and inform future research.

Methods

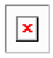
Overview

We used a corpus-based approach to discourse analysis [24,25]. Corpus linguistics is a methodology and a field of research that uses computational and statistical techniques to examine linguistic patterns in large, digitized bodies of naturally occurring language use [26,27]. The data analyzed in a corpus linguistic study is referred to as the "corpus" (pl. "corpora"). A corpus comprises a large body of naturally occurring language designed to represent a language or specific context of language use. The language or texts in a corpus are typically digitized to enable computational analysis. With computational assistance, those using corpus linguistic methods can identify recurrent patterns of language use in vast corpora quickly and reliably, and also easily perform complex statistical procedures to measure the significance and strength of the patterns observed.

The corpus we constructed for this study—referred to henceforth as the "methamphetamine corpus"—–represents interactions taking place in a subreddit for self-identified users of methamphetamine. We used Python (version 3.0; Python Software Foundation) to extract the 1000 most recent posts and all associated comments made prior to July 6, 2022, the date of data collection. Posts that were posted and subsequently deleted were not available for analysis (n=137). In total, 863 complete posts (370,452 words; see Textbox 1) were analyzed. The subreddit requires posters to solve a CAPTCHA before posting to prove they are human and not a bot.

At the time of sampling, the subreddit had approximately 107,000 members. To ensure that subreddit members' identities are protected as far as possible, no usernames or references to any other personally identifying information will be reproduced in the data extracts cited in this paper.

The corpus-based approach to discourse analysis used in this study is inductive and draws upon a combination of techniques such as keyword analysis, collocation analysis, and manual examination of concordances. We used Sketch Engine [28] to compile a list of keywords for the methamphetamine corpus. Keywords are essentially words that appear with a higher than expected frequency in the study corpus when compared to a reference corpus. The reference corpus typically represents some "benchmark" of language use against which the language in the study corpus can be compared. In this case, we obtained keywords for the methamphetamine corpus by comparing it against enTenTen20, a 36 billion-word corpus representing the English language used on the internet [29]. We used the simple math method of calculating keyness scores and we included all keywords with a keyness score ≥50 in our analysis. The simple math method of calculating keywords contains a smoothing parameter.



where $fpm_{rmfocus}$ is the normalized (per million) frequency of the word in the focus corpus; $fpm_{rmref}$is the normalized (per million) frequency of the word in the reference corpus, and N is a smoothing parameter.

We used N=1 as the smoothing parameter, which is the default value in Sketch Engine, for this analysis. As noted, we used keyword analysis inductively—as a "way in" to identify the characteristic language use and various themes and objects of the discourse produced by the subreddit users in their posts. As will be demonstrated in the Results section, the keywords produced through this step, when inspected manually in their contexts of use, could be grouped into a series of thematic and lexical categories, which reflect the kinds of themes and language use that were characteristic of the interactions in the threads in our corpus relative to general English language use on the internet.

We then examined the use of several keywords of interest using collocation analysis. Collocation is the linguistic device, whereby pairs of words (and sometimes wider networks of words) become bearers of meaning through repeated patterns of co-occurrence in natural language [26,27]. Through their collocational relationships, then, a word can develop meanings and exhibit particular "discourse prosodies" or carry particular ideologies [30]. Discourse prosodies broadly refer to the relationships between words and the textual contexts in which they are embedded. Collocation analysis involves identifying patterns of word co-occurrence, which are more frequent in the corpus than would be expected by chance. Collocational pairings were examined within a 5-word window on either side of the search term (or "node") and ranked using the cubed mutual information statistic. We analyzed collocations using *WordSmith Tools* (Lexical Analysis Software). We also analyzed collocational relationships between words through lexical bundles or n-grams [31]. These are fixed multiword expressions where 2 or more words occur together.

Finally, we followed up our identification of collocational pairings and n-grams with a qualitative discourse analysis, using concordancing, where we examined instances of a particular word or n-gram in context in the study corpus. For this portion of the study, we relied on the critical analysis of the concept of addiction employed by Keane [32]. This approach problematizes the binary opposition of health and disease and recognizes that addiction is not "a universal feature of human existence, but a culturally specific way of understanding, classifying and regulating particular problems of individual conduct." Following Fraser et al [33], we also draw upon the notion of collateral realities described by John Law. Collateral realities refer to realities that are the product of an ontological politics and assert that realities are done in practice and must be "done and redone again to remain stable." This approach regards addiction as an assemblage—a heterogeneous collection of objects, forces, affects, and tendencies—that is unstable and characterized by relations of exteriority.

The outputs of the corpus linguistic procedures were initially interpreted by the first author (AL), with the interpretations then being checked for plausibility by the second author (GB; as this was a plausibility check, this did not involve generating a formal intercoder reliability score for thematic interpretation). The discursive themes and wider patterns reported in the next section were agreed upon by both researchers. The first author (AL) is a psychiatrist. The second (GB) has training in linguistics (including corpus linguistics methodology) and health care communication.

Textbox 1. Quantitative profile of the methamphetamine corpus.

- Posts sampled (n=863)
- Total comments (n=8599)
- Total words (tokens) (n=370,452)
- Total unique words (types) (n=20,190)

Ethical Considerations

All the data used in our analysis were posted in a public space, which are available to any internet user without having to subscribe or login to the site. The subreddit permits users to contribute anonymously with a pseudonymous username that is not linked to their offline identities. Our examination of the posts constitutes what Eysenbach and Till refer to as "passive analysis" [34]. The institutional review board at The Centre for Addiction and Mental Health (University of Toronto) reviewed the proposed study and opined that it did not require formal approval.

Results

Overview

As described in the previous section, our analysis began with the identification of keywords. Table 1 presents the keywords obtained from our corpus through comparison with EnTenTen20, which is manually grouped into thematic and lexical categories based on inspection of their uses in the contexts of the subreddit posts. All *P* values were <.001.

The most statistically marked keyword in the corpus is *meth*. It occurs 1826 times in the corpus with a relative frequency of 4923 per million tokens. The most frequent bigram containing *meth* is the copular "meth is" which occurs 99 times. Examination of concordance lines containing this bigram reveals considerable ambivalence among group contributors regarding the relative merits and hazards of methamphetamine. When describing its harms, contributors focus on the embodied harms of the substance. Note that examples were selected because they were deemed by the researchers to be representative of the patterns being described. To retain their integrity, extracts are reproduced with the contributors' original spelling and punctuation.

> *That's how fucked **meth** is. We are it's victim for life.*
>
> *Even pharmaceutical **meth** is bad for you, the shit we use on the street is pretty much poisonous*
>
> *I would stop snorting and hot railing for a while. Your nose is very irritated, and **meth** is worse for your nose than most drugs. Pretty caustic*
>
> *In my opinion/experience, **meth** is more addictive and with higher abuse potential than adderall dexedrine, etc, not because it produces highs adderall cannot, but bc 1) its cheaper and easier to get, so you'll have a larger stash, and less incentive to not waste it since you won't be forced to wait till end of month to re up. 2) you can smoke it.*

In other instances, contributors use the "meth is" bigram to highlight the corporeal pleasures of methamphetamine use.

> *....**meth** is fun and leads to good sex and lots of energy is all I know*
>
> *There have been studies that found as long as you don't let it fuck with sleeping, eating and taking care of yourself in general, **meth** is no worse for you than caffeine is in adults with no predisposed health issues.*
>
> *I definitely wouldn't say anyone SHOULD do meth, but from what I know about coke - which admittedly, I haven't done coke - **meth** is a lot more sustainable.*

Meth is by far the most common lexicalization used by contributors to describe methamphetamine but a richly textured lexicon permits nuanced descriptions of substances. Although synonymous with meth, the longer and more scientific-sounding synonym *methamphetamine* is sometimes invoked by exponents of a more formal, neuroscientific kind of discourse.

> *Probably because **methamphetamine** does not cause hallucinations. They come from lack of sleep or from stimulant psychosis not from the pharmacokinetics of **methamphetamine**.*
>
> *I know that they sometimes actually literally will prescribe **methamphetamine** HCl in pill form for certain ADHD patients, that is, Desoxyn in two 20-25mg spaced doses.*

Even less frequently, contributors use *desoxyn* as a synonym for *methamphetamine*. Desoxyn is a brand name for pharmaceutical methamphetamine. Almost all instances of this lexicalization of the substance refer to pharmaceutical and other medical uses, frequently in the context of treatment for attention-deficit/hyperactivity disorder (ADHD).

> *ADHD is treated with desoxyn in like 5 mg - 40 mg doses or some shit like that, once per day.*

> *Went to docs and got diagnosed lol Adderall is good but **desoxyn** is the shit! Unfortunately its not easy to get.*

The 3 keywords *meth*, *methamphetamine*, and *desoxyn* refer to the same chemical compound. However, these synonyms allow contributors to invoke or obscure a neurochemical discourse and a more or less formal register. Although the material referent is identical, the sense of each lexicalization is distinct. These divergent forms of expression permit nuanced relations between the material elements of the assemblage in question.

The use of different lexical types to describe methamphetamine highlights the different senses. As Keane notes, "In Plato's Pharmacy Derrida highlights the undecidability and ambiguity of the Greek word pharmakon, which can mean both poison and cure" [32]. The tension evoked by the same compound functioning as both a toxin and a therapeutic or pleasure-causing agent is managed by using different lexical types, each of which exhibits a distinct discourse prosody.

Finally, the keyword *dope* is also used to refer to methamphetamine in the study corpus. However, of all lexicalizations of methamphetamine, *dope* is most highly polysemous, contested, and at times, ambiguous.

In a thread where a contributor refers to methamphetamine as *dope*, an interlocutor responds: "meth ≠ dope. dope is heroin or other strong downers."

To this, another poster replies:

> *What region? Here in the Midwest, meth is absolutely 100% **dope**. **Dope** is definitely not heroin or any downers. **Dope** = shit = crystal, at least here. And that's the standard In every circle I've been involved with - old heads, heroin addicts, tweakers, teachers, shooters, smokers (although people who only fuck w weed/occasionally psychedelics like to call weed **dope**, and if pressed, will usually agree with you. But that's because they literally have no idea, tho)*

Depending, then, on the context of the interaction, *dope* can refer to methamphetamine, but it can also refer to marijuana or to cocaine, heroin, or other opioids. In some instances, *dope* does not refer to a substance at all, but to a positive quality as in:

> *That color is so dope. Nice.*

> *They look dope, very nice lamp-work*

Adderall also appears as a keyword and occurs 90 times in the corpus. Like desoxyn, *Adderall* is a commercial trade name and refers to a preparation of dextroamphetamine and racemic amphetamine. It is used therapeutically in the treatment of ADHD and narcolepsy. Table 2 shows the collocates of this keyword.

Collocation analysis of *Adderall* reveals that *meth* is its only lexical (ie, nongrammatical) collocate. This pairing occurs because contributors compare *Adderall* and *meth:*

> *Dude, I flat out have conversations with voices, it's like my favorite musicians and shit telepathically speaking to me, I enjoy to hear them, and meth and **adderall** all are similar but I feel way more euphoric on meth.*

> ***Adderall** where I was, was going for fucking $10 for the 30mg shit. A gram of good meth is $40. Met someone who noticed I took the addies recreationally as well to study. Asked. Told me that **Adderall** is just smart meth.*

Table 1. Keywords grouped into thematic or lexical categories and ranked by keyness score.

Thematic or lexical category	Associated keywords (raw frequency, keyness score)
Substances and chemicals	*meth* (1826, 1802); *adderall* (90, 179); *dope* (211, 121); *benzo* (51, 116); *amphetamine* (65, 100); *fent* (37, 99); *dopamine* (96, 87); *stim* (38, 80); *opiate* (66, 75); *methamphetamine* (65, 73); *n-iso* (24, 66); *fentanyl* (48, 65); *coke* (118, 58); *desoxyn* (21, 57); *Seroquel* (20, 50); *stimulant* (49, 50)
Psychological states	*psychosis* (150, 150); *comedown* (46, 117); *sober* (159, 89); *euphoria* (40, 50)
Descriptions of consumers of methamphetamine	*tweaker* (118, 288); *addict* (213, 63)
Mode/implements of administration	*roa* (68, 141); *snort* (178, 136); *bong* (78, 118); *boof* (36, 94); *crackback* (34, 92); *shard* (102, 87); *boofing* (29, 79); *tho* (141, 78); *redose* (20, 55); *hotrail* (19, 52); *vape* (35, 52); *smoke* (675, 52)
Internet-related lexis	*lmao* (137, 230); *idk* (108, 156); *subreddit* (42, 86); *lol* (630, 84); *Reddit* (84, 79); *cuz* (76, 59), *tbh* (46, 69); *lmfao* (25, 61); *wtf* (57, 53); *ur* (109, 52); *haha* (99, 50)
Swear words	*shit* (952, 146); *fuck* (975, 88); *fuckin* (66, 86)
Forms of address	*bro* (168, 107); *homie* (44, 89); *bruh* (29, 75)

Table 2. Collocates of *Adderall* ranked by cubed mutual information.

Rank	Collocate	Collocation frequency	Cubed mutual information
1	*and*	27	13.04
2	*was*	11	11.53
3	*meth*	11	11.30
4	*would*	7	10.68
5	*for*	11	10.54

Affective Assemblages

As noted by Duff [22], "[a]ll drug assemblages should be regarded as affective entities in as much as affective processes are at least partially responsible for the formation of the assemblage." Used in this sense, *affect* refers to emotional states and also refers to a body's capacity to act upon or be acted upon, by other bodies. Our keyword analysis identified 4 lexical types which, when analyzed in their wider contexts of use, seemed to correspond to this broad definition of affect: *psychosis*, *comedown*, *euphoria*, and *sober*.

Psychosis

Table 3 shows the collocates of the first of the affect keywords, *psychosis*.

Collocation analysis of *psychosis* reveals that *induced* is its most frequent collocate, typically occurring in the L1 position (ie, immediately preceding the node) to form the bigram induced *psychosis*. The term *drug-induced psychosis* is a clausal nominalization in that it converts a process into a noun and identifies a drug as the cause of psychosis. This diagnostic category and attendant causal explanation are consistent with both the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR) and International Classification of Diseases 11th Revision (ICD-11) diagnostic classification systems. The function of this is to reproduce, within this context, the way in which methamphetamine and addiction are discursively represented in some research and policy practices [33].

> *Finally found medications and treatment that worked and officially diagnosed schizophrenic, since i was sober when the long lasting fully documented episode happend it wasnt considered drug induced psychosis.*

Recent editions of the DSM have strived to be atheoretical [35] and to engage in "pure" description without invoking etiological explanations. However, attempts at description necessarily invoke assumptions regarding pathology whether implicit or explicit. In the case of substance-induced psychosis, the etiological formulation is included in the name. The use of the term *drug-induced psychosis* enacts the collateral reality that drugs cause psychosis.

By using this terminology, this excerpt provides an example of cases in which users enact the collateral reality that methamphetamine use causes psychosis. At the same time, it challenges that reality by alluding to the performative function of psychiatric diagnosis by using the modalizing mental verb *considered*. In so doing, this post challenges the reification of the diagnostic category and acknowledges that diagnoses are at least partially made in practice.

The causes of psychosis are also contested in the corpus. A medicalizing discourse of psychosis views it as a simple cause and effect, whereby methamphetamine use inevitably leads to psychosis. In contrast, contributors view the development of psychosis as resulting from a more complex assemblage that includes excessive methamphetamine use along with other habits and practices.

> *I used to get way too high and around day 3 of being up and overamped, I'd sit there and be unable to move. It's just a matter of moderation and how many days somebody is up. I eventually went into a bad **psychosis** from not staying hydrated, not eating, not sleeping, and not moving. Went deep into my head. But tonight for example I'm moderating what I smoke and I've done all my laundry and cleaned my house. For me, less is more.*
>
> *I'm off the shit finally, but I know that **psychosis** is caused by nutrition and rest deficiency.*
>
> *I think the longest run I've done is 11 days without sleep; but I was able to fend off **psychosis** by eating three meals a day and shutting down for 4-5 hours a day where it's laying in bed with low stimuli, so even if I'm not sleeping, I'm resting my spine, joints, and mind.*
>
> *As said a million times over, the best way to fend off **psychosis** is regular sleep and healthy meals.*

These posts present psychosis as the result of a complex array of corporeal factors, such as hydration, food, sleep, and activity level. They formulate psychosis as part of an assemblage that narrows and limits the range of activities that a body can do. Descriptions of positive experiences of methamphetamine use endorse an expansion of capacity. However constructions of experiences and activities situate a "bad psychosis" as being part of a diminished capacity to act. Contributors also use the verbal construction *fend off* to invoke the metaphor of battling an assailant. Metaphor is pervasive in the language used to reify mental states and forms of distress [36].

The notion that sleep deprivation is a necessary condition for methamphetamine use to cause psychosis is highly prevalent in our corpus. This belief is directly at odds with generally accepted medical knowledge, and a recently published scholarly review of methamphetamine-induced psychosis unequivocally stated that "methamphetamine and amphetamine can result in a paranoid psychosis…and the syndrome is not due to sleep deprivation” [37].

Although most descriptions of psychosis are negative, a few are positive:

> *That is cool psychosis and a lot of people don't say it if you get sleep I'm like you I like day three and four the best if you learne to handle it you enjoy the* ***psychosis*** *it's fun for me anyway the whole place turns into what I call non-comporeal beings.*

> *When I started using Meth, I enjoyed the* ***psychosis****.*

Table 3. Collocates of *psychosis*, ranked by cubed mutual information.

Rank	Collocate	Collocation frequency	Cubed mutual information
1	*induced*	6	14.55
2	*the*	43	14.21
3	*experiencing*	5	13.99
4	*and*	25	12.02
5	*from*	12	11.81

Comedown

Contributors use the term *comedown* to refer to the affective and corporeal experiences that follow a period of methamphetamine use. In describing the nature of the methamphetamine comedown, contributors refer to other substances as a point of comparison.

> ***COMEDOWN****: first let's state that coke made me feel empty and anxious, molly made me feel depressed and meth makes me feel very on edge and paranoid and very nervous.*

> *You get this weird gut anxiety which is different compared to other comedowns and it can be off setting but what goes up just comedown especially at this level.*

Contributors discuss various strategies to mitigate the intensity of the comedown. Some contributors recommend vitamins, energy drinks, or nutritional supplements.

> *In future, oral dosing lasts a lot longer, the* ***comedown*** *is a lot smoother, and it's a lot less obsessive (you just eat 20mg instead of spending hours upon hours chasing the rush).*

> *You need some* ***comedown*** *candy's and to accept that fact that. 1. Sleep is really the only thing that will help you 2. You're going down for a day or two.*

> *I'm fucking spun, I grabbed some Tina, some raw coke, and a Xan for the* ***comedown****.*

> *I'm interested in knowing since as ssri it makes serotonin relapse slower, may it increase the high or make the* ***comedown*** *easier when taken single dose after meth in not that high doses?*

While accounts of methamphetamine addiction enacted in research and policy emphasize risks and risk mitigation, consumer accounts emphasize the role of pleasure in narratives of psychological addiction to methamphetamine [38]. Contributors used the keyword *euphoria* to describe the intense pleasure associated with methamphetamine ingestion. In response to a thread asking: “Should I sleep or load one more bowl .”

One contributor responds:

> *Man if you've been up for more than a day or two already it should always be sleep. Redosing isn't even fun at that point, just keeps you awake and any feelings of* ***euphoria*** *have gone to bed.*

This response navigates the tension between binging on methamphetamine, which lies outside the responsible ideals of normal life and the more responsibilized agency and subjectivity of such normative temporal routines as sleep. Contributors recruit discourses of neuroscience invoking the action of neurotransmitters to account for *euphoria* of the lack thereof,

> *You don't feel any* ***euphoria*** *and only feel tweaky because all the dopamine has been converted into adrenaline.”*

Another contributor invokes a neuroscientific discourse to account for the euphoric effects of methamphetamine:

> *But initially, and if taking a healthy dose, its similar to most add meds just with a longer half life. Also, meth penetrates the BBB way faster and harder so you get the added euphoria.*

By invoking the term *half life*, borrowed from pharmacology, and a reference to the blood–brain barrier, this post invokes a neuroscientific discourse in explaining the euphoric effect of methamphetamine use. As previously noted [33], consumer accounts of methamphetamine use enact several collateral realities, one of which is that methamphetamine use affects the brain. It also positions the contributor as belonging to a community of knowledge that includes consumers as well as scientists [33].

> *There is* ***euphoria*** *but it's not the main aspect of it all to me. More than anything, meth feels like I want to keep doing whatever it is I am currently doing for the whole day.*

This post highlights that the motivation for methamphetamine use extends beyond an experience of pleasure as a benefit. It apprehends the “dynamics of space, embodiment and practice” [39]. The post speaks to the pleasure of practice or activities enabled by methamphetamine use. Although methamphetamine use is characterized by a sought-after, intensely pleasurable corporeal state, in order to construct the assemblage of methamphetamine use, it must occur in concert with practice.

Several of the keywords including *snort, boof, hotrail, vape,* and *smoke* refer to different modes of administering methamphetamine. *Snort* refers to nasal insufflation of methamphetamine crystal. *Boof* refers to rectal administration. *Hotrail* refers to nasal inhalation of methamphetamine smoke. To *vape* refers to the inhalation of a vapor formed by heating methamphetamine to the point that it evaporates but not so hot that it burns. Finally, *smoking* refers to inhaling smoke formed by the combustion of methamphetamine.

Redose refers to administering methamphetamine sometime, typically several hours, after a previous dose, when the effects of the earlier dose begin to wear off.

> *I am losing control. I compulsively **redose** and neglect important human needs such as hydration, hygiene, nourishment, among other things.*
>
> *I usually just IV now. It goes from pure excitement to HOLY FUCK INTENSE as I shoot up, and it stays intense or an hour or two, tapers off for a good while and then I'm straight. I do have the urge to **redose** after the intense part wears off.*
>
> *Beware though, the transformation into a mechanized tweakazoid **redose** robot isn't limited to the ritual of freebasing.*

The implication that *redose* involves the repetitive ingestion of methamphetamine hints at the habitual pattern of some methamphetamine use. These posts suggest that using habitually opens the door to a pattern of use where using overshadows other social and occupational activities, leading to a problematic pattern of use.

Contributors associate the corporeal practices associated with different routes of administration with variations in the experience of intoxication. One post reads:

> *I am gonna stop smoking this shit.*
>
> *It's such a waste and it's very fiendish. I can easily smoke 2 grams in 1 night. Plus I am waay to obsessed with forming a perfect crack back. It fucks with my head tbh.*

A response to the thread offers:

> *Yup. It turns you into a re-dosing robot, long after the **euphoria** is gone. Don't be a slave to the drug. Consider taking it orally in one dose instead.*

The suggestion connects the corporeal practice of administration with the affective experience of intoxication.

Lexicalizing Methamphetamine Users

Contributors use the term *addict* and *tweaker* to refer to people who habitually use methamphetamine.

The most statistically marked collocate of an *addict* is *functional*, occurring almost exclusively in the L1 position to form the bigram, *functional addict* (see Table 4).

> *First thing you gotta do is get a morning routine. It's possible to be a **functional addict** you just gotta force yourself to take breaks so you can realize things like this.*
>
> *Sure it's possible to be a **functional addict**, or even an occasional user.*

Here again, contributors emphasize that stabilizing routines tend to promote an assemblage of less destructive use.

Tweaker is used to refer to people who habitually use methamphetamine. It refers to a more stigmatized subject than an addict. Posts referencing tweakers equate this identity with mental illness:

> *I had a **tweaker** girl tell me some wild story about her first hand experience with the fent cut meth, but she seemed like a nutball so I took what she said with a grain of salt at the time.*
>
> *thank you. i didn't make it though, my friend who is like the most retarded, mentally ill, **tweaker** out there somehow made this thing. if i ever get around to making my own (which i probably should bc i dont trust anything that comes from him) then i'll use that tip ;)*
>
> *Ya paranoid **tweaker** fuck.*

Posts referencing tweaker also focus on the embodied stigmata of chronic methamphetamine use including psychomotor agitation, grooming, and skin lesions:

> *Just don't look like a **tweaker** and you're set.*
>
> *If you've ever met a **tweaker** who sleeps you'll see that they're actually not some twitchy dude looking over his shoulders at 1,000Mph itching his friends head because his back has bugs coming out of his skin.*
>
> *That's what separates a thriving user from a **tweaker** or junkie, self preservation and self respect.*

Table 4. Collocates of *addict* ranked by cubed mutual information.

Rank	Collocate	Collocation frequency	Cubed mutual information
1	*functional*	12	17.26
2	*meth*	16	12.80
3	*drug*	8	12.05
4	*full*	5	11.99
5	*and*	18	11.16

Discussion

Principal Findings

The discursive patterns identified in this corpus of interactions problematize several of the collateral realities enacted by pre-existing scholarly research on methamphetamine use. As noted by Dwyer and Moore [40], extant drug education and research texts adopt a "pharmacology as destiny" approach to the relationship between methamphetamine and psychosis. This approach privileges pharmacology as the key determinant of experienced effects and discounts the importance of set, setting [41], and other contextual factors in drug experiences generally and psychosis in particular. It foregrounds a Humean constant conjunction [42] between methamphetamine use and psychosis. These discursive practices construct conventional ontological understandings of methamphetamine as anterior, singular, stable, and definite, and of its psychological effects as indicative of pathology. In contrast, the first-hand descriptions of psychosis provided by contributors conceptualize psychosis as existing as part of a complex assemblage. They posit that when methamphetamine use occurs in a complex of behaviors that includes regular sleep, healthy nutrition, meaningful activities, and a moderate degree of stimulation, then it does not necessarily include psychosis. This perspective is at odds with that of the existing scholarly literature on amphetamine-induced psychosis [43]. This insight may be helpful in guiding harm reduction efforts relating to methamphetamine psychosis by directing attention to the activities that occur alongside methamphetamine use, rather than solely toward the quantity used or the route of administration. This also has important implications for future research which may attempt to disentangle the complex relationship between methamphetamine use, psychosis, and the array of other factors that bear upon them.

The reasons for this apparent conflict in perspectives are not clear and merit further investigation. Studies concluding that amphetamines cause psychosis were conducted in laboratory settings with controlled conditions. In the real world, methamphetamine use is a complex, multifaceted, and changing phenomenon. As suggested by Law [44], experimental methods are ill-suited to address things that are emotional, ephemeral, elusive, or indistinct. We submit that real-world drug use, intoxication, and addiction are messy phenomena that defy clear and definite characterization or explanation.

This study found that methamphetamine users can choose from a rich and varied lexicon to refer to methamphetamine and other substances. In particular, lexical choices highlight that amphetamines can be seen as a life-giving elixir, an invigorating tonic, a prescription medication sanctioned by medical authorities, or a poison that harms and robs those who use it of agency. Contributors perform these different amphetamine objects by using different lexical items and by situating them in purposive discursive contexts. These include informal terms such as *meth* and *dope*, as well as more formal terms including *methamphetamine*. Use of the lexical item *dope* can be used to signal "in group" status in the community. Users also invoked terms derived from the commercial production, marketing, and sale of stimulants. In doing so, they frequently invoked a more formal neuroscientific register and positioned themselves as members of a community of knowledge that includes scientists as well as drug consumers. This may be a feature of the kind of web-based context represented in the corpus data assembled for this study since such environments have provided the contextual basis for observations of the discursive enactment of so-called "expert patient" identities in previous studies of digital health-related communication [25,45,46].

Contributors described affective states associated with the use and the cessation of methamphetamine use. Affective states such as euphoria represent the bliss of amphetamine ingestion, while comedown represents its opposite. Contributors construed these two affective states as existing in tension, each opposite being a necessary counterpoint. Psychosis was generally framed as undesirable and contributors traded in strategies to avoid psychosis. However, a minority countervailing viewpoint challenged that characterization and identified that psychosis can be an exciting novel experience.

Contributors also use lexical choices to construct symbolic boundaries [47] between various types of methamphetamine users. As noted by Copes et al [48], people who use methamphetamine are highly stigmatized and make distinctions between different types of users based on their ability to maintain their physical appearance, mental health, and daily obligations. The use of terms such as *functional* addict is presented as desirable in contrast to *tweakers* who are presented in stigmatized terms as having lost control of their use. This echoes previous findings that individuals negotiate the boundaries of acceptable drug use by aligning their identity with the virtues of autonomy, control, and responsibility and distancing themselves from such stigmatized subject positions as addict or junkie [2].

Limitations

A limitation of our analysis is that we have only studied 1 subreddit. Future research should consider a wider range of interactional contexts—including other sites and sites using languages other than English—in order to ascertain the extent to which the patterns observed in this study might be a feature of the culture and subculture (ie, this particular subreddit) we studied, and the extent to which they might be considered part of a broader discourse around methamphetamine use. Furthermore, it is possible that other subreddits on meth use exist on topics that may have different positive and negative sentiments about use and behavioral-related topics or may have different types of contributing users. This brings us to another limitation, while the anonymity of the site we studied was, on one hand, an advantage for the purposes of our analysis (potentially giving rise to more candid disclosures, as discussed earlier), a limitation of such anonymity is that it is not possible for us to reliably assess the sociodemographic make-up of the users represented by our data, nor to draw comparisons between groups. Addressing these limitations would require to use of different—possibly even purposively designed—data sets. It is our intention that the findings reported in this paper will provide a springboard for such work and a wider program of research to better understand the discursive dynamics of addictive substance use.

Conclusions

In summary, this study analyzed digital discussions among methamphetamine users, which finds that the contributions making up the subreddit under study frequently challenged a preexisting research view that links methamphetamine use directly with psychosis. The contributors whose discourse we studied presented a complex array of factors including regular sleep, nutrition, meaningful activities, and stimulation, which, when combined with methamphetamine use, may not necessarily lead to psychosis. This "set and setting" perspective offers a fresh viewpoint on drug-induced psychosis and can guide future harm reduction strategies and research. Many previous studies might thus be viewed as overlooking the real-world complexities of methamphetamine use, perhaps due to the use of controlled experimental settings. Actual drug use, intoxication, and addiction are complex, multifaceted, and elusive phenomena that defy straightforward characterization. Notably, we found evidence of lexical choices whose use indicated a delineation of symbolic boundaries between types of methamphetamine users, establishing a societal hierarchy within the user community.

Conflicts of Interest

None declared.

References

1. Methamphetamine DrugFacts. National Institute on Drug Abuse. 2019. URL: https://nida.nih.gov/publications/drugfacts/methamphetamine [accessed 2023-09-08]
2. Green R, Moore D. "Meth circles" and "pipe pirates": crystal methamphetamine smoking and identity management among a social network of young adults. Subst Use Misuse 2013;48(9):691-701. [doi: 10.3109/10826084.2013.782046] [Medline: 23581505]
3. Methamphetamine research report: overview. National Institute on Drug Abuse. 2021. URL: https://nida.nih.gov/publications/research-reports/methamphetamine/overview [accessed 2023-09-08]
4. Kim B, Yun J, Park B. Methamphetamine-induced neuronal damage: neurotoxicity and neuroinflammation. Biomol Ther (Seoul) 2020;28(5):381-388 [FREE Full text] [doi: 10.4062/biomolther.2020.044] [Medline: 32668144]
5. Glasner-Edwards S, Mooney LJ. Methamphetamine psychosis: epidemiology and management. CNS Drugs 2014;28(12):1115-1126 [FREE Full text] [doi: 10.1007/s40263-014-0209-8] [Medline: 25373627]
6. Simon SL, Domier C, Carnell J, Brethen P, Rawson R, Ling W. Cognitive impairment in individuals currently using methamphetamine. Am J Addict 2000;9(3):222-231. [doi: 10.1080/10550490050148053] [Medline: 11000918]
7. Salo R, Flower K, Kielstein A, Leamon MH, Nordahl TE, Galloway GP. Psychiatric comorbidity in methamphetamine dependence. Psychiatry Res 2011;186(2-3):356-361 [FREE Full text] [doi: 10.1016/j.psychres.2010.09.014] [Medline: 21055832]
8. Shetty V, Mooney LJ, Zigler CM, Belin TR, Murphy D, Rawson R. The relationship between methamphetamine use and increased dental disease. J Am Dent Assoc 2010;141(3):307-318 [FREE Full text] [doi: 10.14219/jada.archive.2010.0165] [Medline: 20194387]
9. Weiland N, Sanger-Katz M. Overdose deaths continue rising, with fentanyl and meth key culprits. The New York Times. 2022. URL: https://www.nytimes.com/2022/05/11/us/politics/overdose-deaths-fentanyl-meth.html [accessed 2023-09-08]
10. Fraser S, Pienaar K, Dilkes-Frayne E, Moore D, Kokanovic R, Treloar C, et al. Addiction stigma and the biopolitics of liberal modernity: a qualitative analysis. Int J Drug Policy 2017;44:192-201. [doi: 10.1016/j.drugpo.2017.02.005] [Medline: 28366599]
11. Pavalanathan U, De Choudhury M. Identity management and mental health discourse in social media. New York, NY, US: Association for Computing Machinery; 2015 Presented at: Proceedings of the 24th International Conference on World Wide Web; May 18-22, 2015; Florence, Italy p. 315-321. [doi: 10.1145/2740908.2743049]
12. Suler J. The online disinhibition effect. Cyberpsychol Behav 2004;7(3):321-326. [doi: 10.1089/1094931041291295] [Medline: 15257832]
13. Barratt MJ. Discussing illicit drugs in public internet forums: visibility, stigma, and pseudonymity. New York, NY, US: Association for Computing Machinery; 2011 Presented at: Proceedings of the 5th International Conference on Communities and Technologies; 29 June 2011-2 July 2011; Brisbane, Australia p. 159-168. [doi: 10.1145/2103354.2103376]
14. Hunt D, Brookes G. Corpus, Discourse and Mental Health. London: Bloomsbury Publishing; 2020.
15. De Choudhury M, De S. Mental health discourse on reddit: self-disclosure, social support, and anonymity. : Association for the Advancement of Artificial Intelligence; 2014 Presented at: Proceedings of the International AAAI Conference on Web and Social Media; June 1-4, 2014; Ann Arbor, Michigan, USA p. 71-80. [doi: 10.1609/icwsm.v8i1.14526]
16. Barak A, Boniel-Nissim M, Suler J. Fostering empowerment in online support groups. Comput Hum Behav 2008;24(5):1867-1883. [doi: 10.1016/j.chb.2008.02.004]

17. Pérez-Mañá C, Castells X, Torrens M, Capellà D, Farre M. Efficacy of psychostimulant drugs for amphetamine abuse or dependence. Cochrane Database Syst Rev 2013(9):CD009695 [FREE Full text] [doi: 10.1002/14651858.CD009695.pub2] [Medline: 23996457]
18. Paulus MP, Stewart JL. Neurobiology, clinical presentation, and treatment of methamphetamine use disorder: a review. JAMA Psychiatry 2020;77(9):959-966 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.0246] [Medline: 32267484]
19. Bhatt M, Zielinski L, Baker-Beal L, Bhatnagar N, Mouravska N, Laplante P, et al. Efficacy and safety of psychostimulants for amphetamine and methamphetamine use disorders: a systematic review and meta-analysis. Syst Rev 2016;5(1):189 [FREE Full text] [doi: 10.1186/s13643-016-0370-x] [Medline: 27842569]
20. Guerin AA, Nestler EJ, Berk M, Lawrence AJ, Rossell SL, Kim JH. Genetics of methamphetamine use disorder: a systematic review and meta-analyses of gene association studies. Neurosci Biobehav Rev 2021;120:48-74 [FREE Full text] [doi: 10.1016/j.neubiorev.2020.11.001] [Medline: 33217458]
21. Oksanen A. Deleuze and the theory of addiction. J Psychoactive Drugs 2013;45(1):57-67. [doi: 10.1080/02791072.2013.763563] [Medline: 23662332]
22. Duff C. Assemblages, territories, contexts. Int J Drug Policy 2016;33:15-20. [doi: 10.1016/j.drugpo.2015.10.003] [Medline: 26631273]
23. Pedersen W, Tutenges S, Sandberg S. The pleasures of drunken one-night stands: assemblage theory and narrative environments. Int J Drug Policy 2017;49:160-167. [doi: 10.1016/j.drugpo.2017.08.005] [Medline: 28870499]
24. Baker P. Using Corpora in Discourse Analysis, Second Edition. London: Bloomsbury Publishing; 2023.
25. Brookes G. Insulin restriction, medicalisation and the internet. CAM 2019;15(1):14-27. [doi: 10.1558/cam.33067]
26. Brookes G, McEnery T. Correlation, collocation and cohesion: a corpus-based critical analysis of violent jihadist discourse. Discourse & Society 2020;31(4):351-373 [FREE Full text] [doi: 10.1177/0957926520903528]
27. Brookes G, McEnery T. Corpus linguistics. In: Adolphs S, Knight D, editors. The Routledge Handbook of English Language and Digital Humanities. Abingdon, Oxon: Routledge; 2020:378-404.
28. Kilgarriff A, Baisa V, Bušta J, Jakubíček M, Kovář V, Michelfeit J, et al. The Sketch Engine: ten years on. Lexicography 2014;1(1):7-36. [doi: 10.1007/s40607-014-0009-9]
29. Jakubíček M, Kilgarriff A, Kovář V, Rychlý P, Suchomel V. The TenTen corpus family. 2013 Presented at: 7th International Corpus Linguistics Conference CL; 22-26 July 2013; Lancaster University, UK p. 125-127.
30. Stubbs M. Words and Phrases: Corpus Studies Of Lexical Semantics. Oxford: Blackwell Publishers; 2001.
31. Biber D, Conrad S, Cortes V. If you look at ...: lexical bundles in university teaching and textbooks. Appl Linguist 2004;25(3):371-405. [doi: 10.1093/applin/25.3.371]
32. Keane H. What's Wrong with Addiction?. Carlton South, Victoria, Austrlia: Melbourne University Press; 2002.
33. Fraser S, Moore D, Keane H. Habits: Remaking Addiction. New York, NY: Palgrave Macmillan; 2014.
34. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001;323(7321):1103-1105 [FREE Full text] [doi: 10.1136/bmj.323.7321.1103] [Medline: 11701577]
35. Aftab A, Ryznar E. Conceptual and historical evolution of psychiatric nosology. Int Rev Psychiatry 2021;33(5):486-499. [doi: 10.1080/09540261.2020.1828306] [Medline: 33047992]
36. Brookes G, Harvey K. Examining the discourse of mental illness in a corpus of online advice-seeking messages. In: Pickering L, Friginal E, Staples S, editors. Talking at Work: Corpus-Based Explorations of Workplace Discourse. London: Palgrave Macmillan; 2016:209-234.
37. McKetin R. Methamphetamine psychosis: insights from the past. Addiction 2018;113(8):1522-1527. [doi: 10.1111/add.14170] [Medline: 29516555]
38. Duff C. The pleasure in context. Int J Drug Policy 2008;19(5):384-392. [doi: 10.1016/j.drugpo.2007.07.003] [Medline: 17768037]
39. Duff C. Towards a theory of drug use contexts: space, embodiment and practice. Addict Res Theory 2009;15(5):503-519. [doi: 10.1080/16066350601165448]
40. Dwyer R, Moore D. Enacting multiple methamphetamines: the ontological politics of public discourse and consumer accounts of a drug and its effects. Int J Drug Policy 2013;24(3):203-211. [doi: 10.1016/j.drugpo.2013.03.003] [Medline: 23540297]
41. Zinberg N. Drug, Set and Setting: The Basis for Controlled Intoxicant Use. New Haven, CT: Yale University Press; 1984.
42. Hume D. An Enquiry Concerning Human Understanding. Oxford: Oxford University Press; 2008:9780199549900.
43. Hsieh JH, Stein DJ, Howells FM. The neurobiology of methamphetamine induced psychosis. Front Hum Neurosci 2014;8:537 [FREE Full text] [doi: 10.3389/fnhum.2014.00537] [Medline: 25100979]
44. Law J. After Method: Mess in Social Science Research. London and New York: Routledge; 2004.
45. Brookes G. Corpus linguistics in illness and healthcare contexts: a case study of diabulimia support groups. In: Demjen Z, editor. Applying Linguistics in Illness and Healthcare Contexts. London: Bloomsbury Publishing; 2020.
46. Fox NJ, Ward KJ, O'Rourke AJ. The 'expert patient': empowerment or medical dominance? The case of weight loss, pharmaceutical drugs and the internet. Soc Sci Med 2005;60(6):1299-1309. [doi: 10.1016/j.socscimed.2004.07.005] [Medline: 15626525]

47. Lamont M, Molnár V. The study of boundaries in the social sciences. Annu Rev Sociol 2002;28(1):167-195. [doi: 10.1146/annurev.soc.28.110601.141107]
48. Copes H, Tchoula W, Kim J, Ragland J. Symbolic perceptions of methamphetamine: differentiating between ice and shake. Int J Drug Policy 2018;51:87-94. [doi: 10.1016/j.drugpo.2017.11.007] [Medline: 29227843]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
DSM-5-TR: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision
ICD-11: International Classification of Diseases 11th Revision

Edited by T Mackey; submitted 09.05.23; peer-reviewed by J Guendouzi, C Laranjeira; comments to author 06.06.23; revised version received 27.07.23; accepted 06.09.23; published 29.09.23.

Please cite as:
Lustig A, Brookes G
Corpus-Based Discourse Analysis of a Reddit Community of Users of Crystal Methamphetamine: Mixed Methods Study
JMIR Infodemiology 2023;3:e48189
URL: https://infodemiology.jmir.org/2023/1/e48189
doi:10.2196/48189
PMID:37773617

Original Paper

Reproductive Health Experiences Shared on TikTok by Young People: Content Analysis

Isha Nair[1], BS; Sophia P Patel[2], BS; Ashley Bolen[1], BS; Samantha Roger[3], BS; Kayla Bucci[4]; Laura Schwab-Reese[4], MA, PhD; Andrea L DeMaria[4], MS, PhD

[1]School of Health Sciences, Purdue University, W Lafayette, IN, United States
[2]Department of Biological Sciences, Purdue University, W Lafayette, IN, United States
[3]School of Nursing, Purdue University, W Lafayette, IN, United States
[4]Department of Public Health, Purdue University, W Lafayette, IN, United States

Corresponding Author:
Andrea L DeMaria, MS, PhD
Department of Public Health
Purdue University
812 W State St
W Lafayette, IN, 47907-2060
United States
Phone: 1 765 494 8300
Email: ademaria@purdue.edu

Abstract

Background: TikTok is a popular social media platform that allows users to create and share content through short videos. It has become a place for everyday users, especially Generation Z users, to share experiences about their reproductive health. Owing to its growing popularity and easy accessibility, TikTok can help raise awareness for reproductive health issues as well as help destigmatize these conversations.

Objective: We aimed to identify and understand the visual, audio, and written components of content that TikTok users create about their reproductive health experiences.

Methods: A sampling framework was implemented to narrow down the analytic data set. The top 6 videos from each targeted hashtag (eg, #BirthControl, #MyBodyMyChoice, and #LoveYourself) were extracted biweekly for 16 weeks (July-November 2020). During data collection, we noted video characteristics such as captioning, music, likes, and cited sources. Qualitative content analysis was performed on the extracted videos.

Results: The top videos in each hashtag were consistent over time; for example, only 11 videos appeared in the top 6 category for #BirthControl throughout the data collection. Most videos fell into 2 primary categories: personal experiences and informational content. Among the personal experiences, people shared stories (eg, intrauterine device removal experiences), crafts (eg, painting their pill case), or humor (eg, celebrations of the arrival of their period). Dancing and demonstrations were commonly used in informational content.

Conclusions: TikTok is used to share messages on myriad reproductive health topics. Understanding users' exposure provides important insights into their beliefs and knowledge of sexual and reproductive health. The study findings can be used to generate valuable information for teenagers and young adults, their health care providers, and their communities. Producing health messages that are both meaningful and accessible will contribute to the cocreation of critical health information for professional and personal use.

(JMIR Infodemiology 2023;3:e42810) doi:10.2196/42810

KEYWORDS

TikTok; social media; reproductive health; women's health; health outcome; content analysis; health information; sexual health; web-based information; COVID-19; health message

Introduction

Background

Social media platforms allow users from various age groups to share and communicate information in an accessible manner. Prevalent social media platforms, such as Facebook, Instagram, Twitter, and TikTok, differ in layout and structure [1]. For instance, Twitter derives content from text; Instagram, from images; and Facebook, from a combination of both [1]. The increasing popularity of these platforms has created the potential for the spread of sensitive and stigmatizing topics regarding reproductive and sexual health promotion [2]. The public chooses the platform that best suits their needs based on how the layout and structure fit the context. Another factor is the stratification of social media platform users by age; younger people appear to prefer different social media platforms than those from previous generations. For example, a recent poll of 3000 Americans showed that 26% of Generation Z members chose TikTok as the app they would choose to use forever if they could only use one app, making it the top response for the entire age group [3]. This is in considerable contrast to older generations, which chose Facebook as their favorite by far (51% of all other ages of social media users) [3]. The relatively new platform TikTok launched in China in 2016 and went worldwide in 2018 with up to 800 million active users following the COVID-19 pandemic and global lockdowns [4]. TikTok, which allows users to create 15-second to 3-minute–long videos, has approximately 37 billion video views each month [5]. The platform is most popular among users aged between 13 and 24 years [5]. However, during the COVID-19 pandemic, some health care professionals and the World Health Organization turned to TikTok to facilitate the spread of reliable public health information [5].

Digital Storytelling

Digital storytelling refers to how the public shares their story through various forms of media; using social media to do this has become popular. People often use social media to share their reproductive health experiences, whether they are stories of sexual assault, political statements, or health-related ordeals [2]. From #FreeThePill to #MyBodyMyChoice, early narratives being shared were crucial to large movements. One of the most well-known examples of such prominent social media movements is #MeToo. The phrase was initially used on the social media platform MySpace in 2006, but in 2017, it became very popular on Twitter, igniting the movement [6]. The goal was to promote, support, and reduce the stigma around sharing experiences of sexual assault. The hashtag quickly went viral, thrusting the topic into the spotlight [7]. Since then, the way reproductive health experiences are shared on the web has changed, with social media becoming a more prominent way to contribute to the narrative.

People find new, unique ways to share their stories every day, especially during the COVID-19 pandemic when they felt most isolated. This makes users feel more connected to each other and the world [8]. TikTok's growing popularity has made it an ideal platform for young adults to share their experiences in a quick yet far-reaching manner [9], allowing users to share their health experiences uniquely through videos [10]. Even before TikTok, many social media users relied on platforms such as Twitter, YouTube, Facebook, and Instagram to share various reproductive health experiences [10-14]. For example, many nonprofessionals have shared overwhelmingly positive posts regarding intrauterine devices (IUDs), in contrast to the posts using more negative keywords from health care organizations and far more negative posts by law organizations, highlighting pain, issues with insertion and removal, and other side effects [11,15]. However, among the YouTube posts regarding IUDs, approximately one-third of the IUD testimonials contained inaccurate information that could skew a viewer's opinion on IUDs and other contraceptives [13]. Miscarriage has also been discussed on social media, particularly on Instagram. The trending hashtag #IHadAMiscarriage provided women a way to cope and find support [12]. Abortion is another popular topic that has been discussed on social media. Through the popular hashtags #StandWithPP, #ShoutYourAbortion, and #NotoriousRBG, users can show their support and help normalize abortions [14]. These platforms have similar digital narratives; however, TikTok's unique user interface and content-sharing methods are drawing attention. The platform exclusively shares videos and does so continuously, using an algorithm to ensure that each individual is seeing content curated specifically for them and adapting the stream accordingly [10,16]. On the basis of the content that viewers engage with the most, they receive similar videos [17]. Videos that are proven to be successful by some viewers will be pushed to others who might like similar content [16]. A study by Li et al [18] investigated the different video formats and types that TikTok provided specifically related to COVID-19 pandemic–related content and how other video formats and types contributed to user engagement. These video formats included video length, subtitle, text, spoken language, captions, and music, whereas the video types included whether they provided acting, an animated infographic, documentary, news, oral speech, a pictorial slideshow, or a TikTok dance [18].

Many popular platforms such as Facebook, Instagram, and Twitter focus predominantly on text or still images [10]. In contrast, the content on TikTok, often includes more detail with text, music, a continuous stream of content, and response videos to original TikTok content [10]. TikTok also allows users to combine several layers of overlapping content to create a unique narrative that differs from that of other platforms [19]. In addition, TikTok is trendiest among young people, especially in India, the United States, and Turkey, the countries with the largest monthly user base [20]. In fact, 41% of all TikTok users are aged 16 to 24 years, and in the United States, the top user ages fall between 18 and 24 years (42%) and 13 and 17 years (27%) [20]. Most TikTok users are young people, so the way content is spread on the platform conveys much about how young people interact with the information on it. Because TikTok rarely features content from familiar users, this narrative empowers the younger demographic to connect with new individuals in a more accessible way than other well-known social media platforms [19]. Although it is understood that TikTok's digital narrative differs from other platforms, there is currently minimal research on how reproductive health content is spread on this growing platform.

Social Media in Health Care

The growing prevalence of social media has also provided health care professionals with a platform to effectively share health information and public health interventions with over 1 billion users worldwide [21]. Because it is such a prominent way that important information is being spread, it is essential to understand how younger people, especially those disproportionately impacted by negative health outcomes (13-24 years), use the app to talk about reproductive health and what topics most interest them. In the same way in which people create videos and partake in trends to share health information involving periods, birth control, infertility, and more, users also make TikTok content about all types of relationships: romantic, platonic, familial, and professional [9,22]. As relationships play a large role in how people view topics and make decisions about topics such as reproductive and sexual health, it is an important aspect to study. However, there is not much research about how relationships are discussed on TikTok. This study can provide crucial pilot data for future research and program development proposals that use technology to reach young people. Analyzing teenagers' and young adults' behavior on TikTok when discussing and engaging in reproductive health experiences can not only generate valuable information for them, their health care providers, and their communities but also help produce health messages that are both meaningful and accessible. The insight we gained from this investigation can allow for contributions to the cocreation of critical health information for professional and personal use. Therefore, this study aims to identify and understand the content TikTok users are sharing about their everyday reproductive health experiences and the audio, visual, and written techniques they use to do so.

Methods

Data Collection

This study used an approach based on the principles of program science to inform the planning, implementation, and evaluation of complex health interventions [23]. This iterative approach consists of 3 primary phases, providing a framework for moving research and practice forward when an evidence-driven solution does not exist. The strategic planning stage focuses on developing a solid understanding of the current situation to enable informed decisions, the program implementation stage focuses on developing and refining the specifics of the program, and the program management and evaluation stage uses an ongoing and iterative quality improvement approach based on the evolution of the program. Future strategic planning that further refines the program will be informed by the knowledge generated through the program management and evaluation phase. This study focused primarily on the strategic planning phase, in which an understanding of the current situation was developed to aid future program development and implementation.

A qualitative content analysis approach was used, consistent with our prior research [24-29]. TikTok has ≥1 billion active users. Therefore, an appropriate sampling framework was identified to reduce the scope of the analytic data set. At the beginning of the project period, a number generator was used to identify 2 random days for each week of data collection. On these identified days, the top 6 videos, in terms of viewer engagement, for several different hashtags were extracted. TikTok's algorithm implies that the top videos in these hashtags are the videos young people interact with the most. This video extraction occurred biweekly for 16 weeks from July to November 2020 in the United States. To analyze how creators present information on reproductive health, some popular hashtags were identified and selected based on the study's research questions, a preliminary review of hashtag use, and alignment with the funding mechanism: #MyBodyMyChoice (2.2 billion views), #BirthControl (3.5 billion views), #PeriodProblems (1.4 billion views), #OBGYN (4.0 billion views), #WomensHealth (5.9 billion views), and #FreeThePill (1.4 million views). The hashtags #LoveYourself (21.1 billion views), #CoupleGoals (236.2 billion views), #BestFriend (81.1 billion views), and #Relationship (148.4 billion views) were studied to analyze how relationships were represented on TikTok. Although these hashtags may have multiple potential permutations, the versions that were most viewed were those used.

Data Analysis

After data collection was completed, the analysis phase was conducted. As with prior research, a codebook to analyze the TikTok videos was created using a content analysis framework [30-33]. Content analysis is a term used to describe a range of qualitative or quantitative methodologies. For this project, conventional content analysis, as defined by Hsieh and Shannon [31], was used where the coding frame was directly derived from the text (Multimedia Appendix 1). This approach is used frequently to compensate for the limited research on a subject [31]. The code-generation process was based on the adaptation of Schreier [33] on grounded theory. We began the analysis by reviewing the extracted TikTok videos. During a second review of the TikTok posts, notes were taken on emerging concepts with a particular focus on concepts that repeated or diverged between posts, making a note of recurring songs (if music was used), and the number of likes, views, comments, and shares on each post, all metrics of viewer engagement. In addition to the content, the length of the posts, whether said content was typed or spoken and whether it was part of a duet or challenge, was also recorded. This framework ensured broad insight into what users were creating, engaging with, and reproducing. A draft codebook was developed based on these notes and discussed among the entire research team to apply shared perspectives and finalize the codebook based on consensus. Once the codebook was finalized, it was applied to 5% of the available TikTok posts. On the basis of this pilot study, the codebook was revised and implemented to define the codes and categories. This process continued until the team determined that the codebook captured and defined the various domains and theoretical propositions under study and the team agreed on the applied codes [29,34]. As coding progressed, the team communicated regularly to discuss coding and resolved any uncertainties via discussion.

Then, segments were reviewed in relative isolation after the codes were applied. The segmentation process described by Schreier [33] was followed. In general, 1 study found that the use of video made HIV/AIDS content more engaging and more emotionally impactful [35]. As TikTok content come in video form, visual content, whether written or physical, is essential for a video's success. However, according to the platform itself, audio is just as important if not more. TikTok notes that 88% of its viewers consider audio vital to the experience and 50% say that music makes content more engaging and energizing [36]. It was also found that recall increases by 8 times the baseline when distinctive sounds are used [36]. The audio, visual, and written aspects were analyzed separately for each video. The audio segment was coded by IN, the visual segment was coded by AB, and the written segment was coded by SPP. For some codes, patterns and themes were identified in the content to describe the results qualitatively. For other codes, categorical variables (eg, primary topic and the sentiment of comments) were created, allowing the basic characteristics of the analytic data set to be summarized. The codes used for the audio aspect of each video involved topic, goal, tone, speed, duration, perceived gender of the speaker, points of view, the number of audio sources used, and the type of audio source used. It must also be noted that tone is a subjective code. Although insightful, it depends on the coders' perceptions and may be considered less reliable for that reason. Similarly, the speed of speech was also judged comparatively by the coders. For the visual aspect of each video, things such as the attire of the user, content type, methods and features used, number of people, perceived gender, presence of the user, props used, TikTok user, and video location. Similarly, for the written aspect of each video, the number of individual text boxes, text duration, the color of text, font of the text, goal of the text, caption content, number of hashtags present in the caption, hashtag content, presence of "for you page" hashtags, and the overall importance of writing in the video to understand the TikTok was analyzed. This process allowed us to identify differences in patterns and themes based on the characteristics of the conversations in the TikTok videos.

Ethics Approval

This study was approved by Purdue University's institutional review board (IRB-2020-803).

Results

A few key aspects comprise every piece of TikTok content—audio, visual, and written components. The results demonstrate the most popular ways users incorporate these essential aspects, as well as how health care professionals and other health practitioners may best use them to discuss this content with their patients and community members. In terms of content (Table 1), there was extensive overlap between different themes among the reproductive health hashtags observed. This caused 10 of the 100 videos to be duplicates. Most of these (7/100, 70%) came from the hashtag #WomensHealth. This resulted in 90 videos being coded. Numerous types of reproductive health topics were shared on this platform. Almost one-quarter of the videos (22/90, 25%) discussed relationships, making this the most discussed topic. The next most common topics were equal in prevalence—birth control and hygiene in health, with 15% (13/90) of videos being about them. Many videos were about periods or political situations (eg, abortion and punishments for sexual assault). The remaining topics occurred less frequently.

Table 1. Content results (n=90).

Topic	Occurrence, n (%)
Relationships	22 (25)
Birth control	13 (15)
Hygiene and health	13 (15)
Periods	11 (12)
Political situations	12 (13)
Birth and pregnancy	4 (4)
Sex education	2 (2)
Assault and safety	1 (1)
Body positivity	1 (1)
Other	11 (12)

Audio Aspect of Videos

There are several ways in which audio (Table 2) may convey a specific message on TikTok. Users make many choices in terms of audio to better suit the purposes of their TikTok. Of the 90 videos coded, there were some common trends. For example, popular videos seemed to have only one speaker at a time when disseminating reproductive health information. Most videos (66/90, 73%) came from a single individual speaking or music by a female singer. While the remaining videos may have multiple audio sources, approximately 99% (89/90) of the videos included only one perspective—no duets or stitches (ie, methods of adding to and interacting with a video that has already been made).

Table 2. Audio results (n=90).

Codes and subcodes	Occurrence, n (%)
Type of audio	
Just music	36 (40)
Just speaking	43 (48)
Both music and speaking	11 (12)
Perceived speaker or singer	
Singular female individual	62 (68)
Both female and male individuals	17 (19)
Multiple female individuals	6 (7)
Singular male individual	4 (4)
Multiple male individuals	1 (1)
Perceived tone	
Comedic	32 (36)
Enthusiastic	25 (28)
Angry	8 (9)
Sad	5 (6)
Informative or monotonous	20 (22)
Speed	
Fast	12 (13)
Slow	42 (47)
Duration	
Whole TikTok	87 (97)
≥75% of TikTok	2 (2)
25%-50% of TikTok	1 (1)
Purpose of audio	
Tell a story	19 (21)
State an opinion	8 (9)
Informational	13 (14)
Provide additional context or commentary	14 (16)
Participate in a trend	11 (12)
No connection or background	20 (22)
Other	5 (6)
Number of audio sources	
1	88 (98)
2	2 (2)
Multiple points of view	
Duet or stitch	1 (1)
None	89 (99)

There is slightly more variability with the rest of the results. For example, 48% (43/90) of the collected videos only had speaking as the audio source while 40% (36/90) only had music. The remaining 12% (11/90) had both. Of the 54 videos that contained speaking, there were variations in tone and speed. Of the speakers, 47% (42/90) spoke noticeably slower than the remaining speakers. The sample of videos analyzed contained a broad range of topics and used different tones. For example, 36% (32/90) of the coders perceived the videos as comedic and 28% (25/90), as enthusiastic. Some participants (20/90, 22%) seemed to have no emotion and a more informative, monotonous tone. The remaining videos contained intense emotions such as sadness or anger. TikTok creators used audio for many other purposes. From the data collected, some videos had audio entirely unconnected to their message (20/90, 22%), while others were telling a story (19/90, 21%) or stating an opinion (8/90, 9%). Other prominent purposes were to provide information, following a TikTok trend or using their audio as a supplement that provided additional context and commentary on what they were trying to say. Of the videos coded, 6% (5/90) were more unique to the reproductive health realm, such as debating specific political issues, and had audios that could not be characterized as having any of the purposes listed earlier.

Visual Aspect of Videos

The visual content (Table 3) in each of the 90 TikTok videos was also analyzed and collected to evaluate how influencers presented reproductive health information to TikTok viewers. After compiling the data, the types of videos that were popularized in terms of visual content could be seen. A prominent feature in all videos was the presence of text at some point. Text was shown in the video 43% (39/90) of the time. Text was coded if there were visual words on the video. Captions were not considered as text. Other codes with common trends that displayed popular creator choices were as follows: the user's attire, number of people, perceived gender, TikTok user type, video location, and features used. Looking at the user's attire, most (75/90, 83%) individuals did not dress in medical uniforms (eg, scrubs or a laboratory coat). Regarding the type of TikTok users, 80% (72/90) were general users: ordinary people who were neither famous nor presenting as medical professionals. Nearly one-fifth (17/90, 19%) of the participants were medical professionals. In the videos, it was common to see only 1 individual present (62/90, 69%). Regarding video location, more than half (51/90, 57%) were filmed in what appeared to be a personal residence. No videos were located in a hospital, and the remaining videos were in other locations, such as an office or a car.

Table 3. Visual results (n=90).

Codes and subcodes	Occurrence, n (%)
Presence of text	
Present	39 (43)
Not present	51 (57)
Attire	
Casual	48 (53)
Scrubs	13 (15)
Laboratory coat	1 (1)
Combination	1 (1)
Other	27 (30)
Number of individuals present	
1	62 (69)
Other	28 (31)
Type of creator	
General	72 (80)
Medical professional	17 (19)
Celebrity	1 (1)
Location	
Personal residence	51 (57)
Hospital	0 (0)
Office	4 (4)
Outdoors	14 (16)
Other	21 (23)

Written Aspect of Videos

Another video aspect analyzed was the written content (Table 4) presented. Of the videos analyzed, 41% (37/90) had no textboxes and the rest had varying numbers of textboxes. The color and font of the text present in the TikTok content also varied, with no clear trends. Overall, 27% (24/90) of the videos used text as a description for audio or visuals. The rest used text for different reasons, such as to reply to comments, to label visuals, and subtitles, to name a few. Most videos used at least 1 textbox, caption, or hashtag as a secondary form of communication after visual and audio, and over half of the videos used textboxes, where their primary use was to describe the audio or visuals in the video. Captions were another way to present information. Captions were used to further describe the video content in 48% (44/90) of videos. An additional 26% (23/90) of captions presented information on health-related topics such as abortion, birth control, friendship, menstruation, relationship with self, romantic relationship, sex, and other health content. Moreover, 8% (7/90) of the captions presented hashtags only, while the rest had completely unrelated captions that served the purpose of giving credit to trend creators or asking their viewers to interact. Hashtags were present in nearly all (83/90, 92%) of the videos, with most (66/90, 73%) containing ≥5 hashtags. In 28% (23/90) of the videos, video creators used hashtags that were directly related to their video content. In 57% (51/90) of the videos, the videos contained hashtags that were both related and unrelated to the content. In addition, less than half of the hashtags in the captions were "#ForYouPage," "#ForYou," or "#FYP."

Table 4. Written results (n=90).

Codes and Subcodes	Occurrence, n (%)
Textboxes present	
No textboxes	37 (41)
1-4 textboxes	25 (28)
≥5 textboxes	28 (31)
Color and font	
Black or white font	31 (34)
Colored font	22 (24)
Both black or white font and colored font	7 (8)
Default font	48 (53)
Multiple fonts	3 (3)
Typewriter font	1 (1)
Purpose of text	
Reply to comments	2 (2)
Subtitles	6 (7)
Description for audio or visuals	24 (27)
Clarify myths	2 (2)
Give personal advice	2 (2)
Label visuals	4 (4)
Provide information	10 (11)
Purpose of captions	
Describe content further	44 (48)
Ask viewers to interact	4 (4)
Give credit to creator of trend	1 (1)
Present information unrelation to health content	11 (12)
Hashtags only	7 (8)
Present information related to health topics	23 (26)
Hashtags present	
Present	83 (92)
Not present	7 (8)
Number of hashtags	
≥5	66 (73)
≤5	24 (27)
Relation of hashtags to reproductive health content	
Hashtags related to one topic addressed in video	23 (28)
Hashtags related to multiple topics addressed in video	23 (26)
Hashtags both related and unrelated to content	51 (57)

Discussion

Principal Findings

This study aimed to better understand the distinct components of content that TikTok users incorporate to create reproductive health-related information and share their experiences on this social media platform. The results demonstrate that there were some key aspects of videos that appeared to contribute to success. Although the number or style of the textboxes seems insignificant, audio appears to be essential for success. In addition, having some kind of tone or emotion likely increases engagement, and general content creators seem to connect better and have more reach with the public than a health care provider does.

Several reproductive health topics were discussed by users. Although some videos did not fall under any of these categories, the most popular subjects, in order of decreasing popularity, were relationships, birth control, hygiene and health, political situations (eg, abortion and punishments for sexual assault), periods, birth and pregnancy, education on sexual intercourse, assault and safety, body positivity. This suggests that these are the topics viewers are most curious about and the areas health care providers and other health practitioners might want to address when educating their patients and communities.

We assessed the role of audio, visual, and written approaches in the sharing of reproductive health-related information. In doing so, we found many commonalities among the most viewed videos. The common trends observed within the audio choices suggest a potential correlation between the way audio is used in TikTok and the video reach. Platforms such as Twitter, Instagram, and Facebook allow videos with audio but also often rely on pictures with no audio [10]. The fact that all videos on TikTok had some form of audio implies that audio may be essential to any TikTok's success—something any user, including health experts, should remember when trying to use the platform to reach young people. In addition, most TikTok content had only 1 speaker, if any. They also mostly came from one perspective, with no stitches or duets. Although other videos across the platform commonly used the feature that allows you to add another video as a form of reply to continue a conversation, not one of the reproductive health videos we studied did so. This may be due to the sensitivity of some of the subjects, as well as the lack of universal knowledge on them. Popular videos used only one audio source at a time, such as a song or spoken content rather than both together, suggesting that more than one source may be too much of a distraction when trying to process the content in a meaningful manner.

Most videos appeared to have some kind of tone, such as anger, sadness, or excitement. They also had slower-paced spoken words. This is probably a tactic to connect with viewers more easily, and it seems to be something to which they respond well. Slower speech may be easier to understand and process, getting the message across more efficiently, but slow speech in conjunction with a monotonous tone would likely bore audiences who care more about what creators have to say when it emotionally impacts them. In fact, according to a study conducted by Holub [37] on the effects of intonation, it was found that monotony has a negative impact on audience comprehension. However, some TikTok content is purely informational. Although it might not make sense for those videos to use tone in an emotional manner, an enthusiastic tone or the inclusion of upbeat music and dancing goes a long way to making the video noticeable and therefore allowing it to deliver information to viewers.

Analyzing the visual content showed that content creators who wear casual clothing, film in a personal residence, and come off as general content creators rather than health care providers reach more TikTok users about reproductive health. It is possible that viewers find this information more credible when it comes from someone they can relate to and find these informal settings more comfortable. These factors are all things that providers can consider when creating an appealing environment to spread information in their videos.

Following Comp et al [10], we found that the written aspects of TikTok videos were used to either supplement audio and visual information or to guide viewers to audio and visual information rather than to serve as the primary form of communicating information to viewers during the COVID-19 pandemic. We found that the number or style of textboxes used in each video had no apparent pattern. The primary purpose of the captions was to guide the viewer and provide additional information. Overall, although written content seemed to be a popular tool to enhance the video and provide elaboration, video creators did not have a specific way to present written content.

Hashtags served an overall greater purpose in comparison with textboxes and captions. They connected viewers to a video so that they could learn about women's reproductive and sexual health through the video's primary forms of communication: audio and visual content. Furthermore, many of the findings of Zappavigna [38] and the conclusions of Herrman [39] on the use of hashtags were echoed in this study, where hashtags served a classificatory function, providing a way for viewers to identify content based on a specific topic, as opposed to presenting additional information on the content as captions did. Hashtags also allowed creators to connect themselves and their content to other TikTok creators and viewers who may share similar stories, content, or opinions on reproductive and sexual health. The use of multiple hashtags for one video led to some overlap in data collection. The occurrence of numerous hashtags could have been related to the creator's desire to increase the searchability of their content and increase their range of audience, which should be considered by experts and educators when creating videos and looking to increase reach.

Limitations

Although this study provides insight into how information on reproductive health is shared on TikTok, it has its limitations. As discussed earlier, one of the things that makes studying TikTok so important is that it is constantly changing; therefore, some new aspects cannot be encompassed in this research. For example, when data collection began, the longest a TikTok video could be was 1 minute. Since then, an update to the app has increased the amount of time a video can run from 1 minute to 3 minutes. Thus, our research cannot examine how this change affects what are now the most popular videos under these hashtags. This could have a huge impact on popular creative choices and is an aspect we could not explore. There may be other similar changes that were not anticipated that continue to affect popular trends. Some of these codes are also quite subjective, such as the speed, tone, and perceived gender of the speaker, which depend on the coder's perspective. Although these are vital aspects of each video that should be examined, the reliability of our study was affected by the lack of intercoding in each segment, a change that will be made for future studies.

This study and future studies could benefit from a larger sample size with data collected over a longer period. The level of overlap in our data decreased the number of videos analyzed. Although this says a lot about the popularity of certain visual or audio aspects of these videos, there was not necessarily enough material to explore things that varied more, such as written content, to the extent it could have been. Further analysis of a larger video sampling can provide more insight into the written content trends and provide more validity to the study.

Despite these limitations, this project is vital because it sheds light on how reproductive health information is shared among young people on a major social media platform. It focuses on exactly how reproductive health information is shared on TikTok, showing the great variety of topics shared but also how there are so many similarities in how that information is shared. Many trends were identified that can guide health care providers and other health practitioners to build a toolkit to better disseminate their information to a larger audience.

Implications

TikTok is an expanding digital app that continuously engages more young adult users every day. It is a platform in which many content creators have access to different formatting tools to captivate viewers by spreading information regarding reproductive and sexual health topics. In the age of social media, TikTok is a tool to spread information that many people have already been engaging with. Even before TikTok emerged, there have been studies examining the impact of social media and its effect on the medical field.

One specific study by Omurtag and Turek [40] discussed how to provide guidance for reproductive health professionals and how to use social media to effectively disseminate information, specifically on infertility and in vitro fertilization. A key aspect they found that makes social media appealing to individuals is anonymity—individuals who engage on these platforms do not have to include any contact information, allowing for privacy [40]. Although both studies promote the use of social media by medical professionals, our study goes into depth specifically on what topics users interact with and how they view them, providing insight that medical professionals can use to make videos that are both meaningful and accessible to young adults. Our study was able to summarize the common aspects of videos that were appealing, such as having a casual setting and a creator perceived to be a woman as the main messenger. Relatability was key, and videos with these features overall showed a higher chance of being viewed.

The videos analyzed had specific hashtags related to reproductive health, including #MyBodyMyChoice, #BirthControl, #PeriodProblems, #OBGYN, #WomensHealth, and #FreethePill. Many of these topics deal with individuals' personal stories or experiences. It is important to see what the users are drawn to understand how to best reach them regarding important information about women's reproductive health. Future research could examine each specific hashtag. Individuals may respond uniquely to the content within each hashtag and its associated topics or create content using a specific methodology. The study by Li et al [18] found that for COVID-19 pandemic content, dancing videos were favored significantly more than any other method when presenting the information. This was not the case in our study (7/90, 7%), likely because our content did not include all COVID-19 pandemic content but rather reproductive health content shared at the height of the pandemic. Our study did, however, have similar findings on what other features users seem to engage with most. These interesting features included captions, hashtags, humorous tones, primarily visible human subjects, and some background audio. The features common to both studies allow a greater understanding of the information provided in the video and contribute to the increased engagement of users. The data collected in our study provide new information on what individuals engage with the most to guide health practitioners in creating impactful and viewable content. We found that some kind of tone, humorous or not, was helpful, and popular women's health–themed hashtags were a major way to ensure a larger audience. Looking further at a specific topic in women's health might allow deeper insight into what viewers engage best within each topic and allow professionals to provide meaningful information. Another area of research related to this is examining the methods of each video after observing the number of shares, likes, and comments.

Conclusions

This study sought to understand how young people shared reproductive health information on TikTok, which was captured during the COVID-19 pandemic. Most videos presented content that was either helpful or relatable to the viewer. In addition, videos in which a single individual was dressed informally and videos that used audio for the entire duration were viewed more often. The written aspect of the videos was supplemental to the visual and audio content. If health care professionals present content in this way, young people will continue to be open and willing to learn about women's reproductive and sexual health on social media platforms such as TikTok.

TikTok is a continuously growing platform. Currently, we are in the age of social media and many young people share and receive information, including information related to reproductive health. Health care professionals and other health practitioners have an opportunity to provide reproductive health information through social media, especially through TikTok. Because of this study, a guideline on how to make the most meaningful and engaging videos can be created. These guidelines may include the use of captions, hashtags, a humorous tone, a primarily visible female creator, and background audio. By following these guidelines, creators can develop and share videos that are meaningful and engaging to TikTok users worldwide.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Codes and subcodes determined for audio, visual, and written codebooks used to analyze video creation choices of TikTok users discussing reproductive and sexual health.
[DOCX File , 19 KB - infodemiology_v3i1e42810_app1.docx]

References

1. Masciantonio A, Bourguignon D, Bouchat P, Balty M, Rimé B. Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. PLoS One 2021 Mar 11;16(3):e0248384 [FREE Full text] [doi: 10.1371/journal.pone.0248384] [Medline: 33705462]
2. Gabarron E, Wynn R. Use of social media for sexual health promotion: a scoping review. Glob Health Action 2016 Sep 19;9(1):32193 [FREE Full text] [doi: 10.3402/gha.v9.32193] [Medline: 27649758]
3. 26% of Gen Z users ‘chose’ TikTok as their no. 1 social media platform: reporter. Yahoo. URL: https://finance.yahoo.com/video/26-gen-z-users-chose-211528709.html [accessed 2023-03-28]
4. TikTok: the story of a social media giant. BBC News. 2020 Aug 04. URL: https://www.bbc.com/news/technology-53640724 [accessed 2023-03-28]
5. 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]
6. Jaffe S. The collective power of #MeToo. Dissent 2018;65(2):80-87 [FREE Full text] [doi: 10.1353/dss.2018.0031]
7. Xiong Y, Cho M, Boatwright B. Hashtag activism and message frames among social movement organizations: semantic network analysis and thematic analysis of Twitter during the #MeToo movement. Public Relat Rev 2019 Mar;45(1):10-23 [FREE Full text] [doi: 10.1016/j.pubrev.2018.10.014]
8. Nabity-Grover T, Cheung CM, Thatcher JB. Inside out and outside in: how the COVID-19 pandemic affects self-disclosure on social media. Int J Inf Manage 2020 Dec;55:102188 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102188] [Medline: 32836645]
9. Meisenzahl M. The wildly popular TikTok app is becoming an unlikely place for women to talk about infertility and reproductive health - and it's a clear sign the platform is for more than just teens. Business Insider. URL: https://www.businessinsider.com/tiktok-infertility-ivf-community-growing-women-obgyns-2019-11 [accessed 2020-04-28]
10. Comp G, Dyer G, Gottlieb M. Is TikTok the next social media frontier for medicine? AEM Educ Train 2021 Oct 21;5(3):aet2.10532 [FREE Full text] [doi: 10.1002/aet2.10532] [Medline: 34095694]
11. Jung C, Richter F, Lunde B, Chen K. IUDs in the Twittersphere: sentiment analysis of IUD-related Tweets [23F]. Obstet Gynecol 2017;129:67 [FREE Full text] [doi: 10.1097/01.aog.0000514852.66590.1f]
12. Mercier RJ, 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]
13. Nguyen BT, Allen AJ. Social media and the intrauterine device: a YouTube content analysis. BMJ Sex Reprod Health 2018 Jan 23;44(1):28-32 [FREE Full text] [doi: 10.1136/bmjsrh-2017-101799] [Medline: 29170151]
14. Whaley N, Brandt J. Claiming the abortion narrative: a qualitative exploration of mainstream and social media reflections on abortion. In: Choudhury S, Erausquin JT, Withers M, editors. Global Perspectives on Women's Sexual and Reproductive Health Across the Lifecourse. Cham, Switzerland: Springer; 2018:159-171.
15. Wu J, Trahair E, Happ M, Swartz J. TikTok, #IUD, and user experience with intrauterine devices reported on social media. Obstet Gynecol 2023 Jan 01;141(1):215-217. [doi: 10.1097/AOG.0000000000005027] [Medline: 36473194]
16. Hern A. How TikTok’s algorithm made it a success: ‘it pushes the boundaries’. Guardian. 2022 Oct. URL: https://www.theguardian.com/technology/2022/oct/23/tiktok-rise-algorithm-popularity [accessed 2023-06-25]
17. Smith B. How TikTok reads your mind. New York Times. 2021 Dec 06. URL: https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html [accessed 2023-06-25]
18. Li Y, Guan M, Hammond P, Berrey LE. Communicating COVID-19 information on TikTok: a content analysis of TikTok videos from official accounts featured in the COVID-19 information hub. Health Educ Res 2021 Jul 12;36(3):261-271 [FREE Full text] [doi: 10.1093/her/cyab010] [Medline: 33667311]
19. Unni Z, Weinstein E. Shelter in place, connect online: trending TikTok content during the early days of the U.S. COVID-19 pandemic. J Adolesc Health 2021 May;68(5):863-868 [FREE Full text] [doi: 10.1016/j.jadohealth.2021.02.012] [Medline: 33824071]
20. Elad B. TikTok statistics 2022 - users, usage, trends and facts. Enterprise Apps Today. URL: https://www.enterpriseappstoday.com/stats/tiktok-statistics.html [accessed 2022-12-30]
21. Eghtesadi M, Florea A. Facebook, Instagram, Reddit and TikTok: a proposal for health authorities to integrate popular social media platforms in contingency planning amid a global pandemic outbreak. Can J Public Health 2020 Jun 09;111(3):389-391 [FREE Full text] [doi: 10.17269/s41997-020-00343-0] [Medline: 32519085]
22. Galchen G. TikTok's take on love. 34th Street. 2022 Feb 14. URL: https://www.34st.com/article/2022/02/tiktok-love-trend-viral-relationship-social-media-happiness [accessed 2023-01-02]

23. Blanchard JF, Aral SO. Program Science: an initiative to improve the planning, implementation and evaluation of HIV/sexually transmitted infection prevention programmes. Sex Transm Infect 2011 Feb 01;87(1):2-3. [doi: 10.1136/sti.2010.047555] [Medline: 21126962]
24. Sundstrom B, Smith E, Vyge K, Miletich A, Benigni G, Delay C, et al. Moving oral contraceptives over the counter: theory-based formative research to design communication strategy. J Health Commun 2020 Apr 02;25(4):313-322. [doi: 10.1080/10810730.2020.1752334] [Medline: 32306857]
25. Schwab-Reese L, Kanuri N, Cash S. Child maltreatment disclosure to a text messaging-based crisis service: content analysis. JMIR Mhealth Uhealth 2019 Mar 25;7(3):e11306 [FREE Full text] [doi: 10.2196/11306] [Medline: 30907745]
26. Cash SJ, Murfree L, Schwab-Reese L. "I'm here to listen and want you to know I am a mandated reporter": understanding how text message-based crisis counselors facilitate child maltreatment disclosures. Child Abuse Negl 2020 Apr;102:104414 [FREE Full text] [doi: 10.1016/j.chiabu.2020.104414] [Medline: 32070934]
27. DeMaria AL, Delay C, Sundstrom B, Wakefield AL, Avina A, Meier S. Understanding women's postpartum sexual experiences. Cult Health Sex 2019 Oct 09;21(10):1162-1176. [doi: 10.1080/13691058.2018.1543802] [Medline: 30624136]
28. DeMaria AL, Hollub AV, Herbenick D. Using genital self-image, body image, and sexual behaviors to predict gynecological exam behaviors of college women. J Sex Med 2011 Sep;8(9):2484-2492. [doi: 10.1111/j.1743-6109.2011.02379.x] [Medline: 21718451]
29. Demaria AL, Sundstrom B, Ferrara M, Meier S, Higa M. Fulfilling the promise of telemedicine: a case study of south Carolina. Health Behav Policy Rev 2019 Mar 01;6(2):140-151 [FREE Full text] [doi: 10.14485/hbpr.6.2.3]
30. Graneheim U, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
31. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
32. Krippendorff K. Content Analysis: An Introduction to Its Methodology. 3rd edition. Thousand Oaks, CA: Sage Publications; 2012.
33. Schreier M. Qualitative Content Analysis in Practice. Thousand Oaks, CA: Sage Publications; 2012.
34. Yin RK. Case Study Research: Design and Methods. Thousand Oaks, CA: Sage Publications; 2013.
35. Yadav A, Phillips MM, Lundeberg MA, Koehler MJ, Hilden K, Dirkin KH. If a picture is worth a thousand words is video worth a million? Differences in affective and cognitive processing of video and text cases. J Comput High Educ 2011 Feb 11;23(1):15-37 [FREE Full text] [doi: 10.1007/s12528-011-9042-y]
36. Hutchinson A. TikTok shares new insights into the importance of audio in brand marketing. Social Media Today. 2022 Jan. URL: https://www.socialmediatoday.com/news/tiktok-shares-new-insights-into-the-importance-of-audio-in-brand-marketing/616633/ [accessed 2023-07-01]
37. Holub E. Does intonation matter? The impact of monotony on listener comprehension. OpenstarTs. 2010. URL: https://www.openstarts.units.it/handle/10077/4753 [accessed 2023-01-04]
38. Zappavigna M. Searchable talk: the linguistic functions of hashtags. Social Semiotics 2015;25(3):274-291 [FREE Full text] [doi: 10.1080/10350330.2014.996948]
39. Herrman J. How TikTok Is rewriting the world. New York Times. 2019. URL: https://www.nytimes.com/2019/03/10/style/what-is-tik-tok.html [accessed 2021-01-19]
40. Omurtag K, Turek P. Incorporating social media into practice: a blueprint for reproductive health providers. Clin Obstet Gynecol 2013 Sep;56(3):463-470. [doi: 10.1097/GRF.0b013e3182988cec] [Medline: 23722919]

Abbreviations

IUD: intrauterine device

Edited by T Mackey; submitted 19.09.22; peer-reviewed by L Porsch, A Zheluk, EWJ Lee; comments to author 14.12.22; revised version received 04.01.23; accepted 30.09.23; published 13.11.23.

Please cite as:
Nair I, Patel SP, Bolen A, Roger S, Bucci K, Schwab-Reese L, DeMaria AL
Reproductive Health Experiences Shared on TikTok by Young People: Content Analysis
JMIR Infodemiology 2023;3:e42810
URL: https://infodemiology.jmir.org/2023/1/e42810
doi:10.2196/42810
PMID:37831780

Original Paper

COVID-19–Associated Misinformation Across the South Asian Diaspora: Qualitative Study of WhatsApp Messages

Anjana E Sharma[1,2], MAS, MD; Kiran Khosla[3], BA; Kameswari Potharaju[4], BA; Arnab Mukherjea[5], MPH, DrPH; Urmimala Sarkar[2,6], MPH, MD

[1]Department of Family and Community Medicine, University of California San Francisco, San Francisco, CA, United States

[2]Center for Vulnerable Populations at Zuckerberg San Francisco General Hospital, University of California San Francisco, San Francisco, CA, United States

[3]School of Public Health, Boston University, Boston, MA, United States

[4]University of California, Berkeley, CA, United States

[5]Department of Public Health, California State University East Bay, Hayward, CA, United States

[6]Division of General Internal Medicine at Zuckerberg San Francisco General Hospital, University of California San Francisco, San Francisco, CA, United States

Corresponding Author:
Anjana E Sharma, MAS, MD
Department of Family and Community Medicine
University of California San Francisco
1001 Potrero Avenue
Zuckerberg San Francisco General Hospital
San Francisco, CA, 94143
United States
Phone: 1 628 206 4943
Email: anjana.sharma@ucsf.edu

Abstract

Background: South Asians, inclusive of individuals originating in India, Pakistan, Maldives, Bangladesh, Sri Lanka, Bhutan, and Nepal, comprise the largest diaspora in the world, with large South Asian communities residing in the Caribbean, Africa, Europe, and elsewhere. There is evidence that South Asian communities have disproportionately experienced COVID-19 infections and mortality. WhatsApp, a free messaging app, is widely used in transnational communication within the South Asian diaspora. Limited studies exist on COVID-19–related misinformation specific to the South Asian community on WhatsApp. Understanding communication on WhatsApp may improve public health messaging to address COVID-19 disparities among South Asian communities worldwide.

Objective: We developed the COVID-19–Associated misinfoRmation On Messaging apps (CAROM) study to identify messages containing misinformation about COVID-19 shared via WhatsApp.

Methods: We collected messages forwarded globally through WhatsApp from self-identified South Asian community members between March 23 and June 3, 2021. We excluded messages that were in languages other than English, did not contain misinformation, or were not relevant to COVID-19. We deidentified each message and coded them for one or more content categories, media types (eg, video, image, text, web link, or a combination of these elements), and tone (eg, fearful, well intentioned, or pleading). We then performed a qualitative content analysis to arrive at key themes of COVID-19 misinformation.

Results: We received 108 messages; 55 messages met the inclusion criteria for the final analytic sample; 32 (58%) contained text, 15 (27%) contained images, and 13 (24%) contained video. Content analysis revealed the following themes: "community transmission" relating to misinformation on how COVID-19 spreads in the community; "prevention" and "treatment," including Ayurvedic and traditional remedies for how to prevent or treat COVID-19 infection; and messaging attempting to sell "products or services" to prevent or cure COVID-19. Messages varied in audience from the general public to South Asians specifically; the latter included messages alluding to South Asian pride and solidarity. Scientific jargon and references to major organizations and leaders in health care were included to provide credibility. Messages with a pleading tone encouraged users to forward them to friends or family.

Conclusions: Misinformation in the South Asian community on WhatsApp spreads erroneous ideas regarding disease transmission, prevention, and treatment. Content evoking solidarity, "trustworthy" sources, and encouragement to forward messages may increase the spread of misinformation. Public health outlets and social media companies must actively combat misinformation to address health disparities among the South Asian diaspora during the COVID-19 pandemic and in future public health emergencies.

(JMIR Infodemiology 2023;3:e38607) doi:10.2196/38607

KEYWORDS

misinformation; COVID-19; South Asians; disparities; social media; infodemiology; WhatsApp; messages; apps; health information; reliability; communication; Asian; English; community; health; organization; public health; pandemic

Introduction

Misinformation, or false and inaccurate information, is a major public health challenge during the COVID-19 pandemic. The World Health Organization has identified COVID-19 information as an "infodemic"—an overabundance of COVID-19–related information, including deliberate attempts to foment misinformation [1]. Many formal definitions of misinformation exist [2-4]; misinformation is sometimes distinguished from disinformation (ie, false information with intent to harm) and malinformation (ie, facts used out of context with intent to harm). For the purpose of this paper, we use the term "misinformation" as an umbrella term to comprise false information, where the intent is not apparent [5]. Social media is an important channel of distribution of COVID-19 misinformation; false news diffuses more quickly than truth [6], and low-credibility web sources are shared more frequently than any single high-credibility news source on Facebook or Twitter [7].

The South Asian diaspora, defined as communities with origins from Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, and Sri Lanka [8], is highly active on social media platforms. One common platform is WhatsApp. Unlike public social media platforms, such as Facebook, Twitter, and Instagram, WhatsApp is a private messaging platform allowing users to send information through text, photos, and videos to one person directly or in groups of up to 1023 individuals. Almost 400 million Asian Indians use WhatsApp, comprising the greatest number of platform users worldwide [9]. WhatsApp messages are readily forwarded with limited capacity to determine the original author or provide factual checks, enabling misinformation to spread easily.

Misinformation is harmful for public health. For example, myths circulating about various remedies, such as highly concentrated alcohol, ingested sanitizer, or Datura seeds led to cases of illness, blindness, and death [10]. Belief in COVID-19 vaccine conspiracy theories is associated with vaccine hesitancy [11]. Areas with higher exposure to news media denying COVID-19 severity are associated with greater COVID-19 case rates and deaths [12]. Given that misinformation related to COVID-19 may drive COVID-19 morbidity and mortality, understanding the ways COVID-19 misinformation spreads within specific communities is paramount. This is concerning given the disproportionate burden of COVID-19 infection, hospitalization [13], and death experienced by South Asians globally [14].

Understanding culturally specific misinformation may inform policy-making and targeted public health messaging efforts. As part of the COVID-19–Associated misinfoRmation On Messaging apps (CAROM.) study, we analyzed COVID-19 misinformation circulated within the South Asian diaspora via WhatsApp.

Methods

Procedure

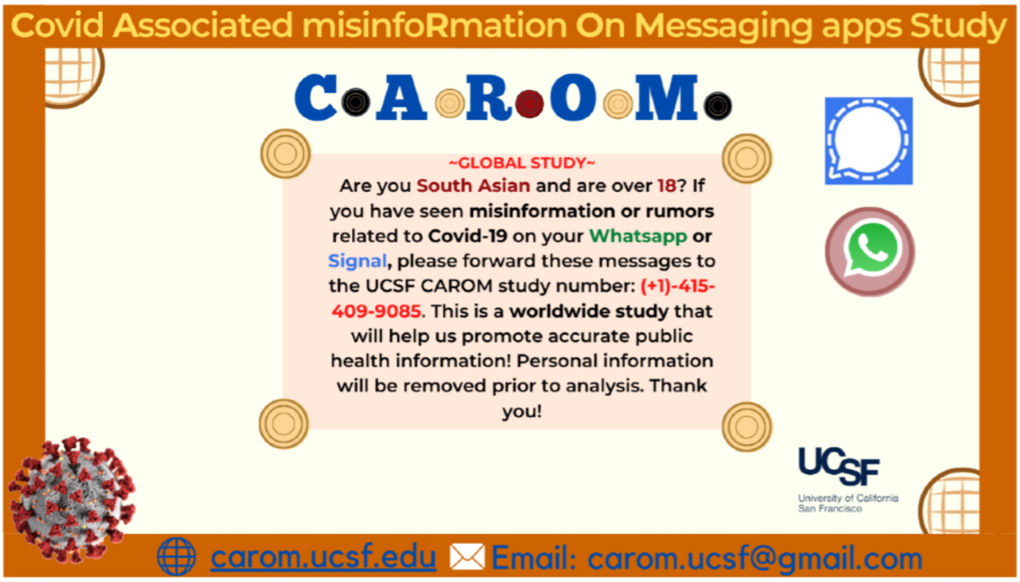
Individuals who self-identified as members of the South Asian community older than 18 years of age anywhere globally were eligible for inclusion in the study. We specifically chose WhatsApp for sampling due to its significance as one of the most widely used messaging programs among the South Asian Diaspora [9]; it is therefore widely recognized and familiar in the South Asian community. We recruited participants via English-language outreach on Twitter, Facebook, and WhatsApp using web-based flyers with a QR code, hashtags, direct messages to community leaders, blog posts [15], and emails to South Asian organizations. We named the study after a beloved game that is familiar across the diaspora (Figure 1); additional recruitment materials are available in Multimedia Appendix 1.

We asked individuals who self-identified as members of the South Asian or Desi community to forward deidentified screenshots of WhatsApp messages containing what they perceived to be "misinformation or rumors" related to COVID-19 to an official study phone number. This allowed us to receive messages being transmitted or forwarded within WhatsApp without any personal or identifiable information included. We chose this method to allow individual WhatsApp users to share what was being transmitted in their private feeds in a deidentified manner, as a means to access traditionally closed communications in a way that protected the privacy of individual users. We requested WhatsApp messages with potential COVID-19 misinformation to be forwarded to an official study phone number. The research team advertised the study starting March 23, 2021, and collected messages until June 3, 2021. The team deidentified screenshots and media files if they were not already deidentified by the study participant. Study team members reviewed each message content to determine if inclusion criteria were met. Specifically, they assessed the relevance to COVID-19 and whether or not the information in the message was factual based on the team's medical and scientific knowledge as well as web-based fact-checking when appropriate. Messages not in English;

received after June 3, 2021; not relevant to COVID-19; or not containing misinformation were excluded.

The team developed an abstraction form to identify media format (eg, written text, picture, video, URL, or a combination of these formats), country or location mentioned, content category, and tone in REDCap (Research Electronic Data Capture; Vanderbilt University; Multimedia Appendix 2). Two team members (KK and KP) abstracted the messages and conducted open coding, using content analysis of messages and categorizing messages into thematic groups of specific "content types" of COVID-19 misinformation. The entire research team cross-checked coding and the thematic analysis until the team arrived at consensus.

Figure 1. COVID-19–Associated misinfoRmation On Messaging apps (CAROM) Study web-based recruitment flyer.



Ethics Approval

This study was approved by the Internal Review Board of University of California, San Francisco (20-32758).

Results

We collected 108 messages, deduplicated to 96, of which 55 messages met the inclusion criteria. Message formats included plain text, images, and videos, or a combination of these formats (Table 1). India was most commonly mentioned (21/55, 38%); followed by China (7/55, 13%), the United States (4/55, 7%), and Italy (3/55, 5%).

The content fit into one or more of the following thematic categories: (1) community transmission, (2) prevention, (3) treatment, and (4) products (Table 2). "Community transmission" messages included conspiracy theories about the origins and spread of COVID-19; one example described India's second wave as suspicious following rising political tensions with China. Many "prevention" messages proposed ways to prevent coronavirus infection through home-based strategies, such as inhaling steam, eating a banana every day, or consuming alkaline foods. Messages about "treatment" offered self-treatments for COVID-19, such as drinking "a teaspoon of pepper powder, two teaspoons of honey, and ginger juice." Lastly, messages about "products" publicized commercial treatments or cures for COVID-19, such as a nasal spray purported to offer protection from SARS-CoV-2.

Messages ranged from containing "universal" misinformation addressed to the general public to "South Asian–specific" content containing cultural references. One universal message was a video of an alleged Irish scientist describing purported mortality risk with messenger RNA vaccines. In comparison, a South Asian–specific reference discussed traditional natural remedies, such as Ayurveda and homeopathy. A few messages appealed to ethnic or national pride, with statements such as "we Indians are built to last" or "proud to be an Indian."

Some messages contained information that was entirely false; for example, one image claiming to be published by UNICEF (United Nations International Children's Emergency Fund) stated that the coronavirus will be killed if it is exposed to a temperature of 26-27 °C, and therefore, encouraged drinking hot water and increasing sun exposure. However, other messages shared a mixture of true and false information. One message combined evidence-based recommendations for preventing COVID-19 infection, such as social distancing and wearing a face mask, while also encouraging behaviors without evidence, such as eating vegetarian food and removing belts and rings.

Message tone included fear- or panic-based encouragement to share purportedly useful information and pseudoscientific expertise. Pseudoscientific messages contained scientific jargon unfamiliar to a nonscientific audience, such as "anticoagulants" and "ground-glass opacities," in combination with references to reputable organizations, such as the World Health Organization and the Indian Council on Medical Research, as well as individual experts, such as doctors and scientists. A third of messages (18%-33%) used a pleading or encouraging tone to promote dissemination, asking recipients to "share with all your family and friends" or "send to all your groups."

Table 1. Summary characteristics of the sample (N=55).

Characteristics	Values, n (%)
Message format	
Text only	32 (58)
Image only	15 (27)
Video only	13 (24)
Link	6 (11)
Other	1 (2)
Countries mentioned	
India	21 (38)
China	7 (13)
United States	4 (7)
Italy	3 (5)
Japan	2 (4)
Australia	1 (2)
Bangladesh	1 (2)
Bhutan	1 (2)
Ireland	1 (2)
Nepal	1 (2)
Pakistan	1 (2)
Spain	1 (2)
Sri Lanka	1 (2)
Switzerland	1 (2)
Taiwan	1 (2)
Tone	
Good intentions	27 (50)
Pleading or call to action	18 (33)
Warning or fear-based	10 (18)
Blame	4 (7)
Other	6 (15)

Table 2. Main thematic domains of misinformation.

Level	Domains	Subthemes	Definition	Examples
Individual				
	Prevention or diagnosis	Self-diagnosis and self-remedies	Content had to do with preventing COVID-19 disease, exposure to Sars-CoV-2, and screening for COVID-19.	• •
	Treatment	Ayurveda, homeopathy, or natural remedies	Content promoting nonevidence-based means to treat or cure COVID-19	•
	Products or services		For-sale devices or products marketed to reduce risk, prevent, or treat COVID-19 infection	• •
Population				
	Community transmission	Conspiracy theories	Content explains how and why COVID-19 is spreading at the local or international level	•
	Cultural pride		Speaking to positive or robust aspects of South Asian identity, value of South Asian attributes, or direct contributions to efforts to fight COVID-19	•

Discussion

Principal Findings

We found misinformation circulating through South Asian networks is largely aimed at providing alternative explanations for the etiology and spread as well as alternative treatment and prevention methods for COVID-19.

The South Asian diaspora comprises the largest diaspora population in the world [16], with high social media use, and transnational communication via messaging means that misinformation can have global reach almost instantaneously. Our analysis adds to the literature by providing a window into the nature of these closed group conversations. A prior narrative review of misinformation across Asian American communities [17] and a study of Twitter misinformation in Hindi [18] both highlighted religious-based content, including Islamophobic messaging, as the major thematic finding for South Asians. We instead found a focus on understanding, preventing, and treating the spread of the virus using alternative or Ayurvedic methods. Our study is the first to our knowledge to assess this topic within the broader global South Asian diaspora.

COVID-19 misinformation, disinformation, and malinformation have persisted due to a perfect storm of evolving uncertainty of the disease, malintent of actors pushing political or business interests, and the well-meaning intentions of community members who may have low health or media literacy [19]. The South Asian community has experienced a disproportionate burden of COVID-19 with less media attention, particularly extensive collective trauma after the 2021 Delta variant surge, which may have killed more than 3 million people [20]. South Asian audiences may thus be particularly receptive to promoting messages that provide a sense of clarity, trustworthiness, and personal control in an uncertain time.

Misinformation circulates more readily among homogenous groups or “echo chambers” [21], and misinformation with culturally specific language to promote “in-group identity” may receive higher engagement [22,23]. Messages broadcasting cultural pride may therefore be more readily amplified within relatively insular groups of South Asian users. Messages included “name checks” and logos of reputable organizations or individuals, enhancing trust [24]. Factual information was blended with false statements, mimicking credibility. Lastly, participants were often exhorted to share messages with others, encouraging the spread of misinformation under the well-meaning intention of promoting community safety.

The closed, trusted groups in WhatsApp and similar platforms, such as Viper, Weibo, and Signal, may actually foment misinformation [25,26]. WhatsApp currently puts the burden on individual users to stop its spread [27]. In epidemiological modeling, limits on the number of message forwards slows the spread of misinformation but does not ultimately stop the spread of viral content [28]. Clear countermessaging to identify and correct misinformation can be effective [29]. when promoted across a multitude of platforms; however, more research is needed to identify how to best countermessage COVID-19–related misinformation without causing unintended backfire [30,31]. Social media corporations must do more to monitor, detect, and possibly delete or flag dangerous misinformation. Fact-checking organizations can also perform this role and circulate countermessaging [32,33]. South Asian–specific, culturally literate public health experts and community organizations should collaborate with Fact-checking organizations to create and disseminate countermessaging.

Given the complexity of misinformation, multiple counteracting strategies beyond countermessaging will be needed. This will include the use of machine learning to flag or identify misinformation more easily, policy changes to increase legal accountability for harmful misinformation, and heightened scrutiny and investigation of organizations and web-based influencers who are frequent sources or spreaders of misinformation [30].

These web-based myths can have very real consequences. The spread of culturally specific misinformation may lead to unsafe health behaviors [34] and contribute to preventable burdens of COVID-19 among South Asian communities [35,36]. Given the vast diversity of ethnic groups included by the term “Asian,” more funding and research to promote disaggregated data collection and analysis (eg, for specific East Asian, South Asian, Southeast Asian, and other groups) [37] is greatly needed to understand and hopefully counteract culturally specific misinformation regarding COVID-19 and future pandemics.

Strengths and Limitations

Study strengths include a transnational sample and focus on a social media platform not publicly visible for analysis. We were, however, unable to collect demographic information about original message senders or the recipients who forwarded the messages. We relied on participants to self-identify as members of the South Asian diaspora; participants may have had varying definitions of this identity. Our focus on English-language messages likely limited the content of the final sample, given the linguistic diversity of the South Asian diaspora. Our study focused on adult participants who could give informed consent; therefore, our analytic sample may not be generalizable for COVID-19 misinformation disseminated among children and adolescents. We could not gauge whether specific messages were shared with intent to harm, as with disinformation or malinformation, for purposes of “collective fact-checking” [38], or shared because they were genuinely believed. As our data collection relied on users identifying potential misinformation, our team may have not received the misinformation that users viewed in their WhatsApp feed but believed to be true. A potential limitation of our methods is that although the individual participant consented to share anonymous content from their WhatsApp, there were no means to obtain permission for analysis from the original poster or others who had forwarded the message. We ensured both the individual participant and any other WhatsApp users had zero identifiable information collected. Moreover, WhatsApp does not have any way to identify original posters, nor prior forwarders, of a specific message. Although our ethical approach was to maximize public health benefit and avoid any individual harm, interdisciplinary gold standards for social media research are still needed [39]. Given that messages were sent on a voluntary basis from individuals who chose to participate, this research cannot be viewed as a systematic analysis of misinformation shared across the South Asian diaspora on private messaging platforms. The sample should be considered hypothesis generating; larger samples may provide greater insights into messaging among the global South Asian diaspora.

Conclusions

We found that COVID-19–related misinformation from WhatsApp messages within the South Asian diaspora relate to four themes: transmission, prevention, treatment, and product or service promotion. Tactics to enhance credibility and spread of messages included use of jargon, blending of true and false information, mention of reputable organizations and expert credentials, and references to ethnic pride. Encouragement to share misinformation messages among personal networks makes it urgent to find ways to interrupt misinformation in real time so as not to exacerbate COVID-19 disparities. Novel public health strategies, including culturally specific fact-checking, will be needed to counteract misinformation among the South Asian diaspora.

Acknowledgments

We thank the University of California, San Francisco (UCSF) Department of Family and Community Medicine research division for providing startup funding for this work. We would also like to thank our fellow members of the global South Asian community who participated in the COVID-19–Associated misinfoRmation On Messaging apps (CAROM) study and promoted it among their networks. Thanks to Sarah Lisker for institutional review board support, Jessica Davidman for citation management, Sylvia Chou for contextual advice, and Marika Dy for background research and formatting assistance.

AES is supported by the Agency for Healthcare Research and Quality (K08HS028477-01).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Recruitment materials for COVID-19–Associated misinfoRmation On Messaging apps (CAROM) Study.

[DOCX File , 1691 KB - infodemiology_v3i1e38607_app1.docx]

Multimedia Appendix 2

Abstraction form.
[PDF File (Adobe PDF File), 45 KB - infodemiology_v3i1e38607_app2.pdf]

References

1. WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU. Managing the COVID-19 infodemic: promoting healthy behaviours and mitigating the harm from misinformation and disinformation. UNICEF. 2020 Sep 23. URL: https://www.unicef.org/press-releases/managing-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-harm [accessed 2022-04-08]
2. Journalism, 'fake news' and disinformation: a handbook for journalism education and training. UNESCO. 2018. URL: https://en.unesco.org/fightfakenews [accessed 2022-11-18]
3. Gabarron E, Oyeyemi SO, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bull World Health Organ 2021 Jun 01;99(6):455-463A [FREE Full text] [doi: 10.2471/BLT.20.276782] [Medline: 34108756]
4. Mis, dis, malinformation. Cybersecurity & Infrastructure Security Agency. URL: https://www.cisa.gov/mdm [accessed 2022-11-18]
5. Ormel I, Stalteri R, Goel V, Horton S, Puchalski-Ritchie L, Milaney K, et al. Misinformation and disinformation in relation to COVID-19. CanCOVID Issue Note 2021 Apr 07:1-38 [FREE Full text] [doi: 10.4324/9781003004431-5]
6. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 09;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
7. Yang K, Pierri F, Hui P, Axelrod D, Torres-Lugo C, Bryden J, et al. The COVID-19 Infodemic: Twitter versus Facebook. Big Data & Society 2021 May 05;8(1):205395172110138. [doi: 10.1177/20539517211013861]
8. Pulipaka A. South Asian Diaspora. NYU Libraries. URL: https://guides.nyu.edu/southasiandiaspora/Home [accessed 2022-04-08]
9. Jain S. Indians Most Active On WhatsApp With 390.1 million monthly active users in 2020. Forbes India. 2021 Aug 27. URL: https://www.forbesindia.com/article/news-by-numbers/indians-most-active-on-whatsapp-with-3901-million-monthly-active-users-in-2020/70059/1 [accessed 2022-04-08]
10. Islam MS, Sarkar T, Khan SH, Mostofa Kamal A, Hasan SMM, Kabir A, et al. COVID-19-Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis. Am J Trop Med Hyg 2020 Oct;103(4):1621-1629 [FREE Full text] [doi: 10.4269/ajtmh.20-0812] [Medline: 32783794]
11. Sallam M, Dababseh D, Eid H, Al-Mahzoum K, Al-Haidar A, Taim D, et al. High rates of COVID-19 vaccine hesitancy and its association with conspiracy beliefs: a study in Jordan and Kuwait among other arab countries. Vaccines (Basel) 2021 Jan 12;9(1):42 [FREE Full text] [doi: 10.3390/vaccines9010042] [Medline: 33445581]
12. Bursztyn L, Rao A, Roth C, Yanagizawa-Drott D. Misinformation during a pandemic. SSRN Journal 2020:1-114. [doi: 10.2139/ssrn.3580487]
13. Kalyanaraman Marcello R, Dolle J, Tariq A, Kaur S, Wong L, Curcio J, et al. Disaggregating Asian race reveals COVID-19 disparities among Asian American patients at New York City's public hospital system. Public Health Rep 2022;137(2):317-325 [FREE Full text] [doi: 10.1177/00333549211061313] [Medline: 34965776]
14. Mathur R, Rentsch CT, Morton CE, Hulme WJ, Schultze A, MacKenna B, OpenSAFELY Collaborative. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. Lancet 2021 May 08;397(10286):1711-1724 [FREE Full text] [doi: 10.1016/S0140-6736(21)00634-6] [Medline: 33939953]
15. Ramachandran V. Viral virus texts. Red, White and Brown. 2021 Apr 08. URL: https://redwhiteandbrown.substack.com/p/viral-virus-texts [accessed 2022-11-18]
16. Singh R. Origin of world's largest migrant population, India seeks to leverage immigration. Migration Policy Institute. 2022 Mar 09. URL: https://www.migrationpolicy.org/article/india-migration-country-profile [accessed 2022-11-18]
17. Chong SK, Ali SH, Đoàn LN, Yi SS, Trinh-Shevrin C, Kwon SC. Social Media Use and Misinformation Among Asian Americans During COVID-19. Front Public Health 2022 Jan 14;9:764681 [FREE Full text] [doi: 10.3389/fpubh.2021.764681] [Medline: 35096736]
18. Park S, Han S, Kim J, Molaie MM, Vu HD, Singh K, et al. COVID-19 discourse on Twitter in four Asian countries: case study of risk communication. J Med Internet Res 2021 Mar 16;23(3):e23272 [FREE Full text] [doi: 10.2196/23272] [Medline: 33684054]
19. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Roundtable on Health Literacy. Addressing health misinformation with health literacy strategies: proceedings of a Workshop—in brief. In: National Academies Press. 2020 Dec 17 Presented at: Roundtable on Health Literacy; July 29; Virtual p. 1-8 URL: https://nap.nationalacademies.org/read/26021/chapter/1 [doi: 10.17226/26021]
20. Jha P, Deshmukh Y, Tumbe C, Suraweera W, Bhowmick A, Sharma S, et al. COVID mortality in India: National survey data and health facility deaths. Science 2022 Feb 11;375(6581):667-671 [FREE Full text] [doi: 10.1126/science.abm5154] [Medline: 34990216]
21. Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, et al. The spreading of misinformation online. Proc Natl Acad Sci U S A 2016 Jan 19;113(3):554-559 [FREE Full text] [doi: 10.1073/pnas.1517441113] [Medline: 26729863]
22. Chen K, Jin Y, Shao A. Social Media + Society 2022 Mar 11;8(1):205630512210770. [doi: 10.1177/20563051221077019]

23. Chen K, Shao A, Jin Y, Ng A. I am proud of my national identity and I am superior to you: the role of nationalism in knowledge and misinformation. SSRN Journal 2021 Feb 18:1-34. [doi: 10.2139/ssrn.3758287]
24. Trivedi N, Krakow M, Hyatt Hawkins K, Peterson E, Chou W. "Well, the message is from the institute of something": exploring source trust of cancer-related messages on simulated Facebook posts. Front Commun 2020 Feb 28;5(12):1-9 [FREE Full text] [doi: 10.3389/fcomm.2020.00012]
25. Johns A, Cheong N. The affective pressures of WhatsApp: from safe spaces to conspiratorial publics. Continuum 2021 Oct 18;35(5):732-746. [doi: 10.1080/10304312.2021.1983256]
26. Banaji S, Bhat R, Agarwal A, Passanha N, Sadhana PM. WhatsApp vigilantes: an exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India. The London School of Economics and Political Science. London, UK; 2019 Nov 11. URL: https://eprints.lse.ac.uk/104316/1/Banaji_whatsapp_vigilantes_exploration_of_citizen_reception_published.pdf [accessed 2022-04-08]
27. WhatsApp Help Center. How to prevent the spread of misinformation. URL: https://faq.whatsapp.com/general/security-and-privacy/how-to-prevent-the-spread-of-misinformation/?lang=en [accessed 2022-04-08]
28. Melo PDF, Vieira C, Garimella K, de MP, Benevenuto F. Can WhatsApp counter misinformation by limiting message forwarding? ArXiv Preprint posted online Sep 18, 2019 [FREE Full text] [doi: 10.48550/arXiv.1909.08740]
29. Vraga EK, Bode L. Correction as a Solution for Health Misinformation on Social Media. Am J Public Health 2020 Oct;110(S3):S278-S280. [doi: 10.2105/AJPH.2020.305916] [Medline: 33001724]
30. Calo R, Coward C, Spiro ES, Starbird K, West JD. How do you solve a problem like misinformation? Sci Adv 2021 Dec 10;7(50):eabn0481 [FREE Full text] [doi: 10.1126/sciadv.abn0481] [Medline: 34878833]
31. Wood T, Porter E. The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence. Polit Behav 2018 Jan 16;41(1):135-163. [doi: 10.1007/s11109-018-9443-y]
32. MisinformationDisinformation: TCAISCOVID1FCS. Misinformation and disinformation: thinking critically about information sources. CSI Library. URL: https://library.csi.cuny.edu/c.php?g=619342&p=7605649 [accessed 2022-04-08]
33. The CoronaVirusFacts/DatosCoronaVirus Alliance Database. Poynter. URL: https://www.poynter.org/ifcn-covid-19-misinformation/ [accessed 2022-04-08]
34. Kim HK, Tandoc EC. Consequences of online misinformation on COVID-19: two potential pathways and disparity by ehealth literacy. Front Psychol 2022 Feb 14;13:1-11 [FREE Full text] [doi: 10.3389/fpsyg.2022.783909] [Medline: 35237207]
35. Chakraborty T, Subbiah G, Kulshrestha R, Subbiah K, Swathi G, Damade Y. Impact of the COVID 19 Infodemic on Knowledge, Attitude and Social Behavior in India: A Mixed Method Survey. Int J Intell Secur Public Aff 2021 Sep 30;23(3):197-215. [doi: 10.1080/23800992.2021.1956777]
36. Kanozia R, Arya R. "Fake news", religion, and COVID-19 vaccine hesitancy in India, Pakistan, and Bangladesh. Media Asia 2021 May 17;48(4):313-321 [FREE Full text] [doi: 10.1080/01296612.2021.1921963]
37. Shimkhada R, Scheitler AJ, Ponce NA. Capturing racial/ethnic diversity in population-based surveys: data disaggregation of health data for Asian American, Native Hawaiian, and Pacific Islanders (AANHPIs). Popul Res Policy Rev 2021 Jan 16;40(1):81-102 [FREE Full text] [doi: 10.1007/s11113-020-09634-3]
38. Metzger MJ, Flanagin AJ, Mena P, Jiang S, Wilson C. From dark to light: the many shades of sharing misinformation online. Media and Communications 2021 Feb 03;9(1):134-143 [FREE Full text] [doi: 10.17645/mac.v9i1.3409]
39. Pagoto S, Nebeker C. How scientists can take the lead in establishing ethical practices for social media research. J Am Med Inform Assoc 2019 Apr 01;26(4):311-313 [FREE Full text] [doi: 10.1093/jamia/ocy174] [Medline: 30698793]

Abbreviations

CAROM: COVID-19–Associated misinfoRmation On Messaging apps
UNICEF: United Nations International Children's Emergency Fund

Edited by T Mackey; submitted 11.04.22; peer-reviewed by L Heidi, F Medina; comments to author 22.10.22; revised version received 18.11.22; accepted 30.11.22; published 05.01.23.

Please cite as:
Sharma AE, Khosla K, Potharaju K, Mukherjea A, Sarkar U
COVID-19–Associated Misinformation Across the South Asian Diaspora: Qualitative Study of WhatsApp Messages
JMIR Infodemiology 2023;3:e38607
URL: https://infodemiology.jmir.org/2023/1/e38607
doi:10.2196/38607
PMID:37113380

Original Paper

Measuring the Burden of Infodemics: Summary of the Methods and Results of the Fifth WHO Infodemic Management Conference

Elisabeth Wilhelm[1*], MA; Isabella Ballalai[2*], MD; Marie-Eve Belanger[3*], PhD; Peter Benjamin[4*], MPH, MD; Catherine Bertrand-Ferrandis[5*], DVM; Supriya Bezbaruah[6*], PhD; Sylvie Briand[6*], MD, PhD; Ian Brooks[7*], PhD; Richard Bruns[8*], PhD; Lucie M Bucci[9*], MA; Neville Calleja[10*], MD, PhD; Howard Chiou[1,11*], MD, PhD; Abhinav Devaria[12*], MS; Lorena Dini[13*], MScIH, PhD; Hyjel D'Souza[12*], BSc; Adam G Dunn[14*], PhD; Johannes C Eichstaedt[15,16*], PhD; Silvia M A A Evers[17*], LLM, PhD; Nina Gobat[18*], PhD; Mika Gissler[19*], MSocSci, DrPhil; Ian Christian Gonzales[20*], MD; Anatoliy Gruzd[21*], PhD; Sarah Hess[6*], MPH; Atsuyoshi Ishizumi[6*], MSc; Oommen John[12*], MD, PhD; Ashish Joshi[22*], MBBS, MPH, PhD; Benjamin Kaluza[23*], PhD; Nagwa Khamis[24*], MD; Monika Kosinska[25*], MA; Shibani Kulkarni[1*], MPH, PhD; Dimitra Lingri[26*], LLM; Ramona Ludolph[6*], MPH, PhD; Tim Mackey[27*], MAS, PhD; Stefan Mandić-Rajčević[28*], MD, PhD; Filippo Menczer[29*], PhD; Vijaybabu Mudaliar[12*], BPharma, PGDCRM, MS; Shruti Murthy[12*], BDS, MPH, PhD; Syed Nazakat[30*], MA; Tim Nguyen[6*], MS; Jennifer Nilsen[31*], MPP; Elena Pallari[32*], PhD; Natalia Pasternak Taschner[33,34*], PhD; Elena Petelos[35,36*], MPH, PhD; Mitchell J Prinstein[37,38*], PhD; Jon Roozenbeek[39*], PhD; Anton Schneider[40*], AB; Varadharajan Srinivasan[12*], MPhil; Aleksandar Stevanović[28*], MD; Brigitte Strahwald[41*], MSc, Mmel; Shabbir Syed Abdul[12,42*], PhD; Sandra Varaidzo Machiri[43*], MPH; Sander van der Linden[39*], PhD; Christopher Voegeli[44*], MPH, PhD; Claire Wardle[45*], PhD; Odette Wegwarth[46*], PhD; Becky K White[6*], PhD; Estelle Willie[47*], MPH; Brian Yau[6*], BSc; Tina D Purnat[6*], MSc

[1]US Centers for Disease Control and Prevention, Atlanta, GA, United States
[2]Brazilian Inmunization Society, São Paulo-SP, Brazil
[3]Department of Political Science and International Relations, Université de Genève, Geneva, Switzerland
[4]HealthEnabled, Cape Town, South Africa
[5]Olylo, Paris, France
[6]Department of Epidemic and Pandemic Preparedness and Prevention, World Health Organization, Geneva, Switzerland
[7]Center for Health Informatics, School of Information Sciences, University of Illinois, Champaign, IL, United States
[8]Johns Hopkins Center for Health Security, Baltimore, MD, United States
[9]Immunize Canada, Canadian Public Health Association, Ottawa, ON, Canada
[10]Directorate for Health Information and Research, Ministry for Health, Valletta, Malta
[11]US Public Health Service Commissioned Corps, Rockville, MD, United States
[12]The George Institute for Global Health, New Delhi, India
[13]Working Group Health Policy and Systems Research and Innovation, Institute for General Practice, Charité Universitätsmedizin Berlin, Berlin, Germany
[14]Biomedical Informatics and Digital Health, Faculty of Medicine and Health, University of Sydney, Sydney, Australia
[15]Department of Psychology, Stanford University, Stanford, CA, United States
[16]Institute for Human-Centered AI, Stanford University, Stanford, CA, United States
[17]Department of Health Services Research, Maastricht University, Maastricht, Netherlands
[18]Department of Country Readiness Strengthening, World Health Organization, Geneva, Switzerland
[19]Department of Knowledge Brokers, THL Finnish Institute for Health and Welfare, Helsinki, Finland
[20]Field Epidemiology Training Program, Epidemiology Bureau, Department of Health, Manila, Philippines
[21]Ted Rogers School of Management, Toronto Metropolitan University, Toronto, ON, Canada
[22]Department of Epidemiology and Biostatistics, Graduate School of Public Health and Health Policy, City University of New York, New York, NY, United States
[23]Department Technological Analysis and Strategic Planning, Fraunhofer Institute for Technological Trend Analysis INT, Euskirchen, Germany
[24]Infection Prevention and Control Department, Children's Cancer Hospital Egypt-57357, Ain Shams University Specialized Hospital, Cairo, Egypt
[25]Department of Social Determinants, World Health Organization, Geneva, Switzerland
[26]European Healthcare Fraud and Corruption Network, Aristotle Universtity of Thessaloniki, Brussels, Belgium
[27]Global Health Program, Department of Anthropology, University of California, San Diego, CA, United States

[28]Institute of Social Medicine, Faculty of Medicine, University of Belgrade, Belgrade, Serbia
[29]Observatory on Social Media, Luddy School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, United States
[30]DataLEADS (Health Analytics Asia), New Delhi, India
[31]Technology and Social Change Project, Harvard University, Cambridge, MA, United States
[32]Health Innovation Network, Guy's and St Thomas' Hospital, London, United Kingdom
[33]Center of Science and Society, Columbia University, New York, NY, United States
[34]Instituto Questão de Ciência, São Paulo, Brazil
[35]Department of Health Services Research, Care and Public Health Research Institute, Maastricht University, Maastricht, Netherlands
[36]Clinic of Social and Family Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece
[37]American Psychological Association, Washington DC, DC, United States
[38]Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[39]Department of Psychology, University of Cambridge, Cambridge, United Kingdom
[40]Bureau for Global Health, Office of Infectious Disease, United States Agency for International Development, Washington DC, DC, United States
[41]Pettenkofer School of Public Health, Ludwig-Maximilians-Universität München, Munich, Germany
[42]Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei, Taiwan
[43]African Field Epidemiology Network, Harare, Zimbabwe
[44]Office of the Director, National Center for Immunization and Respiratory Diseases, US Centers for Disease Control and Prevention, Atlanta, GA, United States
[45]Information Futures Lab, School of Public Health, Brown University, Providence, RI, United States
[46]Heisenberg Chair for Medical Risk Literacy & Evidence-Based Decisions, Charite – Universitätsmedizin Berlin, Berlin, Germany
[47]Communications, Policy, Advocacy, The Rockefeller Foundation, New York, NY, United States
[*]all authors contributed equally

Corresponding Author:
Tina D Purnat, MSc
Department of Epidemic and Pandemic Preparedness and Prevention
World Health Organization
Avenue Appia 20
Geneva, 1211
Switzerland
Phone: 41 22 791 21 11
Email: purnatt@who.int

Abstract

Background: An infodemic is excess information, including false or misleading information, that spreads in digital and physical environments during a public health emergency. The COVID-19 pandemic has been accompanied by an unprecedented global infodemic that has led to confusion about the benefits of medical and public health interventions, with substantial impact on risk-taking and health-seeking behaviors, eroding trust in health authorities and compromising the effectiveness of public health responses and policies. Standardized measures are needed to quantify the harmful impacts of the infodemic in a systematic and methodologically robust manner, as well as harmonizing highly divergent approaches currently explored for this purpose. This can serve as a foundation for a systematic, evidence-based approach to monitoring, identifying, and mitigating future infodemic harms in emergency preparedness and prevention.

Objective: In this paper, we summarize the Fifth World Health Organization (WHO) Infodemic Management Conference structure, proceedings, outcomes, and proposed actions seeking to identify the interdisciplinary approaches and frameworks needed to enable the measurement of the burden of infodemics.

Methods: An iterative human-centered design (HCD) approach and concept mapping were used to facilitate focused discussions and allow for the generation of actionable outcomes and recommendations. The discussions included 86 participants representing diverse scientific disciplines and health authorities from 28 countries across all WHO regions, along with observers from civil society and global public health–implementing partners. A thematic map capturing the concepts matching the key contributing factors to the public health burden of infodemics was used throughout the conference to frame and contextualize discussions. Five key areas for immediate action were identified.

Results: The 5 key areas for the development of metrics to assess the burden of infodemics and associated interventions included (1) developing standardized definitions and ensuring the adoption thereof; (2) improving the map of concepts influencing the burden of infodemics; (3) conducting a review of evidence, tools, and data sources; (4) setting up a technical working group; and (5) addressing immediate priorities for postpandemic recovery and resilience building. The summary report consolidated group input toward a common vocabulary with standardized terms, concepts, study designs, measures, and tools to estimate the burden of infodemics and the effectiveness of infodemic management interventions.

Conclusions: Standardizing measurement is the basis for documenting the burden of infodemics on health systems and population health during emergencies. Investment is needed into the development of practical, affordable, evidence-based, and systematic methods that are legally and ethically balanced for monitoring infodemics; generating diagnostics, infodemic insights, and recommendations; and developing interventions, action-oriented guidance, policies, support options, mechanisms, and tools for infodemic managers and emergency program managers.

(JMIR Infodemiology 2023;3:e44207) doi:10.2196/44207

KEYWORDS

COVID-19; infodemic; burden of infodemic; infodemic management; infodemic metrics; World Health Organization; technical consultation; infodemiology

Introduction

The Challenge That Infodemics Pose to Health System Response in Emergencies

An infodemic is excess information of varying quality, including false/misleading information or ambiguous information or both, that spreads in digital and physical environments during a health emergency [1,2]. Infodemics are more complex than just the amplification and spread of mis- and disinformation; they spread across a wider information landscape where population questions, concerns, and information voids can lead to misinformation growth and spread, particularly in societies undergoing digital transformation. The COVID-19 pandemic has been accompanied by an unprecedented global infodemic that has led to confusion about the benefits of medical and public health interventions, with substantial impact on risk-taking and health-seeking behaviors, eroding trust in health authorities and compromising the effectiveness of public health responses and policies [3].

There are several key concepts that are integral to discussing infodemics and how they link to health authority responses, including the online information environment; the channels, formats, and quality of health information people are exposed to; individual-level literacy; the psychology of emergencies; and the multifaceted aspects of trust and how they influence perception and behavior in health. Many of these areas have bodies of research and literature and measures associated with them in specific fields of study, such as psychology, governance and policy, and digital user experience, but they are usually not connected in a systematic, causal way that is applicable to how health systems act in emergencies.

The Information Environment and Accessing Health Information

As the world becomes more digitized, the digital information environment increasingly influences social dynamics between people and across communities, influencing health decisions and behaviors [4-6]. The accessibility and availability of health information, the algorithms of social media platforms, the architecture of online communities and news channels, and format all impact how individuals receive and act on health information [4,6]. Creating and updating credible, accurate health information for dissemination to different audiences is within the purview of health authorities and tends to be formalized in policies for health care delivery and in public health matters, especially in emergencies [7]. However, health information is often shared through unofficial and unregulated channels and made available in a wider variety of formats and for channels not typically used by the health system or for communication of public health guidance, creating a gap between which communities have access to official and credible information and those that do not [8]. For example, TikTok and closed messaging networks, such as WhatsApp and Telegram, have increasingly been used to share health information and misinformation [9].

Literacies Related to Health, Infodemics, and Emergencies

Simply having access to health information is insufficient for instigating positive behavior change [10]. Health, digital, media, science, information, and influence literacies all play a role at the individual level, mediating between the availability of health information and the individual ability to process, understand, and act on it [11]. However, in emergencies, people seek, process, and act on information differently, looking for information to protect themselves and their families, even though information may be scarce, and looking for alternate sources of information, while tending to believe the first thing they hear. There are examples related to noncommunicable diseases, such as tobacco cessation campaigns, that aim to address health literacy gaps and counter harm from misinformation [12]. Teaching critical thinking skills to individuals to be able to identify and rebut health misinformation can broadly inoculate against specific misinformation narratives and is one promising intervention for building resilient individuals and communities against misinformation. Therefore, building skills and resilience against misinformation and other infodemic harms and improving the use of health information during times of calm are not sufficient alone to help people during emergencies, where traditional health communications pathways, such as communicating with a primary health care provider, may have been interrupted.

Building Trust to Prevent Erosion During Emergencies

Building trust in health authorities is critical before emergencies strike, because infodemics can quickly erode trust, especially when there is low trust at baseline. Trust contributes toward the willingness to accept and adopt necessary measures and can be the deciding factor in terms of how successful the implementation of a sound public health strategy will be—for example, in the context of implementing public health and social measures to control a disease as part of containment strategies. Trust can be eroded by what the public may perceive as conflicting guidance and mixed messages, information released late, multiple experts with divergent opinions, paternalism, and political infighting [13]. Infodemics can further add friction by promoting misinformation and more destructive forms, such as disinformation or conspiracy theories; not addressing people's questions and concerns in a timely manner; or leaving people struggling to access accurate, credible, and up-to-date health information [14].

Why Measure the Burden and Cost of Infodemics?

Due to the multifaceted nature of infodemics affecting individuals, communities, societies, economies, and health systems during emergencies, it can be difficult to know how to prepare for infodemics, determine when and where to intervene, and understand how to more effectively reduce harm to population health. Globally applicable infodemic measurements and metrics are needed. The true cost of infodemics has not been robustly measured but has been anecdotally reported, with impacts such as stigma, violence against health workers, overdoses of nonrecommended treatments or stockouts, refusal by individuals or communities to wear masks or get vaccinated, and frivolous lawsuits against health systems and health care workers [15]. One academic brief suggested that COVID-19 misinformation cost US $50-$300 million a day at the height of the pandemic in the United States [16]. Without measures or costing, it is difficult to develop effective interventions and advocate for supportive policies. More innovation in how measurements and metrics are developed is needed due to the multilevel nature of the phenomenon and the sheer diversity of disciplines and in-depth expertise required to measure or estimate different aspects of infodemics.

Spurring the Development of Metrics to Measure the Burden of Infodemics and Interventions as Part of the WHO Public Health Research Agenda on Infodemiology

Since the beginning of the COVID-19 pandemic, the World Health Organization (WHO) has expanded the concept of infodemiology beyond the use of data produced and consumed on the web to inform public health officials, agencies, and policy into a multidisciplinary scientific field. Interventions must account for an information environment where information flows online and offline, highly tailored to people's information diets, and their responses can lead to nonprotective behaviors and poor health outcomes offline [1,2,17]. Building harmonized measures and cohesive interventions requires an amalgamation of cross-disciplinary and mixed methods approaches to inform the health emergency response and routine health system–strengthening efforts online and offline [17].

Early in the COVID-19 response, the First WHO Infodemiology Conference in June-July 2020 brought together experts from a range of disciplines to begin a global conversation about the science of infodemiology and establish a public health research agenda for managing infodemics, recognizing that each discipline has a different perspective on the problems of infodemics, different ways of measurement, and a different vocabulary [17,18]. Although previous conferences have expanded our understanding of infodemic drivers [19] and social listening approaches [20], the Fifth WHO Infodemic Management Conference aimed to collaboratively develop a proposed action plan to foster implementation for work stream 1 of the WHO public health research agenda for managing infodemics: the development of metrics and indicators for measuring the burden of an infodemic and related interventions. The full conference report is available on the WHO website [21]. In this paper, we summarize the conference structure, proceedings, outcomes, and proposed actions.

Methods

Overview

The conference used an iterative human-centered design (HCD) approach in line with the purpose-outcome-process (POP) model, a tool for focusing actions on creating results [22,23]. Held in the context of the COVID-19 pandemic and with travel restrictions in place, the meeting necessarily took place online via videoconference. The virtual discussions took place over four 3-hour meetings during 2 weeks in November 2021, resulting in a summary report and recommended actions to advance 5 key areas for the development of metrics to assess the burden of infodemics. The summary report consolidated the participants' input for a common vocabulary, concepts, standardized study designs, measures, and tools to estimate the burden of infodemics or the effectiveness of infodemic management interventions.

Ethical Considerations

Institutional Review Board review was not sought because the work described in this paper was based on observation of discussions at the conference and focused on the synthesis of expert opinion following the Chatham House Rule [24]. No personal information was collected from participating experts.

Design Approaches to Promote Effective Interdisciplinary Discussion

The organizers used an HCD approach to intentionally facilitate engaging and effective conference deliberations [23]. First, the conference format was designed to offer a level playing field for all participants who were encouraged to contribute their knowledge in an environment where most participants came from extremely diverse disciplines, backgrounds, country settings, and professional experiences. Second, the conference structure was designed with the help of the POP model [22] to provide a structured output to conference deliberations. Third, sessions were scheduled on different days, allowing the organizing team to synthesize inputs and prepare for the next session and adapt the deliberations and format. Consideration was given to what and how essential information was shared with participants before and during the conference sessions, to the emotional pacing of the interactions that would support intense cross-disciplinary expert deliberations, and to the environment that would support the participant behaviors and discussions toward actionable recommendations.

First Approach: Designing for Consensus on Outcomes and Recommendations

The 4 conference sessions were structured over the following thematic areas for a cumulative duration of 12 hours. Facilitators directed participant discussions to arrive at actionable recommendations on the last day. Ahead of each session, the design of the proceedings and the group discussion tasks and visuals were also tested with the cochairs of the conference and a group of experts, and the feedback was used to set clear discussion tasks and discussion aids. This approach aimed to prepare each session discussions by building on the collective knowledge from previous sessions and to effectively facilitate technical discussions despite complex multidisciplinary topics.

The meeting schedule was designed as follows: Ahead of each of the 3-hour virtual meetings, the organizing team prepared introductory talks to set the task of the day, defined discussion questions, developed visual aids, and designed the discussion process. During the session, the outcomes of discussions were recorded by facilitators and note takers on Miro boards that were used during the session. After each session, debriefs with breakout group facilitators reflected on the group dynamics and technical discussion. The organizing team used all this information to adapt and refine the preparation of the next session, including the discussion questions and discussion inputs on Miro. Moreover, after each session and debrief, the organizing team updated the concept map on a summary Miro board to capture the progressive discussions and made it available for asynchronous review and comments by conference participants.

Synthesis of discussions using thematic analysis by the organizing team led to the identification of 5 key areas for immediate action for the development of metrics to assess the burden of infodemics and associated interventions. They were summarized alongside a participant-generated list of proposed actions and concrete next steps for each area for implementation.

Second Approach: Using the Purpose-Outcome-Process Model for the Conference

The purpose of the meeting was to determine how to measure the burden of infodemics associated with the information mix people access and the associated drivers for people's behaviors over time and to discuss new ways to characterize information exposure and health outcomes that support this measurement. The expected outcomes of the meeting were to synthesize collective feedback and arrive at concrete next steps on (1) a concept map on the main pathways on the wider effects of infodemics (individual, society, health system, and policy); (2) a list of principles for ranking and prioritization of concepts and indicators to be used; (3) a prioritized list of actions, study designs, and metrics that need development; (4) the establishment of collaborations to advance the work. Because the expected outcomes were ambitious for the planned total 12 hours of deliberations, careful consideration was given to how the conference outcomes could best benefit from the expertise of participating senior academics and policy makers.

Third Approach: Designing for Emotional Pacing, Engagement, and Behaviors Supportive of Desired Conference Outcomes

Experience from previous WHO infodemic management meetings has shown that infodemiology discussions often require a design that helps overcome barriers in differences in the language, terminology, and focus of the actions or aims of research between researchers from different disciplines and practitioners from different health programs or evidence-informed policy functions in health authorities [17,18,25]. Several meeting design features aimed to address this:

- The concept map and lightning talks by experts at the beginning of the day were used to highlight perspectives from different scientific disciplines or public health practice on the discussion task of the day.
- Facilitators of small group discussions were coached and provided with facilitator guides with prompts to help them move the discussion toward the task and were given a space on the discussion boards, where they could record suggestions tangential to the task at hand.
- The schedule deliberately emphasized more discussion time in smaller groups in comparison to in-plenary to allow for maximum participation and exchange of experience.
- The synthesis of collective discussion was used to prepare for the next session. This was a resource-intensive activity that aimed to learn as much as possible from participants, while keeping them interested, engaged, and motivated to provide further input in the next session.
- The organizing team reflected back to the group not only a technical summary of the discussions but also the observations on the discussions—for example, the development of a common understanding of vocabulary and small group identities.
- Because the discussions were highly technical and required intense engagement, breaks were designed to be playful. Music videos on the topics of public health and science were played at the beginning of the meeting and during breaks to set the tone of interactions at the conference.

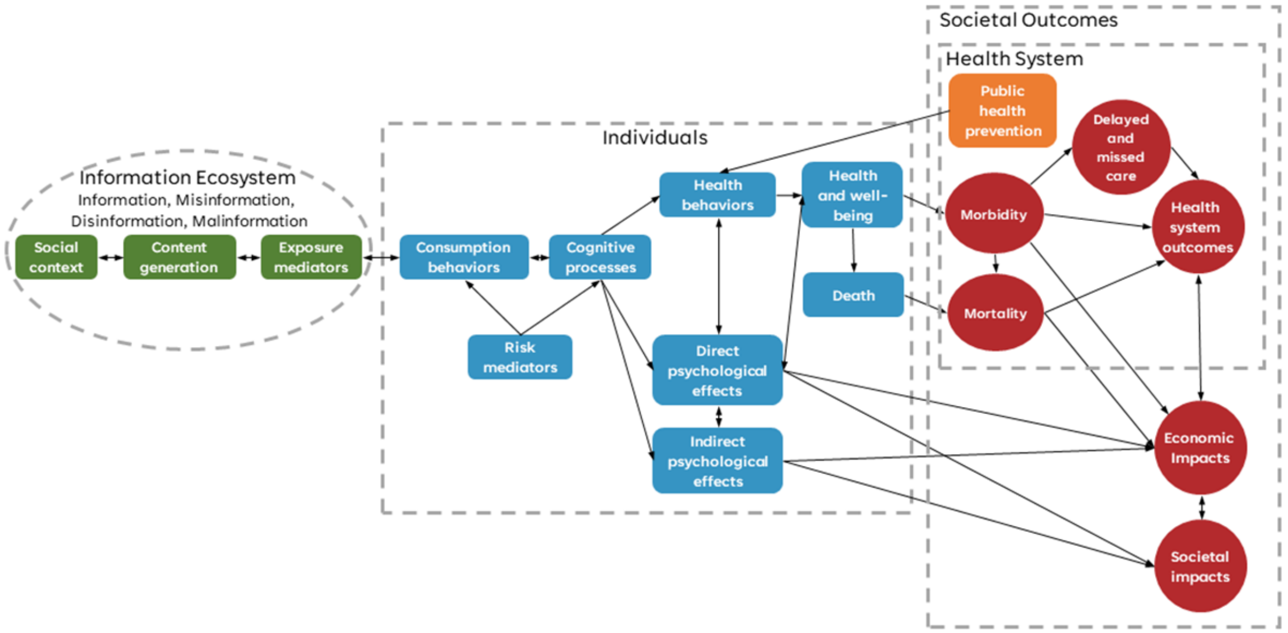
Profile of Participants

The 86 invited participants included academics and public health practitioners from 48 organizations, including voices from 28 countries across 18 time zones, with a 56%:44% gender split in favor of women (n=48 females vs n=38 males). In addition, 48 additional invited academics and policy makers were not available to participate. The conference participants were academics selected by the organizers for the relevance of their publication record in the past 2 years for the purpose of this meeting or practitioners who were working in health metrics, measurement, and health program implementation. The participants also included 5 observers from civil society and global public health implementing partners. Conflicts of interest were reviewed in accordance with WHO procedures for the management of the declaration of interest for expert consultations [26]. An extended conference-organizing team comprising 32 members was drawn from across the WHO, the US Centers for Disease Control and Prevention (US CDC), and the George Institute for Global Health (TGI), India. More information about the structure and methodology of the conference is detailed in Multimedia Appendix 1.

Framing Discussions With a Concept Map of the Wider Impacts of Infodemics

A map of concepts of the wider effects of infodemics was developed and used as a structured aid to facilitate streamlined discussions during the conference (Figure 1). The map itself was organized into 4 sections, representing elements relating to the hypothetical influence of the information environment and their potential effects on individual, health, and societal impacts. Further details of the burden of the infodemic concept map can be found in Multimedia Appendix 2.

Figure 1. Concept map of the burden of infodemics, as discussed at the conference. It was organized across 4 domains: (1) in green, the level of the information ecosystem (online and offline content, social context, and structures that affect the dynamics of information consumption and transmission); (2) in blue, the individual level (behaviors and psychological mediators that determine exposure and susceptibility to the information characteristics of infodemics, as well as the proximal physical and psychological outcomes after this exposure); (3) in red, the level of health system impacts focused on metrics and outcomes specific to health care delivery and public health systems; and (4) in red, the societal level (infodemic impacts and ultimate outcomes that affect groups of individuals).



Concept mapping is a technique from the social and natural sciences to represent hypotheses about how elements affect one another [27,28]. These maps are meant to be preliminary frameworks—for example, concept maps typically start in a highly qualitative form, similar to mind mapping or causal mapping techniques. Although concept maps may eventually inform the basis of quantitative research, such as structural equation modeling, highly qualitative concept maps can be helpful for nascent problems to provide a system-level visualization of potential causal links, which, in turn, informs strategies for their investigation.

A brief review of the literature did not yield any comprehensive existing frameworks to discuss the whole complexity of the infodemic. Therefore, a concept map was developed to help participants from different backgrounds have a common frame and vocabulary for discussion.

The draft concept map was based on theoretical expectations, drawing from existing models from multiple disciplines, including anthropology, psychology, sociology, and informatics. The concept map sought to apply exposure or dose-relationship models from medicine and public health toward infodemic impacts and drew from socioecological models to consider interactions between individuals and broader societal factors. It sought to provide a system-level visualization representing hypotheses about how key factors may affect outcomes in an infodemic. A synthetic map was needed as the majority of research to date has focused only on limited facets of the system. For example, 1 study sought to estimate the total monetized cost of decisions not to receive a COVID-19 vaccination based on misinformation or disinformation [16]. Another study focused on the incremental health costs due to additional COVID-19 cases caused by misinformation, as well as the impact on the gross domestic product due to government restrictions needed to address the infection growth rate attributable to the impact of misinformation [29]. Directionality and potential causal links between different concepts on the map would be a point of discussion during the meeting.

The concept map was designed to help overcome challenges associated with bringing together such a diverse group of participants from diverse fields and areas of public health practice and policy making. As research on infodemiology remains emergent, significant variations in how infodemics and their impacts are conceptualized exist. Any research seeking to measure the predictors, mediators, and impacts of either health behaviors or human cognition is intrinsically complex. The interdisciplinary nature of infodemiological research draws interest from a wide variety of diverse disciplines ranging from the social sciences to health informatics. Moreover, experts working in infodemiology vary in professional settings ranging from public health action to academic research.

The concept map was prepared by expert members of the organizing team, was shared with participants ahead of the conference, and was referred to through all deliberations. Several map limitations were communicated to the participants ahead of time. First, the map was used as a discussion tool, and its primary purpose was not considered a formal model. Second, elements that were likely to be challenging to measure were included in the map to foster discussion. Third, the model was based on theoretical expectations and not a systematic review of the literature. Fourth, the model was not comprehensive and should not be used to inform intervention design or quantitative modeling.

Results

Key Areas for Action

The meeting was oriented to formulate practical actions that could be taken in the future in the context that in November 2021, the COVID-19 pandemic continued to cause massive disruptions and the first countries were beginning to roll out COVID-19 vaccines as fast as possible. There were major concerns that the basic inputs that would underpin the burden of infodemic measurement were not yet in place, such as a common language, concepts, and thorough evidence and literature reviews. This was difficult to achieve due to the cross-disciplinary nature of the challenge. Therefore, practical, immediate actions were prioritized to strengthen the foundation for measuring the burden of infodemics.

There were many rich discussions on concepts and frameworks, and participants worked together to reach recommendations that would work toward coherence across disciplines. Together, we identified 5 key areas for immediate action toward the development of metrics to assess the burden of infodemics and associated interventions over the 4 sessions. The richness and evolution of discussions could not be fully reflected in the summary of the action areas, but we reflect on them broadly here. The concrete actions are summarized next, and more details of each of the action areas are provided in Multimedia Appendix 3.

First, participants noted that currently, although often referred to, no established and widely accepted definition exists of what exactly characterizes infodemics and related aspects (eg, misinformation) and thus urged to establish the development of standardized definitions related to infodemic measurement and management. This could be achieved through the establishment of a working group aimed at developing working definitions, which could later be validated using a Delphi method. Participants assessed this task as a priority since the term “infodemic” was conceptually conflated, was often overworked, and was currently used to refer to different concepts in different fields or country settings. A glossary of terms associated with the measurement of infodemics—examples include “information exposure,” “overload,” “risk mediators of individual effects,” and “delayed care due to infodemic”—with standardized definitions was urgently needed to aid infodemiology research as well as public discourse.

Second, participants proposed the establishment of a multidisciplinary working group to review and build on the concept map to reflect and reconcile different perspectives and disciplines that look at the information ecosystem, the individual, the health system, and societal factors contributing to the infodemic. A Delphi method was recommended to be used to validate the concept map. Efforts to improve the concept map should be closely coordinated with the technical working groups responsible for developing standardized outcomes (area 1) and with the group conducting a desk review of the evidence, tools, and data sources (area 3). This is essential as the definition of the appropriate elements in the map will be in association with the terminology being developed. Similarly, evidence from the literature reviews will be vital to arriving at relevant connections/associations between the elements in the map. Participants assessed this task to be a priority and voted to retain the infodemic burden concept map. However, participants warned against following any concept map too closely, as it might lead to disregarding critical elements that were not already elaborated on the map. They agreed on its value in identifying the various inputs and outcomes, as well as the confounding factors that determine the contours of a complex object of scientific inquiry, such as an infodemic.

Third, participants proposed the establishment of a working group to draft a protocol for conducting a review of evidence, tools, and data sources related to infodemic measurement. The working group would also explore options and partnerships that could implement the review. Participants assessed this task to be a priority. Given the emerging contours of infodemiology, its scope would extend beyond that of a traditional review. While drawing on tools for systematic reviews of ongoing and upcoming research, it would, for instance, also involve searches within the gray literature.

Fourth, participants suggested the establishment of a working group to review and improve different policy, practice, and research priorities on a rolling basis and work toward the alignment of infodemic management efforts at the global level by different stakeholders. Additionally, this group would support mainstreaming of infodemic management into public health practice, policy, and capacity building. This core group would be complemented by a wider array of related groups, leveraging expertise in specific areas in a Delphi method to reach consensus on various items discussed in the group. Participants assessed this task to be a priority.

Fifth, participants identified 4 urgent aspects of COVID-19 infodemic management needing attention in the short term. Additionally, participants ranked them in order of priority and offered inputs on their potential modification and expansion: (1) development of harmonized tools for the measurement of information diet/exposure and establishment of a global research collaboration to use them; (2) development of behavioral/process models that can be used for the development and evaluation of interventions; (3) measurement of the economic cost of the COVID-19 infodemic and related spill-over effects; and (4) identification of data sources and measures following the concept map, which can be used for defining global open data sets to facilitate modelling and research.

Participants agreed that the pandemic response, health system recovery, and resilience building remain key priorities for most health authorities and continue as a research focus for academicians. In addition to the 5 key areas of action, several additional themes of conversations were identified during the discussions (see details in Multimedia Appendix 4), in general reflecting on the barriers and enablers to assessing and measuring the burden of infodemics.

Discussion

Principal Findings

The meeting was started with the aim to discuss and arrive at a concrete action plan, but the discussion proved to be so rich that it was important to reflect on cross-disciplinary considerations for the burden of infodemic metric development. Therefore, the concrete action areas reflect the wider context that needs to be considered when discussing measuring the burden of infodemics. Participants reflected on the inherent tension between discussing abstract concepts and research gaps compared to the need to develop practical actions to move toward better measurement of the burden of infodemics quickly enough to assist in the current global crisis. Several considerations recurred in the discussions, cutting across all meeting days, which should be kept in mind when discussing frameworks for measuring the burden of infodemics.

- To successfully respond to infodemics and integrate infodemic management into health systems and health policies, it is crucial to be able to measure the burden of infodemics on society. The conference discussions reaffirmed that there is an urgent need for infodemiology research to be fast-tracked and oriented in directions that are most effective for infodemic management in public health. Efforts to identify metrics for assessing the burden and evaluating interventions related to infodemics will benefit if they proceed in a parallel manner. Metrics that are feasible to measure and implement across a wide range of public health programmatic settings should continue to be prioritized. Standard indicators already used for measuring health, population, and economy should be given priority over the invention of new ones.
- The identification of sources and metrics from established and routine health and data systems should be rigorously prioritized over the formulation of new ones. Integrating insights from online and offline sources of information would be essential to an objective infodemic burden assessment.
- Despite the efforts focused on characterizing misinformation, little research in the area has been designed to measure population-level associations between (mis-)information exposure and attitudes, such as vaccine hesitancy, or behaviors, such as nonadherence to public health practices [30]. Research in data-driven infodemiology has mainly focused on identifying the types of misinformation that appear online and their prevalence, often limiting itself to a single social media platform [31]. With a few exceptions [32], research designs do not associate information exposure with individual outcomes (eg, attitudes, practices, or behaviors) and thus cannot be used to assess the burden of infodemics [33,34]. This results in the absence of solid evidence that could support effective design of public health interventions.
- However, the difficulties in harmonized measurement of the burden of infodemics should not pause the efforts in public health practice to introduce evidence-based interventions through rigorous implementation research and adaptable health programming. For example, lessons should be drawn from how policies to address the burden of noncommunicable diseases on populations evolved over time. Measurements, such as monitoring of blood sugar levels, became standard practice and indicators before science was able to unequivocally link them to health outcomes and the burden of disease.
- Although the WHO Member States have recognized the perils of health misinformation [35], WHO, Member States, civil society, and other stakeholders have different roles to play in infodemic management and response. To be effective, management and response activities need to understand where the greatest risks are and rapidly capture the positive impact of responses without having to develop new, robust evaluation programs for every activity. Observational studies that simply report on the prevalence of misinformation make recommendations based on biased data and without measuring associations with behavior. For example, it was assumed that bots were important for disseminating misinformation, but research could not prove the real impact on the attitudes of social media users [36]. Studies that do not directly link information exposure to behavior can lead to wasted effort and unintended consequences. Understanding the mediating role that the social determinants of health play in individuals' susceptibility to misinformation should be investigated.
- An infodemic causes harm on many levels, and it is by its nature a complex problem. Assessing its burden on health and society will require rethinking not only the frameworks, pathways, and protocols for measurement but also how the data are collected in a sustainable manner. WHO is developing activities to support pandemic preparedness and to mitigate the current pandemic, and several WHO preparedness activities rely on the development of new technologies and tools. The deployment of standardized tools for measuring how population-level differences in exposure to information risk factors explain the differences in behaviors after accounting for demographic differences is a challenge. New forms of global collaborations are needed to collect harmonized data through distributed collective measurement of the burden of infodemics. Moreover, research and data collection should consider using participatory research methods with communities and infodemic managers where the generation of metrics is paired up with interventions.
- Infodemics can be best addressed using a multidisciplinary approach and grounding in public health practice [17]. The currently emergent stage of the science of infodemiology, combined with the heterogeneity of academic expertise and

professional backgrounds of the participants at the conference, offered rich opportunities for multifaceted technical discussions on metrics related to infodemics. At the same time, the meeting reconfirmed that conversations across diverse backgrounds must be prepared carefully to facilitate discussion across different scientific terminologies and approaches, as well as the differences between research methods and public health practice considerations.

- The lack of trust or mistrust toward health authorities can compromise adherence, compliance, and, ultimately, the overall success of the public health response, with all that these imply in terms of adverse outcomes on individual and population-wide levels. Identifying public health and social indicators for measuring and monitoring the impact of infodemics on health behaviors is now a priority for many health authorities that require evidence for planning, implementing, and evaluating interventions and policies. Trust metrics should be incorporated into infodemic metrics and modeling because these concepts are so interlinked.
- Currently, there are few published studies that reflect how policies foster or hinder infodemic-related outcomes; without measures that are identified that can be acted on by health systems, it will be difficult to institute more supportive and effective policies to mitigate the effects of infodemics on health.
- The way information access, exposure, and engagement are estimated for individuals is inconsistent across studies and often restricted to single social media platforms, limiting the value of the research. Furthermore, it remains unclear whether data from social media and web platforms can be used as proxy measures for a person's broader information diet and whether these data capture differences in how people make sense of that information in terms of attention, trust, and prior beliefs. Ultimately, understanding how a person's interaction with an increasingly individually attenuated and complex information ecosystem affects their health behaviors should be better studied to understand linkages to their online interactions.

However, there are ways forward to advance the measurement of infodemic harms and impacts and the use of infodemic management interventions. The 5 conclusions and 5 key actions from the conference represent the convergence of many of the limitations and opportunities mentioned before for the field and propose a roadmap for advancing the field for WHO.

Other Policy Developments That Will Affect the Measurement of the Burden of Infodemics

Previously, many efforts in research and coordination in the misinformation space have focused on individual, societal, or media-related domains in a siloed manner. Now is the time to firmly center the health system in the infodemic management conversation when it comes to health emergencies. Strengthening preparedness, prevention, and resilience aspects to health systems in infodemic management will mean moving from defining terms and metrics to routinizing infodemic measures in routine data collection and decision-making in "peacetime" preparedness work and ramp up engagement, grounded in policy and enabled by sufficient workforce capacity and resources, during emergency activations of incident management structures.

In an attempt to reduce siloed approaches, multidisciplinary research and partnerships between public health, academic, media and civil society institutions should be fostered to identify interconnections which could provide basis for such assessments. Convenings similarly patterned on HCD principles may be well-suited to further discussing and establishing frameworks for interdisciplinary areas of health that are identified as priorities following emergencies and outbreaks, even if science and policy surrounding the topic is only emergent. This could include focus on burgeoning areas of governance, privacy and ethics in infodemic management and even in other health areas affected by infodemic harms. For example, WHO is convening a WHO ethics panel to deliberate on ethical considerations of social listening and infodemic management.

Ultimately, a successful infodemic response will lead to informed policies and promote healthy behaviors by individuals and communities. To do this, it identifies and addresses individuals' and communities' questions, concerns and information voids on health topics; reduces the spread and impact of misinformation; and refines public health engagement strategies (ie, promoting health equity, addressing scientific uncertainty and promoting culturally relevant risk communication and education) and health system response to more effectively promote healthy behaviors. To support countries, WHO has fostered development of tools to provide an evidence-based response to the infodemic and strengthen epidemic and pandemic preparedness activities [37,38]. These complement efforts by governments, media and factchecking organizations, civil society organizations and academic groups to develop valuable tools and resources to develop stronger methods for evidence-based decision-making for infodemic management. As the COVID-19 response has shown, all emergencies and pandemics in the future will be accompanied by infodemics that will be better addressed with the tools and insights developed today.

Health authorities seeking instructive policies or global technical guidance on infodemic management as the global epidemiological picture changes. WHO is working to establish a technical working group to support development of technical guidance that will be relevant to different country contexts, emergencies and outbreaks. A policy brief for COVID-19 infodemic management has also been published, outlining key recommendations for policy makers to integrate infodemic management in COVID-19 response and strengthen preparedness for other emergencies [39].

Countries are seeking solutions—interventions to stem current and future infodemics. Since the conference, WHO has commissioned an evidence gap map (EGM) exercise to analyze and visually map areas where there is evidence, the strength and applicability of that evidence of infodemic management interventions in the time of COVID-19 to the wider field, and where there are evidence gaps [40]. In conjunction with the conference outcomes and priorities identified by participating experts, this EGM can aid in prioritizing where investments in research and interventions should be directed.

Conclusion

Infodemics now constitute a condition of our times and are here to stay, even it is extremely difficult to measure them precisely. To advocate evidence-based interventions for use in preparedness, prevention, and emergency response, a thorough assessment of infodemics' impact and burden on society is required. This, however, requires to first reach consensus about what we exactly mean when we talk about infodemics and also about their moderating determinants. When definitions are set, formulating an adequate methodology—relevant in various health care settings and contexts—can be pursued that helps measure and eventually express the damaging effects of infodemics by using standard indicators. This conference was the first global step toward achieving these objectives.

We are standing on the shoulders of giants as diverse knowledge can be transferred from other disciplines and contexts into infodemic management for emergencies. Yet, we need further research and innovation to address some of the longstanding questions and bring about a truly multidisciplinary effort that serves both academic research and public health emergency preparedness, prevention, and response.

Acknowledgments

The authors wish to acknowledge and thank all the participants for their time, work, and contributions to the event, which has greatly helped expand the knowledge base. The authors are grateful to the organizing and extended organizing team's contributions for the smooth functioning of the main event and supportive activities; these include country, regional, and international levels of the World Health Organization (WHO), including the WHO Information Network for Epidemics (EPI-WIN) team, a range of external partners, and the Demand for Immunization Team of the US Centers for Disease Control and Prevention (US CDC) Global Immunization Division. Special thanks are also due to the conference support team: Romana Rauf, Petros Ginkyo, Zerthun Alemu Belay, Sam Bradd, André Buell, and Isabelle Jacobs.

Data Availability

Raw notes from the conference proceedings are stored offline and are available upon request. Cleaned up versions of notes that were also shared with participants are available in the appendices.

Conflicts of Interest

SB, SBr, NG, SH, AI, MK, RL, TN, TDP, and BY are staff of the World Health Organization (WHO); CBF and BKW are consultants with WHO; SK and CV are staff of the US Centers for Disease Control and Prevention (US CDC). These authors alone are responsible for the views expressed in this paper, and they do not represent the views of their organizations. The conflicts of interest were reviewed and managed as per WHO procedures. AD declared that his university received research support on information diet measurement by WHO for the product owned by WHO. He was not part of the consensus driving during the closing session of the meeting. TM is the current Editor-in-Chief of JMIR Infodemiology and declared ownership interest in a company for work not related to the deliberation in this publication. LMB works for Immunize Canada/the Canadian Public Health Association, which has received educational grants/funding from Merck Canada, Pfizer Canada, Pfizer Global, Moderna Canada, Seqirus, Sanofi Canada, GSK Canada, and the Public Health Agency of Canada (PHAC). These funds are not related to the paper. CW was executive director of the nonprofit organization First Draft, which received funds for research and advocacy work from Google, and research project support on the effectiveness of SMS-based social inoculation from WHO. She chaired the first 3 days of the conference but was not part of the consensus driving during the closing session of the meeting. EP declared receiving conference stipends, training fees, and publication fees from the Medical Research Council. He was not part of the consensus driving during the closing session of the meeting. IB is director of the WHO Collaborating Center on information systems for health, which supports WHO with broader digital health analytics and policy analysis. The center has supported the Pan American Health Organization (PAHO)/WHO with infodemic analytics during COVID-19. SMR declared receiving consultancy fees from the EURO Health Group research consortium and is currently a consultant in infodemic management for WHO. JR and SVDL declared that their university received research funding from NATO Strategic Communications Centre of Excellence, Google Jigsaw, WhatsApp, British Academy, the Economic and Social Research Council (ESRC), the UK Cabinet Office, and EU Horizon 2020. They were not part of the consensus driving during the closing session of the meeting. AG declared that his university received research funds from the Canadian Institutes of Health Research (CIHR). AS declared receiving consultancy fees from Euro Health Group A/S – Denmark for services unrelated to the topic of the conference. PB is founder and CEO of HealthEnabled, which received past funding from Gavi, the Vaccine Alliance, to conduct digital social listening. JN declared employment with Harvard University, working in the field of medical misinformation. MG and MEB declared no conflicts of interest for this paper.

Multimedia Appendix 1
Conference meeting: methodology and structure.

[DOCX File , 16 KB - infodemiology_v3i1e44207_app1.docx]

Multimedia Appendix 2
Concept map.
[DOCX File , 72 KB - infodemiology_v3i1e44207_app2.docx]

Multimedia Appendix 3
Discussions about action areas.
[DOCX File , 24 KB - infodemiology_v3i1e44207_app3.docx]

Multimedia Appendix 4
Additional themes of conversations identified during the discussions.
[DOCX File , 17 KB - infodemiology_v3i1e44207_app4.docx]

References

1. 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]
2. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
3. Gruzd A, De Domenico M, Sacco PL, Briand S. Studying the COVID-19 infodemic at scale. Big Data Soc 2021 Jun 10;8(1):205395172110211. [doi: 10.1177/20539517211021115]
4. Lutz C. Digital inequalities in the age of artificial intelligence and big data. Hum Behav Emerg Tech 2019 Apr 26;1(2):141-148. [doi: 10.1002/hbe2.140]
5. Ries M. The COVID-19 infodemic: mechanism, impact, and counter-measures—a review of reviews. Sustainability 2022 Feb 23;14(5):2605. [doi: 10.3390/su14052605]
6. Global trends to 2030: identities and biases in the digital age. European Strategy and Policy Analysis System. 2018. URL: https://espas.secure.europarl.europa.eu/orbis/sites/default/files/generated/document/en/Ideas%20Paper%20Digital%20Identities%20ESPAS-EPSC_V08.pdf [accessed 2023-02-07]
7. Adini B, Singer S, Ringel R, Dickmann P. Earlier detection of public health risks - health policy lessons for better compliance with the International Health Regulations (IHR 2005): Insights from low-, mid- and high-income countries. Health Policy 2019 Oct;123(10):941-946 [FREE Full text] [doi: 10.1016/j.healthpol.2019.06.007] [Medline: 31288952]
8. Clark‐Ginsberg A, Petrun Sayers E. Communication missteps during COVID‐19 hurt those already most at risk. J Contingencies Crisis Manag 2020 Jul 19;28(4):482-484 [FREE Full text] [doi: 10.1111/1468-5973.12304]
9. Limaye RJ, Holroyd TA, Blunt M, Jamison AF, Sauer M, Weeks R, et al. Social media strategies to affect vaccine acceptance: a systematic literature review. Expert Rev Vaccines 2021 Aug 30;20(8):959-973. [doi: 10.1080/14760584.2021.1949292] [Medline: 34192985]
10. Betsch C, Wieler LH, Habersaat K. Monitoring behavioural insights related to COVID-19. Lancet 2020 Apr;395(10232):1255-1256. [doi: 10.1016/s0140-6736(20)30729-7]
11. El Benny M, Kabakian-Khasholian T, El-Jardali F, Bardus M. Application of the eHealth literacy model in digital health interventions: scoping review. J Med Internet Res 2021 Jun 03;23(6):e23473 [FREE Full text] [doi: 10.2196/23473] [Medline: 34081023]
12. Yang W, Zhang J, Ma R. The prediction of infectious diseases: a bibliometric analysis. Int J Environ Res Public Health 2020 Aug 27;17(17):6218 [FREE Full text] [doi: 10.3390/ijerph17176218] [Medline: 32867133]
13. Balakrishnan V, Ng W, Soo M, Han G, Lee C. Infodemic and fake news - a comprehensive overview of its global magnitude during the COVID-19 pandemic in 2021: A scoping review. Int J Disaster Risk Reduct 2022 Aug;78:103144 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.103144] [Medline: 35791376]
14. Bunker D. Who do you trust? The digital destruction of shared situational awareness and the COVID-19 infodemic. Int J Inf Manag 2020 Dec;55:102201 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102201] [Medline: 32836649]
15. Stewart R, Madonsela A, Tshabalala N, Etale L, Theunissen N. The importance of social media users' responses in tackling digital COVID-19 misinformation in Africa. Digit Health 2022 Mar 18;8:20552076221085070 [FREE Full text] [doi: 10.1177/20552076221085070] [Medline: 35321021]
16. Bruns R, Hosangadi D, Trotochaud M, Kirk ST. COVID-19 vaccine misinformation and disinformation costs an estimated $50 to $300 million each day. The Johns Hopkins Center for Health Security. 2021. URL: https://tinyurl.com/2p8azs7b [accessed 2023-02-07]
17. Calleja N, AbdAllah A, Abad N, Ahmed N, Albarracin D, Altieri E, et al. A public health research agenda for managing infodemics: methods and results of the First WHO Infodemiology Conference. JMIR Infodemiology 2021 Sep 15;1(1):e30979 [FREE Full text] [doi: 10.2196/30979] [Medline: 34604708]

18. WHO public health research agenda for managing infodemics. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240019508 [accessed 2023-02-07]
19. WHO third global infodemic management conference: whole-of-society challenges and approaches to respond to infodemics. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240034501 [accessed 2023-02-07]
20. 4th Virtual WHO Infodemic Management Conference: advances in social listening for public health. World Health Organization. 2021. URL: https://www.who.int/news-room/events/detail/2021/05/04/default-calendar/4th-virtual-who-infodemic-management-conference-advances-in-social-listening-for-public-health [accessed 2023-02-07]
21. 5th Virtual WHO Infodemic Management Conference meeting report: steps towards measuring the burden of infodemics. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240047174 [accessed 2023-02-07]
22. Grass R. Tools for transformation: the fabulous POP model. Social Transformation Project. 2013. URL: http://stproject.org/wp-content/uploads/2014/11/the-fabulous-pop-model.pdf [accessed 2023-02-07]
23. Roumani N. Design engaging in-person and remote meetings. Stanford Design School. 2020. URL: https://dschool.stanford.edu/resources/design-engaging-in-person-and-remote-meetings [accessed 2023-02-07]
24. Chatham House rule. Chatham House. URL: https://www.chathamhouse.org/about-us/chatham-house-rule [accessed 2022-05-07]
25. Rubinelli S, Purnat TD, Wilhelm E, Traicoff D, Namageyo-Funa A, Thomson A, et al. WHO competency framework for health authorities and institutions to manage infodemics: its development and features. Hum Resour Health 2022 May 07;20(1):35 [FREE Full text] [doi: 10.1186/s12960-022-00733-0] [Medline: 35525924]
26. Conflict of interest. World Health Organization. 2021. URL: https://www.who.int/about/ethics/declarations-of-interest [accessed 2023-02-07]
27. Trochim WM, McLinden D. Introduction to a special issue on concept mapping. Eval Program Plann 2017 Feb;60:166-175. [doi: 10.1016/j.evalprogplan.2016.10.006] [Medline: 27780609]
28. van Bon-Martens MJ, van de Goor IA, van Oers HA. Concept mapping as a method to enhance evidence-based public health. Evaluat Program Plan 2017:213-228 [FREE Full text] [doi: 10.1016/j.evalprogplan.2016.08.014]
29. The cost of lies: assessing the human and financial impact of COVID-19 related online misinformation on the UK – January 2021. London Economics. 2021. URL: https://londoneconomics.co.uk/blog/publication/the-cost-of-lies-assessing-the-human-and-financial-impact-of-covid-19-related-online-misinformation-on-the-uk-january-2021/ [accessed 2023-02-07]
30. Pierri F, Perry BL, DeVerna MR, Yang K, 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 [FREE Full text] [doi: 10.1038/s41598-022-10070-w] [Medline: 35474313]
31. Muhammed TS, Mathew SK. The disaster of misinformation: a review of research in social media. Int J Data Sci Anal 2022 Feb 15;13(4):271-285 [FREE Full text] [doi: 10.1007/s41060-022-00311-6] [Medline: 35194559]
32. 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 05;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
33. Wegwarth O, Kurzenhäuser-Carstens S, Gigerenzer G. Overcoming the knowledge-behavior gap: the effect of evidence-based HPV vaccination leaflets on understanding, intention, and actual vaccination decision. Vaccine 2014 Mar 10;32(12):1388-1393 [FREE Full text] [doi: 10.1016/j.vaccine.2013.12.038] [Medline: 24486360]
34. Hoffmann TC, Del Mar C. Clinicians' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA Intern Med 2017 Mar 01;177(3):407-419. [doi: 10.1001/jamainternmed.2016.8254] [Medline: 28097303]
35. Seventy-third World Health Assembly: COVID-19 response. World Health Organization. 2020. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA73/A73_R1-en.pdf [accessed 2023-02-07]
36. Dunn AG, Surian D, Dalmazzo J, Rezazadegan D, Steffens M, Dyda A, et al. Limited role of bots in spreading vaccine-critical information among active Twitter users in the United States: 2017–2019. Am J Public Health 2020 Oct;110(S3):S319-S325. [doi: 10.2105/ajph.2020.305902]
37. COVID-19 research and innovation. Powering the world's pandemic response – now and in the future. World Health Organization. 2022. URL: https://www.who.int/publications/m/item/covid-19-research-and-innovation---powering-the-world-s-pandemic-response-now-and-in-the-future [accessed 2023-02-07]
38. How global research can end this pandemic and tackle future ones. World Health Organization. 2022. URL: https://www.who.int/publications/m/item/how-global-research-can-end-this-pandemic-and-tackle-future-ones [accessed 2023-02-07]
39. WHO policy brief: COVID-19 infodemic management, 14 September 2022. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Policy_Brief-Infodemic-2022.1 [accessed 2023-02-07]
40. Purnat T, John O, Pundir P, Ishizumi A, Murthy S, Rajwar E. Protocol: evidence and gap map of interventions for infodemic management during health emergencies: a case of COVID-19 pandemic. Open Science Framework. 2022. URL: https://osf.io/7r3a8/?view_only=f749274dcae74c43a754aa51f3df93c0 [accessed 2023-02-07]

Abbreviations

HCD: human-centered design
US CDC: US Centers for Disease Control and Prevention
WHO: World Health Organization

Edited by A Mavragani; submitted 10.11.22; peer-reviewed by M Cinelli, AL Lohiniva; comments to author 01.12.22; revised version received 09.01.23; accepted 26.01.23; published 20.02.23.

Please cite as:
Wilhelm E, Ballalai I, Belanger ME, Benjamin P, Bertrand-Ferrandis C, Bezbaruah S, Briand S, Brooks I, Bruns R, Bucci LM, Calleja N, Chiou H, Devaria A, Dini L, D'Souza H, Dunn AG, Eichstaedt JC, Evers SMAA, Gobat N, Gissler M, Gonzales IC, Gruzd A, Hess S, Ishizumi A, John O, Joshi A, Kaluza B, Khamis N, Kosinska M, Kulkarni S, Lingri D, Ludolph R, Mackey T, Mandić-Rajčević S, Menczer F, Mudaliar V, Murthy S, Nazakat S, Nguyen T, Nilsen J, Pallari E, Pasternak Taschner N, Petelos E, Prinstein MJ, Roozenbeek J, Schneider A, Srinivasan V, Stevanović A, Strahwald B, Syed Abdul S, Varaidzo Machiri S, van der Linden S, Voegeli C, Wardle C, Wegwarth O, White BK, Willie E, Yau B, Purnat TD
Measuring the Burden of Infodemics: Summary of the Methods and Results of the Fifth WHO Infodemic Management Conference
JMIR Infodemiology 2023;3:e44207
URL: https://infodemiology.jmir.org/2023/1/e44207
doi:10.2196/44207
PMID:37012998

Original Paper

Potential Impact of the COVID-19 Pandemic on Public Perception of Water Pipes on Reddit: Observational Study

Zihe Zheng[1*], BSc; Zidian Xie[2*], MSc, PhD; Maciej Goniewicz[3], PharmD, PhD; Irfan Rahman[4], PhD; Dongmei Li[2], MSc, PhD

[1]Goergen Institute for Data Science, University of Rochester, Rochester, NY, United States

[2]Department of Clinical and Translational Research, University of Rochester Medical Center, Rochester, NY, United States

[3]Department of Health Behavior, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States

[4]Department of Environmental Medicine, University of Rochester Medical Center, Rochester, NY, United States

[*]these authors contributed equally

Corresponding Author:
Dongmei Li, MSc, PhD
Department of Clinical and Translational Research
University of Rochester Medical Center
265 Crittenden Boulevard Cu 420708
Rochester, NY, 14642-0001
United States
Phone: 1 5852767285
Email: Dongmei_Li@urmc.rochester.edu

Abstract

Background: Socializing is one of the main motivations for water pipe smoking. Restrictions on social gatherings during the COVID-19 pandemic might have influenced water pipe smokers' behaviors. As one of the most popular social media platforms, Reddit has been used to study public opinions and user experiences.

Objective: In this study, we aimed to examine the influence of the COVID-19 pandemic on public perception and discussion of water pipe tobacco smoking using Reddit data.

Methods: We collected Reddit posts between December 1, 2018, and June 30, 2021, from a Reddit archive (PushShift) using keywords such as "waterpipe," "hookah," and "shisha." We examined the temporal trend in Reddit posts mentioning water pipes and different locations (such as homes and lounges or bars). The temporal trend was further tested using interrupted time series analysis. Sentiment analysis was performed to study the change in sentiment of water pipe–related posts before and during the pandemic. Topic modeling using latent Dirichlet allocation (LDA) was used to examine major topics discussed in water pipe–related posts before and during the pandemic.

Results: A total of 45,765 nonpromotion water pipe–related Reddit posts were collected and used for data analysis. We found that the weekly number of Reddit posts mentioning water pipes significantly increased at the beginning of the COVID-19 pandemic ($P<.001$), and gradually decreased afterward ($P<.001$). In contrast, Reddit posts mentioning water pipes and lounges or bars showed an opposite trend. Compared to the period before the COVID-19 pandemic, the average number of Reddit posts mentioning lounges or bars was lower at the beginning of the pandemic but gradually increased afterward, while the average number of Reddit posts mentioning the word "home" remained similar during the COVID-19 pandemic ($P=.29$). While water pipe–related posts with a positive sentiment were dominant (12,526/21,182, 59.14% before the pandemic; 14,686/24,583, 59.74% after the pandemic), there was no change in the proportion of water pipe–related posts with different sentiments before and during the pandemic ($P=.19$, $P=.26$, and $P=.65$ for positive, negative, and neutral posts, respectively). Most topics related to water pipes on Reddit were similar before and during the pandemic. There were more discussions about the opening and closing of hookah lounges or bars during the pandemic.

Conclusions: This study provides a first evaluation of the possible impact of the COVID-19 pandemic on public perceptions of and discussions about water pipes on Reddit.

(JMIR Infodemiology 2023;3:e40913) doi:10.2196/40913

KEYWORDS

water pipes; Reddit; COVID-19; COVID-19 pandemic; public perception

Introduction

Water pipe tobacco, also known as "hookah," is a combustible tobacco product usually used in a group setting [1]. A previous systematic review found that the major motivations for water pipe smoking are socialization, relaxation, pleasure, and entertainment [2]. Water pipe tobacco smoking (WTS) often involves the use of an apparatus that heats the tobacco and passes the smoke through water before it can be inhaled through a hose by the user. Data from wave 1 to 3 of the Population Assessment of Tobacco and Health (PATH) study showed that young adults (aged 18-24 years) had a higher prevalence of WTS than youth and adults aged 25 years or older [1]. The prevalence of past-30-day WTS in young adults was 9.2%, compared to 0.7% in youth and 1.2% in adults aged 25 years or older, according to PATH wave 3 data collected between 2015 and 2016 [1]. A previous online survey study of US adults aged 18 to 30 years found that positive attitudes and perceived peer acceptability of WTS were significantly associated with WTS in young adults [3].

Similar to many other tobacco products, WTS is related to many health issues, including lung cancer, respiratory illness, low birth weight, and periodontal disease, as well as bronchitis, metabolic syndrome, cardiovascular disease, and mental health [4,5]. Many water pipe users prefer flavored tobacco [6]. The most popular flavors are fruit flavors, followed by sweets, spice, alcohol, and other beverages [6]. It has been reported that the main motivations for WTS are socialization, relaxation, pleasure, and entertainment [2]. Homes and hookah lounges are the most common places where people use water pipe tobacco [7].

It has been shown that smoking behaviors were affected by the recent COVID-19 pandemic [8]. COVID-19 is an infectious disease caused by SARS-CoV-2 [9]. The first case of COVID-19 was diagnosed in December 2019 in Wuhan, China, and it was declared a global pandemic on March 11, 2020, by the World Health Organization (WHO) [10]. To mitigate the spread of the disease, many countries, including the United States, India, and China, enforced lockdowns on the economy and cities [11,12]. Studies have shown many positive outcomes of lockdowns, such as a significant decrease in the growth rate of confirmed cases, as well as improved global air quality and lower pollution [11,13]. On the other hand, the lockdowns caused mental health problems, such as anxiety, depression, loneliness, sleep difficulties, and hyperarousal, as well as a higher tendency to overeat and experience obesity [14-16]. Some studies showed that psychiatric emergency admissions increased during the lockdowns, raising a debate on how the pandemic might have affected mental health [17,18]. The COVID-19 pandemic and the accompanying lockdown policies have been proven to have influenced tobacco smoking [19,20]. Smoking prevalence was shown to have decreased in urban counties in the United States [20]. In addition, the vaping rate among youth and young adults declined during the pandemic in the United States [19]. A recent online survey study conducted among 1223 US adults in 2020 showed that a more severe perception of smoking-related COVID-19 risks was associated with a higher likelihood of smoking reduction and quit attempts [21]. A cross-sectional household survey study conducted in England in 2020 showed that a minority of e-cigarette users attempted to quit vaping because of COVID-19 [22]. Another survey study of Israeli smokers showed increases in both the number of smokers and attempts to quit smoking due to COVID-19 [23]. Given water pipes are often smoked in groups during social gatherings, it is possible that the lockdowns during the COVID-19 pandemic affected water pipe use.

As of January 2021, Reddit had more than 50 million daily active users, 100,000 active communities, and 13 billion posts and comments [24]. Reddit has been widely used for discussing many public health events [25]. Due to its increasing popularity, Reddit has been used to study public perceptions and discussions of tobacco products. For example, to understand reasons why people with mental health problems smoke e-cigarettes, a group of scientists analyzed 3263 posts on Reddit and found that the main reasons for e-cigarette use included self-medication, freedom and control, and motivation from caregivers and online communities [26]. Several studies have used Reddit data to study public perceptions of flavored e-cigarette use and related health symptoms [27-29]. Another study analyzed Reddit posts related to oral nicotine pouches and found that people generally had a positive attitude toward oral nicotine pouches [30]. Since the pandemic started, active discussions on Reddit have made it a useful resource to study the influence of the COVID-19 pandemic. Recently, Reddit has been used to find patterns of posts that imply mental health problems and identify at-risk users on the platform during the COVID-19 pandemic [31].

In this study, we aimed to understand how the COVID-19 pandemic influenced public perceptions and discussions of water pipe tobacco smoking on Reddit through interrupted time series (ITS) data analyses, sentiment analyses, and topic modeling. More importantly, we aimed to investigate whether the COVID-19 pandemic had an impact on water pipe smoking behaviors, such as reducing use in hookah lounges or bars during the pandemic, given water pipes are commonly smoked during social events. Our study provides an initial but important evaluation of the potential impact of the COVID-19 pandemic on water pipe perception and discussion, as well as potential water pipe behavior changes, through social media data mining.

Methods

Data Collection and Preprocessing

Reddit posts (comments) from December 1, 2018, to June 30, 2021, were downloaded from a Reddit archive (PushShift). We extracted posts related to water pipes using a set of keywords from a previous study [32], including *water pipe*, *hookah*, *shisha*, *narghile*, *argileh*, *hubble-bubble*, *goza*, *borry*, *qaylan*, *mada'a*, *mouassal*, *jurak*, *tumbak*, *hooka*, *sheesha*, and *hubblebubble*. In total, 62,699 water pipe–related Reddit posts were obtained.

A multi-filter process was used to preprocess the Reddit data. First, we applied a filter to obtain all the posts written in English. Second, we applied additional filters to ensure that all posts were related to water pipes. For example, Reddit posts that contained "shisha octane," "waterpipe shotgun," and "hookah attack" were discussing games like *Rust* or *Rainbow Six Siege* instead of water pipe smoking. To eliminate such noise, we removed Reddit posts that contained the above combinations of keywords and Reddit posts from game subreddits, including /ps4, /xbox, /playrust, /rainbow6, /boombeach, /siegeacademy, /r6proleague, and /valorant. Third, commercial Reddit posts were removed if they contained keywords such as *discount*, *deal*, and *dealer*, or if their usernames included keywords such as *dealer*, *water pipe*, or *hookah*.

Location Analysis

To determine whether the COVID-19 pandemic affected water pipe smoking behavior due to restrictions on social gatherings, we conducted a location analysis. We constructed a location data set for Reddit posts that mentioned specific locations according to the most common locations of water pipe smoking mentioned in the posts. First, we performed an item count that included single words, bigrams, and trigrams. Then, we manually examined the items with high frequency to identify location-related items. We classified these items into 2 main categories: *home* and *lounge/bar*. The *home* category included *home*, *house*, *living room*, and *dining room* while the *lounge/bar* category included *lounge*, *lounges*, *bar*, *bars*, *cafe*, *cafes*, *coffee shop*, *coffee shops*, *strip club*, and *strip clubs*. The location data set consisted of 9344 Reddit posts mentioning either *home* or *lounge/bar*.

Temporal and ITS Analysis

To study trends in the discussion of water pipe tobacco on Reddit, we calculated the number of water pipe–related Reddit posts per week, as well as the number of Reddit posts that mentioned either *home* or *lounge/bar* per week. An ITS analysis was used to determine if trends before the COVID-19 pandemic were different from trends during the COVID-19 pandemic, for either *home* or *lounge/bar*. In an ITS analysis, time series are segmented by the intervention point and segmented regression is used to evaluate the changes in level and slope before and after the intervention point [33]. In our study, we set the intervention point as March 11, 2020, which is the day that the WHO declared COVID-19 to be a global pandemic [10]. We identified 155 posts that were comments on a single popular post made on February 19, 2020, about a husband smoking water pipe tobacco while his wife was pregnant. After carefully examining these posts, we manually removed all of them, since they were not related to WTS. The ITS analyses were conducted using SAS (version 9.4; SAS Institute). The significance level of the test was set at 5% for 2-sided tests.

Sentiment Analysis

We generated a sentiment score for each post using VADER (Valence Aware Dictionary and Sentiment Reasoner). VADER is designed for qualitative sentiment analysis of social media using a list of lexical features combined with rules about conventions for expressing emotions [34]. A post is considered to have a positive attitude if it has a score equal to or higher than 0.05, a negative attitude if the score is equal to or lower than –0.05, and a neutral attitude if the score is between –0.05 and 0.05. To determine if there was any change in the proportion of posts with different sentiments before and during the COVID-19 pandemic, we performed a 2-proportion *z* test with a significance level of .05.

Topic Modeling

To identify and compare the topics discussing water pipes on Reddit, we performed topic modeling using the latent Dirichlet allocation (LDA) model on posts made before and during the COVID-19 pandemic. The LDA model is a generative statistical model that can be used to find topics in documents [35]. The algorithm first calculates the probabilities of each word appearing in each topic and then defines each topic with the words that have the highest possibility of appearing in that topic. We chose the optimal number of topics based on the maximum coherence score. Sentences from the posts were transformed to lowercase letters, and stop words, such as *the*, *am*, and *you*, were removed using the Natural Language Toolkit (NLTK Team) in Python. Additionally, the words were lemmatized using *spacy* (ExplosionAI GmbH) [36] in Python.

Ethics Approval

Only publicly available Reddit posts were used for this study. There was no identifying information on Reddit users in this study. To protect human subjects included in this study, this study was reviewed and approved by the Research Subjects Review Board of the Office for Human Subject Protection at the University of Rochester (STUDY00006570).

Results

Discussion of Water Pipes on Reddit

From 62,699 Reddit posts extracted from the Reddit archive based on water pipe–related keywords, we identified 56,462 English Reddit posts, among which 51,387 were related to water pipes. Further removal of promotion posts resulted in 45,765 Reddit posts related to water pipes, which were used for further analysis. To understand trends in the discussion of water pipes on Reddit over time, we examined the number of posts related to water pipe tobacco per week from December 1, 2018, to June 30, 2021 (Figure 1). The vertical line marks March 11, 2020, the starting date of the COVID-19 pandemic. As shown in Figure 1, ITS analysis showed that the discussion of water pipe tobacco on Reddit was significantly increasing before the COVID-19 pandemic (P<.001). After the announcement of the COVID-19 pandemic, the popularity of water pipe–related Reddit posts significantly decreased (P<.001). ITS analysis further showed that the average number of water pipe–related posts per day during the pandemic was significantly higher than before the pandemic (P<.001).

To examine whether the pandemic had any impact on the location of water pipe tobacco use, we first identified posts mentioning the 2 most common locations for water pipe use, homes and lounges or bars (Figure 2). In total, we identified 2194 posts mentioning *home/house*, and 7150 posts mentioning *lounge/bar*. Further ITS analysis showed that discussion of smoking water pipe tobacco at home remained stable over the period of study (P=.29). In contrast, discussion about smoking water pipes at lounges or bars significantly decreased in the time leading up to the pandemic (P=.004), then significantly increased after the announcement of the pandemic (P<.001). Compared to before the pandemic, the average number of posts mentioning smoking water pipes at lounges or bars was significantly lower during the pandemic (P<.001). In addition, we identified a peak in November 2020 that resulted from 70 posts related to a news story about closing hookah lounges in Saskatoon, Saskatchewan.

Figure 1. Longitudinal trend in Reddit posts related to water pipes.

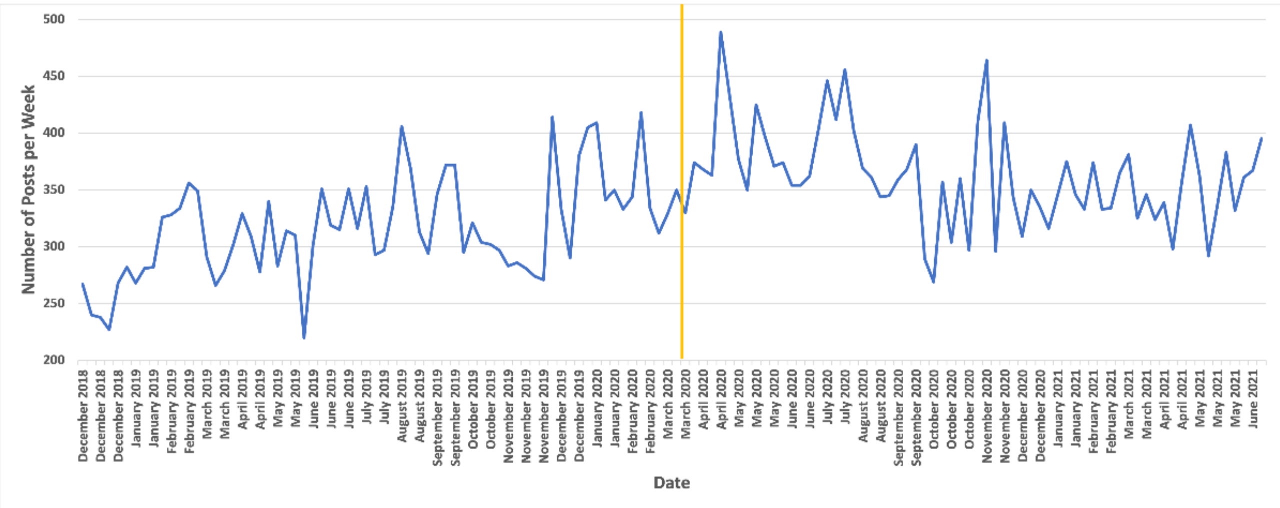
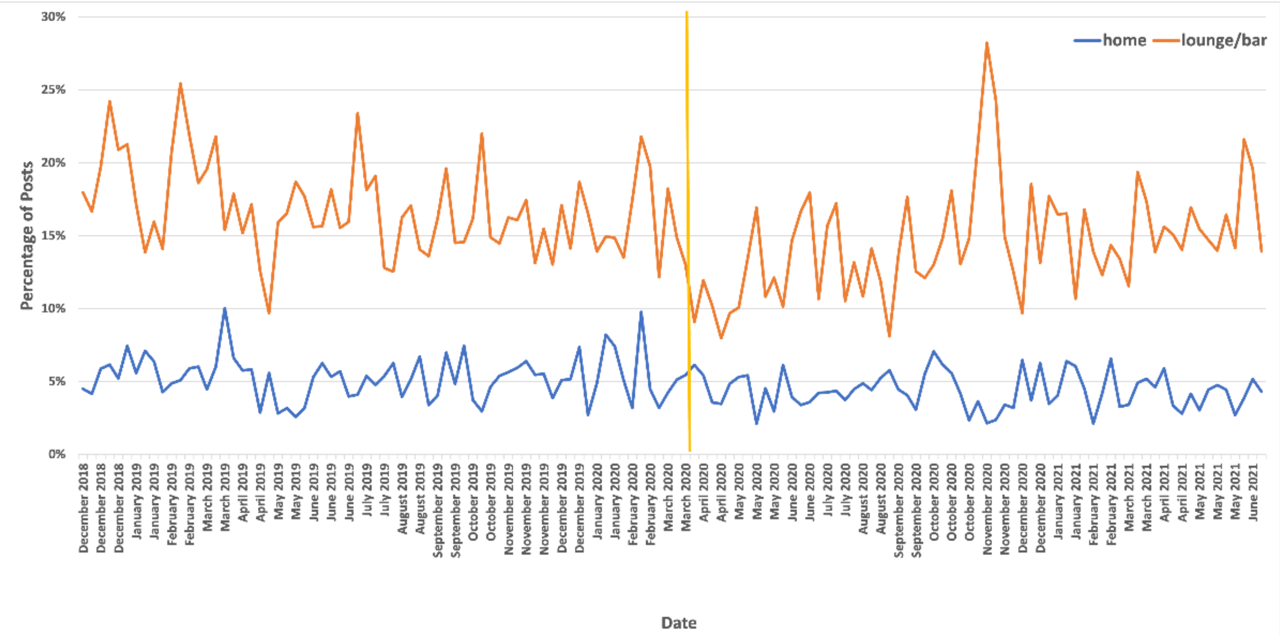


Figure 2. Longitudinal trend in the proportion of water pipe–related Reddit posts mentioning either *home* or *lounge/bar*.



Sentiment Changes in Water Pipe–Related Posts Before and During the COVID-19 Pandemic

To understand whether the COVID-19 pandemic had any impact on the sentiments of water pipe–related posts, we performed a sentiment analysis of water pipe–related posts before and during the COVID-19 pandemic. Before the pandemic, 59.14% (12,526/21,182) of posts had a positive attitude, 21.44% (4541/21,182) of posts had a negative attitude, and 19.43% (4115/21,182) of posts had a neutral attitude. After the pandemic, 59.74% (14,686/24,583) of posts had a positive attitude, 21.01% (5164/24,583) of posts had a negative attitude, and 19.25% (4733/24,583) of posts had a neutral attitude. Further statistical analysis using 2-proportion z tests showed that there was no significant change in the proportion of posts with positive attitude (P=.19), negative attitude (P=.26), or neutral attitude (P=.65) before and after the pandemic.

Topics Discussed in Water Pipe–Related Reddit Posts

LDA topic modeling was used to identify popular topics related to water pipes on Reddit before and during the pandemic. As shown in Table 1, water pipe–related posts (n=21,182) had 6 major topics before the pandemic, including “friends spending time together” (3220, 15.2%), “hookah culture in different countries” (2284, 10.78%), “discussion about waterpipe accessories” (3064, 14.46%), “getting bad feelings when using waterpipe tobacco with other substances" (3755, 17.73%), "smoking hookah at hookah bars/lounges" (4114, 19.42%), and "discussion about coal and shisha flavor" (4745, 22.4%). The 6 most popular topics in Reddit posts after the pandemic started (n=24,583; Table 2) included "friends spending time together" (3453, 14.05%), "getting bad feelings when using waterpipe tobacco with other substances" (3651, 14.85%), "discussion about coal and shisha flavor" (3662, 14.9%), "opening and closing hookah bars/lounges" (4430, 18.02%), "discussion about waterpipe accessories" (4841, 19.69%), and "good feelings about hookah" (4546, 18.49%). The keywords and associated example Reddit posts are also included in both Tables 1 and 2.

Table 1. Topics discussed in water pipe–related Reddit posts (n=21,182) before the pandemic.

Topic	Percentage of tokens, n (%)	Keywords	Example quotations
Friends spending time together	3220 (15.2)	friend, time, back, guy, start, day, feel, leave, year, end	• "Looking through this thread and seems to be like I'm in the minority. I do, and I do it a lot. Granted, only with my close friends, but I have quite a few of those. We call them "heart to hearts", and they are usually accompanied by whiskey, hookah, and petting my too-needy-but-cute-as-shit pittie." • "I took the truck back after he was done camping one night and went to the hookah bar with a friend and to the location and back windows were wide open and we had our shirts over our mouths as we were yelling in fear the entire drive..."
Hookah culture in different countries	2284 (10.78)	people, man, thing, give, post, life, find, world, show, country	• "In Bosnia, hookah is known as Shisha. A lot of Bosnians have moved to Germany, more than any other European country, so maybe that's why." • "I'm from the middle east. Our culture includes: Hookah, Thobes.... Making houses out of fur (usually for camping). The shit stuff you're thinking about is islam."
Discussion about waterpipe accessories	3064 (14.46)	water, make, work, put, pipe, long, hit, water pipe, bit, bong	• "Solo 2 is so good... get the 14mm water pipe adaptor for it and it's a vong beast!" • "When washing your bowl (if you're like me who smokes 100 flavor on same bowl and same hookah). Fully wash the bowl with hot water, use normal towel to dry it, and place it on stove to fully dry it."
Getting bad feelings when using waterpipe tobacco with other substances	3755 (17.73)	smoke, hookah, tobacco, cigarette, smoking, vape, bad, weed, time, nicotine	• "Almost sitting *[sic]* cigarettes, down to 1 or 2 a day from 7-8...by next week it should be once in 2 days or less. But I don't smoke a cig today and did sheesha and went to the gym straight. IT WAS HORRIBLE…Never again. And not even going to do sheesha ever." • "My girlfriend made me try weed before I tried LSD, even though I had no interest in weed at all. I really don't like weed at all. It feels like it clouds my mind, and I feel like I am too heavy…I have smoked hookah for years, so when I went to smoke weed, I kept clearing the bowl in one hit... Apparently that isn't good to do for a first timer."
Smoking hookah at hookah bars/lounges	4114 (19.42)	bar, hookah, good, lounge, place, lot, hooka, great, love, pretty	• "Hound dogs pizza! If you're a smoker hookah bars are usually open really late. Also diners are a good option, fitzys on shrock rd is 24 hours and waffle house and steak and shake." • "More than Fumari and Starbuzz I see. I wonder if it is allowed to be smoked at the hookah lounges in US."
Discussion about coal and shisha flavor	4745 (22.4)	hookah, bowl, good, shisha, coal, buy, hose, flavor, clean, brand	• "Sorry, never personally tried Starbuzz Carbine, cannot really compare the draw differences between them..." • "Serbetli amazing shisha, strong flavor and thick clouds!"

Table 2. Topics discussed in water pipe–related Reddit posts (n=24,583) during the pandemic.

Topic	Percentage of tokens, n (%)	Keywords	Example quotations
Friends spending time together	3453 (14.05)	friend, people, guy, talk, play, back, life, call, time, show	• "Before I moved here from Omaha I used to frequent the hookah lounge all the time!... My boyfriend and I were thinking about going to one of them the other night but didn't know which one was gonna be the best option at the time. I'll definitely have to check out both though!" • "This was just a few weeks ago. Met a girl through hanging with mutual friends at a hookah bar. We spent a couple hours talking to each other, and she asked our friend to give her my number..."
Getting bad feelings when using waterpipe tobacco with other substances	3651 (14.85)	smoke, time, people, feel, day, year, smoking, bad, cigarette, weed	• "I think it can be done. I do hookah sometimes. That doesn't mean I'm ever smoking again! I even, can't stand the smell of cigarettes anymore. In your experience and your strength, there are no rules!" • "Smoke a cigarette or cigar or hookah or vape-tobacco, marijuana, herbal, or ANYTHING!! YUCK! Smelly! Stinky! Obnoxious! And Cancer!... did I mention lung cancer??"
Discussion about coal and shisha flavor	3662 (14.9)	tobacco, bowl, smoke, hookah, coal, shisha, flavor, water, taste, heat	• "Anything by Mason Shishaware. The Gravyl bowl is the best hookah bowl I've ever smoked hands down." • "It depends on what do you like (I like mixing dark leaf with white, using neutral flavors of dark leaf (dark side tobacco (as base (grapefruit, orange, banana (and add strong tastes of daily hookah (exotic fruits, different cocktails, mint, Red Bull)"
Opening and closing hookah bars/lounges	4430 (18.02)	bar, lounge, place, night, open, home, back, sit, room, close	• "til not all non-essential businesses were told to close my distant cousins hookah cafe in a city centre in jabodetabek is apparently still open" • "I really miss Hookah since I usually go out to those shops to smoke em. How much would it cost to buy home a hookah set?"
Discussion about waterpipe accessories	4841 (19.69)	Hookah, hose, buy, pipe, make, glass, clean, quality, base	• "By whip, do you mean like a hookah hose? If so, I LOVE concentrates through one. ZC Glass makes some sweet parts if you ever need splitters or new mouthpieces." • "Glass Lung sells hookah style mason jar bongs that have 2 14 mm female adapters so you can attach a whip and a bowl to. They're pretty cool."
Good feelings about hookah	4546 (18.49)	Hookah, good, work, love, pretty, great, lol, hit, hooka, big	• "Wow, I love the way this hookah looks. Especially the 'vase'! Hot damn I want that!" • "That's pretty awesome! I have a Starbuzz Carbine that has the 360 degree swivel...very nice feature!"

Discussion

Principal Findings

In this study, we showed that the discussion of water pipes on Reddit was gradually increasing until the beginning of the COVID-19 pandemic, and then gradually decreased during the pandemic. There was more discussion about water pipes on Reddit during the pandemic than before the pandemic. While the proportion of posts mentioning water pipe use at home did not change during the study period, the proportion of posts mentioning water pipe use at lounges or bars significantly decreased at the beginning of the pandemic, and gradually increased. Positive water pipe–related posts were dominant, and this did not change with the COVID-19 pandemic. The discussion on water pipes was similar before and during the pandemic. There was more discussion about the opening and closing of hookah bars and lounges during the pandemic.

By examining trends in all water pipe–related posts on Reddit and performing an ITS analysis, we showed that the number of posts related to water pipe tobacco had a growing trend before the pandemic. The announcement of COVID-19 as a global pandemic had a positive effect on this increasing trend, which might be due to the discussion of pandemic lockdowns. After the sharp increase at the beginning of the pandemic, the number of related posts started to decrease.

Comparison With Prior Work

Our findings are aligned with those of a study on smoking in Saudi Arabia that concluded that the use of cigarettes and water pipe tobacco has slightly decreased while e-cigarette use has significantly increased since the pandemic [37]. Our study focused on noncommercial water pipe smoking–related Reddit posts. It would be interesting to see how the sales of water pipe tobacco and accessories changed during the pandemic; this was beyond the scope of this study and will be explored in the future.

In this study, we used location-specific keywords to explore the impact of the COVID-19 pandemic on mentions of different locations (hookah bars/lounges or home). This may be the first study to use social media data to examine how the pandemic affected water pipe use. Our temporal analysis showed that the proportion of discussions about hookah lounges and bars significantly decreased at the beginning of the pandemic and slowly increased during the pandemic. This trend aligns with the timeline for the pandemic: major lockdowns started in March 2020, when the state of California issued a stay-at-home order;

then, in May 2020, the US Centers for Disease Control and Prevention released guidance for reopening the country, followed by the gradual reopening of the economy in the United States [38]. The number of people under confinement worldwide reached its highest point on April 5, 2020, and then started to decrease [39]. Therefore, our results suggest that the pandemic might have had some potential impact on water pipe use, based on mentions of different locations (such as bars and lounges) on Reddit during the pandemic. Due to the lockdown at the beginning of the pandemic, many hookah bars or lounges were closed or not accessible, which might have led to fewer mentions of hookah bars and lounges on Reddit. With fewer restrictions during the later lockdowns, some hookah bars and lounges started to open, and people began searching for possible social activities, such as hanging out at hookah bars and lounges with their friends, which might have resulted in more mentions of hookah bars and lounges on Reddit. The proportion of discussions about using water pipe tobacco at home was low, which indicates that fewer people smoked water pipe tobacco at home in general.

Most of the Reddit posts related to water pipes had positive sentiments in our study. Sentiment analysis of water pipe–related posts on Twitter showed that 59.5% (352,116/591,792) of tweets had a positive attitude, while 30% (177,537/591,792) had a negative attitude, and 10.5% (62,139/591,792) had a neutral attitude [40]. We further showed that there was no change in the sentiment of water pipe–related posts before and during the pandemic, suggesting that the pandemic did not impact the public perception of water pipes.

By comparison, we showed that the most popular topics in water pipe–related posts were similar before and during the pandemic, including “friends spending time together while smoking waterpipe tobacco,” discussions about waterpipe-related products,” and “getting bad feelings when using waterpipe tobacco with other substances like cigarettes and weed.” However, we did notice that while discussion of hookah bars and lounges was present in posts both before and during the pandemic, the focus of the discussion shifted to the opening and closing of hookah bars and lounges after the pandemic started. In addition, before the pandemic people frequently posted and discussed hookah culture around the world, and this became less popular during the pandemic. The great number of travel bans and restrictions caused by the pandemic might be one of the possible reasons for this change [41].

Limitations

There are several limitations to our study. First, as the majority of Reddit users are from North America, our data set may not be representative of the global discussion about water pipes during the pandemic [25]. Given the unique social context of water pipe smoking in the United States, the findings are not generalizable to all countries, especially those countries where water pipe smoking may primarily occur alone or in private homes [42]. Due to the lack of detailed geolocation information for Reddit users, we could not compare the potential impact of the COVID-19 pandemic on water pipe tobacco smoking in Canada and the United States. Therefore, it will be important in the future to examine how the pandemic might affect the use of water pipe or other tobacco products in different countries with different lockdown policies, especially how tobacco product users changed their user behaviors. Second, the Reddit data that we used in this study were historical data, so some water pipe–related Reddit posts might have been deleted, which would have introduced bias to our results. Third, although we examined temporal trends in the percentage of discussions about water pipes and specific locations, we could not distinguish posts about actually smoking water pipes in hookah bars and lounges from discussions about reopening hookah bars and lounges. Fourth, some water pipe users who searched for other locations during the pandemic might not have mentioned it on Reddit, which could also have brought bias into our results. Finally, the sentiment analysis was performed at the post level, so it may not have reflected the actual attitude of Reddit users toward water pipes.

Conclusions

Our study provides a thorough analysis of the potential influence of the COVID-19 pandemic on public perceptions and discussions about water pipes on Reddit by performing temporal analysis and comparing sentiments and topics discussed before and during the pandemic. Our findings show that during the pandemic, especially during lockdowns, mentions of opening or closing of hookah bars and lounges on Reddit gradually increased, suggesting that people were searching for collaborative activities, such as water pipe tobacco smoking in hookah bars and lounges with their friends. Our study shows the potential impact of the pandemic on water pipe tobacco smoking, such as the closing of hookah bars and lounges; this might create an opportunity for public health authorities to communicate with the public during lockdowns about what kind of health collaborative activities they should search for instead of water pipe tobacco smoking. Our study also provides another valid data source obtained from social media for studying the pandemic and any other important public health issues, considering the increasing prevalence of social media use in the modern world.

Acknowledgments

The research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health (NIH) and the Food and Drug Administration (FDA) 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 NIH or the FDA.

Data Availability

The data sets and scripts generated during and/or analyzed during the current study are available on request from the corresponding author (DL).

Authors' Contributions

ZX, MG, and DL conceived and designed the study. ZZ analyzed the data. ZZ and ZX wrote the manuscript. ZX, MG, IR, and DL assisted with interpretation of analyses and edited the manuscript. All authors have approved the final article.

Conflicts of Interest

MG received a research grant from Pfizer and also served as a member of the scientific advisory board of Johnson & Johnson.

References

1. Sharma E, Bansal-Travers M, Edwards KC, Halenar MJ, Taylor KA, Kasza KA, et al. Longitudinal pathways of exclusive and polytobacco hookah use among youth, young adults and adults in the USA: findings from the PATH Study Waves 1-3 (2013-2016). Tob Control 2020 May;29(Suppl 3):s155-s162 [FREE Full text] [doi: 10.1136/tobaccocontrol-2020-055625] [Medline: 32321849]
2. Akl EA, Jawad M, Lam WY, Co CN, Obeid R, Irani J. Motives, beliefs and attitudes towards waterpipe tobacco smoking: a systematic review. Harm Reduct J 2013 Jul 02;10:12 [FREE Full text] [doi: 10.1186/1477-7517-10-12] [Medline: 23816366]
3. Sidani JE, Shensa A, Yabes J, Fertman C, Primack BA. Waterpipe tobacco use in college and non-college young adults in the USA. Fam Pract 2019 Mar 20;36(2):103-109 [FREE Full text] [doi: 10.1093/fampra/cmy037] [Medline: 29741621]
4. Akl EA, Gaddam S, Gunukula SK, Honeine R, Jaoude PA, Irani J. The effects of waterpipe tobacco smoking on health outcomes: a systematic review. Int J Epidemiol 2010 Jul;39(3):834-857. [doi: 10.1093/ije/dyq002] [Medline: 20207606]
5. Waziry R, Jawad M, Ballout RA, Al Akel M, Akl EA. The effects of waterpipe tobacco smoking on health outcomes: an updated systematic review and meta-analysis. Int J Epidemiol 2017 Mar 01;46(1):32-43 [FREE Full text] [doi: 10.1093/ije/dyw021] [Medline: 27075769]
6. Smith-Simone S, Maziak W, Ward K, Eissenberg T. Waterpipe tobacco smoking: knowledge, attitudes, beliefs, and behavior in two U.S. samples. Nicotine Tob Res 2008 Mar;10(2):393-398 [FREE Full text] [doi: 10.1080/14622200701825023] [Medline: 18236304]
7. Kassem NOF, Kassem NO, Liles S, Jackson SR, Posis AIB, Chatfield DA, et al. Levels of urine cotinine from hookah smoking and exposure to hookah tobacco secondhand smoke in hookah lounges and homes. Int J High Risk Behav Addict 2018 Mar;7(1):e67601 [FREE Full text] [doi: 10.5812/ijhrba.67601] [Medline: 29805963]
8. Bommele J, Hopman P, Walters BH, Geboers C, Croes E, Fong G, et al. The double-edged relationship between COVID-19 stress and smoking: Implications for smoking cessation. Tob Induc Dis 2020;18:63 [FREE Full text] [doi: 10.18332/tid/125580] [Medline: 32733178]
9. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med 2020 Mar 26;382(13):1199-1207 [FREE Full text] [doi: 10.1056/NEJMoa2001316] [Medline: 31995857]
10. Cucinotta D, Vanelli M. WHO declares COVID-19 a pandemic. Acta Biomed 2020 Mar 19;91(1):157-160 [FREE Full text] [doi: 10.23750/abm.v91i1.9397] [Medline: 32191675]
11. Lau H, Khosrawipour V, Kocbach P, Mikolajczyk A, Schubert J, Bania J, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J Travel Med 2020 May 18;27(3):taaa037 [FREE Full text] [doi: 10.1093/jtm/taaa037] [Medline: 32181488]
12. Sameer A, Khan M, Nissar S, Banday M. Assessment of mental health and various coping strategies among general population living under imposed COVID-lockdown across world: a cross-sectional study. Ethics Med Public Health 2020;15:100571 [FREE Full text] [doi: 10.1016/j.jemep.2020.100571] [Medline: 32838000]
13. Loh HC, Looi I, Ch'ng ASH, Goh KW, Ming LC, Ang KH. Positive global environmental impacts of the COVID-19 pandemic lockdown: a review. GeoJournal 2022;87(5):4425-4437 [FREE Full text] [doi: 10.1007/s10708-021-10475-6] [Medline: 34316088]
14. Holingue C, Badillo-Goicoechea E, Riehm KE, Veldhuis CB, Thrul J, Johnson RM, et al. Mental distress during the COVID-19 pandemic among US adults without a pre-existing mental health condition: Findings from American trend panel survey. Prev Med 2020 Oct;139:106231 [FREE Full text] [doi: 10.1016/j.ypmed.2020.106231] [Medline: 32758507]
15. Zhang S, Sun L, Zhang D, Li P, Liu Y, Anand A, et al. The COVID-19 pandemic and mental health concerns on Twitter in the United States. Health Data Sci 2022;2022:9758408. [doi: 10.34133/2022/9758408] [Medline: 36408202]
16. Robinson E, Boyland E, Chisholm A, Harrold J, Maloney NG, Marty L, et al. Obesity, eating behavior and physical activity during COVID-19 lockdown: A study of UK adults. Appetite 2021 Jan 01;156:104853 [FREE Full text] [doi: 10.1016/j.appet.2020.104853] [Medline: 33038479]

17. Ambrosetti J, Macheret L, Folliet A, Wullschleger A, Amerio A, Aguglia A, et al. Psychiatric emergency admissions during and after COVID-19 lockdown: short-term impact and long-term implications on mental health. BMC Psychiatry 2021 Oct 24;21(1):465 [FREE Full text] [doi: 10.1186/s12888-021-03469-8] [Medline: 34560856]
18. Rodrigues CA, Rodrigues N, Nascimento M, Oliveira-Silva J. Patterns of adult and youth inpatient admissions before and after the COVID-19 pandemic in a psychiatric ward: an observational study. BMC Health Serv Res 2022 Aug 17;22(1):1048 [FREE Full text] [doi: 10.1186/s12913-022-08374-8] [Medline: 35978322]
19. Kreslake JM, Simard BJ, O'Connor KM, Patel M, Vallone DM, Hair EC. e-Cigarette use among youths and young adults during the COVID-19 pandemic: United States, 2020. Am J Public Health 2021 Jul;111(6):1132-1140. [doi: 10.2105/AJPH.2021.306210] [Medline: 33856888]
20. Gaffney A, Himmelstein DU, Woolhandler S. Smoking prevalence during the COVID-19 pandemic in the United States. Ann Am Thorac Soc 2022 Jul;19(6):1065-1068 [FREE Full text] [doi: 10.1513/AnnalsATS.202110-1184RL] [Medline: 35044900]
21. Nyman AL, Spears CA, Churchill V, Do VV, Henderson KC, Massey ZB, et al. Associations between COVID-19 risk perceptions and smoking and quitting behavior among U.S. adults. Addict Behav Rep 2021 Dec;14:100394 [FREE Full text] [doi: 10.1016/j.abrep.2021.100394] [Medline: 34869823]
22. Tattan-Birch H, Perski O, Jackson S, Shahab L, West R, Brown J. COVID-19, smoking, vaping and quitting: a representative population survey in England. Addiction 2021 May;116(5):1186-1195 [FREE Full text] [doi: 10.1111/add.15251] [Medline: 32918300]
23. Bar-Zeev Y, Shauly M, Lee H, Neumark Y. Changes in smoking behaviour and home-smoking rules during the initial COVID-19 lockdown period in Israel. Int J Environ Res Public Health 2021 Mar 17;18(4):1931 [FREE Full text] [doi: 10.3390/ijerph18041931] [Medline: 33671203]
24. Reddit by the Numbers. Reddit. URL: https://www.redditinc.com/press [accessed 2023-03-31]
25. Proferes N, Jones N, Gilbert S, Fiesler C, Zimmer M. Studying Reddit: A systematic overview of disciplines, approaches, methods, and ethics. Soc Media Soc 2021 May 26;7(2) [FREE Full text] [doi: 10.1177/20563051211019004]
26. Sharma R, Wigginton B, Meurk C, Ford P, Gartner C. Motivations and limitations associated with vaping among people with mental illness: A qualitative analysis of Reddit discussions. Int J Environ Res Public Health 2016 Dec 22;14(1):7 [FREE Full text] [doi: 10.3390/ijerph14010007] [Medline: 28025516]
27. Chen L, Lu X, Yuan J, Luo J, Luo J, Xie Z, et al. A social media study on the associations of flavored electronic cigarettes with health symptoms: Observational study. J Med Internet Res 2020 Jul 22;22(6):e17496 [FREE Full text] [doi: 10.2196/17496] [Medline: 32568093]
28. 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 Jul 24;22(6):e17280 [FREE Full text] [doi: 10.2196/17280] [Medline: 32579123]
29. Luo J, Chen L, Lu X, Yuan J, Xie Z, Li D. Analysis of potential associations of JUUL flavours with health symptoms based on user-generated data from Reddit. Tob Control 2021 Oct;30(5):534-541 [FREE Full text] [doi: 10.1136/tobaccocontrol-2019-055439] [Medline: 32709604]
30. Shao Y, Zou J, Xie Z, Mayne RG, Ossip DJ, Rahman I, et al. Perceptions of oral nicotine pouches on Reddit: Observational study. J Med Internet Res 2022 Jul 15;24(7):e37071 [FREE Full text] [doi: 10.2196/37071] [Medline: 35838764]
31. Low DM, Rumker L, Talkar T, Torous J, Cecchi G, Ghosh SS. Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on Reddit during COVID-19: Observational study. J Med Internet Res 2020 Oct 12;22(10):e22635 [FREE Full text] [doi: 10.2196/22635] [Medline: 32936777]
32. Grant A, O'Mahoney H. Portrayal of waterpipe (shisha, hookah, nargile) smoking on Twitter: a qualitative exploration. Public Health 2016 Dec;140:128-135. [doi: 10.1016/j.puhe.2016.07.007] [Medline: 27520707]
33. Penfold RB, Zhang F. Use of interrupted time series analysis in evaluating health care quality improvements. Acad Pediatr 2013;13(6 Suppl):S38-S44. [doi: 10.1016/j.acap.2013.08.002] [Medline: 24268083]
34. Hutto C, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the International AAAI Conference on Web and Social Media. 2014 Presented at: Eighth International AAAI Conference on Weblogs and Social Media; June 1-4, 2014; Ann Arbor, MI p. 216-225. [doi: 10.1609/icwsm.v8i1.14550]
35. Blei D, Ng A, Jordan M. Latent Dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text]
36. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, et al. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimed Tools Appl 2018 Nov 28;78(11):15169-15211. [doi: 10.1007/s11042-018-6894-4]
37. BinDhim NF, Althumiri NA, Basyouni MH, AlMousa N, AlJuwaysim MF, Alhakbani A, et al. Exploring the impact of COVID-19 response on population health in Saudi Arabia: results from the "Sharik" health indicators surveillance system during 2020. Int J Environ Res Public Health 2021 May 16;18(10):5291 [FREE Full text] [doi: 10.3390/ijerph18105291] [Medline: 34065706]
38. CDC Museum COVID-19 Timeline. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/museum/timeline/covid19.html [accessed 2023-03-31]

39. Bates AE, Primack RB, Moraga P, Duarte CM. COVID-19 pandemic and associated lockdown as a "Global Human Confinement Experiment" to investigate biodiversity conservation. Biol Conserv 2020 Aug;248:108665 [FREE Full text] [doi: 10.1016/j.biocon.2020.108665] [Medline: 32549587]
40. Allem J, Ramanujam J, Lerman K, Chu K, Boley Cruz T, Unger JB. Identifying sentiment of hookah-related posts on Twitter. JMIR Public Health Surveill 2017 Oct 18;3(4):e74 [FREE Full text] [doi: 10.2196/publichealth.8133] [Medline: 29046267]
41. Gössling S, Scott D, Hall CM. Pandemics, tourism and global change: a rapid assessment of COVID-19. J Sustain Tour 2020 Apr 27;29(1):1-20 [FREE Full text] [doi: 10.1080/09669582.2020.1758708]
42. El Shahawy O, Park SH, Rogers ES, Shearston JA, Thompson AB, Cooper SC, et al. Hookah use patterns, social influence and associated other substance use among a sample of New York City public university students. Subst Abuse Treat Prev Policy 2020 Aug 28;15(1):65 [FREE Full text] [doi: 10.1186/s13011-020-00283-5] [Medline: 32859230]

Abbreviations

FDA: Food and Drug Administration
ITS: interrupted time series
LDA: latent Dirichlet allocation
NIH: National Institutes of Health
VADER: Valence Aware Dictionary and Sentiment Reasoner
WTS: water pipe tobacco smoking

Edited by T Mackey; submitted 08.07.22; peer-reviewed by M Nali, R Kobayashi, N Seeman; comments to author 11.01.23; revised version received 01.02.23; accepted 27.03.23; published 20.04.23.

Please cite as:
Zheng Z, Xie Z, Goniewicz M, Rahman I, Li D
Potential Impact of the COVID-19 Pandemic on Public Perception of Water Pipes on Reddit: Observational Study
JMIR Infodemiology 2023;3:e40913
URL: https://infodemiology.jmir.org/2023/1/e40913
doi:10.2196/40913
PMID:37124245

Original Paper

Global Misinformation Spillovers in the Vaccination Debate Before and During the COVID-19 Pandemic: Multilingual Twitter Study

Jacopo Lenti[1,2], MSc; Yelena Mejova[3], PhD; Kyriaki Kalimeri[3], PhD; André Panisson[2], PhD; Daniela Paolotti[3], PhD; Michele Tizzani[3], PhD; Michele Starnini[2,4], PhD

[1]Department of Computer, Control, and Management Engineering Antonio Ruberti, Sapienza University of Rome, Rome, Italy
[2]CENTAI Institute S.p.A., Turin, Italy
[3]Institute for Scientific Interchange Foundation, Turin, Italy
[4]Departament de Fisica, Universitat Politecnica de Catalunya, Barcelona, Spain

Corresponding Author:
Michele Starnini, PhD
Departament de Fisica, Universitat Politecnica de Catalunya
Campus Nord B4
Barcelona, 08034
Spain
Phone: 34 934 01 62 00
Email: michele.starnini@upc.edu

Abstract

Background: Antivaccination views pervade online social media, fueling distrust in scientific expertise and increasing the number of vaccine-hesitant individuals. Although previous studies focused on specific countries, the COVID-19 pandemic has brought the vaccination discourse worldwide, underpinning the need to tackle low-credible information flows on a global scale to design effective countermeasures.

Objective: This study aimed to quantify cross-border misinformation flows among users exposed to antivaccination (no-vax) content and the effects of content moderation on vaccine-related misinformation.

Methods: We collected 316 million vaccine-related Twitter (Twitter, Inc) messages in 18 languages from October 2019 to March 2021. We geolocated users in 28 different countries and reconstructed a retweet network and cosharing network for each country. We identified communities of users exposed to no-vax content by detecting communities in the retweet network via hierarchical clustering and manual annotation. We collected a list of low-credibility domains and quantified the interactions and misinformation flows among no-vax communities of different countries.

Results: The findings showed that during the pandemic, no-vax communities became more central in the country-specific debates and their cross-border connections strengthened, revealing a global Twitter antivaccination network. US users are central in this network, whereas Russian users also became net exporters of misinformation during vaccination rollout. Interestingly, we found that Twitter's content moderation efforts, in particular the suspension of users following the January 6 US Capitol attack, had a worldwide impact in reducing the spread of misinformation about vaccines.

Conclusions: These findings may help public health institutions and social media platforms mitigate the spread of health-related, low-credibility information by revealing vulnerable web-based communities.

(JMIR Infodemiology 2023;3:e44714) doi:10.2196/44714

KEYWORDS

vaccination hesitancy; vaccine; misinformation; Twitter; social media; COVID-19

Introduction

Background

The COVID-19 pandemic has extended vaccination from the purview of parents and health-compromised individuals to the purview of the broader public. Restrictions around vaccination created an additional potential to impact one's personal freedom and the world economy, as well as one's health. However, vaccination hesitancy continues to limit the impact of this highly effective intervention [1]: hundreds of thousands of lives were lost to COVID-19 that could have been prevented with vaccinations in the United States alone [2].

Vaccination hesitancy is a complex issue that has been associated with science denial [3], alternative health practices [4], and belief in conspiracy theories [5]. Among the many factors contributing to vaccine hesitancy is the spread of misinformation, especially on the web [6,7]. The impact of antivaccination content on online social media (OSM) may be compounded by the so-called echo chamber effect [8], in which users' beliefs are reinforced through interactions with like-minded peers [9-11]. Created by the interplay of (1) homophily between users' interactions and (2) polarization of the debate, echo chambers arise from a combination of the psychological tendencies of confirmation bias and selective exposure [12-14] together with algorithmic optimization for greater engagement at the cost of content diversity [15]. Importantly, echo chambers have also been found on OSM in the discussions around vaccination [16-19].

Thus far, scientific studies of the debate around vaccination on OSM have focused on specific countries [17,18,20,21] or English-speaking users [19]. Nevertheless, the COVID-19 pandemic has brought the vaccination discourse to a global scale [22], creating a deluge of international news around the development and deployment of COVID-19 vaccines, including low-quality content and misinformation [23]. The danger of this "infodemic" was acknowledged in mid-2020 by the United Nations and World Health Organization, which called for the member states to develop and implement the necessary action plans [1,24]. Thus, it is imperative to understand the flow of antivaccine—or no-vax—information not only nationally but also internationally to obtain a bird's-eye view on the topic and inform effective communication campaigns.

To address this need, in this work, we focused on the Twitter (Twitter, Inc) platform by leveraging 316 million tweets related to vaccines in 18 different languages from a pre–COVID-19 pandemic era to April 2021 to quantify misinformation flows among users in no-vax communities across national borders and identify which countries are central in the global vaccination debate. To this end, we first investigated (1) how polarized, in terms of echo chambers phenomenon, the vaccination debate is in different countries over time to identify users in no-vax communities and (2) how susceptible, in terms of circulation of information, these no-vax communities are to low-quality information. We proposed a flexible, language-neutral community detection approach and combined it with human-in-the-loop expert knowledge to track polarization and echo chambers in different countries and time periods. We show that communities in which no-vax content was shared (1) increased in number during the pandemic, (2) became less isolated in the national vaccination debate, and (3) displayed much stronger cross-border connections than the rest of the users. Alarmingly, users in these communities tend to heavily rely on low-credibility information sources and to spread it across national borders, resulting in international spillovers of misinformation through a global no-vax network.

Related Works

Vaccination deliberation on Twitter has been studied mainly in English and in the United States [25-27]. However, recently, the platform has gained attention from researchers also focusing on European countries. Before the pandemic, an analysis of the Dutch Twitter revealed an antagonistic relationship between an "anti-establishment" community and the community of journalists and writers, reinforcing the "arrogance of the elite" world view in the former [28]. On the Italian Twitter, the debate around vaccination revealed polarization in terms of retweets (RTs), where vaccine skeptics often mentioned vaccine advocates (generally in attacks), whereas the advocates seemed to ignore the skeptics altogether [17]. Outside Europe, a randomized study on Indonesian Twitter showed the importance of celebrity endorsement in message engagement and that the inclusion of the information source is associated with decreased propagation [29]. The COVID-19 pandemic has spurred increased attention to this topic. A recent examination of vaccine-critical actors on Francophone Twitter found that their place in discussions on vaccines has remained relatively constant during the pandemic compared with the mainstream media [20]. Furthermore, Crupi et al [18] studied the Italian Twitter during the rollout of the COVID-19 vaccinations, showing greater engagement across vaccine-supporting and hesitant communities in terms of mentions and similarity between the communities in the topics discussed.

Attempts to study the flows of vaccination discussion across borders have thus far been limited to dyadic relationships and English. A study of Canadian Twitter users found that most misinformation circulating on Twitter that was shared by Canadian accounts was retweeted from US-based accounts and that increased exposure to US-based information on Twitter is associated with an increased likelihood to post misinformation [30]. Beyond Twitter, Ng et al [22] examined news articles about COVID-19 from 20 countries, identifying the shift in narratives as the pandemic occurred. However, the data were limited to the English language and failed to capture the local language coverage. Unlike the previous studies, our study tracked the vaccination debate in the native languages of numerous countries to systematically study the flow of information (and potential misinformation) across national borders.

The most concerning aspect of the vaccination debates studied here is that misinformation may damage the confidence in the procedure. Controlled exposure studies have shown that web-based misinformation—especially misinformation that sounds scientific—negatively impacts vaccination intent in participants in the United States, United Kingdom [31], and New Zealand [32]. A panel study of US Twitter users found that the risk of average users occasionally sharing misinformation was alarmingly high, despite social bots' contribution to misinformation sharing being "surprisingly low" [33]. Although some efforts have been made toward using high-quality, manually annotated data sets for identifying misinformation [34], the quality of the cited URL domains is often used as a gauge of the quality of the tweet's content [35,36]. In this study, we used a similar approach by combining lists of low-credibility domains from several languages and countries.

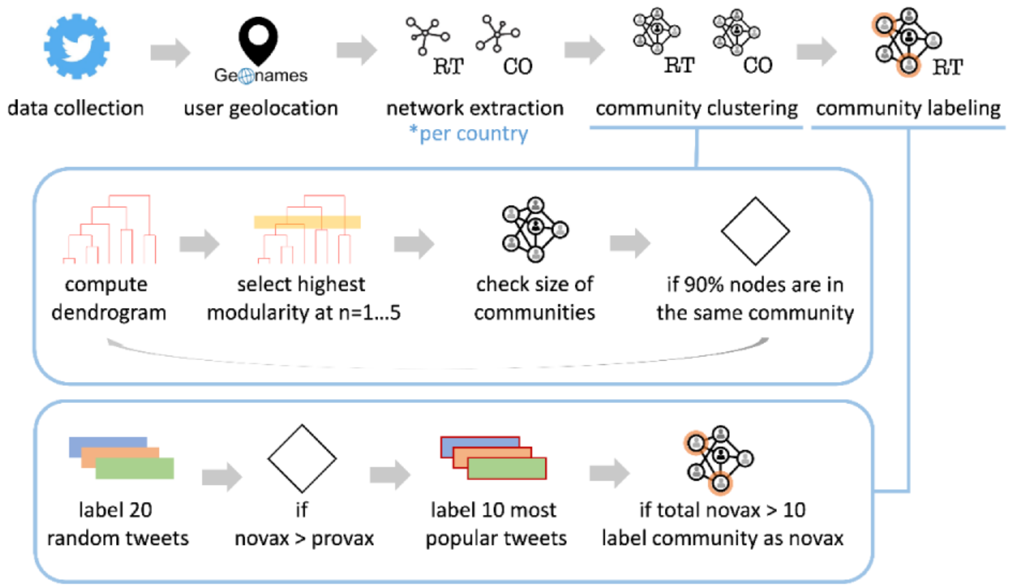
Beyond content analysis, an important aspect of information and misinformation spread is the network structure underlying such dynamic processes. Echo chambers in the Twitter debate around the impeachment of former Brazilian President Dilma Rousseff have been shown to alter the diffusion of information between the supporters and opponents of the impeachment [37]. A similar methodology has been used to compare different topics across social media [8], highlighting that Facebook (Meta Platforms, Inc) shows a higher segregation of news consumption than Reddit (Reddit Inc). Along the same research line, the Random Walk Controversy (RWC) score [38] quantifies how controversial the topics discussed over a certain social network are as the probability of an average user being exposed to information from their own side versus from the opposing side. Although several studies address the presence of echo chambers on social media and their effect on information diffusion, little to no efforts have been devoted to understanding the echo chamber effects within cross-border information spreading, which we examined in this study.

Methods

Overview

The methodology of the data processing pipeline is outlined in the flowchart in Figure 1. First, we used the Twitter Streaming application programming interface (API) to collect a multilingual data set, which we geolocated using the GeoNames database [39]. To identify potential misinformation, we found lists of low-credibility domains in different languages. For the selected countries, we built 2 networks, RT and cosharing (CO; identified by users sharing the same URLs), and applied clustering to find communities. We then manually labeled (in 2 stages) samples of tweets from these communities to identify communities in which users were likely to encounter no-vax content. Finally, we computed several measures to quantify network polarization and information CO, as well as the intensity of cross-national interactions among no-vax communities.

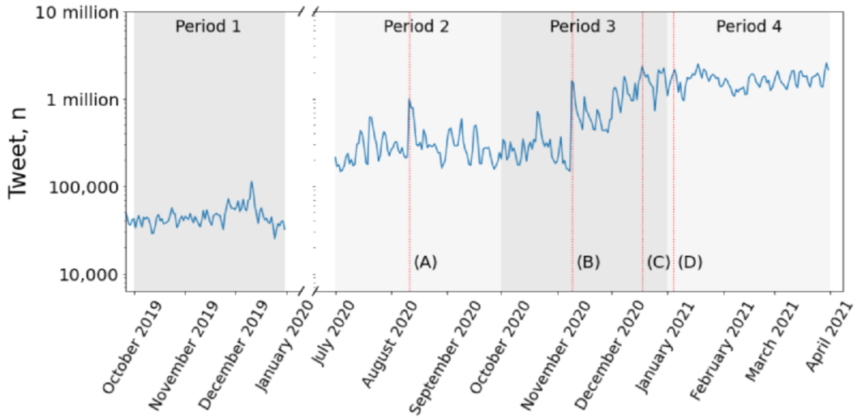
Figure 1. Flowchart of data processing, network extraction, community clustering, and community labeling. RT: retweet; CO: cosharing; novax: number of tweets labelled as discrediting vaccines per community; provax: number of tweets labelled as supporting vaccines per community.



Data Set

We began by assembling a list of vaccine-related words translated into 18 different languages (vaccine, novax, measles, MMR, vaccinated, etc), obtaining a set of 459 keywords (see queries here [40]). An existing list from previous work [17] was expanded by iterative querying of Twitter and expanding the list until no new keywords could be found. Native speakers were then recruited to translate the words into other languages and were instructed to include different common grammatical variations or local relevant keywords. For each language, we query the Twitter Streaming API [41] for the tweets containing the keywords in that language (translated by volunteer native speakers) and keywords in English by applying a language filter. For analysis, we chose four 3-month periods: (1) pre–COVID-19 pandemic period, from October 1, 2019, to December 31, 2019; (2) prevaccine period, from July 1, 2020, to September 30, 2020; (3) vaccine development period, from October 1, 2020, to December 31, 2020; and (4) vaccine rollout period, from January 1, 2021, to March 31, 2021. Figure 2 presents a summary of the daily volume of the data set. The volume increased by 2 orders of magnitude during the pandemic, from 6 million tweets in the 3-month pre–COVID-19 pandemic period to 39 million tweets in the prevaccine period to 91 million tweets in the vaccine development period to 178 million tweets in the vaccine rollout period. To check the completeness of our data, we ran an Historical API [42] in the pre–COVID-19 pandemic period with the same keywords. Owing to account suspension or post removal by the users themselves, a wide fraction of the tweets (72%) was not retrieved by this API, showing that such a data set cannot be retrieved by a retrospective search. Moreover, we took advantage of the passage of time to revisit the most notable accounts (present in the networks described subsequently) using the Twitter Get User API call [43] to check on their status, specifically noting whether the accounts have been suspended by the platform or deleted by the users.

Figure 2. Volume of the vaccination debate on Twitter. Some external events with a substantial impact: (A) August 11, 2020: Sputnik V vaccine announced; (B) November 9, 2020: Pfizer-BioNTech vaccine announced; (C) December 18, 2020: Moderna vaccine announced; and (D) January 4, 2021: first AstraZeneca vaccine inoculation.



Geolocation

To capture country-specific dynamics of the social networks, we geolocated the users: we matched the location they provided in their description with the geographical database of locations from GeoNames. Manually verifying the matching accuracy, we filtered out >500 words often associated with nonlocations in this field. To further limit incorrect geolocations, we (1) removed the geolocation of users who changed their country locations during the observed period and (2) manually inspected users responsible for >50% of RTs between 2 pairs of countries in 1 period, assuming that a user who is heavily retweeted from another country is more likely to be wrongly geolocated. Under these conditions, we geolocated 48.7% of the users. This then allowed us to select countries for the study (as the focus was on the Western languages, we selected countries from Europe, North America, South America, and Oceania). To this end, we filtered countries with >2000 unique users in each period, obtaining 28 countries spanning 11 languages. Figure S1 in Multimedia Appendix 1 provides further details on the volume of tweets per language. The total number of different users geolocated in the chosen countries is 14.9 million, corresponding to 39.4% of the total number of users of the data set.

Low-Credibility Domains

Following the previous literature on misinformation tracking [44,45], we collected a list of low-credibility domains. As sources of low-credibility websites, we relied on Bufale (Italian) [46], Wikipedia (English) [47], Media Bias/Fact-Check (English) [48], Le Monde (French) [49], and dwrean (Greek) [50], obtaining a list of 1732 domains. The fact that we were unable to find lists for less-used languages is an important limitation of this work, which we discuss in the *Discussion* section.

Network Reconstruction

For each country, for each period, we built an RT network and a CO network. To limit the number of geolocation mismatches and filter users belonging to debates in other countries, we constrained the tweets considered for each country to the most common language in our data from that country among the languages spoken in the country (according to Wikipedia). The RT network is a directed weighted graph, where each node is a user, and the weight of the directed link *ij* is the number of times that user *i* retweeted user *j*. The CO network is an undirected weighted graph, where each node represents a user, and the weight of the undirected link *ij* is the number of unique URLs shared by both the users. In order to alleviate the computational cost of the network analyses, we filtered out the edges with weight equal to 1 (just 1 retweet) for the networks with more than 200.000 nodes. This filter affects the RT and CO networks in the United States (all periods), Brazil (PD, VD, and VR periods), Great Britain (VD and VR), and Spain and Mexico (VR period). This filter affects only the country-specific analyses of the RT networks, without influencing the later cross-country analysis and the findings about suspended accounts. When considering the constructed networks, we focused on the Giant Connected Component (GCC). On average, the GCC of the RT network contains 92% of its nodes, while the GCC of the CO network contains 76% of its nodes. In addition, we measured the Overlap Coefficient (OC) between the sets of users in the RT and CO networks. The OC is defined as the ratio of the size of the intersection of 2 sets, A and B, to the size of the smaller set, that is, $OC(A,B) = |A \cap B| / \min(|A|, |B|)$. During the vaccine rollout period, the OC between the sets of users in the RT and CO networks increases from 0.72 (pre–COVID-19 pandemic period) to 0.86, indicating that more people are sharing URLs. The total number of users in the reconstructed RT and CO networks is 2.7 million.

Hierarchical Clustering

Next, we applied a community detection algorithm to cluster the users of the RT and CO social networks. Because the goal was to find a small number of large groups of users, we adopted hierarchical clustering, instead of unsupervised algorithms (eg, Louvain), which finds the optimal partitioning with a very large number of often small communities. We used Paris, an agglomerative hierarchical clustering algorithm induced by the probability of sampling node pairs [51]. Next, we cut the dendrogram to have a reasonable number of communities that are not too unbalanced following the steps listed here:

1. Build the dendrogram of the hierarchical clustering.

2. Compare the partitions obtained with cutoffs at heights 2, 3, 4, and 5 (ie, having this number of communities).

3. Pick the partition with the highest modularity.

4. If >90% of the nodes are in the same community, compare the partitions with cutoffs at the following 5 heights and repeat from step 3.

Using this procedure, we ensured that 90% of the users were partitioned into at least 2 communities but no >5 communities (although it is possible to have many small communities that comprise <10% of the nodes). Using this method, we found, on average, 6.7 communities in the RT networks and 5.0 communities in the CO networks, with a maximum of 20.

Labeling

To identify communities in which users were exposed to no-vax content, we labeled a sample of tweets shared in each community. First, we filtered out small communities by considering only those with >1% of the users of the network, resulting in 400 communities. Next, we randomly sampled 20 tweets from each community, resulting in a total of 8000 tweets. A total of 12 people were involved in labeling, all of whom had a background in vaccine debate and knowledge of the language used in the tweet to label. Furthermore, we translated all tweets into English using Google Translate to allow for cross-checking. Each person labeled between 600 and 1000 tweets, with an overlap of 20 tweets with other annotators. The tweets were labeled as "pro-vax," "no-vax," or "other." We labeled tweets as pro- or no-vax only if they were clearly supporting or discrediting vaccines, respectively. Therefore, more than half of the labels were "other," comprising nonrelevant posts, posts with unclear positions, discussions on other policies, and all generic pieces of news that did not express a stance. The task of distinguishing between pro vaccination and antivaccination stances proved to be fairly easy, with Cohen κ computed on an overlapping set at κ=0.84 (only 3% received different labels). By contrast, the task involving the "other" label proved to be more difficult, with κ=0.51 for the 3-class setting (disagreement of 26%), mainly because of the confusion between "other" and "pro-vax" labels (disagreement of 20%). However, note that we were only interested in distinguishing between the antivaccination stance and the rest.

To improve the quality of the labels, we then proceeded to a second round of annotation, focusing on the communities that have a majority of content with a no-vax stance. Specifically, we chose the communities with a majority of no-vax tweets and annotated the 10 most popular tweets in each (excluding the 50 most popular tweets in the whole network). The second labeling stage encompassed 82 communities, totaling in 820 tweets. At this stage, the Cohen κ for the 3 classes was 0.64. Finally, we defined a community as no-vax if the total number of "no-vax" labels in the rounds was >10, resulting in 58 communities. Because some networks had >1 antivax community, we had 52 networks with a no-vax community, that is, a community where users were substantially exposed to no-vax content (Figure S2 in Multimedia Appendix 1).

Clustering Robustness

Next, we assessed the robustness of our approach to determine if our methodology influenced the results. To do so, we compared the communities previously identified in the RT networks with those obtained using 2 alternative partitioning algorithms: Louvain and Paris hierarchical clustering with a cutoff of 10. We chose the Louvain algorithm for its popularity in community detection problems in social network analysis and Paris with a cutoff of 10 for comparing the results obtained with a different parameter in the cutoff of the same dendrogram. We quantified the number of labeled tweets shared by users in the new clustering and categorized communities as "no-vax" using 2 different thresholds: the "majority threshold" and the "strict threshold." The former was applied when "no-vax" labels outnumbered "pro-vax" labels, while the latter was used when "no-vax" labels surpassed both "pro-vax" and "other" labels. This yielded 4 alternatives to our method: Louvain partitioning with majority threshold, Louvain partitioning with strict threshold, hierarchical clustering with majority threshold, and hierarchical clustering with strict threshold.

To evaluate the robustness of our methodology, we calculated the accuracy for each network as the proportion of users classified in the same group as the previous method (either "no-vax" or "not no-vax"). Our results demonstrate a high level of robustness, with average accuracies of 0.90 and 0.94 using Louvain partitioning with majority and strict thresholds (SD 0.15 and 0.10) and 0.92 and 0.95 using hierarchical clustering (SD 0.15 and 0.10), respectively. These findings support the consistency of the results presented in this paper, which are not overly dependent on the methodology used to detect and label communities.

RWC Score

Following previous literature [10,17,52], we used the RWC score to quantify the polarization between the communities labeled as no-vax and the rest of the network. Given an RT network, partitioned into 2 clusters X and Y, RWC is calculated as $P_{XX}\,P_{YY}-P_{XY}\,P_{YX}$, where P_{XY}=P (a random walk ended in Y started in X). Intuitively, it represents the difference in probability for an average user in the network to be exposed to information from their own side versus that from the opposing side. Spanning (0, 1), an RWC close to 1 represents a polarized social network with 2 distinct groups that do not endorse each other's opinions, whereas an RWC close to 0 represents a noncontroversial topic where both opinions are equally likely to be received.

Normalized Mutual Information

We quantified the echo chamber effect by measuring the extent to which users from different RT communities shared the same sources of information (as quantified by the CO network), as a proxy for the information siloing in an echo chamber. To do this, we used normalized mutual information (NMI) [53] to gauge the similarity between the RT and CO communities obtained by hierarchical clustering, using the normalized_mutual_info_score module in the Python package *scikit-learn*. Spanning (0, 1), an NMI of 1 means that the community structure is the same between the networks, whereas a low NMI indicates different community structures. Note that this metric did not use the opinion leaning determined in the labeling step and thus could be computed for any country/period network, regardless of whether it had a no-vax community.

Normalized RT Volume

To assess the extent to which one country retweets another, we computed a normalized retweeting volume for each pair of countries. To this aim, we started by the total number of RTs from country i to country j (which we have indicated as a_{ij}). Then, we divided it by the total number of RTs by users in country i (indicated as s_i^{out}) and the total number of RTs to users in country j (indicated as s_j^{in}), and we multiplied it by the total number of RTs by all countries (indicated by W). The normalized retweet volume n_{ij} is thus equal to the following:

Note that , and is the expected number of RTs from country i to country j in the random graph with the same node strengths. Hence, n_{ij} >1 if i retweets j more than it would in a random baseline context. As the vast majority of RTs were within the same countries, if , otherwise . To focus on cross-border interactions, we consider n_{ii} = 0, for any i.

Cross-Border Interactions Between No-Vax Communities

We measured the strength of ties among the users in no-vax communities in different countries by comparing the number of RTs among these users with the number of RTs among the rest of the users in the same countries. In particular, we define V_i^K as the set of users in communities with stance K in a country i, where K can be A (antivax) or O (others). We define W_{ij}^K as the number of RTs from users in V_i^K to users in country j with the same stance K, V_j^K. Thus, one can measure the density, the ratio of observed RTs; and the total possible pairs between sets V_i^K and V_j^K, that is, the probability that 2 random users in V_i^K and V_j^K, respectively, are connected. For each pair of countries, we analyzed if , it means that the probability that 2 random users in no-vax communities in countries i and j are connected is higher than the probability that 2 random users in the rest of the network of the same countries are connected.

Ethical Considerations

Although the data were collected using the platform's own API, resulting only in posts that were posted publicly, it is possible that some users were unaware of the scope of their potential audience. Thus, we follow the platform's terms of service and share only the IDs of the tweets so that when the data are recollected, deleted content will not be available (notably limiting the reproducibility of any Twitter-based study). Thus, the data that are shared do not contain any identifiable information about the poster or any other information except the numeric ID of the tweet, preserving the privacy of the user to the extent that they choose to keep their posts public on the platform. This will affect the reproducibility of the study, as some content may be deleted by the users over time. Furthermore, the multinational nature of the data captured wildly varying biases in the way people around the world are able to access the internet or Twitter specifically. Local barriers to access to the internet and local blocks of the platform itself shape the communities captured in this study. For instance, dissidents or those who wished to remain anonymous would likely not have shared their location information on their profile and would not have been captured as being a part of a country's discussion. The construction of retweet and cosharing networks also necessitates enough activity by the user to be included in the analysis, biasing our results to those who are more active in the conversation, especially in retweeting and sharing URLs. Moreover, the data may have captured vulnerable groups, including those who experienced or who were at risk of specific health conditions, those who had financial barriers to health care, and even those who were more susceptible to misinformation. Despite the negative connotation around "no-vax" communities, users found to propagate harmful information may first and foremost endanger themselves by following faulty advice. Thus, we would discourage the future researchers from publishing verbatim tweet text to preserve user privacy.

Finally, in this paper, we present tools that may be used to track and profile groups of Twitter users around a topic. These tools may then be used by both the platform and the government. However, such tools may also be used to target communities for harassment, doxing (providing private user information to harm or intimidate the person), and other abuses. On the one hand, it is the responsibility of the platforms and their communities to uphold the civil code of conduct and block the abusers. On the other hand, we call for the research community, as well as corporate and governmental actors, to use these tools ethically, with minimal harm to the participants.

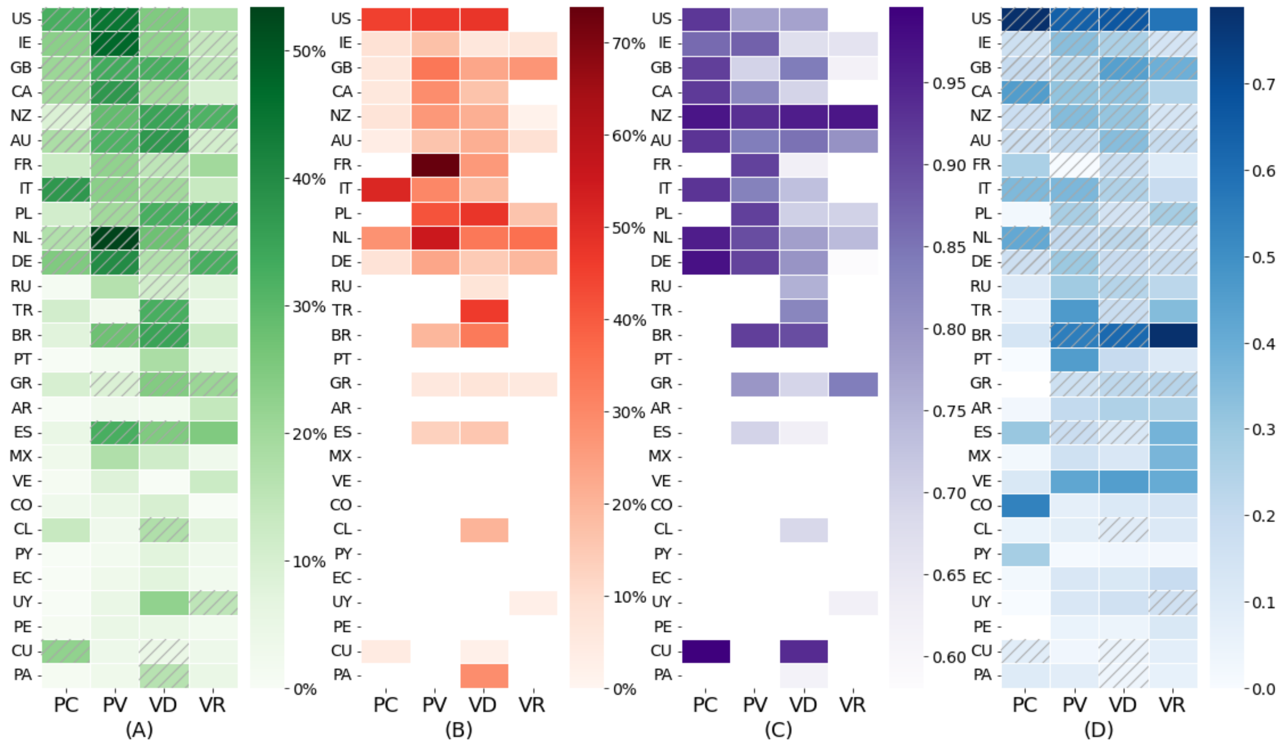
Results

Polarization of the Vaccination Debate

We began by examining different measures of polarization and no-vax activity in different countries over the 4 periods. Figure 3A shows that a high presence of no-vax tweets in a certain country and period is often associated with the presence of a community labeled as no-vax (dashed lines). This implies that no-vax content is generally clustered and not homogeneously distributed in the RT network, suggesting that the debate is polarized, as illustrated subsequently. Furthermore, we found that no-vax communities were generally present in the English-speaking countries (eg, compared with the Spanish-speaking ones). However, some of the relatively largest country-specific no-vax communities appeared in France, Italy, the Netherlands, Poland, and the United States (Figure 3B). No-vax communities were particularly present in the prevaccine and vaccine development periods, where they also spanned a larger fraction of users compared with the other periods. Turning to potential echo chambers in these networks, we found that the

RWC score was overall very high (Figure 3C), indicating that the vaccination debate was generally highly polarized. However, it decreased substantially over time, suggesting that the users in no-vax communities became less isolated in the vaccination discourse during the COVID-19 pandemic. Furthermore, we investigated whether the users in the no-vax communities were exposed to information sources that were different from those that the rest of the users were exposed to by considering NMI. Despite the NMI being independent of the labeling of the communities, Figure 3D shows that, on average, the NMI of the networks with a no-vax community was higher than that of the others (0.27 vs 0.22, *P*<.05), indicating that the users in no-vax communities tended to have common information sources. Some countries, such as the United States and Brazil, showed an especially high NMI, indicating that the polarization in the RT network was reflected in the different content shared. The Spanish-speaking countries, conversely, were less polarized than the English-speaking countries (average 0.15 vs 0.33, *P*<.001).

Figure 3. Characterization of no-vax communities for each country and period considered via retweet (RT) and cosharing (CO) networks. (A) Proportion of tweets labeled as no-vax. (B) Proportion of users in no-vax communities with respect to the size of the RT networks (average in the 4 periods: 16.9% [SD 0.18], 30.9% [SD 0.18], 23.1% [SD 0.14], and 13.7% [SD 0.12]). (C) Random Walk Controversy between no-vax communities and the rest of the networks (average in the 4 periods: 0.94 [SD 0.04], 0.84 [SD 0.08], 0.76 [SD 0.10], and 0.73 [SD 0.12]). (D) Normalized mutual information between RT and CO communities. Countries with no-vax communities are marked with dashed lines. PC: pre–COVID-19 period; PV: prevaccine period; VD: vaccine development period; VR: vaccine rollout period. AR: Argentina; AU: Australia; BR: Brazil; CA: Canada; CL: Chile; CO: Colombia; CU: Cuba; DE: Germany; EC: Ecuador; ES: Spain; FR: France; GB: Great Britain; GR: Greece; IE: Ireland; IT: Italy; MX: Mexico; NL: Netherlands; NZ: New Zealand; PA: Panama; PE: Peru; PL: Poland; PT: Portugal; PY: Paraguay; RU: Russia; TR: Turkey; US: United States; UY: Uruguay; VE: Venezuela.



Characterizing the Users in No-Vax Communities

Considering the behavior of the users in no-vax communities, we found that they were more likely to retweet (Figure 4A) and share URLs (Figure 4B), especially URLs to YouTube (Figure 4C), than the other users. Furthermore, the URLs they posted were much more likely to have been from low-credibility domains (Figure 4D) compared with those posted in the rest of the networks. The difference is remarkable: 26% of the domains shared by no-vax communities came from lists of known low-credibility domains versus only 2.4% of those cited by the other users came from lists of known low-credibility domains (*P*<.001). The most common low-credibility websites among the no-vax communities were *Zero Hedge*, *LifeSiteNews*, *Daily Mail* (considered right-biased and questionably sourced), and *Children's Health Defense* (conspiracy/pseudoscience). These findings extend the existing literature on English-language vaccination rhetoric to a multilingual, international scope by confirming the elevated social engagement in antivaccination communities [54] and provide additional evidence of the misleading nature of the popular COVID-19 pandemic–related YouTube videos [55].

Next, we investigated the effects of content moderation by Twitter on the vaccination debate. We found that the average proportion of suspended accounts in no-vax communities was much larger than that among the rest of the users for each country and period considered (average 13.3% vs 1.8%, *P*<.001; Figure 5A). The highest proportions of suspended accounts were found in the English-speaking countries, Germany, and the Netherlands, which also showed a larger presence of no-vax content, than in the other countries. Furthermore, a large portion of suspensions came after the January 2021 US Capitol attack in Washington, DC [56] (Figure 5B). The proportion of suspended accounts from the United States increased from 38% before January 1 to 77% during the days around the Washington riots (January 1-12). Note that (1) 89% of the US users who were suspended belonged to the no-vax community in the vaccine development period; (2) the account realDonaldTrump (suspended on January 8) was one of the most popular accounts among the no-vax communities of the first 3 periods; and (3) in the last period, a no-vax community was not present in the US RT network, indicating that the suspension of US accounts following the Washington riots heavily impacted the vaccination debate on Twitter. These findings suggest that political leaning is often associated with strong stances taken in the vaccination debate (in line with previous literature [17,21]) and that actions taken in the political domain may greatly impact the quality of the public health discourse.

Figure 4. Behavior of users in no-vax communities versus those of other users. (A) Average retweets, (B) average URLs, (C) average YouTube URLs, and (D) proportion of low-credibility domains shared by users. Note that low-credibility domains were collected only in Italian, French, English, and Greek; therefore, the plots refer to countries speaking these languages. No-Vax: discrediting vaccines. Other: non-discrediting vaccines.

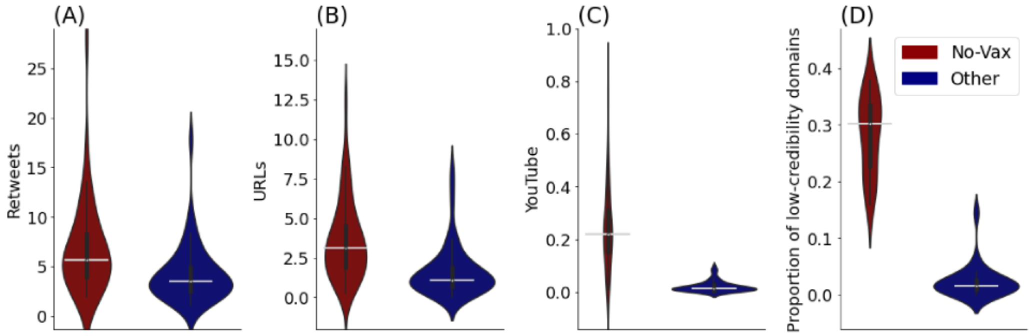
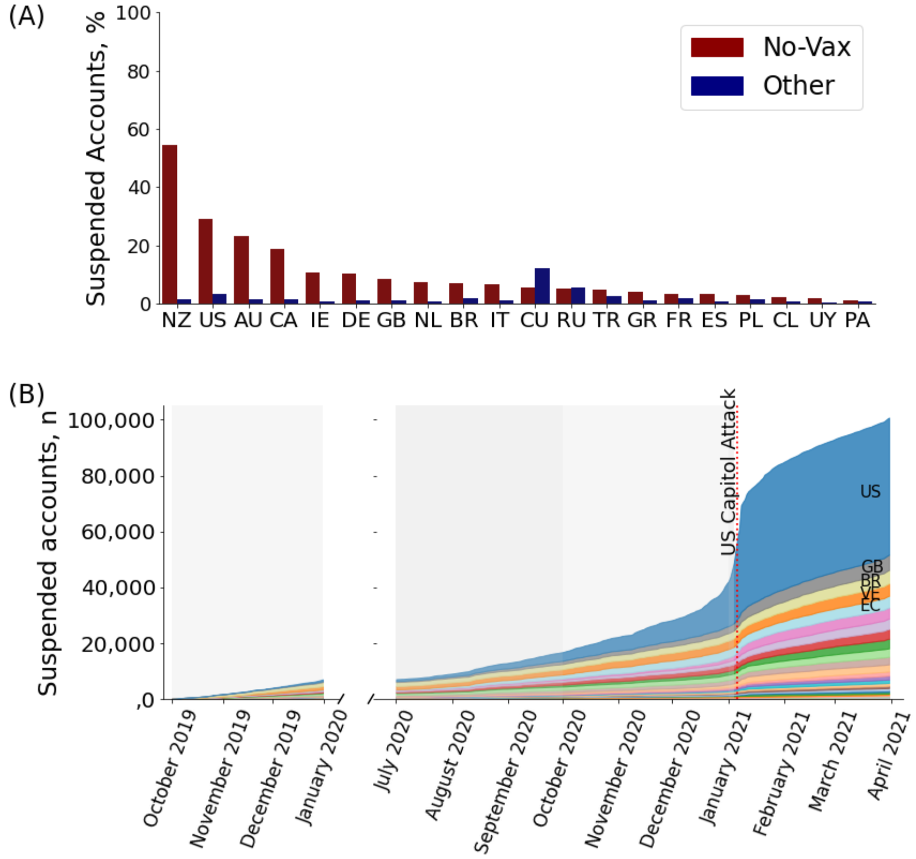


Figure 5. Suspended users per country in no-vax communities. (A) Average proportion of suspended accounts per country in the period in which no-vax community has been detected, computed separately for no-vax side and rest of users. (B) Number of suspended accounts as a function of the date they posted their last tweet, colored by country. No-Vax: discrediting vaccines. Other: non-discrediting vaccines. AR: Argentina; AU: Australia; BR: Brazil; CA: Canada; CL: Chile; CO: Colombia; CU: Cuba; DE: Germany; EC: Ecuador; ES: Spain; FR: France; GB: Great Britain; GR: Greece; IE: Ireland; IT: Italy; MX: Mexico; NL: Netherlands; NZ: New Zealand; PA: Panama; PE: Peru; PL: Poland; PT: Portugal; PY: Paraguay; RU: Russia; TR: Turkey; US: United States; UY: Uruguay; VE: Venezuela. No-Vax: plots refer to users in no-vax communities. Other: plots refer to users in other communities.



Cross-Border Information Spillover in the Global Vaccination Debate

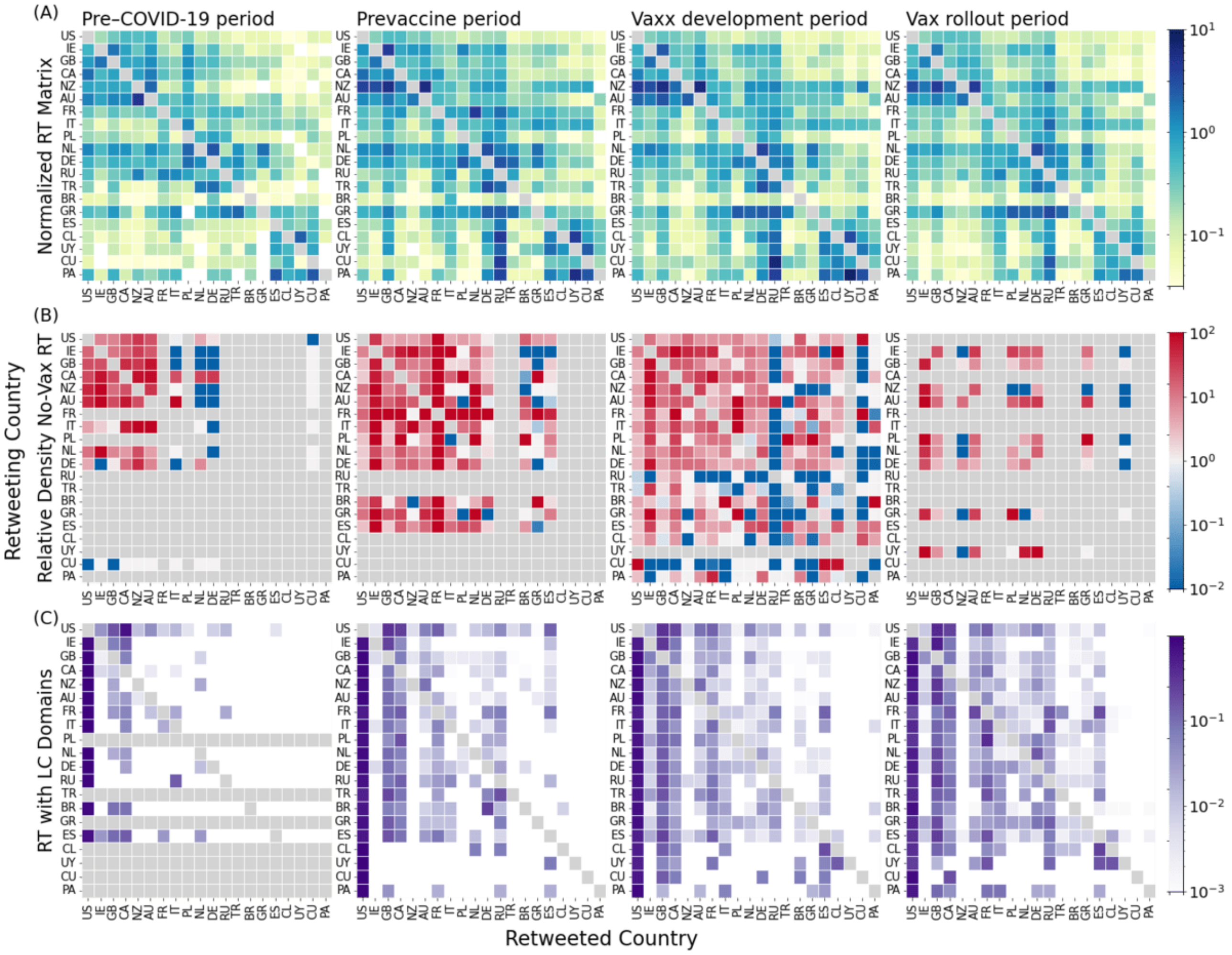
Next, we quantified the information spillover across countries by considering the number of RTs from one country to another, normalized by the total number of RTs produced and received in the 2 countries (see Normalized RT Volume in the Methods section; Figure 6A). First, one can observe language homophily, indicated by the darker regions in the top left (English) and bottom right (Spanish) of the panels, as well as the pair Portugal-Brazil, in all periods. The darker patches corresponding to the interactions between Germany and the Netherlands and those between Germany and Turkey also reflect possible cultural or expat relationships. Second, the cross-border interaction matrices are not symmetrical: information generally flows in a preferred direction. For instance, the Spanish-speaking countries retweeted the English-speaking countries much more than the English-speaking countries retweeted the Spanish-speaking countries. Note that the United States is central in the global information flow, being a net exporter of information to the rest of the world when comparing inflows versus outflows of information for each country (Figure S3 in Multimedia Appendix 1). Interestingly, from the prevaccine period, Russia also became a net exporter, especially to South American countries (Figure S3 in Multimedia Appendix 1): some of the most used hashtags in the prevaccine and vax development periods are *#sputnikesesperanza* and *#sputnikparaelpueblo*.

In Figure 6B, we quantified the strength of the cross-border interactions among the users in no-vax communities compared with that among the rest of the users (see Cross-Border Interactions Between No-Vax Communities in the Methods section). We found that cross-border interactions among the users in no-vax communities were generally much stronger, sometimes by orders of magnitude, than the interactions among the rest of the users, creating a tightly knit global no-vax network. In particular, the users in the no-vax communities of the English-speaking countries, Germany, and the Netherlands were tightly connected in all periods. By contrast, the users in the no-vax communities from Cuba and Russia were isolated (adding to their unusual user suspension statistics). Again, cross-border interactions can be asymmetrical: for instance, in the pre–COVID-19 pandemic period, the users in the no-vax communities in Germany and the Netherlands retweeted the users in the other countries, but not vice versa.

Finally, we focused on misinformation flows across countries by considering the fraction of low-credibility domains imported per country (Figure 6C), that is, the fraction of tweets pointing to low-credibility URLs, over the total number of RTs from one country to another. We stress that we considered flows of low-credibility information across borders spread by both humans and bots, without engaging in the difficult task of distinguishing them, as we were interested in quantifying how exposed a certain country A is to misinformation coming from country B. As in the previous case, the matrices show a clear asymmetry. The US users were responsible for exporting a large fraction of misinformation to the rest of the world: 68% of all the low-credibility URLs retweeted worldwide came from the United States (average over the 4 periods), a proportion much higher than the total volume of URLs (42%) retweeted from the United States.

Interestingly, the fraction of low-credibility URLs from the United States dropped from 74% in the vax development period to 55% in the vax rollout period. This large decrease can be directly ascribed to Twitter's moderation policy: 46% of cross-border RTs of US users linked to low-credibility websites in the vax development period came from accounts that were suspended following the US Capitol attack (Figure S5A in Multimedia Appendix 1). Note that Twitter's account purge substantially impacted the misinformation spread worldwide: the proportion of low-credibility domains in the URLs retweeted from the United States dropped from 14% to 7%. Finally, despite not having a list of low-credibility domains in Russian, Russia was central in exporting potential misinformation in the vax rollout period, especially to Latin American countries. In these countries, the proportion of low-credibility URLs coming from Russia increased from 1% in the vax development period to 18% in the vax rollout periods (Figure S5B in Multimedia Appendix 1).

Figure 6. Cross-border information flows in the global vaccination debate. (A) Normalized retweet (RT) matrix. Normalized number of retweets (excluding diagonal elements, colored in gray). (B) The ratio between probability of interactions between users in no-vax communities and rest of the networks from the same pair of countries (see the Methods section). Darker red (blue) elements of the matrices represent higher (lower) tendency of cross-border interactions between users in no-vax communities with respect to other users. Countries without no-vax communities are colored in gray. (C) Proportion of URLs that come from the retweeted country among the low-credibility (LC) domains imported from the retweeting country. Countries importing ≤10 low-credibility URLs are colored in gray. Diagonal elements are also gray, as our focus was on cross-country interactions. Element aij of each matrix represents the information flow from country j to country i. In the plots, only the countries with at least 1 no-vax community in the 4 periods are represented. The extended version is Figure S4 in Multimedia Appendix 1.



Discussion

Principal Findings

The international, multilingual nature of the data we present here supports the ongoing efforts in monitoring the non-English debate around the topic of vaccination [57-59]. Using this information, we reveal the increasingly globalized nature of the vaccination debate as the COVID-19 vaccines were proposed, developed, and deployed. This increased globalization had a marked impact on vaccine-hesitant discourse: not only did the prominence of the no-vax communities increase within individual countries but their cross-border connections also strengthened around the world. We showed that the users in these communities are much more prone to sharing potential misinformation than other users, even across national borders.

Furthermore, the real-time nature of the data collection allowed us to capture Twitter's content moderation efforts, which proved to be uneven both across countries and time. The users blocked immediately following the January 6 Washington riots were responsible for a substantial amount of misinformation spread—both within the United States and, crucially, internationally. Thus, we paint a picture of a "global no-vax Twitter network" that calls for the international collaboration of both public health and technology experts.

Limitations

Our study has several important limitations. First and foremost, it is well known that Twitter users are not a representative sample of the real population but are biased toward more educated, urban, younger, and male individuals [60]. Furthermore, Twitter use wildly differs among the countries under consideration, so cross-country comparisons should be taken with caution. With respect to this, the geolocation task also introduced some bias in the results due to the different fraction of missing accounts across countries. However, to the best of our knowledge, there are no means to collect real-time, representative data at this spatial and temporal scale. Note that we did not engage in bot detection, as this task is notoriously difficult [61], and, most importantly, misinformation can be spread by complex interactions by bots and humans [44]. Moreover, our study was limited to 11 chosen languages and to the 4 languages for which low-credibility domains were collected, a limitation necessary to control the cultural heterogeneity of the data analyzed. However, the fact that low-credibility domains in other languages were not found (for instance, it was challenging to find a reputable list of low- or high-credibility domains in Russian) means that the potential misinformation flows presented here are a lower bound—one which should be expanded using additional resources. Furthermore, the content considered for this study is limited to the keywords and the query processing of the Twitter search engine. Note that we did not check for spelling errors, which may lead to underdetection of some tweets. However, we did perform a relevance check on a random selection of 300 tweets, resulting in 7% of pet-related tweets, 6% of nonrelevant tweets, and the remaining tweets relevant to vaccination. As we did not modify the keywords as events unfolded—notably when new vaccines were developed—in the aim of keeping a consistent methodology allowing for comparison over time, some content pertaining to time-specific keywords was missed (although we were still able to capture a large amount of discussion around these developments with existing keywords).

Future work should also be devoted to including countries from Africa and Asia, as well as to update and extend the list of low-credibility information sources to other languages. For the latter task, one could leverage the identification of no-vax communities—more susceptible to share low-credibility information—proposed in this study. Other possible limitations of this study include the method used to identify no-vax communities, hierarchical clustering of the RT network and labeling of the popular tweets in the resulting communities, which may have been sensitive to the thresholds adopted. However, note that this method was not aimed at detecting the stance of single users about vaccination but at identifying large clusters of users exposed to a certain kind of antivaccine narrative.

Broader Impact

Despite the platform's tweet flagging and removal policies around COVID-19 [62,63], it is the bout of account suspensions around the Washington riots that made the most impact on the national and international spread of vaccine-related misinformation, suggesting that political concerns elicit much stronger curbing of the freedom of speech than health concerns. It is possible that the effects of this event changed the social media landscape itself, with platforms such as Truth Social appearing in the aftermath of the event. More documentation of the causal link between web-based misinformation and adverse health outcomes may provide a more solid ground for making censorship decisions for both the platforms and the politicians governing them. For instance, a randomized controlled trial in the United Kingdom and the United States showed that "relative to factual information, recent misinformation induced a decline in intent of 6.2 percentage points" [31].

Further, the Centers for Disease Control and Prevention and the Kaiser Family Foundation estimate that the lack of action early in the pandemic may have contributed to deaths of hundreds of thousands by June 2021 [2]. Furthermore, this study illustrates the impact of 1 social media platform's editorial policies on the international public health discourse, especially when the country involved is as culturally influential as the United States. Without examining in detail the content shared by the suspended accounts, we cannot be certain that the accounts indeed were sharing harmful content. Monitoring the censorship activities of major platforms (triggered by either internal policies or governments' requests) is important for assessing the users' freedom of speech. For instance, the Electronic Frontier Foundation has recently criticized social media platforms for blocking political dissidents who a decade ago used the same platforms to "push for political change and social justice" [64]. Fortunately, "de-platforming as censorship" is a topic of ongoing deliberation at the European Union's Internet Governance Forum involving civil society and government representatives [65].

An international perspective may also benefit the tracking of malicious actors, such as semiautomated or fully automated accounts, networks of colluding agents, and sources of poor-quality content. It has been shown that accounts identified as Russian trolls were more likely to tweet about vaccination before the pandemic [66]. During the pandemic, Russian trolls often posted misinformation concerning the personal dangers of vaccines, purported civil liberty violations, and vaccine conspiracies [67]. Since the beginning of the Ukraine war, it has been noted that antivaccine content has diminished dramatically, potentially because of the additional blocking of Twitter in Russia and refocusing of the conspiratorial attention on Ukraine [68]. Our findings suggest that changes in governance and censorship may encourage or discourage the flow of potential misinformation from states with known affinities.

Acknowledgments

The authors acknowledge the support from the Lagrange Project of the Institute for Scientific Interchange Foundation funded by Fondazione Cassa di Risparmio di Torino. This study also received funding from the European Union's Horizon 2020 Research and Innovation Program—project EpiPose (grant) and project PANDEM-2 (grant). This study reflects only the authors' views. The European Commission is not responsible for any use that may be made of the information this study contains. Most of this study was conducted while MS and AP were at the Institute for Scientific Interchange Foundation.

Data Availability

All codes generated in the project have been deposited to GitHub [69]. All data collected in the project have been deposited to Zenodo [70].

Authors' Contributions

All authors contributed to the development of the main ideas of this study. JL analyzed the data and prepared the figures. MT contributed to the data collection. YM and MS led the interpretation of the results and preparation of the manuscript with contributions from all the authors. All the authors critically reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional figures for the data selection, annotation and analysis.
[DOCX File , 1124 KB - infodemiology_v3i1e44714_app1.docx]

References

1. Ten threats to global health in 2019. World Health Organization. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2023-04-27]
2. Amin J, Ortaliza J, Cox C, Michaud J, Kates J. COVID-19 mortality preventable by vaccines. Health System Tracker. 2022 Apr 21. URL: https://www.healthsystemtracker.org/brief/covid19-and-other-leading-causes-of-death-in-the-us/ [accessed 2023-04-27]
3. Browne M, Thomson P, Rockloff MJ, Pennycook G. Going against the herd: psychological and cultural factors underlying the 'Vaccination Confidence Gap'. PLoS One 2015 Sep 1;10(9):e0132562 [FREE Full text] [doi: 10.1371/journal.pone.0132562] [Medline: 26325522]
4. Kalimeri K, Beiró MG, Urbinati A, Bonanomi A, Rosina A, Cattuto C. Human values and attitudes towards vaccination in social media. In: Companion Proceedings of The 2019 World Wide Web Conference. 2019 May Presented at: WWW '19; May 13-17, 2019; San Francisco, CA, USA p. 248-254. [doi: 10.1145/3308560.3316489]
5. Jolley D, Douglas KM. The effects of anti-vaccine conspiracy theories on vaccination intentions. PLoS One 2014 Feb 20;9(2):e89177 [FREE Full text] [doi: 10.1371/journal.pone.0089177] [Medline: 24586574]
6. Burki T. Vaccine misinformation and social media. The Lancet Digit Health 2019 Oct;1(6):e258-e259. [doi: 10.1016/s2589-7500(19)30136-0]
7. Carrieri V, Madio L, Principe F. Vaccine hesitancy and (fake) news: quasi-experimental evidence from Italy. Health Econ 2019 Nov;28(11):1377-1382 [FREE Full text] [doi: 10.1002/hec.3937] [Medline: 31429153]
8. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M. The echo chamber effect on social media. Proc Natl Acad Sci U S A 2021 Mar 02;118(9):e2023301118 [FREE Full text] [doi: 10.1073/pnas.2023301118] [Medline: 33622786]
9. Del Vicario M, Vivaldo G, Bessi A, Zollo F, Scala A, Caldarelli G, et al. Echo chambers: emotional contagion and group polarization on Facebook. Sci Rep 2016 Dec 01;6:37825 [FREE Full text] [doi: 10.1038/srep37825] [Medline: 27905402]
10. Garimella K, De Francisci Morales G, Gionis A, Mathioudakis M. Political discourse on social media: echo chambers, gatekeepers, and the price of bipartisanship. In: Proceedings of the 27th World Wide Web Conference. 2018 Presented at: WWW '18; April 23-27, 2018; Lyon, France. [doi: 10.1145/3178876.3186139]
11. Zollo F, Novak PK, Del Vicario M, Bessi A, Mozetič I, Scala A, et al. Emotional dynamics in the age of misinformation. PLoS One 2015 Sep 30;10(9):e0138740 [FREE Full text] [doi: 10.1371/journal.pone.0138740] [Medline: 26422473]
12. Bakshy E, Messing S, Adamic LA. Political science. Exposure to ideologically diverse news and opinion on Facebook. Science 2015 Jun 05;348(6239):1130-1132. [doi: 10.1126/science.aaa1160] [Medline: 25953820]
13. Cinelli M, Brugnoli E, Schmidt AL, Zollo F, Quattrociocchi W, Scala A. Selective exposure shapes the Facebook news diet. PLoS One 2020 Mar 13;15(3):e0229129 [FREE Full text] [doi: 10.1371/journal.pone.0229129] [Medline: 32168347]
14. Quattrociocchi W, Scala A, Sunstein CR. Echo chambers on Facebook. SSRN J 2016 Jun 15. [doi: 10.2139/ssrn.2795110]
15. Schmidt AL, Zollo F, Del Vicario M, Bessi A, Scala A, Caldarelli G, et al. Anatomy of news consumption on Facebook. Proc Natl Acad Sci U S A 2017 Mar 21;114(12):3035-3039 [FREE Full text] [doi: 10.1073/pnas.1617052114] [Medline: 28265082]
16. 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]
17. Cossard A, De Francisci Morales G, Kalimeri K, Mejova Y, Paolotti D, Starnini M. Falling into the echo chamber: the Italian vaccination debate on Twitter. Proc Int AAAI Conf Web Soc Media 2020 May 26;14(1):130-140. [doi: 10.1609/icwsm.v14i1.7285]

18. Crupi G, Mejova Y, Tizzani M, Paolotti D, Panisson A. Echoes through time: evolution of the Italian COVID-19 vaccination debate. Proc Int AAAI Conf Web Soc Media 2022 May 31;16(1):102-113. [doi: 10.1609/icwsm.v16i1.19276]
19. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. Polarization of the vaccination debate on Facebook. Vaccine 2018 Jun 14;36(25):3606-3612. [doi: 10.1016/j.vaccine.2018.05.040] [Medline: 29773322]
20. Faccin M, Gargiulo F, Atlani-Duault L, Ward JK. Assessing the influence of French vaccine critics during the two first years of the COVID-19 pandemic. PLoS One 2022 Aug 4;17(8):e0271157 [FREE Full text] [doi: 10.1371/journal.pone.0271157] [Medline: 35925962]
21. 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 May 01;53(2):335-342. [doi: 10.1017/s0008423920000396]
22. Ng R, Chow TY, Yang W. News media narratives of COVID-19 across 20 countries: early global convergence and later regional divergence. PLoS One 2021 Sep 1;16(9):e0256358 [FREE Full text] [doi: 10.1371/journal.pone.0256358] [Medline: 34469446]
23. Rovetta A, Bhagavathula AS. Global infodemiology of COVID-19: analysis of Google web searches and Instagram hashtags. J Med Internet Res 2020 Aug 25;22(8):e20673 [FREE Full text] [doi: 10.2196/20673] [Medline: 32748790]
24. Managing the COVID-19 infodemic: promoting healthy behaviours and mitigating the harm from misinformation and disinformation: joint statement by WHO, UN, UNICEF, UNDP, UNESCO, UNAIDS, ITU, UN Global Pulse, and IFRC. World Health Organization. 2020 Sep 23. 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 2023-04-27]
25. Blankenship EB, Goff ME, Yin J, Tse ZT, Fu KW, Liang H, et al. Sentiment, contents, and retweets: a study of two vaccine-related twitter datasets. Perm J 2018;22:17-138 [FREE Full text] [doi: 10.7812/TPP/17-138] [Medline: 29911966]
26. Kennedy A, Lavail K, Nowak G, Basket M, Landry S. Confidence about vaccines in the United States: understanding parents' perceptions. Health Aff (Millwood) 2011 Jun;30(6):1151-1159. [doi: 10.1377/hlthaff.2011.0396] [Medline: 21653969]
27. Meadows CZ, Tang L, Liu W. Twitter message types, health beliefs, and vaccine attitudes during the 2015 measles outbreak in California. Am J Infect Control 2019 Nov;47(11):1314-1318. [doi: 10.1016/j.ajic.2019.05.007] [Medline: 31266661]
28. Lutkenhaus RO, Jansz J, Bouman MP. Mapping the dutch vaccination debate on Twitter: identifying communities, narratives, and interactions. Vaccine X 2019 Apr 11;1:100019 [FREE Full text] [doi: 10.1016/j.jvacx.2019.100019] [Medline: 31485580]
29. Alatas V, Chandrasekhar AG, Mobius M, Olken BA, Paladines C. When celebrities speak: a nationwide twitter experiment promoting vaccination in Indonesia. National Bureau of Economic Research. 2019 Feb. URL: https://www.nber.org/papers/w25589 [accessed 2023-04-27]
30. Bridgman A, Merkley E, Zhilin O, Loewen PJ, Owen T, Ruths D. Infodemic pathways: evaluating the role that traditional and social media play in cross-national information transfer. Front Polit Sci 2021 Mar 29;3:20. [doi: 10.3389/fpos.2021.648646]
31. 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]
32. Thaker J, Subramanian A. Exposure to COVID-19 vaccine hesitancy is as impactful as vaccine misinformation in inducing a decline in vaccination intentions in new zealand: Results from pre-post between-groups randomized block experiment. Front Commun 2021 Aug 19;6:721982. [doi: 10.3389/fcomm.2021.721982]
33. Teng X, Lin Y, Chung W, Li A, Kovashka A. Characterizing user susceptibility to COVID-19 misinformation on Twitter. Proc Int AAAI Conf Web Soc Media 2022 May 31;16(1):1005-1016. [doi: 10.1609/icwsm.v16i1.19353]
34. Weinzierl M, Hopfer S, Harabagiu SM. Misinformation adoption or rejection in the era of COVID-19. Proc Int AAAI Conf Web Soc Media 2021 May 22;15(1):787-795. [doi: 10.1609/icwsm.v15i1.18103]
35. 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 Nov 17;7(11):e30642 [FREE Full text] [doi: 10.2196/30642] [Medline: 34653016]
36. Yang KC, Pierri F, Hui PM, Axelrod D, Torres-Lugo CT, Bryden J, et al. The COVID-19 infodemic: Twitter versus Facebook. Big Data Soc 2021 May 05;8(1):205395172110138. [doi: 10.1177/20539517211013861]
37. Cota W, Ferreira SC, Pastor-Satorras R, Starnini M. Quantifying echo chamber effects in information spreading over political communication networks. EPJ Data Sci 2019 Dec 09;8(1):35. [doi: 10.1140/epjds/s13688-019-0213-9]
38. Garimella K, De Francisci Morales G, Gionis A, Mathioudakis M. Quantifying controversy on social media. ACM Trans Comput Syst 2018;1(1):3. [doi: 10.1145/3140565]
39. GeoNames. URL: http://www.geonames.org/ [accessed 2023-04-27]
40. Vaccine multilingual queries. URL: https://tinyurl.com/344rbjcx [accessed 2023-04-27]
41. Stream Tweets in real-time. Twitter Developer Platform. URL: https://developer.twitter.com/en/docs/tutorials/stream-tweets-in-real-time [accessed 2023-04-27]
42. Choosing a historical API. Twitter Developer Platform. URL: https://developer.twitter.com/en/docs/tutorials/choosing-historical-api [accessed 2023-04-27]

43. Follow, search, and get users. Twitter Developer Platform. URL: https://developer.twitter.com/en/docs/twitter-api/v1/accounts-and-users/follow-search-get-users/api-reference/get-users-show [accessed 2023-04-27]
44. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat Hum Behav 2020 Dec;4(12):1285-1293. [doi: 10.1038/s41562-020-00994-6] [Medline: 33122812]
45. Pierri F. The diffusion of mainstream and disinformation news on Twitter: the case of Italy and France. In: Companion Proceedings of the 2020 Web Conference. 2020 Presented at: WWW '20; April 20-24, 2020; Taipei, Taiwan p. 617-622. [doi: 10.1145/3366424.3385776]
46. The Black List. Bufale. URL: https://www.bufale.net/the-black-list-la-lista-nera-del-web/ [accessed 2023-04-27]
47. List of fake news websites. Wikipedia - The Free Encyclopedia. URL: https://en.wikipedia.org/wiki/List_of_fake_news_websites [accessed 2023-04-27]
48. Media Bias / Fact Check. URL: https://mediabiasfactcheck.com/ [accessed 2023-04-27]
49. Sénécat A. Les fausses informations circulent de moins en moins sur Facebook. Le Monde. 2018 Oct 17. URL: https://www.lemonde.fr/les-decodeurs/article/2018/10/17/les-fausses-informations-perdent-du-terrain-sur-facebook_5370461_4355770.html [accessed 2023-04-27]
50. Oikonomidis K. List of Greek websites that have produced Fake News. dwrean.net. 2018 Oct 12. URL: https://www.dwrean.net/2018/12/fake-news-greek-sites.html [accessed 2023-04-27]
51. Bonald T, Charpentier B, Galland A, Hollocou A. Hierarchical graph clustering using node pair sampling. arXiv Preprint posted online June 5, 2018. [doi: 10.48550/arXiv.1806.01664]
52. Garimella K, De Francisci Morales G, Gionis A, Mathioudakis M. Quantifying controversy in social media. In: Proceedings of the 9th ACM International Conference on Web Search and Data Mining. 2016 Feb Presented at: WSDM '16; February 22-25, 2016; San Francisco, CA, USA p. 33-42. [doi: 10.1145/2835776.2835792]
53. Danon L, Díaz-Guilera A, Duch J, Arenas A. Comparing community structure identification. J Stat Mech 2005;2005(9):P09008. [doi: 10.1088/1742-5468/2005/09/p09008]
54. Germani F, Biller-Andorno N. The anti-vaccination infodemic on social media: a behavioral analysis. PLoS One 2021 Mar 3;16(3):e0247642. Retracted in: PLoS One. December 22, 2022. doi: 10.1371/journal.pone.02797964 [FREE Full text] [doi: 10.1371/journal.pone.0247642] [Medline: 33657152]
55. 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 Jun 15;12(6):e8622 [FREE Full text] [doi: 10.7759/cureus.8622] [Medline: 32685293]
56. Twitter Safety. An update following the riots in Washington, DC. Twitter Blog. 2021 Jan 12. URL: https://blog.twitter.com/en_us/topics/company/2021/protecting--the-conversation-following-the-riots-in-washington-- [accessed 2023-04-27]
57. Di Giovanni M, Pierri F, Torres-Lugo C, Brambilla M. VaccinEU: COVID-19 vaccine conversations on Twitter in French, German and Italian. Proc Int AAAI Conf Web Soc Media 2022 May 31;16(1):1236-1244. [doi: 10.1609/icwsm.v16i1.19374]
58. Haouari F, Hasanain M, Suwaileh R, Elsayed T. ArCOV-19: the first Arabic COVID-19 Twitter dataset with propagation networks. In: Proceedings of the 6th Arabic Natural Language Processing Workshop. 2021 Presented at: WANLP '21; April 19, 2021; Kyiv, Ukraine p. 82-91 URL: https://aclanthology.org/2021.wanlp-1.9.pdf
59. Pierri F, Tocchetti A, Corti L, Di Giovanni M, Pavanetto S, Brambilla M, et al. VaccinItaly: monitoring Italian conversations around vaccines on Twitter and Facebook. arXiv Preprint posted online January 11, 2021. [doi: 10.48550/arXiv.2101.03757]
60. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021 Apr 7. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2023-04-27]
61. Martini F, Samula P, Keller TR, Klinger U. Bot, or not? Comparing three methods for detecting social bots in five political discourses. Big Data Soc 2021 Jul;8(2):205395172110335. [doi: 10.1177/20539517211033566]
62. COVID-19 Misinformation. Twitter Transparency. URL: https://transparency.twitter.com/en/reports/covid19.html#2021-jul-dec [accessed 2023-04-27]
63. Crisis misinformation policy. Twitter Help Center. URL: https://help.twitter.com/en/rules-and-policies/crisis-misinformation [accessed 2023-04-27]
64. Kelley J. A decade after the Arab Spring, platforms have turned their backs on critical voices in the Middle East and North Africa. Electronic Frontier Foundation. 2020 Dec 17. URL: https://www.eff.org/deeplinks/2020/12/decade-after-arab-spring-platforms-have-turned-their-backs-critical-voices-middle [accessed 2023-04-27]
65. De-platforming as Censorship Means in the Digital Era. Internet Governance Forum. URL: https://www.intgovforum.org/en/content/igf-2022-ws-52-de-platforming-as-censorship-means-in-the-digital-era [accessed 2023-04-27]
66. Broniatowski DA, Jamison AM, Qi A, AlKulaib L, Chen T, Benton A, et al. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. Am J Public Health 2018 Oct;108(10):1378-1384. [doi: 10.2105/AJPH.2018.304567] [Medline: 30138075]
67. Warner EL, Barbati JL, Duncan KL, Yan K, Rains SA. Vaccine misinformation types and properties in Russian troll tweets. Vaccine 2022 Feb 07;40(6):953-960. [doi: 10.1016/j.vaccine.2021.12.040] [Medline: 35034832]
68. Schreiber M. ‘Bot holiday’: COVID disinformation down as social media pivot to Ukraine. The Guardian. 2022 Mar 4. URL: https://www.theguardian.com/media/2022/mar/04/bot-holiday-covid-misinformation-ukraine-social-media [accessed 2023-04-27]

69. Lenti J. global-vax-misinfo. GitHub. URL: https://github.com/jaclenti/global-vax-misinfo [accessed 2023-05-09]
70. Lenti J. Global misinformation spillovers in the online vaccination debate before and during COVID-19. Zenodo. 2023 Mar 10. URL: https://zenodo.org/record/7716817#.ZFoj--xBxpR [accessed 2023-05-09]

Abbreviations

API: application programming interface
CO: cosharing
NMI: normalized mutual information
OSM: online social media
RT: retweet
RWC: Random Walk Controversy

Edited by T Mackey; submitted 30.11.22; peer-reviewed by KC Yang, V von Wyl; comments to author 25.02.23; revised version received 15.03.23; accepted 27.03.23; published 24.05.23.

Please cite as:
Lenti J, Mejova Y, Kalimeri K, Panisson A, Paolotti D, Tizzani M, Starnini M
Global Misinformation Spillovers in the Vaccination Debate Before and During the COVID-19 Pandemic: Multilingual Twitter Study
JMIR Infodemiology 2023;3:e44714
URL: https://infodemiology.jmir.org/2023/1/e44714
doi:10.2196/44714
PMID:37223965

Original Paper

Establishing Infodemic Management in Germany: A Framework for Social Listening and Integrated Analysis to Report Infodemic Insights at the National Public Health Institute

T Sonia Boender[1], MSc, PhD; Paula Helene Schneider[1], BSc; Claudia Houareau[2], MSc; Silvan Wehrli[3], MSc; Tina D Purnat[4], MSc; Atsuyoshi Ishizumi[4], MSc, MPH; Elisabeth Wilhelm[5], MA; Christopher Voegeli[6], MPH, PhD; Lothar H Wieler[7,8], DVM, Prof Dr, Dr hc mult; Christina Leuker[1], MSc, Dr rer nat

[1]Risk Communication Unit, Robert Koch Institute, Berlin, Germany
[2]Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany
[3]Centre for Artificial Intelligence in Public Health Research, Robert Koch Institute, Berlin, Germany
[4]Health Emergencies Programme, Department of Pandemic and Epidemic Preparedness and Prevention, World Health Organization, Geneva, Switzerland
[5]School of Public Health Information Futures Lab, Brown University, Providence, RI, United States
[6]Centers for Disease Control and Prevention, Atlanta, Georgia
[7]Robert Koch Institute, Berlin, Germany
[8]Digital Global Public Health, Hasso Plattner Institute, University of Potsdam, Potsdam, Germany

Corresponding Author:
T Sonia Boender, MSc, PhD
Risk Communication Unit
Robert Koch Institute
Nordufer 20
Berlin, 13353
Germany
Phone: 49 030 187540
Email: p1@rki.de

Abstract

Background: To respond to the need to establish infodemic management functions at the national public health institute in Germany (Robert Koch Institute, RKI), we explored and assessed available data sources, developed a social listening and integrated analysis framework, and defined when infodemic management functions should be activated during emergencies.

Objective: We aimed to establish a framework for social listening and integrated analysis for public health in the German context using international examples and technical guidance documents for infodemic management.

Methods: This study completed the following objectives: identified (potentially) available data sources for social listening and integrated analysis; assessed these data sources for their suitability and usefulness for integrated analysis in addition to an assessment of their risk using the RKI's standardized data protection requirements; developed a framework and workflow to combine social listening and integrated analysis to report back actionable infodemic insights for public health communications by the RKI and stakeholders; and defined criteria for activating integrated analysis structures in the context of a specific health event or health emergency.

Results: We included and classified 38% (16/42) of the identified and assessed data sources for social listening and integrated analysis at the RKI into 3 categories: social media and web-based listening data, RKI-specific data, and infodemic insights. Most data sources can be analyzed weekly to detect current trends and narratives and to inform a timely response by reporting insights that include a risk assessment and scalar judgments of different narratives and themes.

Conclusions: This study identified, assessed, and prioritized a wide range of data sources for social listening and integrated analysis to report actionable infodemic insights, ensuring a valuable first step in establishing and operationalizing infodemic management at the RKI. This case study also serves as a roadmap for others. Ultimately, once operational, these activities will inform better and targeted public health communication at the RKI and beyond.

(JMIR Infodemiology 2023;3:e43646) doi:10.2196/43646

KEYWORDS

infodemic; social listening; communication; infodemiology; public health; health promotion; misinformation; integrated analysis; infodemic insights

Introduction

The Infodemic

Over the last few decades, our information ecosystem has undergone changes and shifts, where the general public moved away from traditional media and institutions as a primary source of health information to a more decentralized model with many different sources of information [1]. Different groups and generations have their own networks, information sources, and ways of interacting and sharing information in a digitally connected, increasingly polarized world [2]. The COVID-19 pandemic has made these trends increasingly clear and has been accompanied by an *infodemic*—too much information, including false or misleading information in digital and physical environments during a health emergency [3]. The lack of agreement between different information sources, as well as different levels of trust in different sources by different people, can cause uncertainty in the general population and impact the effectiveness of risk communication. There is more room for misinformation and disinformation to spread, for trust in public policy and political actions to be undermined [3], and for public health measures to be jeopardized [4]. Initial studies evaluating the COVID-19 pandemic response have acknowledged the need for greater investment in risk communication and community engagement strategies to foster trust in public health guidance and ultimately improve adherence to public health guidance and health decision-making [5-8]. Consistent with these findings, trust in institutions has been strongly linked to responsiveness and reliability in delivering policies and services [9]. Therefore, one of the key recommendations of the Organisation for Economic Co-operation and Development (OECD) is to connect and engage better with citizens in policy design, delivery, and reform, and ensure the inclusion of people at a higher risk of negative health impacts from infodemics. Infodemic management aims to achieve this through social listening and community engagement, as well as targeted public health messaging.

Infodemic Management

One way to support people in making informed health decisions is to provide responsive, evidence-based, and target-group-specific risk and health communication [10]. These communications must correspond to people's concerns and questions. Moreover, people need to be equipped with the right tools to find reliable information, identify misinformation [11], and assess the quality of (scientific) evidence. Well-planned and executed infodemic management can help develop the right messages for the right target groups at the right time as well as boost people's health and scientific literacy [12]. Although the terms *infodemiology*, *infodemic*, and *infoveillance* have existed for a long time [13,14], the field of infodemic management and its line of research have now been formally acknowledged by public health organizations as a novel, emerging scientific field and a critical area of practice during a pandemic [15,16]. Responding to *narratives about* the virus requires an approach similar to responding to the spread of the virus itself, including (early) detection, diagnosis, and identification of appropriate responses and interventions [17]. As both should occur early and in parallel, the European Centre for Disease Control and Prevention (ECDC) updated its guidance document in 2022 by adding infodemiology and infodemic management to the core competencies in applied infectious disease epidemiology [18]. Moreover, in 2023, the World Health Organization (WHO) is convening panels to develop WHO guidance and ethical considerations for social listening and integrated analysis, as well as WHO guidance on social listening and integrated analysis for public health, with an application to acute respiratory disease.

Social Listening and Integrated Analysis

Risk communication and community engagement are crucial elements of pandemic response [7]. Effective communication starts with listening; therefore, social listening is an essential tool in the infodemic management toolkit. Social listening is defined as monitoring the understanding, questions, concerns, information voids, narratives, misinformation, and disinformation that circulate in both web-based and offline environments (Textbox 1) [19,20]. Although it is a common practice for businesses to engage in digital marketing and monitor social media channels for mentions of their brand, competitors, or products [21], social media monitoring is just starting to find its way into the public sector. The increase in web-based communications has, in combination with computational power and artificial intelligence (AI), enabled real-time social listening, as implemented with the pilot Early AI-supported Response with Social Listening platform that tracks web-based COVID-19 conversations [22,23]. In addition to monitoring web-based conversations, offline social listening (including traditional media and other sources such as user search trends, epidemiological data, and socio-behavioral data) can be used to understand ongoing narratives at the population level [24].

Integrated analysis extends web-based social listening by considering data sources beyond social media (Textbox 1). These include news articles, Google searches, primary research, community dipstick surveys, citizen questions posed via hotlines, monitoring or surveillance reports, epidemiological and behavioral data, surveys and polls, and many more. Any data source that can provide insight into behaviors, questions, concerns, information voids, circulating narratives, misinformation, and disinformation (Textbox 1) within a given population for a given public health event was eligible. In the integrated analysis, different data sources were combined to identify themes and narratives across data sources. One advantage of integrated analysis is that it is less biased toward social media users and includes more diverse population groups. Another advantage is that a specific theme's importance may be judged more easily through triangulation (eg, if the same theme comes across many different sources). The scope of an integrated analysis can be varied based on current challenges and goals and available resources, for example, one could monitor and assess narratives around COVID-19 and monkeypox (mpox) as a whole (WHO Infodemic Insights reports for COVID-19 [23,24] and mpox [25]) or focus on vaccines and vaccine confidence (US Centers for Disease Control and Prevention [US CDC] State of Vaccine Confidence Insights Reports for COVID-19 [26] and mpox [27]).

Identifying and understanding the information voids, narratives, and sentiments behind conversations regarding public health issues through social listening and integrated analysis can help design adapted and targeted risk communication messages. These risk communication messages can have several aims: to prevent the circulation of misinformation by prebunking anticipated misinformation narratives, or responding to them if necessary, counteract stigma against affected groups [28,29], fill information voids, promote resilience, or contribute to behavioral change. Social listening and integrated analysis also have the power to identify research gaps and programmatic bottlenecks that the public perceives as a challenge (including access barriers), as well as guidance that confuses people or could potentially erode trust. These infodemic insights can point out confusion where the health authority is experiencing communication failures with the public, and what policy or programmatic levers can be used to address it beyond risk communication activities.

Textbox 1. Infodemic management terminology used in this work.

- Infodemic: an overabundance of information—some accurate and some not—that occurs during an epidemic [3,15].
- Infodemic management: the systematic use of risk- and evidence-based analysis and approaches to manage the infodemic and reduce its impact on health behaviors during health emergencies. Infodemic management aims to enable good health practices through four types of activities: (1) listening to community concerns and questions; (2) promoting understanding of risk and health expert advice; (3) building resilience to misinformation; and (4) engaging and empowering communities to take positive action [30].
- Infodemiology: the epidemiology of information; describing and analyzing information and communication patterns and their relationship to population health status [13].
- Infodemic insights: findings or conclusions from a data source (report) that has its own analysis plan that is tailored to the data type, source, and context of where the data are collected and the population it covers [24]. It is used to make recommendations for action, for more effective engagement.
- Infoveillance: using infodemiology data for surveillance [13,14].
- Social listening or infodemic surveillance [17], sometimes used as a synonym for “infoveillance”: Monitoring different web-based data (e social media) and offline data (traditional media and other sources such as user search trends, epidemiological data, and socio-behavioral data) sources to understand population understanding, perceptions, concerns and questions, information voids, narratives, misinformation and disinformation, and other relevant information about people’s reactions to a health topic [24].
- Integrated analysis: a planned methodological examination of different types of data sources that combine social listening intelligence with other types of information (eg, health seeking behavior, health service use, epidemiology, fact-checking and information seeking trends, and mobility reports) to produce infodemic insights [24].
- (Infodemic) insights report: a reporting output of integrated analysis that contextualizes findings from social listening and other data sources for use by health authorities to act based on a planned methodological frame for prioritization of actions and interventions. Important elements to include are a diagnosis of barriers and facilitators to desired behavior and how possible recommended actions support desired public health behaviors, which may be internal to the health system and externally facing strategies [24].
- Misinformation: false information, regardless of the intent to mislead [11].
- Disinformation: misinformation that is deliberately disseminated to mislead [11].

The German Context

In Germany, the Federal Centre for Health Education (*Bundeszentrale für gesundheitliche Aufklärung* [*BZgA*]) [27] is tasked with health education and promotion focused on the public. During the COVID-19 pandemic, an increasing number of citizens have also turned to the national public health institute—the Robert Koch Institute (RKI)—as well as to communications by the Federal Ministry of Health for behavioral advice and information on the pandemic. One indicator of this is the number of daily visitors on the RKI website, which has increased from ~30,000 in early February 2020 to an average of ~250,000 to 350,000 visits per day since the end of May 2020. The number of visits peaked on March 16, 2020, with 1,685,000 visits. The RKI’s follower count on Twitter (@rki_de) has increased from 12,000 (January 1, 2020) to 600,000 as of the time of writing (October 2022).

At the end of 2021, the German chancellor convened a scientific expert council of 19 members from different disciplines to develop evidence-based proposals to help curb the spread of the virus and tackle the pandemic [31]. In their fifth statement (January 20, 2022), the council unanimously called for the implementation of coordinated risk and health communication practices [4], which are also consistent with key infodemic management principles: (1) *generating* the best available knowledge to date (eg, through monitoring media and the extent to which the public takes up health-relevant behaviors), (2) *translating* relevant data, statistics, and indicators into behaviorally relevant advice for different target groups (Who is reached via which medium and format? How does information complexity need to be adapted?) and countering misinformation and disinformation; (3) *disseminating* communications via multiple channels, making use of web-based and offline media, influencers (by providing them with adequate materials), eHealth offers (such as web-based consultations) and collaborating with science communicators; and (4) *evaluating* the aforementioned measures and using the results for continuous quality improvement. As the council says, "in a decentralized and pluralistic society such as Germany, there will always be diverse actors that communicate and inform the general public" [4]. In Germany, these actors include political actors (eg, the Federal Ministry of Health), public health institutes (the RKI at the national level and public health institutes in different federal states and federal districts), federal institutes tasked with public communication (BZgA), a diverse range of web-based and offline media, and individuals (eg, influencers, individual scientists, journalists, politicians, and science communicators). In such an environment, it is particularly important to establish infrastructure for coordinated, professional, and evidence-based health communication. The expert council called for setting up such infrastructure quickly and in a sustainable manner to be better prepared for future crises.

Aim and Research Questions

At the RKI, much expertise and information are available for social listening and integrated analysis, but they are not fully leveraged to inform risk and health communication. Developing and testing structures to manage the infodemic is in line with the RKI's strategy and research agenda for the year 2025 [32,33], according to which the institute seeks to develop evidence-based methods for communicating with specific target audiences. To respond to the need for infodemic management in Germany, and specifically at the RKI, in this work, we review and explore opportunities for social listening and integrated analysis to enhance preparedness for future health crises [34,35]. This work focuses on two research questions: At the national public health institute for Germany, the RKI, (1) how can we establish response structures for social listening and integrated analysis? (2) what are the criteria under which these social listening and integrated analyses should be conducted to produce infodemic insights, and the accompanying response structures should be activated? Our case study also aims to serve as a road map for other institutes, within and outside Germany, to follow.

Methods

Approach, Aim, and Objectives

On the basis of a desk review, we gathered the available technical guidance [36] and training documentation [28,37-39] on infodemic management, as well as international examples of social listening and insight reporting [24,26,27]. We verified these sources and received technical assistance from our partners, the WHO and the US CDC. We aligned our aim and objectives with the WHO's public health research agenda for infodemic management [40] stream 1 "Measure and monitor the impact of infodemics during health emergencies." We aimed to establish a framework for social listening and integrated analysis of public health in the German context. The framework can, in turn, serve as a road map for others to establish infodemic management at other institutions within and outside of Germany. Our key objectives were (1) to identify (potentially) available data sources for social listening and integrated analysis at the RKI, (2) to assess these data sources for their suitability and usefulness for integrated analysis at the RKI, (3) to develop a framework and workflow to combine social listening and integrated analysis, to report back actionable infodemic insights for public health communications by the RKI and stakeholders, and (4) to define criteria for activating infodemic insight reporting in the context of a specific health event or health emergency. The reader should note that the actual insights reporting is outside of the scope of this work.

Data Sources and (Automation of) Data Extraction

To identify all potential data sources and tools used for social listening in the context of public health, with relevance for Germany, we reviewed the identified documentation (technical documentation [41], [World Health Organization, unpublished data, November 2022], guidance documents, infodemic training materials [28,37-39]), and the methodology of insight reports [24,26,27]. The review team consisted of a health scientist and field epidemiologist (TSB) and a behavioral scientist (CL), both with training in infodemic management [38,42]; a psychology student assistant (PS) who completed the OpenWHO Infodemic Management 101 training [39]; and a data scientist (SW) with specific expertise in web-based social listening using machine learning techniques, including natural language processing (NLP).

First, we reviewed web-based and social media listening tools and analytics, as well as the available social media data sources through application programming interfaces (APIs). We identified the largest social media platforms in Germany for web-based listening based on studies on media consumption [1,43], and identified the respective tools and analytics available for web-based social listening for these platforms. In addition, we reviewed the available APIs of both data aggregators and specific social media platforms based on their (technical) API documentation. Second, we gathered internal, RKI-specific data sources in consultation with colleagues working in the Department of Infectious Disease Epidemiology (including the Emergency Operations Center), Department of Health Monitoring, Risk Communication Unit, Press Office, and Social Media Task Force. Finally, through desk research, we gathered infodemic insight reports by governmental institutions, academia, and nongovernmental organizations.

Integrated (Data) Analysis to Report Infodemic Insights

Subsequently, we assessed the suitability of these sources for social listening and integrated the analysis of the RKI. Specifically, we assessed how each data source could potentially be analyzed to identify themes and narratives for infodemic insights, how frequently data become updated and available, and the extent to which there may be data protection risks. To best use the available resources at the RKI, we discarded several data sources from the initial list of potential data sources. The initial list of potential data sources and reasons for inclusion and exclusion are presented in Multimedia Appendix 1. Our goal was to gather an as-diverse-as-possible pool of data sources, given limited resources. Tools, analytics, and data sources were ideally open sources and available for noncommercial use, in support of open science, and appropriate for use by public or governmental institutions. For instance, many tools were available to collect and analyze Twitter data, and we initially decided to rely on the freely available tools TweetDeck and epitweetr to cover Twitter data. However, because of recent changes in Twitter, API access (on which these tools rely) now comes with a new cost model [44].

We decided to include all RKI-specific data sources as they reflect questions directed at the RKI that are very different from Twitter data and all infodemic insight data sources that include survey data that are published less frequently as a report and are processed, and thus do not require many additional resources from an RKI-based infodemic insights team (eg, to analyze raw data). These decisions were made within the RKI author group (SB, CL, and PS) after trying out various data sources and tools for practical use, and assessing the costs and benefits of each data source. Thus, the list of data sources presented in this study should be considered as a starting point that covers an as-diverse-set-as-possible, while keeping in mind a reasonable allocation of resources (eg, an initially small team of infodemic managers).

We tabulated the frequency with which each data source generated new data, and thus, the frequency with which it should be analyzed, and the type of data that can be extracted from each data source (ie, different outcome variables or indicators per data source, such as the number of Twitter comments per topic). In addition, we described possibilities within the framework of extending web-based social media listening using recent techniques from the field of NLP. Furthermore, we evaluated each data source based on ethical and data protection considerations (eg, a person writing a private message to the RKI should remain private and not end up in an infodemic insight report). The analysis of the heterogeneous set of identified data sources followed a mixed methods approach to combine qualitative (themes) and quantitative data (analytics). Qualitative data were analyzed through reflexive thematic analysis to identify themes (people's experiences, views, perceptions, and representations) regarding the public health event of interest [45,46], per data source.

Data Protection and Ethics

Many different data sources warrant careful consideration regarding privacy and data protection before they can be used for active social listening and integrated analysis. To formally assess data handling, 2 researchers performed an independent risk assessment of each identified data source using the RKI's standardized data protection questionnaire (version 03/2019). Risk was categorized by the dimensions of low, normal, high, or very high levels of data protection required; potential disagreement was discussed. The RKI's standardized data protection questionnaire facilitates adherence to the European Union regulation 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [47]. Multimedia Appendix 1 is an excerpt of the data protection questionnaire and summarizes what each dimension entails.

Setting Criteria for Activating Infodemic Insights Reporting

To define criteria for activating infodemic insights reporting structures in the context of a specific health event or health emergency, we (CH and TSB) consulted the RKI's preparedness and response group. The Emergency Operating Center is within their portfolio, which is situated within the RKI Department for Infectious Disease Epidemiology. We identified and reviewed the RKI's crisis management structures and preparedness and response plans for Germany and the human resources (including potential surge capacity), to see when and where infodemic response activities could be activated.

Results

Identification and Assessment of Data Sources

Table 1 presents the 16 data sources (including tools and reports) that were identified as suitable starting points for web-based and non-web-based social listening and integrated analysis at the RKI, based on the full list of 42 identified sources provided in the Multimedia Appendices 1-3. These fall into 3 main categories: social media and web-based listening, RKI-specific, and infodemic insights. We included primary data sources, such as social media data, requests addressed to the institute (through the Emergency Operations Center [48] and Press Office), task force meetings, press requests, and secondary data sources (ie, secondary research data and reports [24,49-51]). Of note, 3 infodemic insights were COVID-19 specific (ie, COVID-19 Snapshot Monitoring [COSMO] [49], the German Federal Institute for Risk Assessment [BfR]-Corona-Monitor [50], and Ministry of Health COVID-19 digital emergency operating center weekly briefing [52]), the WHO Infodemic Insights reports both on COVID-19 and mpox [24,25], and the general, non-COVID-19 specific reports were published by the Center for Monitoring, Analysis and Strategy (CEMAS) [51].

For data extraction, a public health taxonomy helps to identify thematic categories in conversations relevant to the public health response. Validated public health taxonomies for social listening, that is, the COVID-19 and mpox taxonomies developed by the WHO [23,53], formed the basis of the generic taxonomy related to public health issue X, as shown in Figure 1. The taxonomy includes topics and subtopics to capture the breadth of these conversations and help identify the structure and changes in narratives within thematic categories relevant to public health response. To create an infodemic insights report at the RKI, we translated the taxonomy into German (Multimedia Appendix 2). For some data sources, the taxonomy could be directly translated into German and used as Boolean search terms. Question marks can be added to Boolean search terms to identify information voids (questions) [15]. The taxonomy provided for public health issue X is applicable to other infectious diseases or health emergencies and will need to be adapted if the nature of the public health emergency in question is very different from COVID-19 and mpox (eg, war, an extreme weather event). In all events, both the inclusion of data sources and the taxonomy will follow an iterative process and need to be updated regularly to reflect changes in the situation and themes that occur (see the next section).

The German version is provided in Multimedia Appendix 2 based on COVID-19 and mpox taxonomies [23,53].

Several data sources were available weekly (Table 1). Most social media and web-based data (analytics) can be collected more frequently and in real time, but subsequent analyses can still take place weekly. Other data sources can only be monitored less frequently, as surveys and reports are published biweekly, monthly, or on an ad hoc basis.

The data protection risk assessment indicated normal-to low-level risks for social media and web-based listening data sources. RKI-specific data sources were assigned various levels of risk from low to high. Social media activity on RKI accounts was assigned normal (for comments) to high (for direct messages) risk; in all circumstances, user comments or sending a message will remain anonymous. Webpage metrics (RKI website traffic data and search patterns) were considered low risk. The data from the task forces were assigned a normal risk level if RKI employees were informed that the information would be used to develop integrated insights (in anonymized form). emails and phone calls from citizens were assigned a high risk, as these are considered private and nonpublic; however, the data were handled in aggregated form (counts per topic only). RKI press conference questions and Freedom of Information Act requests [54] were considered low risk because they are already in the public domain. So-called “small requests” for information from parliamentary groups or members of the German Federal Parliament [55] were assigned normal risk. Infodemic insights based on research, surveys, and reports from other parties had low risk (anonymous, aggregated data).

Table 1. Data sources evaluated based on suitable variables to identify topics and narratives, data availability and analysis frequencies, and data protection risk assessment.

Data source (tool or organizational unit)	Data extraction	Data availability	Data protection risk assessment
Social media and web-based listening			
Data aggregator (eg, Meltwater and Talkwalker)	• Data streaming from available social media platforms via API access. Automatically classify posts with taxonomy and count number of posts per taxonomy category (per platform and across platforms). Measure number of interactions (eg, sum of likes or retweets; depending on platform) and change over time. Alternatively, and depending on the aggregator, suitable tools for predefined reporting could be used.	Real time/weekly	Normal
Twitter (epitweetr [56])	• Signal detection (alerts) of an unusual increase in the number of tweets for a specific time place and topic.	Real time/weekly	Normal
Google searches, web content (Google Alerts)	• Published articles, blogs, etc, in a given time frame based on taxonomy and Boolean search strings. Analysis could be automated with web scraping (access web-based websites and apply taxonomy-based search).	Real time/weekly	Normal
Google searches (Google Trends)	• Compare topic (keywords) to baseline topic (ie, "COVID-19") and compute relative weekly average. Analysis could be automated with web scraping (download and analysis of trends data).	Real time/weekly	Low
Robert Koch Institute-specific			
Social media traffic and activity on the RKI's[a] social media accounts [57]: [58] (TweetDeck); Mastodon [59]; (Instagram Insights [60]); YouTube [61]; LinkedIn [62]	• Count the number of direct messages and comments regarding a topic relative to the overall number of them in the given time frame and questions may be identified to know where information voids exist and new topics may emerge. • Instagram specific (Instagram Insights): interactions with posts (likes, number of comments, amount of times post was saved) by topic and the number of people that saw the post (reach). If stories are posted, use the interactions with them by topic relative to the average number of interactions with stories (sum of likes and shares) during the week. • Twitter specific (TweetDeck): complementary to automated quantitative data analysis (see above); can be used for qualitative data exploration, that is, to scan trending hashtags regarding COVID-19–related topics.	Weekly	Comments: normal; direct messages: high
Webpage metrics (RKI website traffic data and search patterns)	• Use the number of visits on subpages relative to overall visits on the RKI-web pages.	Weekly	Low
Task Forces (RKI departments, emergency operations center)	• Ask an RKI scientist with technical experience: "What do you think is important and needs to be communicated, now and in the next month?" and let them rank these issues, then create an overall ranking across RKI experts.	Weekly	Normal
Emails and phone calls from citizens and journalists (RKI press office and Emergency Operations Center)	• Emails and calls regarding a topic relative to the overall number of emails and calls in the given time frame and questions may be identified to know where information voids exist and new topics may emerge.	Weekly	High
RKI press conferences questions	• Count the questions regarding a topic during the conference relative to overall number of questions and use them to identify information voids.	Ad hoc, depending on the interval of press conferences	Low

Data source (tool or organizational unit)	Data extraction	Data availability	Data protection risk assessment
Freedom of Information Act requests (*FragDenStaat* Portal [54])	• Count requests by topics addressed to the RKI.	Ad hoc, when requests are submitted	Low
"Small requests" from members of the German Federal Parliament [55]	• Topics gain insights into the issues that politicians and their constituency are concerned with.	Ad hoc, when requests are submitted	Normal
Infodemic insights			
COSMO[b] Snapshot monitoring [49]	• Scan report for sections relevant to the taxonomy. The output depends on how questions are framed (eg, "How informed do you feel about vaccinations?" "X % do not feel very informed"). Read the summary: some results about knowledge show information voids. This also depends on how the question was framed in the survey. For relevant questions, use the proportion of respondents not knowing or worrying. • COSMO could also be used for identifying information voids and new topics that need communication.	Ad hoc, when report is published (currently biweekly)	Low
BfR[c]-Corona-Monitor [50]	• Perceived informedness: how informed do you feel regarding topic X? (less informed—more important topic). • Use the proportion of respondents not feeling informed by topic. • This may be used to identify information voids and new topics that emerge and need communication.	Ad hoc, when report is published (biweekly/monthly)	Low
MoH[d] COVID-19 Digital Emergency Operations Center, weekly briefing [52] (Cosmonauts and Kings)	• Browse reports for sections relevant to the taxonomy. The report summarizes findings of existing studies, social media data (posts, comments, analytics), and reports misinformation (Telegram, fact-checking organizations), including narratives regarding the MoH.	Weekly	Low
WHO[e] Infodemic Insights report [24,25] (Marble Global)	• Browse report sections relevant to the taxonomy (questions in English-speaking communities may also be relevant to German-speaking communities; misinformation across countries may be similar).	Weekly	Low
CEMAS[f] [51]	• Browse reports for sections relevant to the taxonomy. The output depends on how questions are framed. • Scan reports and blog posts for relevant topics.	Ad hoc, when report is published	Low

[a]RKI: Robert Koch Institute.

[b]COSMO: COVID-19 Snapshot Monitoring.

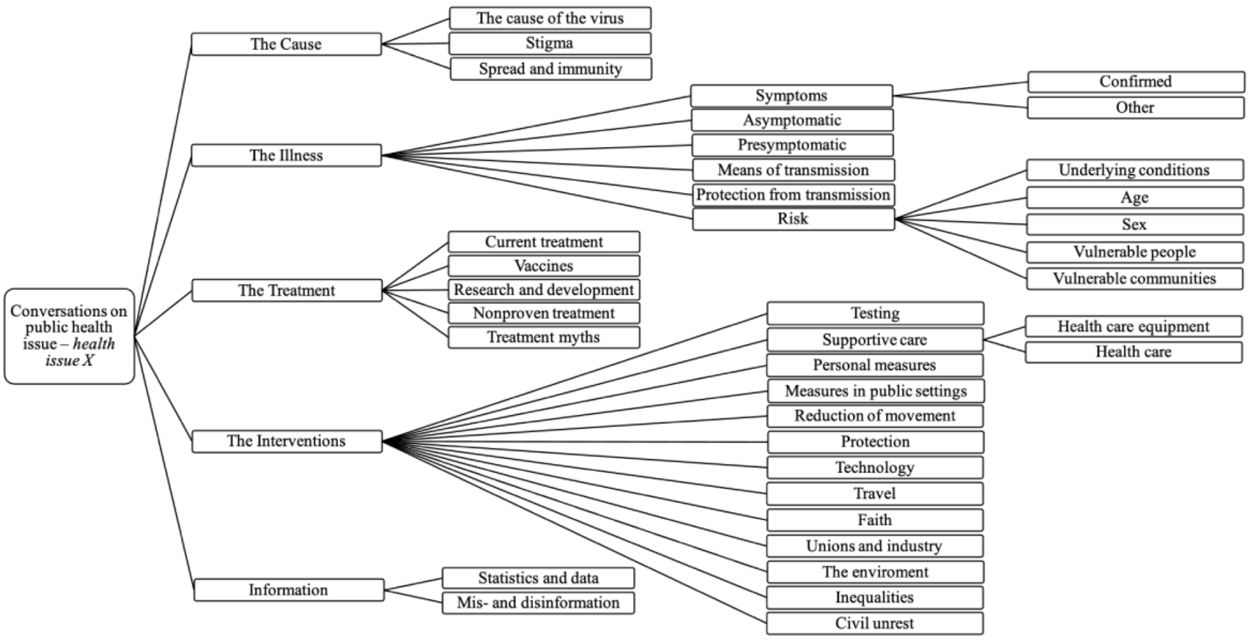
[c]BfR: German Federal Institute for Risk Assessment (Bundesinstitut für Risikobewertung).

[d]MoH: Ministry of Health.

[e]WHO: World Health Organization.

[f]CEMAS: Center for Monitoring, Analysis and Strategy.

Figure 1. Taxonomy to systematically monitor keywords in conversations related to public health issue X within thematic categories relevant to public health response.



Automation of Data Extraction and Analysis From Web-Based Social Media

Web-based social media platforms are constantly producing large amounts of data. For instance, approximately 30 million German tweets were sent every month (not shown, based on an analysis from the Twitter API [44]). Ideally, data analysis should be able to cope with such data streams to derive results that represent the entire data set in a timely manner. In the context of this framework, this can be achieved by collecting data through APIs and deriving quantitative metrics automatically (ie, the number of likes). Typically, APIs offer a type of keyword-based search, allowing the incorporation of taxonomy directly into data collection. The size of data and availability of APIs makes web-based social media listening well-suited for automatization, freeing team members to spend more time on qualitative or more involving data analysis.

Importantly, this lays the foundation for systems with more complex analytical methods, as modern data-driven systems are typically built as automated pipelines (from data collection to analysis). Technological advancements of the last years in AI, particularly in NLP, have led to a wide range of new approaches across different domains, including public health, for text and language analysis [63]. A prominent technique is sentiment analysis, which aims to determine the sentiment expressed in texts, for example, toward vaccination [64]. Furthermore, recent approaches in topic modeling (eg, BERTopic [65]) have demonstrated the ability of language models to cluster texts (such as social media posts) into semantically coherent groups (topics) in large text corpora. These latent topics are usually not known or assumed a priori, and such approaches can help to find unexpected topics in social media conversations. Consequently, data-driven approaches may allow the automatic derivation of an adaptive taxonomy from data that allows topics to emerge over time. This differs from a domain-driven taxonomy (Figure 1), which is built a posteriori and based on the domain knowledge of the involved researchers. Although data-driven approaches offer interesting possibilities, they may be challenging to operate as they require specific technical knowledge (eg, machine learning) and special computer hardware (eg, GPU). In addition, incomprehensible results generated by language models in NLP applications are generally difficult to interpret and require expert knowledge. By contrast, the proposed domain-driven taxonomy (Figure 1) is technically easy to implement and to understand, facilitate, and enable the use of teams with different levels of technical know-how in health institutes.

Integrated Analysis and Workflow to Report Infodemic Insights

The taxonomy (Figure 1) serves as a starting point for deductively coding themes that may emerge in the data and ensures basic comparability across data sources. Additional topics not contained in the predefined taxonomy are inductively coded and can be added if they are not contained in the predefined taxonomy. Topics that emerge across many different data sources may be (relatively) more important than topics that emerge only rarely; however, final judgments of the relative importance of particular topics and the urgency with which they require a response are made based on a risk matrix. Figure 2 shows the proposed workflow for creating an infodemic insight report based on the data sources listed in Table 1. The infodemic insights team lead decides which data sources should be included for analysis based on the availability of the team members and the timeframe. Each team member can be responsible for one or more data sources and independently extract the data and identify themes and topics based on the taxonomy (Figure 1) and add additional themes that emerge, collected in a spreadsheet. In a group meeting, the core team discusses the initial data extraction and identifies potential examples of insights that could be included in the report as illustrative examples of the identified narrative (eg, public comments or posts on Twitter).

The team lead drafts an insight report based on the main themes and topics. The core team then judges each theme in a risk matrix to determine the risk level for each theme, determine which themes to prioritize to be included in the report, and uses a scalar judgment (Figure 3), which is based on the US CDC's Vaccine Confidence Insights Reports [26,27] and adapted for broader public health topics such as chronic diseases, natural disasters, or other emergency responses (such as mpox [53]). The risk matrix is a classic decision matrix, where the first axis is the degree of impact on the uptake of a health-promoting behavior, and the second axis is the frequency it appears in the data sources relative to previous data collection points. High-risk themes can be those that lower health-promoting behaviors, have wide reach, and are pervasive, whereas low-risk themes are concerning, but have limited reach and dissemination. Moderate risk can trigger hesitancy to follow health-promoting behaviors, tend to have moderate reach, and moderate dissemination. Low risk is assigned to themes that can trigger hesitancy but have limited reach and dissemination. No risk themes can include themes that do not concern or even increase health-promoting behaviors. Subsequently, scalar judgment assesses the directionality of the theme over time (eg, since the last report): increasing, stable, and decreasing. Then, the entire team reviews the report, which then undergoes scientific clearance by the RKI's president and is distributed to our stakeholders.

Figure 2. Swim lane graph showing roles and responsibilities across the infodemic management team, as well as a proposed workflow to combine different data sources into an infodemic insights report. Adapted from Kolis and Voegeli [66].

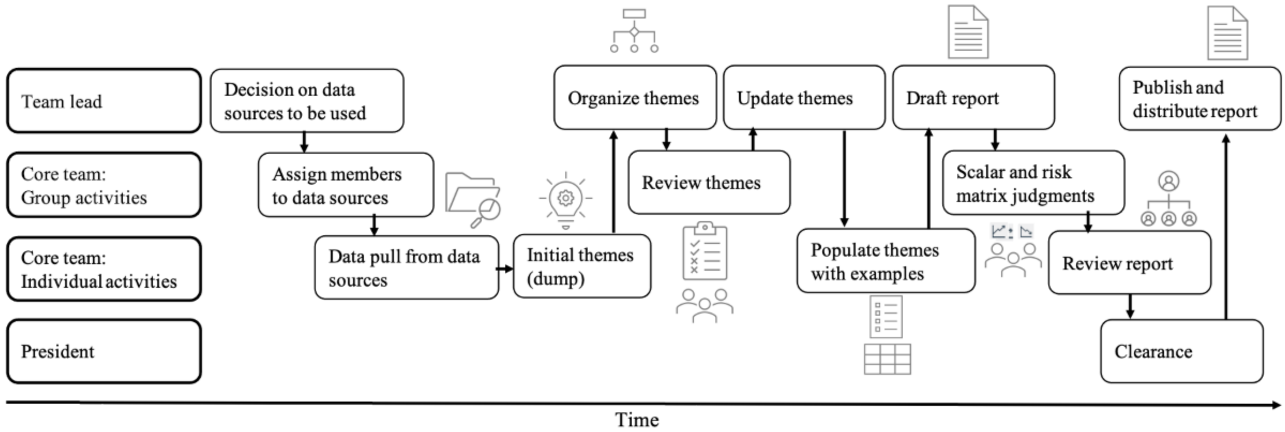
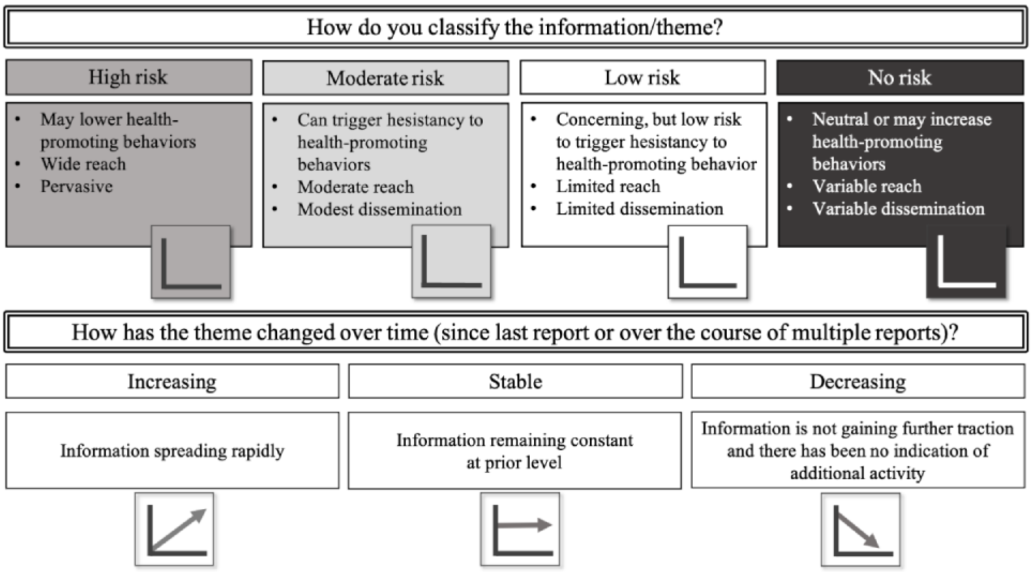


Figure 3. Risk matrix and scalar judgments, adapted from the US Centers for Disease Control and Prevention Vaccine Confidence Report methodology [26,27].



Reporting Back Actionable Infodemic Insights

Communicating infodemic insights and actionable recommendations based on social listening and integrated analysis is essential to support the public health response. The level of reporting details depends on the availability of resources (ie, the number of team members available and the number of hours that can be spent on the project). The output could range from a full-fledged insights report, including actionable recommendations, to a potential set of indicators that can be integrated into existing reports (eg, the RKI's situation reports focused on epidemiological trends and developments). In public health emergencies, speed trump perfection, and depending on the situation, a quick overview in an (epidemiological) situation report may trump a stand-alone (infodemic) report. Nevertheless, careful consideration of the impact of the published report or indicator is needed before it is sent to various audiences or published on the web. To put the integrated analysis results to best use, the infodemic insights report should be shared widely with partners and interested stakeholders who can use these insights for risk communication and community engagement activities. These partners and stakeholders include, but are not limited to, other German public health institutes (state and local), governmental institutions and ministries (Ministry of Health and Federal Centre of Health Education), community and religious organizations, science communicators, journalists (media), and fact-checking organizations.

Criteria for Activating Social Listening and Integrated Analysis Structures

Social listening and integrated analysis structures can be activated in the context of the RKI's crisis management structures [48,67]. Owing to the primary responsibility at the district and federal state levels in dealing with important epidemic situations, the RKI (federal level) usually only becomes active in the case of major or exceptional epidemiological situations [67]. The term important epidemic situations refers to either the local or temporal clustering of threatening communicable diseases, threatening diseases in which pathogens or toxins can be considered as the cause, or the concretely justified possibility that such diseases or illnesses may occur in the near future [67]. The activation of crisis management structures depends on the internal evaluation of the internal workload, number of possibly affected people, disease severity, geographic distribution, and public perception of the situation [48]. However, social listening and integrated analysis structures to report infodemic insights can also be activated for public health emergencies concerning Europe, as a support and prevention of the spread of a communicable disease to Germany, as judged by RKI experts, as our analyses focus on the German-speaking context.

Editable versions of Figures 1-3 and Multimedia Appendix 2 are available in Multimedia Appendix 4.

Discussion

Principal Findings

In this study, we propose a framework to establish social listening and integrated analysis to report infodemic insights at the National Public Health Institute in Germany. We identified and assessed 16 different types of data sources for social listening (at the time of writing, fall or winter 2022/2023) that fall into 3 main categories: social media and web-based listening data, RKI-specific data, and infodemic insights. Monitoring these web-based and non-web-based data sources can help to understand the population's understanding, perceptions, concerns and questions, information voids, narratives, misinformation and disinformation, and other relevant information about people's reactions to a health topic in Germany [24]. Most of these data sources can be analyzed weekly to detect current trends and narratives and to inform a timely response. Emerging data sources can also be included. One forthcoming data source that has the potential to provide key infodemic insights is the platform "RKI Panel—Health in Germany" [67], which plans to repeatedly survey a group of people on various health science topics. Social media and web-based listening data sources are available through different channels such as APIs, commercial data aggregators, or through manual searches. Consequently, obtaining and processing a comprehensive data set is a nontrivial task and is related to both the computational resources and available funds. For example, in the case of web-based social listening, the cost of using a commercially available data aggregator should be weighed against the technical expertise needed to collect and manage data from multiple freely or commercially available sources (ie, social media platforms; Multimedia Appendices 1-3). The selection of data sources used for each public health event might differ, depending on the situation and resources available.

Subsequently, a methodological examination was conducted to produce infodemic insights for the RKI. These insights can point out confusion, where the health authority is experiencing communication failures with the public, and what policy or programmatic levers can be used to address it (including but not limited to risk communication activities). Although there are many reasons for misinformation spreading [68] (eg, individual differences, information voids), identifying and tracking misinformation early can help with prebunking and debunking misinformation. For guidance on when and how to prebunk and debunk, see the Debunking Handbook [11].

The scope and extent of the integrated analysis that is put into place depends on the resources available to the project. We relied on prior experiences by the US CDC [26] to lay out the resources needed for different tasks and responsibilities, such as analyzing specific data sources, identifying common themes across data sources, and finally writing up a structured insight report. The outputs are flexible: either key infodemic insights are added to existing situation reports or a stand-alone report can be published. The primary audience for the infodemic insights reports is the RKI Emergency Operations Center and task forces. In addition, other public stakeholders and communicators involved in acute public health events [67], including but not limited to the Federal Ministry of Health, the Federal Centre of Health Education, the Federal Institute for Risk Assessment, and state- and local-level public health authorities and governmental institutions, could benefit from these reports. Collaboration and exchange with these organizations should be sustained and strengthened through wide sharing of infodemic insights and could also create access to additional data sources for social listening (eg, analysis of hotlines for citizens from the Federal Centre of Health Education).

Finally, we considered different criteria for activating integrated analysis structures and described how these activities could fit into the RKI's existing crisis response structures and Germany's legal framework [48,67]. The infodemic management activities proposed in this work are deemed suitable for addition to the existing preparedness and response structures at the RKI.

As we applied the methods of the WHO's infodemic insights report to the German context on the methodological level, this provides an opportunity to test how robust findings are across languages and geography (eg, compared with findings in the context of the WHO Infodemic Insights report). It is important to note that German speaking does not mean "within Germany," as netizens are widely connected. There is both a German-speaking community outside Germany, Austria, and Switzerland (the DACH [Germany—D, Austria—A, and Switzerland—CH] region) that would be captured well by the analysis, and a non-German-speaking community within Germany that would not be captured well by the proposed analysis. Moreover, our case study for the German context also serves as a roadmap to establish infodemic management at other institutes, both within and outside of Germany.

Limitations

Thus, the proposed activities should be interpreted carefully. The identified data sources include more web-based than non-web-based sources, and all data sources cover different audiences and come with inherent biases. Twitter appears to be a particularly fruitful source, as the data available for analysis are very comprehensive [69]. However, the future of Twitter API for academic research access is uncertain [65]. Moreover, despite being a popular platform, Twitter users are not representative of the general population [67]. Twitter has a major influence on the information ecosystem, for example, through journalists who can bring trending topics to offline media or scientists and politicians who serve as multipliers. Furthermore, not all key social media data sources have API. However, the overall direction points toward open social media data as further social media channels have recently implemented research APIs (eg, TikTok and YouTube; see Multimedia Appendices 1-3). However, the use regulations for these research accesses vary widely in terms of their eligibility. For instance, TikTok's new research program is currently only available for US-based research, and the YouTube research program is only eligible for researchers from higher education institutions (that can grant degrees). These regulations limit the use of research programs for public and governmental institutions, such as the RKI. Data access may only be available via commercial options, either directly from social media platforms or data aggregators.

It is still necessary to include more offline sources such as community dipstick surveys or town hall discussions. This would require additional personnel trained in conducting field studies (eg, anthropology and ethnography). Similarly, we included citizen questions directed to the RKI but not to other public institutions (such as the Federal Ministry of Health or the Federal Centre for Health Education), science communicators, politicians, or other actors. Importantly, even though the public seeks information at the RKI, the RKI predominantly deals with (public) health professionals, which could affect data collection for social listening activities. Public health professionals can, however, still provide valuable insights into ongoing narratives in the general population and serve as an audience for the insights report. Furthermore, there is a trade-off between speed and accuracy. The goal of an integrated analysis is to identify important narratives quickly and respond rapidly (eg, to misinformation). Iterative updates, internal (clearance) procedures, and publishing timeframes can hinder swift publication of infodemic insights. Even ambitious weekly or biweekly reporting may be too slow for a timely operational response to the current narrative, information voids, or an outbreak of misinformation, especially on social media.

Next Steps

To put the proposed framework for social listening and integrated analysis into practice [10], several activities were planned to operationalize social listening and integrated analysis to report infodemic insights at the RKI. First, the proposed setup for data handling will be submitted for ethics and data protection clearance. For data protection clearance, the identified data sources and variables to be obtained are discussed closely with the data protection officer. Second, in collaboration with the RKI's newly established Centre for Artificial Intelligence in Public Health Research, we will seek to further explore the RKI's web-based social listening capacities using artificial intelligence techniques and the data sources identified for social media and web-based listening. Third, the integrated analysis proposed here could potentially be piloted in the form of a field infodemiology project by field epidemiology fellows in Germany, under the supervision of and in collaboration with the RKI's risk communication group and the Department of Infectious Disease Epidemiology, Unit for Preparedness and Response. During this field phase (pilot), the data sources, taxonomy, integrated analysis, and workflow were tested and evaluated in the German context. This will help identify potential difficulties in combining different data sources and in subsequent reporting, particularly as many decisions in the process are subjective. The pilot will also provide insight into the amount of (human) resources needed to operationalize the proposed social listening and integrated analysis activities and their appropriate turnaround and reporting time frame. If the pilot is successful, the analysis can be extended to other health topics (eg, climate crisis). Therefore, novel taxonomies and Boolean search strings need to be developed. The need to constantly analyze narratives surrounding a particular topic (and which) needs to be evaluated and re-evaluated.

Moreover, a continuous and iterative evaluation and re-evaluation of the data sources, infodemic insights reporting, and workflow is required to build sustainable and effective infodemic management activities at the RKI. International exchanges with other public health institutes building experience with social listening [70,71] and communities of practice can foster further advancement in this area. A forthcoming guidance on developing infodemic insight reports will be published in a manual by the WHO and its partners [72]. A final important next step is to involve stakeholders and partners and create appreciation and demand for infodemic insights reporting and integrate this into regular policy making and programmatic decision-making [10]. Actively reaching out to these partners is essential for creating a demand for the report. Conversely, these partners could deliver additional data sources and inputs for future studies. Ultimately, an English version of the findings could be reported to the ECDC and WHO to add to the European and global level of reporting on the infodemic (eg, national surveillance data are being shared through this route, feeding into international surveillance reports).

Conclusions

The RKI identified and assessed a wide range of data sources for social listening and integrated analysis to report actionable infodemic insights, ensuring a valuable first step in establishing and operationalizing infodemic management at the RKI. Setting up the right tools for social media and web-based social listening will help to automate parts of the process. Piloting the proposed work will help refine the proposed workflow and show its value in informing the public health response. Ultimately, this work will provide better and targeted public health communication at the RKI and beyond.

Acknowledgments

The authors thank Auss Abboud and Eunyoung Hwang for their assistance in evaluating epitweetr as a social listening tool. They would also like to thank Wolfgang Scheida and Maud Hennequin for their input and reflections on the available data sources for the Robert Koch Institute and the potential application of the proposed work in the social media task force and press office. The authors thank Jessica Kolis, Kathryn Brookmeyer, and Yulia Chuvileva for creating the original swim lane graph (Figure 2) and risk matrix (Figure 3), which were slightly adapted for this study. We thank Christopher Irrgang for his critical review of the data science aspects of this study.

Authors' Contributions

Conception or design of the work was conducted by TSB and CL. Data collection was done by TSB, CL, PHS, and SW. Data analysis and interpretation was carried out by TSB, CL, PHS, CH, SW, TDP, EW, AI, and CV. TSB, CL, and PHS were involved in drafting the paper. Critical revision of the paper was conducted by TDP, AI, EW, CV, LHW, TSB, CL, PHS, and SW. Final approval of the version to be published was given by all authors.

Conflicts of Interest

EW report no conflict of interest. TSB, CL, PHS, CH, SW and LHW are employed by the Robert Koch Institute, which is an organization that is the subject of inquiry and review. TDP and AI are staff members of the World Health Organization, are responsible for the views expressed in this paper, and do not represent the views of the organization. CV is employed by the US Centers for Disease Control and Prevention; the findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the US Centers for Disease Control and Prevention.

Multimedia Appendix 1
Risk assessment of data protection requirements, based on the data protection questionnaire for new procedures for processing personal data at the Robert Koch Institute (Version 03/2019).
[PDF File (Adobe PDF File), 32 KB - infodemiology_v3i1e43646_app1.pdf]

Multimedia Appendix 2
Taxonomy to systematically monitor keywords in conversations related to public health issue X within thematic categories relevant to public health responses.
[PDF File (Adobe PDF File), 121 KB - infodemiology_v3i1e43646_app2.pdf]

Multimedia Appendix 3
Full list of in- and exclusions of 42 potential data sources for social listening.
[XLSX File (Microsoft Excel File), 18 KB - infodemiology_v3i1e43646_app3.xlsx]

Multimedia Appendix 4
Editable versions of Figures 1-3 and Multimedia Appendix 2.
[PPTX File , 556 KB - infodemiology_v3i1e43646_app4.pptx]

References

1. Newman N, Fletcher R, Robertson CT, Eddy K, Nielsen RK. Reuters Institute Digital News Report 2022. Reuters Institute for the Study of Journalism. 2022. URL: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf [accessed 2022-11-30]
2. Wardle C, Derakhshan H. Information disorder: toward an interdisciplinary framework for research and policy making. Council of Europe report. 2017 Sep 27. URL: https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html [accessed 2022-11-30]
3. 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]

4. 5. Stellungnahme des ExpertInnenrates der Bundesregierung zu COVID-19 - Zur Notwendigkeit evidenzbasierter Risiko- und Gesundheitskommunikation. Informationsamt der Bundesregierung. 2022 Jan 30. URL: https://www.bundesregierung.de/resource/blob/974430/2002168/4251476198ffd2a0af663fd90c29240f/2022-01-30-fuenfte-stellungnahme-expertenrat-data.pdf?download=1 [accessed 2022-11-30]
5. Habersaat KB, Betsch C, Danchin M, Sunstein CR, Böhm R, Falk A, et al. Ten considerations for effectively managing the COVID-19 transition. Nat Hum Behav 2020 Jul;4(7):677-687. [doi: 10.1038/s41562-020-0906-x] [Medline: 32581299]
6. Loss J, Boklage E, Jordan S, Jenny MA, Weishaar H, El Bcheraoui C. Risk communication in the containment of the COVID-19 pandemic: challenges and promising approaches. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2021 Mar;64(3):294-303 [FREE Full text] [doi: 10.1007/s00103-021-03283-3] [Medline: 33564896]
7. COVID-19 National Preparedness Collaborators. Pandemic preparedness and COVID-19: an exploratory analysis of infection and fatality rates, and contextual factors associated with preparedness in 177 countries, from Jan 1, 2020, to Sept 30, 2021. Lancet 2022 Apr 16;399(10334):1489-1512 [FREE Full text] [doi: 10.1016/S0140-6736(22)00172-6] [Medline: 35120592]
8. Savoia E, Piltch-Loeb R, Masterson E, Testa MA, Fantini MP, Tsolova S. Rapid analysis of the first year of the COVID-19 pandemic response for the development of preparedness measures for public communication. SSRN J 2022 Jul 26:1-55 [FREE Full text] [doi: 10.2139/ssrn.4174475]
9. OECD. Building Trust to Reinforce Democracy: Main Findings from the 2021 OECD Survey on Drivers of Trust in Public Institutions. Organisation for Economic Cooperation and Development. 2022 Jul 13. URL: https://www.oecd-ilibrary.org/governance/building-trust-to-reinforce-democracy_b407f99c-en [accessed 2022-11-30]
10. WHO policy brief: COVID-19 infodemic management. World Health Organization. 2022 Sep 14. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Policy_Brief-Infodemic-2022.1 [accessed 2022-11-30]
11. Lewandowsky S, Cook J, Ecker U, Albarracín D, Amazeen MA, Kendeou P, et al. The Debunking Handbook 2020. Center For Climate Change Communication. 2020. URL: https://www.climatechangecommunication.org/wp-content/uploads/2020/10/DebunkingHandbook2020.pdf [accessed 2022-11-30]
12. Herzog SM, Hertwig R. Kompetenzen mit "Boosts" stärken: Verhaltenswissenschaftliche Erkenntnisse jenseits von "Nudging". Appl Psychol 2019;47(4):19-40 [FREE Full text]
13. Eysenbach G. Eysenbach: infodemiology and infoveillance. SlideShare. 2008 Jun 03. URL: https://www.slideshare.net/eysen/eysenbach-infodemiology-and-infoveillance [accessed 2022-11-30]
14. 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]
15. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
16. Eysenbach G. How to fight an infodemic: the four pillars of infodemic management. J Med Internet Res 2020 Jun 29;22(6):e21820 [FREE Full text] [doi: 10.2196/21820] [Medline: 32589589]
17. Scales D, Gorman J, Jamieson KH. The COVID-19 infodemic - applying the epidemiologic model to counter misinformation. N Engl J Med 2021 Aug 19;385(8):678-681. [doi: 10.1056/NEJMp2103798] [Medline: 33979506]
18. Core competencies in applied infectious disease epidemiology in Europe. European Centre for Disease Prevention and Control. 2022 Apr. URL: https://www.ecdc.europa.eu/en/publications-data/core-competencies-applied-infectious-disease-epidemiology-europe [accessed 2022-11-30]
19. Gesser-Edelsburg A. Using narrative evidence to convey health information on social media: the case of COVID-19. J Med Internet Res 2021 Mar 15;23(3):e24948 [FREE Full text] [doi: 10.2196/24948] [Medline: 33674257]
20. 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, WHO, Vaccine Demand Hub, HealthEnabled. 2021. URL: https://www.gavi.org/sites/default/files/2021-06/Finding-the-Signal-Through-the-Noise.pdf [accessed 2022-11-30]
21. Newberry C, Macready H. What is Social Listening, Why it Matters + 14 Tools to Help. Hootsuite Inc. 2022 Dec 01. URL: https://blog.hootsuite.com/social-listening-business/ [accessed 2023-05-03]
22. Purnat TD, Wilson H, Nguyen T, Briand S. EARS - a WHO platform for AI-supported real-time online social listening of COVID-19 conversations. Stud Health Technol Inform 2021 May 27;281:1009-1010. [doi: 10.3233/SHTI210330] [Medline: 34042825]
23. 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]
24. Purnat TD, Nguyen T, Ishizumi A, Yau B, White B, Cecchini S, et al. Delivering actionable infodemic insights and recommendations for the COVID-19 pandemic response. Wkly Epidemiol Rec 2022 Jul 08;97(27):313-324 [FREE Full text] [doi: 10.1093/eurpub/ckac129.645]
25. Monkeypox Infodemic Insights Report. UK: Marble Global; 2022.

26. COVID-19 state of vaccine confidence insights reports. Centers for Disease Control and Prevention. 2022 Nov 10. URL: https://www.cdc.gov/vaccines/covid-19/downloads/sovc-rise-report-11-17-22.pdf [accessed 2022-11-30]
27. CDC's monkeypox state of vaccine confidence insights report. Centers for Disease Control and Prevention. 2022 Oct 05. URL: https://www.cdc.gov/vaccines/covid-19/downloads/sovc-mpx-report-10-18-22.pdf [accessed 2022-11-30]
28. Purnat TD, Vacca P, Burzo S, Zecchin T, Wright A, Briand S, et al. WHO digital intelligence analysis for tracking narratives and information voids in the COVID-19 infodemic. Stud Health Technol Inform 2021 May 27;281:989-993. [doi: 10.3233/SHTI210326] [Medline: 34042821]
29. Monkeypox multi-country outbreak - first update. European Centre for Disease Prevention and Control. 2022 Jul 08. URL: https://www.ecdc.europa.eu/en/publications-data/monkeypox-multi-country-outbreak-first-update [accessed 2022-11-30]
30. Infodemic. World Health Organization. 2022. URL: https://www.who.int/health-topics/infodemic#tab=tab_1 [accessed 2022-11-30]
31. The Expert Council of the Federal Government. Die Bundesregierung. URL: https://www.bundesregierung.de/breg-en [accessed 2022-11-30]
32. Robert Koch institute 2025 strategy (RKI 2025): promoting research and evidence, sharing knowledge, protecting and improving health. Robert Koch institute. 2017 Jun 26. URL: https://www.rki.de/EN/Content/Institute/RKI2025/RKI_2025_strategy.html [accessed 2022-11-30]
33. RKI 2018-2025 summary research agenda. Robert Koch Institute. 2019 Apr 26. URL: https://www.rki.de/EN/Content/Institute/RKI2025/RKI_2025_agenda.html [accessed 2022-11-30]
34. WHO competency framework: building a response workforce to manage infodemics. World Health Organization. 2021. URL: https://apps.who.int/iris/bitstream/handle/10665/345207/9789240035287-eng.pdf?sequence=1&isAllowed=y [accessed 2022-11-30]
35. Rubinelli S, Purnat TD, Wilhelm E, Traicoff D, Namageyo-Funa A, Thomson A, et al. WHO competency framework for health authorities and institutions to manage infodemics: its development and features. Hum Resour Health 2022 May 07;20(1):35 [FREE Full text] [doi: 10.1186/s12960-022-00733-0] [Medline: 35525924]
36. World Health Organization Regional Office for Europe. Monitoring wider effects of the COVID-19 pandemic. In: A Webinar to Learn About Monitoring the Wider Effects of the COVID-19 Pandemic. 2021 Presented at: HSRM '21; June 11, 2021; Virtual.
37. World Health Organization, US Centers for Disease Control and Prevention, Africa Centres for Disease Control and Prevention, Risk communication and community engagement collective service, First Draft. First WHO infodemic manager training. World Health Organization. 2020 Nov. URL: https://cdn.who.int/media/docs/default-source/epi-win/infodemic-management/im-training-booklet-3dec2020.pdf?sfvrsn=d4f5586a_6 [accessed 2022-11-30]
38. US Centers for Disease Control and Prevention, ECDC, UNICEF partners. Second WHO infodemic manager training. World Health Organization. 2021 Jun. URL: https://cdn.who.int/media/docs/default-source/epi-win/who_booklet_08.06.21.pdf?sfvrsn=b87686ab_4 [accessed 2022-11-30]
39. Infodemic management 101. World Health Organization. 2022. URL: https://openwho.org/courses/infodemic-management-101 [accessed 2022-11-30]
40. WHO public health research agenda for managing infodemics. World Health Organization. 2021 Feb 03. URL: https://www.who.int/publications/i/item/9789240019508 [accessed 2022-11-30]
41. Advancing infodemic management in risk communication and community engagement in the WHO European Region: implementation guidance. World Health Organization, Europe. 2022 Nov 08. URL: https://www.who.int/europe/publications/i/item/WHO-EURO-2022-5842-45607-65433 [accessed 2022-11-30]
42. US Centers for Disease Control and Prevention, UNICEF, RCCE collective service. 3rd WHO infodemic manager training. World Health Organization. 2022. URL: https://cdn.who.int/media/docs/default-source/epi-win/who-3rd-training-booklet_final.pdf?sfvrsn=a718d307_9 [accessed 2022-11-30]
43. Leuker C, Hertwig R, Gumenik K, Eggeling LM, Hechtlinger S, Kozyreva A, et al. Wie informiert sich die Bevölkerung in Deutschland rund um das Coronavirus? Umfrage zu vorherrschenden Themen und Gründen, dem Umgang mit Fehlinformationen, sowie der Risikowahrnehmung und dem Wissen der Bevölkerung rund um das Coronavirus. Max-Planck-Institut für Bildungsforschung. 2020 Aug. URL: https://pure.mpg.de/rest/items/item_3247925_11/component/file_3252645/content [accessed 2022-11-30]
44. Twitter API access levels and versions. Twitter Developer Platform. 2023. URL: https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api [accessed 2023-05-04]
45. Spencer L, Ritchie J, O'Connor W. Analysis: practices, principles and processes. In: Ritchie J, Lewis J, editors. Qualitative Research Practice: A Guide for Social Science Students and Researchers. Thousand Oaks, CA, USA: Sage Publications; 2003:198-218.
46. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019;11(4):589-597 [FREE Full text] [doi: 10.1080/2159676x.2019.1628806]
47. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive

95/46/EC. Official Journal of the European Union. 2016 May 04. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 [accessed 2022-11-30]
48. Halm A, Grote U, An der Heiden M, Hamouda O, Schaade L, Rexroth U, RKI-Lagezentrums-Gruppe. Crisis management at the Robert Koch Institute during the COVID-19 pandemic and the exchange between federal and state governments. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2021 Apr;64(4):418-425 [FREE Full text] [doi: 10.1007/s00103-021-03294-0] [Medline: 33666683]
49. WHO Regional Office for Europe. COVID-19 Snapshot Monitoring (COSMO Standard): monitoring knowledge, risk perceptions, preventive behaviours, and public trust in the current coronavirus outbreak - WHO standard protocol. PsychArchives. 2020 Mar 18. URL: https://www.psycharchives.org/en/item/62216bdb-69fa-44e7-92b4-8438b3817341 [accessed 2023-05-05]
50. BfR-Corona-Monitor. Bundesinstitut für Risikobewertung. 2022 Jul 14. URL: https://www.bfr.bund.de/en/bfr_corona_monitor-244792.html [accessed 2023-05-05]
51. CeMAS: Center for Monitoring, Analysis, and Strategy. URL: https://cemas.io/en/ [accessed 2023-05-05]
52. Ministy of Health COVID-19 digital emergeny operating centre, weekly briefing Digitales Lagezentrum - Wöchentliches Briefing. Berlin: Bundesministerium für Gesundheit (BMG); 2022.
53. Public health taxonomy for social listening on monkeypox conversations: for use in infodemic monitoring and insights generation. World Health Organization, Epidemic and Pandemic Preparedness and Prevention Team. 2022 Aug 17. URL: https://cdn.who.int/media/docs/default-source/epi-win/infodemic-monkeypox-taxonomy-27-may-2022-final_updated_v2.pdf?sfvrsn=f0190e04_3&download=true [accessed 2022-11-30]
54. Requests to Robert Koch-Institut. FragDenStaat. URL: https://fragdenstaat.de/en/requests/to/robert-koch-institut/?status=zurueckgezogen [accessed 2022-11-30]
55. Kleine Anfrage. Deutscher Bundestag. URL: https://www.bundestag.de/en [accessed 2022-11-30]
56. Espinosa L, Wijermans A, Orchard F, Höhle M, Czernichow T, Coletti P, et al. Epitweetr: early warning of public health threats using Twitter data. Euro Surveill 2022 Sep;27(39):2200177 [FREE Full text] [doi: 10.2807/1560-7917.ES.2022.27.39.2200177] [Medline: 36177867]
57. Soziale Medien, Newsletter und RSS-Feeds. Robert Koch-Institut. URL: https://www.rki.de/DE/Content/Service/Newsletter/newsletter-abo_node.html [accessed 2022-11-30]
58. Robert Koch-Institut. Twitter. URL: https://twitter.com/rki_de [accessed 2023-05-12]
59. Robert Koch-Institut. Mastodon. URL: https://social.bund.de/@RKI [accessed 2023-05-12]
60. Robert Koch-Institut. Instagram. URL: https://www.secure.instagram.com/rki_fuer_euch/ [accessed 2023-05-12]
61. Robert Koch-Institut. YouTube. URL: https://www.youtube.com/channel/UCDGaa4L_jkBA50D77RLRNJg/videos?view_as=subscriber [accessed 2023-05-12]
62. Robert Koch Institute. Linkedin. URL: https://www.linkedin.com/company/robertkochinstitut/ [accessed 2023-05-12]
63. Baclic O, Tunis M, Young K, Doan C, Swerdfeger H, Schonfeld J. Challenges and opportunities for public health made possible by advances in natural language processing. Can Commun Dis Rep 2020 Jun 04;46(6):161-168 [FREE Full text] [doi: 10.14745/ccdr.v46i06a02] [Medline: 32673380]
64. 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 09;2(1):e32335 [FREE Full text] [doi: 10.2196/32335] [Medline: 35578643]
65. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online March 11, 2022 [FREE Full text]
66. Kolis J, Voegeli C. Integrated analysis of the iceberg. In: Proceedings of the 5th Virtual WHO Infodemic Management Conference. 2021 Apr 22 Presented at: WHO Infodemic '21; November 2-11, 2021; Virtual URL: https://cdn.who.int/media/docs/default-source/epi-win/im_5th_conference_booklet.pdf
67. Allgemeine Verwaltungsvorschrift über die Koordinierung des Infektionsschutzes in epidemisch bedeutsamen Fällen (Verwaltungsvorschrift-IfSG-Koordinierung - IfSGKoordinierungs-VwV). Die Bundesregierung. 2013 Dec 12. URL: https://www.verwaltungsvorschriften-im-internet.de/bsvwvbund_12122013_31945300302.htm [accessed 2022-11-30]
68. Chan MP, Jones CR, Hall Jamieson K, Albarracín D. Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. Psychol Sci 2017 Nov;28(11):1531-1546 [FREE Full text] [doi: 10.1177/0956797617714579] [Medline: 28895452]
69. Twitter API - academic research access. Twitter Developer Platform. URL: https://developer.twitter.com/en/products/twitter-api/academic-research [accessed 2023-05-02]
70. Lohiniva A, Sane J, Sibenberg K, Puumalainen T, Salminen M. Understanding coronavirus disease (COVID-19) risk perceptions among the public to enhance risk communication efforts: a practical approach for outbreaks, Finland, February 2020. Euro Surveill 2020 Apr;25(13):2000317 [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.13.2000317] [Medline: 32265008]
71. Lohiniva AL, Sibenberg K, Austero S, Skogberg N. Social listening to enhance access to appropriate pandemic information among culturally diverse populations: case study from Finland. JMIR Infodemiology 2022 Jul 08;2(2):e38343 [FREE Full text] [doi: 10.2196/38343] [Medline: 37113448]

72. How to develop an infodemic insights report in 6 steps. In: World Health Organization. Geneva: World Health Organization; 2023.

Abbreviations

AI: artificial intelligence
API: application programming interface
BZgA: Bundeszentrale für gesundheitliche Aufklärung
CEMAS: Center for Monitoring, Analysis and Strategy
ECDC: European Centre for Disease Control and Prevention
NLP: natural language processing
OECD: Organisation for Economic Co-operation and Development
RKI: Robert Koch Institute
US CDC: United States Centers for Disease Control and Prevention
WHO: World Health Organization

Edited by T Mackey; submitted 09.12.22; peer-reviewed by P Sacco, C Wardle; comments to author 17.03.23; revised version received 18.04.23; accepted 24.04.23; published 01.06.23.

Please cite as:
Boender TS, Schneider PH, Houareau C, Wehrli S, Purnat TD, Ishizumi A, Wilhelm E, Voegeli C, Wieler LH, Leuker C
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;3:e43646
URL: https://infodemiology.jmir.org/2023/1/e43646
doi:10.2196/43646
PMID:37261891

Original Paper

Using Machine Learning Technology (Early Artificial Intelligence–Supported Response With Social Listening Platform) to Enhance Digital Social Understanding for the COVID-19 Infodemic: Development and Implementation Study

Becky K White[1], PhD; Arnault Gombert[2], MSc; Tim Nguyen[1], MSc; Brian Yau[1], BSc; Atsuyoshi Ishizumi[1], MSc, MPH; Laura Kirchner[2], MSc; Alicia León[2], BA; Harry Wilson[2], BA; Giovanna Jaramillo-Gutierrez[2], PhD; Jesus Cerquides[3], PhD; Marcelo D'Agostino[4], MSc; Cristiana Salvi[5], MS; Ravi Shankar Sreenath[5], MBA; Kimberly Rambaud[5], MSc; Dalia Samhouri[6], MPH, PhD; Sylvie Briand[1], MD, PhD; Tina D Purnat[1], MSc

[1]Department of Epidemic and Pandemic Preparedness and Prevention, World Health Organization, Geneva, Switzerland

[2]Citibeats, Barcelona, Spain

[3]Artificial Intelligence Research Institute, Spanish Council for Scientific Research, Cerdanyola, Spain

[4]Information Systems for Health, Evidence and Intelligence for Action in Health, Pan American Health Organization and World Health Organization Regional Office for the Americas, Washington DC, DC, United States

[5]Risk Communication and Community Engagement Unit, Health Emergencies Division, World Health Organization Regional Office for Europe, Copenhagen, Denmark

[6]Country Health Emergency Preparedness and International Health Regulations (2005), World Health Organization Regional Office for Eastern Mediterranean, Cairo, Egypt

Corresponding Author:
Tina D Purnat, MSc
Department of Epidemic and Pandemic Preparedness and Prevention
World Health Organization
Ave Appia 21
Geneva, 1202
Switzerland
Phone: 41 227912111
Email: purnatt@who.int

Abstract

Background: Amid the COVID-19 pandemic, there has been a need for rapid social understanding to inform infodemic management and response. Although social media analysis platforms have traditionally been designed for commercial brands for marketing and sales purposes, they have been underused and adapted for a comprehensive understanding of social dynamics in areas such as public health. Traditional systems have challenges for public health use, and new tools and innovative methods are required. The World Health Organization Early Artificial Intelligence–Supported Response with Social Listening (EARS) platform was developed to overcome some of these challenges.

Objective: This paper describes the development of the EARS platform, including data sourcing, development, and validation of a machine learning categorization approach, as well as the results from the pilot study.

Methods: Data for EARS are collected daily from web-based conversations in publicly available sources in 9 languages. Public health and social media experts developed a taxonomy to categorize COVID-19 narratives into 5 relevant main categories and 41 subcategories. We developed a semisupervised machine learning algorithm to categorize social media posts into categories and various filters. To validate the results obtained by the machine learning–based approach, we compared it to a search-filter approach, applying Boolean queries with the same amount of information and measured the recall and precision. Hotelling T^2 was used to determine the effect of the classification method on the combined variables.

Results: The EARS platform was developed, validated, and applied to characterize conversations regarding COVID-19 since December 2020. A total of 215,469,045 social posts were collected for processing from December 2020 to February 2022. The machine learning algorithm outperformed the Boolean search filters method for precision and recall in both English and Spanish languages (P<.001). Demographic and other filters provided useful insights on data, and the gender split of users in the platform was largely consistent with population-level data on social media use.

Conclusions: The EARS platform was developed to address the changing needs of public health analysts during the COVID-19 pandemic. The application of public health taxonomy and artificial intelligence technology to a user-friendly social listening platform, accessible directly by analysts, is a significant step in better enabling understanding of global narratives. The platform was designed for scalability; iterations and new countries and languages have been added. This research has shown that a machine learning approach is more accurate than using only keywords and has the benefit of categorizing and understanding large amounts of digital social data during an infodemic. Further technical developments are needed and planned for continuous improvements, to meet the challenges in the generation of infodemic insights from social media for infodemic managers and public health professionals.

(JMIR Infodemiology 2023;3:e47317) doi:10.2196/47317

KEYWORDS

infodemic; sentiment; narrative analysis; social listening; natural language processing; social media; public health; pandemic preparedness; pandemic response; artificial intelligence; AI text analytics; COVID-19; information voids; machine learning

Introduction

Background

From the outset of the COVID-19 pandemic, the infodemic, the excess of information, including misinformation and disinformation that can result in confusion or impact on health attitudes and behaviors during health emergencies, has been of keen interest to those involved in emergency response [1]. The infodemic can have a poor impact on public health outcomes [2,3], and there is evidence that those who are most at risk may be most vulnerable to infodemic [4]. During the pandemic, social and other digital media have allowed for the rapid dissemination of an overwhelming amount of information that can elongate or amplify outbreaks and reduce the effectiveness of epidemic response efforts and interventions [1]. People may feel confused by who to trust; may be confronted by outdated or incorrect information; or may be exposed to, share, or act on misinformation and disinformation. Having a comprehensive infodemic management strategy that includes integrated infodemic insights generation from offline and web-based social listening sources, as well as data sources in the health information systems and outside the health sector, may help build trust in governments and health authorities and help people understand and accept the pandemic response [5]. The World Health Organization (WHO) public health research agenda for infodemiology outlines priority research questions structured under 5 key thematic areas [6]. These themes have recommendations for preparedness and monitoring of the infodemic and for detecting and understanding the spread of infodemics. Better social media analysis tools and metrics are needed to support infodemic managers to understand and respond to an infodemic during a health emergency.

Social Media Analysis for Public Health

Social listening is often understood as an approach from marketing and communication to glean insights from analysis of social media channels [7], whereas in public health, social listening is often more broadly, including using data sources from the health system, sociobehavioral studies, community feedback, as well as nontraditional sources such as mobility data [6,8]. Social listening for public health involves infodemic insights generation, which is analyzed in an integrated manner across data sources from social listening, health information systems, and outside the health system [9]. Social listening and infodemic insight generation include the analysis of social media, traditional media, and other data sources—such as user search trends, epidemiological data, community feedback, and sociobehavioral data—to identify, categorize, and understand perceptions, questions, concerns, information voids, and narratives expressed and circulating in communities [9].

Social listening, integrated analysis of infodemic intelligence, and generation of infodemic insights and infodemic management recommendations are the first steps in providing evidence to manage the infodemic [10]. When analyzing social media, understanding the source, velocity, and volume of global social media information trends can help inform prebunking and debunking initiatives, fill information voids, develop user experience and digital resilience strategies, and inform infodemic responses [5,9]. Misinformation shared on social media can quickly cross international borders and platforms, with the same claim presented in different ways and contexts to users on YouTube or Facebook for example [11].

The WHO has previously reported on using artificial intelligence (AI)–driven social listening to deliver actionable infodemic insights [9] and on the development and validation of a public health social listening taxonomy [12]. The COVID-19 public health taxonomy was designed to provide a practical and structured approach to identifying narratives shared on digital media [13], and taxonomy-driven data analysis and integration have since been applied to other outbreaks, such as mpox [14]. Developed by public health and digital health experts, the taxonomy enables data to be filtered into categories, allowing for the identification of where the global conversation is growing and what the information voids or issues of concern may be.

Faced with millions of data points, AI can help filter data into these specific categories, as well as filter within the categories by questions or demographic identifiers. This enables an analyst to quickly see the signal through noise and obtain meaningful insights. Although there is significant potential in pandemic response, there have been calls for the application of AI to be grounded in ethical and multidisciplinary practice [15]. There is a need for greater cooperation between domain experts and AI practitioners, that tools be adapted from existing tools and aim to reduce, rather than add to, the workload of health care workers; that systems be adapted to the needs of low- and middle-income countries; and that global solutions with local adaptability options are developed [15]. In addition, there is a need for any application of AI to undergo ethical assessment, including users' rights to privacy, protecting journalistic sources, and preventing mass surveillance.

Challenges With Existing Social Media Analysis Systems

Although the need for social understanding amidst the COVID-19 pandemic and infodemic is paramount, traditionally, social media monitoring systems have been established for commercial purposes and brand management rather than providing insights that can inform public health action [16]. There are inherent challenges for public health analysts attempting to navigate these systems to quickly find actionable insights to guide their work. Systems can be costly, which may be a barrier, particularly to people working in lower-resources settings, and may include limited language coverage, data sources, data representatively, privacy, geographical bias, and prioritize "noise" over "signal." Here, we briefly outline these challenges.

Traditionally, social media monitoring tools have the strongest analytical capability in English. In some cases, this is followed by other widely spoken languages, and in many cases, no coverage of the long tail of minority languages, or adaptation to dialects or sociolects or local references. This can be problematic for analysts seeking to understand global narratives. Even if an additional language is included in the analysis tool, there may be many differences in the local references, which may not be captured by pretrained models or predefined dictionaries. This problem is further exacerbated in a fast-changing conversation topic, such as COVID-19, because vocabulary and new terms are constantly entering the conversation in each language as the situation evolves. Although some dedicated research on the COVID-19 infodemic has been conducted on a local level in India [17], Croatia [18], and Malaysia [19], to our knowledge, there have been few broader solutions covering multiple languages.

The diversity of data sources can be a challenge with some platforms prioritizing certain social media platforms, and principally, Twitter, because of data availability, which may have limitations in the representativity of the population [2] as well as experience changes in the quality and content on the platform due to changes in content moderation policies and platform user experience. Alternative text sources, such as surveys, SMS text message responses, call transcripts, and chatbot questions, often cannot be analyzed by commercial systems, meaning valuable data may be missing from the analysis data set. Maintaining user privacy is an important ethical consideration in public health. Within a branding context, identifying individuals to target with messages or advertisements can be a priority; however, identifying individual users may go against privacy protection principles in the public health context [15,20].

Understanding this representatively is important for ensuring community coverage and developing appropriate actions. A common issue is that in many countries, women's voices are underrepresented in the overall sample. The data results from social media analysis may indicate that a given topic is important, but this may be skewed toward perspectives from men [21]. In addition, being able to differentiate data based on country level can help reduce geographical bias, as data can be skewed toward USA-centric data, resulting in an underrepresentation of other countries.

Analytic Approaches to Categorization

A key challenge for public health analysts is how to quickly filter through the "noise"—conversation volume, interactions volume, mentions volume, top influencers, to find the "signal"—the actionable insight that can be identified for the purposes of infodemic management [12]. Finding these signals requires the analyst to look beyond the emphasis on high-engagement posts, rising narrative detection, common questions, sex gaps, and trust indicators. Filtering these narratives through a public health taxonomy enables data to be grouped into relevant categories, enabling analysts to quickly see where the conversation is growing and the directionality.

There are various ways to accomplish this, such as the use of Boolean queries with keywords or natural language processing (NLP). The emergence of NLP models, such as bidirectional encoder representations from transformers (BERT) models [22], and the democratization of those models with open-source hubs, such as the HuggingFace platform [23], enables scientists to implement automatic categorization methods based on machine learning algorithms and deep learning methods. One machine learning approach is supervised learning, which refers to training algorithms from already labeled data. Whereas unsupervised learning relates to classification without any human input, the semisupervised learning paradigm combines a small amount of labeled data provided by a human expert with a large amount of unlabeled data.

Although other studies have reported the use of NLP and automatic categorization for specific components of COVID-19 data [24,25], this is the first study to apply semisupervised machine learning to all COVID-19 narratives. This paper describes the development of the Early AI-Supported Response with Social Listening (EARS) platform [26], data sourcing and collection from December 2020 to February 2022, and how AI was used to filter and categorize data to inform infodemic insights.

Methods

The WHO EARS Platform

The WHO EARS platform was developed and piloted to allow infodemic managers to access real-time information on how people are talking about COVID-19 on the web. The data are collected from multiple sources and combined into a user-friendly platform. Data are categorized as a public health taxonomy developed by experts to enable social understanding of COVID-19 pandemic narratives on social media platforms. This taxonomy was informed by previous work [12,13] but developed bespoke for this project and continually reviewed and revised. The taxonomy has 5 main global topics and 41 different categories. The categories are regularly reviewed and revised. The 5 main topics related to the cause of the virus, illness, treatment, interventions (including prevention), and information and misinformation. Multimedia Appendix 1 lists all categories and definitions.

Data visualizations on the platform include combinations of countries, categories of conversation, questions, and gender segments. Data can be shown by which narratives are most prominent, rising in prominence, or outliers compared with other countries in the platform. The public dashboard and public application programming interface are available for aggregated and anonymized data by all users [27], while a private dashboard is being piloted internally by WHO staff and selected partners, with anonymized and granular-level data. The dashboards are updated daily with new data and are intended to assist public health professionals in understanding the narratives and needs of the public to inform policy, communications decisions, or emergency response recommendations. The project was validated and piloted in 30 countries (Multimedia Appendix 2) and 9 languages (English, Spanish, French, Portuguese, German, Italian, Bahasa Indonesian, Thai, and Arabic). For the purposes of setting up and validating the pilot first version of the platform, languages and countries were selected after an internal review of countries that lack access to high-quality language-specific analysis of languages to COVID-19, which at the same time had sufficient data volume to enable analysis by the EARS system.

Data Sources and Data Collection

The platform collects daily data from web-based conversations in publicly available sources, including Twitter, Facebook Public Pages, and other common web content. However, there are differences in how people use these platforms. Twitter users often share information in real time and react to current events, trends, or topics that are popular in the moment, so information spreads quickly, allowing some events to be detected before they may appear in news outlets. Although Facebook can also be a place for spontaneous discussions, comments tend to be more associated with recent news events or articles, or institutional campaigns. Owing to the country-level analysis, data were collected only where geo-located metadata were available. Examples of common web sources include web-based forums, news comments, blogs, as well as sources such as Reddit, YouTube, and 4Chan. In addition, there are >1000 additional sources of public web- and interest-based communities whose volume varies per country or may be country specific, such as Mumsnet in the United Kingdom, Buriramexpat in Thailand, or Mutdawl in Egypt. These data are collected through the Citibeats platform via their partnerships with data providers, and all analyses focus on aggregated results instead of individual users. A query was developed to capture the COVID-19 data for each country and language. Multimedia Appendix 3 includes the search terms used in each country.

Although Twitter provides most of the data on the platform, efforts have been made to diversify the data collection as much as possible and to apply different methods, such as gender disaggregation, to mitigate any potential biases. To reduce the impact of viral events on Twitter, retweets are not collected, and the aim is to give the same weight and importance to any opinion published on Twitter without granting a more significant position to the most influential and viral voices.

Taxonomy Classification

Overview

After data are collected, they are categorized to the taxonomy using a machine learning algorithm. The main approach to classifying posts relies on a semisupervised learning method that relies on a measure propagation algorithm [28]. Semisupervised learning is a machine learning approach that combines labeled and unlabeled data to train models. It uses additional unlabeled data to improve generalization and leverage the underlying structure of the data. By incorporating unlabeled data, semisupervised learning algorithms can achieve better performance and scalability compared with using only labeled data [29]. This is useful when the labeled data are limited or expensive to obtain.

Measure propagation [28] is a more advanced version of the well-known label propagation algorithm [30]. Label propagation starts with a set of points and a graph structure connecting those points. This algorithm assigns labels to unlabeled data points by propagating the labels in the graph. After the propagation is completed, every data point is assigned to a specific label. Measure propagation is the probabilistic version of label propagation. Instead of propagating labels, measure propagation assigns to each data point a probability distribution over its labels based on the labels of the neighboring nodes in the graph.

We feed the algorithm with texts collected from social media platforms and carefully chosen sets of keywords for each category (eg, a category of fruit would have apple, banana, or grapes as keywords). A keyword can be 1 word, such as "symptom" or several words, such as "patient zero." Once the keywords were defined for the English language, they were translated and adapted to equivalent expressions to the other 8 languages. A list of keywords for each category in Spanish and English is provided in Multimedia Appendix 4.

In training the algorithm, a weighted graph is built from the texts based on the co-occurrences between words. The nodes of the graph are the texts from social posts, and the lines between the nodes represent the similarities between those social posts computed from the co-occurrences of words used between those social posts. Next, the algorithm uses a propagation method across the graph nodes to label the texts that are most likely to belong to this category. In this case, a sentence mentioning 3 fruits would have a high probability to fall into the fruit category. Then, to determine if other texts fall into the category, the algorithm propagates across the graph nodes, updating the categorization by considering the texts it has labeled in the previous step. The algorithm returns to the third step until the classification of the graph nodes converges to a stable state. The propagation step occurs daily.

Where a post does not fall into one of the categories, it is considered as "noise" and excluded from the data set. This occurs when the classification algorithm finds no relationship with any of the categories. Examples of posts considered as noise are posts containing only hashtags or user mentions or posts that contain a word from the taxonomy but are not related to the subject of the study.

Testing the Algorithm

To assess the performance of the algorithm in classifying data into categories, we computed the precision, recall, and F_1 metrics. Recall represents the amount of relevant information retrieved from social media posts, whereas precision refers to how much of the retrieved information is relevant [31]. The F_1-score is the harmonic mean between recall and precision. This metric displays the effectiveness of the retrieval model and considers recall and precision of equal importance and is widely used in machine learning [32-34].

To compare this method with related works [13], we created a baseline based on Boolean queries. We fed the Boolean queries and the machine learning algorithm with exactly the same keywords related to each category and computed the same performance metrics. To calculate the metrics for both approaches, we randomly sampled the data collected from the Citibeats platform and annotated them according to the taxonomy categories. The data were sampled between January and February 2022. For validation, we limited the data to Spanish (from Colombia, Spain, and Mexico) and English (from Kenya, India, and the United Kingdom) languages. For each language, the reviewers manually labeled the data, with a single reviewer leading each language. A second reviewer reviewed a random sample (15% of each language data set) to ensure the quality of annotations; any disagreements were resolved via discussion, and the annotations were rereviewed. The reviewers labeled each text according to the categories they should fall into. For each category, we considered the following text:

1. True positive (TP)—text correctly classified
2. False positive (FP)—text did not belong in the category it was assigned.
3. False negative—text belongs to the category but was not classified into it.

These data were then used to calculate the precision = TP / (TP + FP); recall = TP / (TP + false negative); and finally, the F_1-score = (2 × precision × recall) / (precision + recall). As there were 41 taxonomy categories, we computed this for each category and as an overall average of the scores across all categories. Hotelling T^2 [35] was used to determine whether the difference between the categorization methods (Boolean or algorithm) was statistically significant. As Hotelling T^2 is an *omnibus* test, it indicates whether the combined dependent variables are statistically significantly different in terms of the 2 classification methods. A significance level of $P<.05$ was used.

Demographics and Intents

To enable a more fine-grained analysis and understanding of the representativity of the data, we developed a tool to filter by post origin, poster type (individual or institution), and gender. The ability to detect which posts are from people and which posts are from institutions (including official accounts) is important as institutions' discussions bring relevant information about COVID-19's narratives but may differ from citizens' discussions.

The gender classifier is inferred from the web-based data using a deep learning method. The system uses indicators, such as the author's name and biodescription, and makes a final determination of gender probability. Women are underrepresented in the media [36], and the ability to filter by gender can help isolate and amplify the voices of women.

The tool for gender differentiation extrapolated on existing methods [21,32]. A limitation of the platform is that it currently only supports dominations of men or women, and including nonbinary categories is planned for future integration. For this research, we calculated the known gender proportion of users in our data set and compared it with country data from the study by Hootsuite [37,38] to determine the proportionality of users in our database to the population.

To extract more insightful and actionable information from the analyses, we also implemented a query detector to detect whether citizens' texts contained questions. The detector is a multilingual machine learning algorithm based on multilingual BERT architecture [39] that detects whether social posts carry a question. The classifier discards from the classification all rhetorical questions, quotations, advertisements, newspaper headlines, or questions with an answer. We used Mexico and the United Kingdom as comparator examples for reporting on all demographic filters.

Category Analysis

Calculating the velocity of data change is important for identifying information voids and for the early identification of changes in narratives. Velocity refers to the percentage increase in narratives in a certain category. To identify the velocity change in a category, we compared the weekly volume of social media posts with the moving average volume over the last 4 weeks. Intuitively, the moving average series is smoother, less subject to variation, and represents the trend of volume over time. It also limits FP velocity alerts when we have an alternate series. A category is flagged as a velocity alert when the new weekly volume change is a minimum of 15% higher than the mean of the previous week. A 15% threshold was established based on analysts' experience with velocity data. The formula used to calculate the weekly velocity rates is as follows:

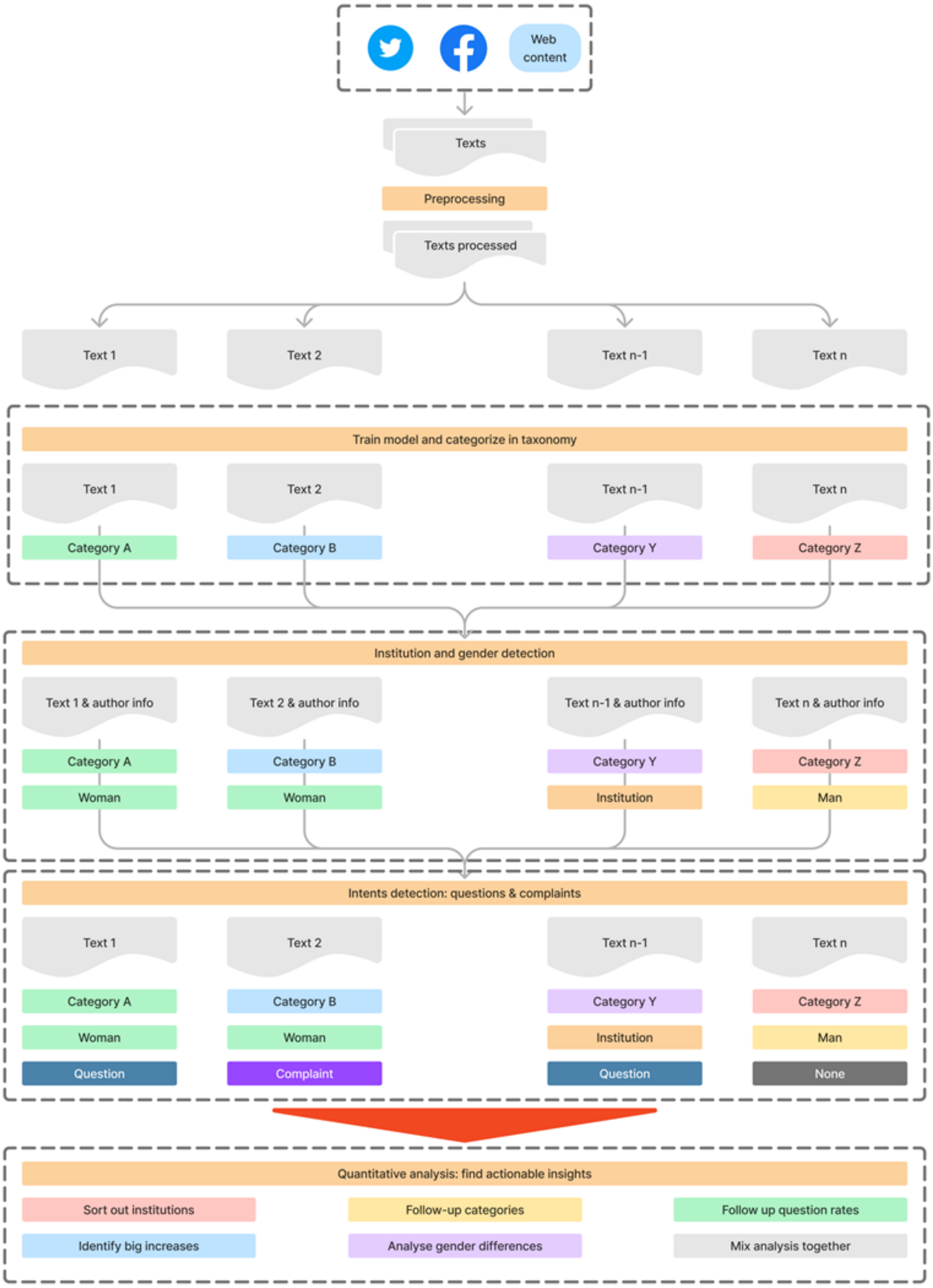
Each week, a velocity (V) is computed according to current and previous weekly data volumes (Vol). The velocity for the week i (V_i) is the relative difference between the current week data volume (Vol_i) and the mean of the data volume across the last 4 weeks.

The 4-week period was chosen because it provided enough time to have a stable comparison while still maintaining recency. The specific number of weeks was defined after testing with higher and lower numbers of weeks and reviewing the relevant changes. After applying data analysis to the different trials per week, the best results were obtained with the 4-week comparison, providing alerts most aligned with the insights needed to inform decision-making regarding COVID-19.

Velocity metrics can be combined with other filters to analyze category velocity through demographic segments, question rates, or combining demographic segments with question rates.

We present the results of the velocity data by filtering for 2021 for Mexico and the United Kingdom. Figure 1 summarizes the entire data processing and analysis pathway, including data collection, categorization into the taxonomy, demographics segmentation, intents detection, and category analysis.

Figure 1. Early Artificial Intelligence–Supported Response With Social Listening (EARS) data processing and analysis method. Texts are first collected from social networks, we then train the semisupervised algorithm to categorize the texts to the taxonomy. Following this, the organization and gender and question intentions are applied, and finally, the data generated is ready for user analysis.



Results

Data Collection

Data were collected for English-, French-, Portuguese-, and Spanish-speaking countries from December 2020 to February 2022. Arabic, Bahasa Indonesian, German, Italian, and Thai languages were added in September 2021; thus, only data from this time point onward are available, resulting in a lower data volume in comparison. In total, 215,469,045 social posts were collected for processing, including 8.5 million posts for Mexico and 29 million for the United Kingdom. Table 1 presents the total number of posts collected before classification into specific categories. Twitter data are overrepresented, accounting for 93.31% (188,644,046/202,177,384) of all data. The United

States, the United Kingdom, and Brazil accounted for the highest data volumes.

Table 1. Total number of data collected per country.

Country	Total posts (N=202,177,384), n (%)	Twitter (n=188,644,046; 93.31%), n (%)	Web content (n=12,208,477; 6.04%), n (%)	Facebook (n=1,324,861; 0.66%), n (%)
Angola	77,666 (100)	39,440 (50.78)	37,349 (48.09)	877 (1.13)
Brazil	21,361,298 (100)	21,120,334 (98.87)	198,439 (0.93)	42,525 (0.2)
Canada	13,521,028 (100)	12,554,897 (92.85)	898,629 (6.65)	67,502 (0.5)
Colombia	3,367,877 (100)	3,239,763 (96.2)	122,244 (3.63)	5870 (0.17)
Democratic Republic of Congo	185,019 (100)	180,488 (97.55)	3731 (2.02)	800 (0.43)
Egypt	745,004 (100)	377,304 (50.64)	339,765 (45.61)	27,935 (3.75)
France	8,754,682 (100)	8,379,214 (95.71)	274,578 (3.14)	100,890 (1.15)
India	10,539,028 (100)	9,876,409 (93.71)	560,071 (5.31)	102,548 (0.97)
Indonesia	4,930,431 (100)	4,681,810 (94.96)	220,653 (4.48)	27,968 (0.57)
Iraq	352,335 (100)	207,738 (58.96)	130,138 (36.94)	14,459 (4.1)
Jordan	301,000 (100)	173,455 (57.63)	93,404 (31.03)	34,141 (11.34)
Kenya	666,320 (100)	587,124 (88.11)	67,273 (10.10)	11,923 (1.79)
Malaysia	2,308,285 (100)	1,607,409 (69.64)	674,606 (29.23)	26,270 (1.14)
Malta	149,348 (100)	74,474 (49.87)	68,065 (45.57)	6809 (4.56)
Mexico	8,542,827 (100)	8,216,404 (96.18)	310,731 (3.64)	15,692 (0.18)
Morocco	293,093 (100)	159,184 (54.31)	127,842 (43.62)	6067 (2.07)
Nicaragua	134,175 (100)	124,389 (92.71)	9062 (6.75)	724 (0.54)
Nigeria	1,203,045 (100)	1,048,254 (87.13)	115,877 (9.63)	38,914 (3.23)
Peru	1,696,642 (100)	1,339,361 (78.94)	345,481 (20.36)	11,800 (0.7)
Philippines	2,116,945 (100)	1,641,901 (77.56)	446,545 (21.09)	28,499 (1.35)
Senegal	142,566 (100)	108,810 (76.32)	24,596 (17.25)	9160 (6.43)
South Africa	3,057,165 (100)	2,697,841 (88.25)	237,888 (7.78)	121,436 (3.97)
Spain	8,904,426 (100)	8,576,075 (96.31)	266,925 (3)	61,426 (0.69)
Switzerland	985,164 (100)	954,406 (96.88)	21,744 (2.21)	9014 (0.91)
Thailand	1,488,541 (100)	1,351,646 (90.80)	95,071 (6.39)	41,824 (2.81)
Trinidad y Tobago	162,215 (100)	140,189 (86.42)	20,652 (12.73)	1374 (0.85)
The United Kingdom	28,955,804 (100)	25,641,890 (88.56)	3,218,910 (11.12)	95,004 (0.33)
The United States	75,975,581 (100)	72,497,816 (95.42)	3,088,538 (4.07)	389,227 (0.51)
Uruguay	1,066,035 (100)	941,735 (88.34)	114,272 (10.72)	10,028 (0.94)
Yemen	193,839 (100)	104,286 (53.8)	75,398 (38.9)	14,155 (7.3)

Taxonomy Classification

Once the data were collected, 34.44% (74,214,770/215,469,045) were categorized into the taxonomy categories, with the rest dismissed as not carrying relevant information or being considered as noise. To test the ability of the algorithm, 3888 social media posts were sourced from both the Spanish and English data sets. For Spanish language, this was an average of 80.9 (SD 65) texts per category, and for English, 86.8 (SD 44.4) texts per category. Precision, recall, and F_1-score are shown as percentages in Table 2 for all categories. We can see that the machine learning approach outperforms the Boolean query for both languages. The Hotelling T^2 test demonstrated a statistically significant difference between the 2 categorization methods for both languages (P≤.001).

The largest difference was observed in precision, with a 16-point difference in English, and a 9.5-point difference for Spanish. The machine learning algorithm disambiguates the meaning of some words that are not possible using Boolean queries. Table 3 shows some examples in which the algorithm correctly categorizes the text.

The same calculations were run for each category, with the full results available in Multimedia Appendix 5. We found that the machine learning algorithm outperformed the Boolean query method for 80% (33/41) of categories for recall, 88% (36/41) for precision, and 93% (38/41) for F_1-scores for the English language. For Spanish, there was a better recall (26/41, 63%), better precision (34/41, 83%), and better F_1-scores (32/41, 78%). Globally, the results were better for the English language when compared with Spanish. The F_1-score was 20 points higher for English than that for Spanish.

Table 2. Precision and recall for algorithm versus Boolean methods per language.

Language and model	Precision (%)	Recall (%)	F_1-score (%)
English			
Algorithm	67.54	81.57	73.90
Boolean	51.45	74.75	60.95
Spanish			
Algorithm	45.28	64.79	53.30
Boolean	35.77	62.36	45.46

Table 3. Examples of categorization errors using Boolean query that are well predicted by the semisupervised algorithm.

Input	Algorithm			
	Semisupervised learning		Boolean queries	
	Category predicted	Result	Category predicted	Result
r aware of 666 is name written in da bible generated frm da word corona now you know antichrist NEED God almighty intervene n manifest Himself only God can save the world from falsehood n demonic china virus	Faith	True positive	The Cause of the Virus	False positive
Our Summary Report is your essential guide to update data in the most promising markets, covering: • Global market stats; • Market intel on key MENA and SEA countries; • Hot topics like online schools & the impact of COVID-19. *Order a copy*	Statistics & Data	True positive	Transmission Settings	False positive
'What, i rarely test myself xx like today spoke with a workmate said a had a sore throat, ear ache and he said take a test, i am like No its not covid'	Other Discussed Symptoms	True positive	Testing	False positive

Demographics and Intent

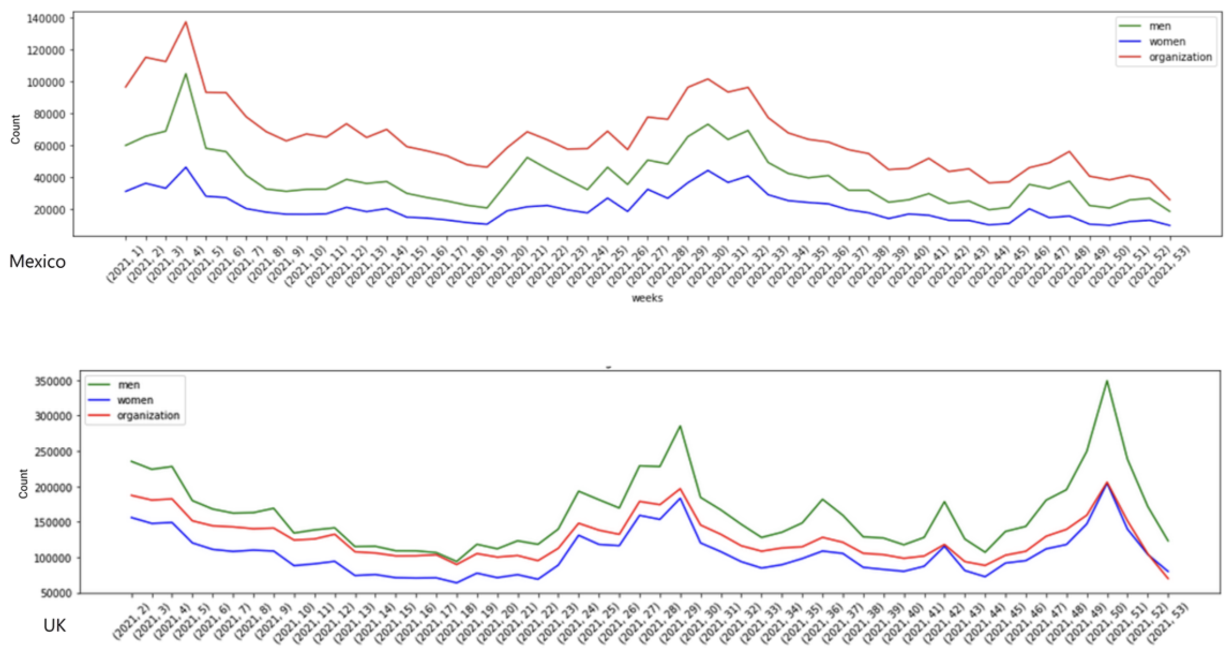
The gender split of data gathered for the EARS platform was largely consistent with population-level data on social media use from Hootsuite [37,38]. Table 4 presents a comparison of the gender proportion output from EARS using the demographics segmentation algorithm across 2 countries, the United Kingdom and Mexico, for 2021 compared with Hootsuite data from surveys on social media in the same year.

For both countries, we mapped the number of social media posts by demographic segment for Mexico and in the United Kingdom (Figure 2). In Mexico, organizations made a higher number of posts than those from either men or women. Data from institutions represented almost half of the discussions in 2021. In the United Kingdom, data from men were more highly represented than those from women by more than 100,000 posts per week. Institutions were also more highly represented than women. We can see from the data that posting patterns are largely similar across demographic types.

Table 4. Gender proportion of social media users in Early Artificial Intelligence–Supported Response with Social Listening (EARS) data set, compared with Hootsuite country-level data.

Country and gender	EARS, n (%)	Hootsuite (%)
Mexico (n=3,961,325 posts)		
Men	2,574,861 (65)	61.1
Women	1,386,463 (35)	38.9
The United Kingdom (n=17,311,716 posts)		
Men	10,508,211 (60.7)	60.1
Women	6,803,504 (39.3)	39.9

Figure 2. Number of social media posts per week over time in Mexico, (top) and United Kingdom (bottom) shown by those posted by men (green), women (blue) and organizations (red).



Category Analysis

Velocity data for Mexico and the United Kingdom were included. Table 5 lists the categories with the highest number of velocity alerts between December 2020 and February 2022. These data were differentiated by gender. Data from all categories over this time are included in Multimedia Appendix 6. The mean number of velocity alerts across all categories in the aggregate is 28.96 for Mexico and 26.67 for the United Kingdom.

The top 3 categories for Mexico were “Youth” (39 weeks with velocity higher than 15%), “General Vaccine Discussion” (38 weeks with velocity higher than 15%), and “Modes of Transmission” (37 weeks with velocity higher than 15%). For the United Kingdom, the categories with the highest velocity alerts were “Immunity” (40 weeks with velocity higher than 15%), “Modes of Transmission” (35 weeks with velocity higher than 15%), and “Civil Unrest” (34 weeks with velocity higher than 15%). In Mexico, there are differences in category velocity between genders. For women, there were high increases in “Modes of Transmission” (21 weeks with velocity higher than 15%), “Youth” (20 weeks with velocity higher than 15%) and “General Vaccine Discussion” (19 weeks with velocity higher than 15%), whereas for men, “Statistics & Data” (19 weeks with velocity higher than 15%), “Youth” (19 weeks with velocity higher than 15%), and “General Vaccine Distinction” (19 weeks with velocity higher than 15%) were the highest. In the United Kingdom, the top velocity categories for men in 2021 were “Immunity” (21 weeks with velocity higher than 15%), “Transmission Settings” (19 weeks with velocity higher than 15%), and “Health Care Workers & Vaccine” (19 weeks with velocity higher than 15%), whereas for women, it was “Immunity” (19 weeks with velocity higher than 15%), “Modes of Transmission” (18 weeks with velocity higher than 15%), and “Civil Unrest” (18 weeks with velocity higher than 15%).

Data can also be filtered by post intent, that is, whether it is a questioning post. Combining these filters—the taxonomy categories, the demographics segments, and the query detector—helps in the quantitative analysis of narrative change. Table 6 presents the number of weeks with a velocity alert for questioning posts for men and women in the United Kingdom and Mexico. These data were restricted to individuals (not institutions). The list of all category velocity changes by question for this time is included in Multimedia Appendix 7. The mean number of velocity alerts across all categories for social media posts carrying questions, for both genders, is 31.83 Mexico and 23.2, the United Kingdom.

Table 6 provides valuable insights into which topics have suddenly raised more questions between December 2020 and February 2022. In Mexico, “Other Discussed Symptoms” (45 weeks with velocity higher than 15% for both genders combined), “Measures in Public Settings” (42 weeks with velocity higher than 15% for both genders combined), and “Vaccine Distribution and Policies on Access” (42 weeks with velocity higher than 15% for both genders combined) generated the most velocity alerts. In the United Kingdom, velocity alerts were oriented toward “Health Technology” (40 weeks with velocity >15% for both genders combined), “Digital Health Technology” (39 weeks with velocity higher than 15% for both genders combined), and “Other Discussed Symptoms” (39 weeks with velocity higher than 15% for both genders combined).

This level of data filtering also provides more visibility by gender in specific areas of concern. In Mexico, the highest velocity question change for men was “General Vaccine

Discussion" (23 weeks with velocity higher than 15% for both genders combined) and "Immunity" (22 weeks with velocity higher than 15% for both genders combined), whereas for women, it was "Myths" (22 weeks with velocity higher than 15% for both genders combined) and "Supportive Care" (21 weeks with velocity higher than 15% for both genders combined). In the United Kingdom, the highest number of changes for men was "Pandemic Fatigue" (18 weeks with velocity higher than 15% for both genders combined), whereas for women, it was "Stigma about the Spread" (18 weeks with velocity higher than 15% for both genders combined).

The table also highlights that some topics raised many more questions at certain points for men than for women and vice versa. For instance, in Mexico, "Pandemic Fatigue" raised twice the number of velocity alerts for men than for women, whereas the "Current Treatments" topic raised 6 more weekly velocity alerts for women. In the United Kingdom, "Reduction of Domestic Movement" was of higher concern for men, whereas "Faith" and "Civic Unrest" were higher among women.

The monitoring system provided valuable insights throughout the pandemic by providing analysts with information to inform further investigations and focus on attention. This included narrative identification at the country level regarding concerns about vaccine side effects; questions about the cause of the virus at different time points, and particularly early on in the pandemic; and concerns and questions about specific measures introduced to reduce movement.

Table 5. Number of velocities alerts by category and gender from December 2020 to February 2022.

Categories	Mexico, n (%)			The United Kingdom, n (%)		
	Total number of velocity alerts	Velocity alerts by gender		Total number of velocity alerts	Velocity alerts by gender	
		Men	Women		Men	Women
Youth	39 (100)	19 (49)	20 (51)	30 (100)	16 (53)	14 (47)
General vaccine discussion	38 (100)	19 (50)	19 (50)	30 (100)	15 (50)	15 (50)
Modes of transmission	37 (100)	16 (43)	21 (57)	35 (100)	17 (49)	18 (51)
Statistics and data	33 (100)	19 (58)	14 (42)	30 (100)	15 (50)	15 (50)
Transmission settings	31 (100)	17 (55)	14 (45)	33 (100)	18 (55)	15 (45)
Pandemic fatigue	31 (100)	18 (58)	13 (42)	24 (100)	13 (54)	11 (46)
Health care workers and vaccine	31 (100)	16 (52)	15 (48)	32 (100)	18 (56)	14 (44)
Digital health technology	30 (100)	13 (43)	17 (57)	20 (100)	11 (55)	9 (45)
Civil unrest	28 (100)	14 (50)	14 (50)	34 (100)	16 (47)	18 (53)
Immunity	36 (100)	18 (50)	18 (50)	40 (100)	21 (52)	19 (48)

Table 6. Number of velocities alerts by category question rate and gender from December 2020 to February 2022.

Categories	Mexico, n (%)			The United Kingdom, n (%)		
	Total number of velocity alerts in questions	Velocity alerts by gender—questions		Total number of velocity alerts in questions	Velocity alerts by gender—questions	
		Men	Women		Men	Women
Other discussed symptoms	45 (100)	25 (56)	20 (44)	35 (100)	18 (51)	17 (49)
Measures in public settings	42 (100)	21 (50)	21 (50)	22 (100)	10 (45)	12 (55)
Vaccine distribution and policies on access	42 (100)	21 (50)	21 (50)	21 (100)	8 (38)	13 (62)
Myths	42 (100)	20 (48)	22 (52)	25 (100)	10 (40)	15 (60)
Digital health technology	40 (100)	19 (48)	21 (52)	39 (100)	21 (54)	18 (46)
Faith	40 (100)	20 (50)	20 (50)	25 (100)	8 (32)	17 (68)
General vaccine discussion	39 (100)	23 (59)	16 (41)	23 (100)	10 (43)	13 (57)
Health technology	39 (100)	19 (49)	20 (51)	40 (100)	19 (48)	21 (52)
Supportive care	38 (100)	17 (45)	21 (55)	18 (100)	9 (50)	9 (50)
Immunity	39 (100)	22 (56)	17 (44)	20 (100)	9 (45)	11 (55)
Stigma around the spread	32 (100)	16 (50)	16 (50)	30 (100)	12 (40)	18 (60)
Civil unrest	32 (100)	16 (50)	16 (50)	30 (100)	12 (40)	18 (60)
Pandemic fatigue	30 (100)	20 (67)	10 (33)	34 (100)	18 (53)	16 (47)
Reduction of domestic movement	29 (100)	13 (45)	16 (55)	18 (100)	13 (72)	5 (28)
Current treatments	26 (100)	10 (38)	16 (62)	24 (100)	10 (42)	14 (58)

Discussion

Principal Findings

The configuration and application of the EARS platform have enabled WHO infodemic managers to access categorized data throughout the pandemic to inform responses.

Compared with other analytical methods that are more manual, required data scientists in the team, or had fewer analytics capabilities, the EARS platform has enabled progress toward more scalable and sustainable analysis of social media for public health action. Real-time data collection and categorization were fully automated, enabling a self-serving model of use. This study has demonstrated the strength of a machine learning algorithm for categorizing COVID-19 narratives into a public health taxonomy. The proposed approach outperformed the Boolean query strategy across all metrics. This approach has not only allowed data to be quickly categorized and useful for infodemic managers but also to change and grow as conversations shift. The algorithm disambiguates the meaning of some words considering co-occurrences and other words in close proximity, considering the context around the keywords rather than just the word. This has enabled the algorithm to include new and emerging words and phrases, which is essential as the infodemic around COVID-19 has moved at such a pace. However, human inputs are still needed for contextualization and translation into actionable insights. The separation of the overall global conversation into country-level, category-specific, intent-specific, and demographically segmented analytics is a significant step toward obtaining more usable and useful data for decision-making.

Other studies have combined machine learning and COVID-19 pandemic data, mainly focusing on 1 component of the information analysis. In 1 study, NLP bots were trained to detect misinformation on the Reddit platform by fine-tuning the BERT model [40]. Another United Kingdom–based study applied sentiment analysis on COVID-19 mental health–related tweets [41]. An NLP study developed to recognize COVID-19 symptoms described in social media posts and used them for disease surveillance and detection reported useful data to identify the prevalence and severity of the symptoms [42]. Other studies explored vaccine-related tweets to identify vaccination-related topics [43,44]. This paper has reported on the categorization of all COVID-19–related narratives, across multiple languages and countries, and adds to the evidence on how infodemic insights can be identified during a health emergency.

The methodology described herein has several strengths. Most social media analysis research is conducted in high-income countries [2], and research and tools that focus on low- and middle-income countries are required. Table 2 presents how overrepresented the American and European regions are in this data set, accounting for 85% of all data collected. By allowing for filtering at the regional and country levels, we can prioritize narratives from other regions to better understand global trends. Although the limitation—that smaller topics of conversation that are gathering velocity and volume are "drowned out" by the major narratives—has been partially addressed by showing relative importance across countries on smaller narratives, more needs to be done on emerging topic detection of topics and increasing diversity of data sources. The language-agnostic approach has enabled scaling and addition of more languages. Filtering by questions and gender allows analysts to quickly find meaning in the data.

Moreover, recent proposals for amendments to International Health Regulations emphasize the need to expand the use of digital applications for public health and call for increasing country infrastructure and human capacities to use technology to support public health [45,46]. This highlights the increased burden that individual infodemics have placed on emergency response structures and health systems and requires intensified investment in systematic strategies for infodemic management.

There are clear areas for future research and work. The area of network analysis to identify narrative amplifiers (eg, users or communities sharing misinformation narratives or sharing information across thematic and interest communities) or social structures about how information flows in the network is a promising tool for infodemic managers. However, this approach would need to carefully balance the potential trade-off between valuable actionable information and user privacy. At an aggregated level, further segmentation of groups can be accomplished (such as studying conversations by health care professionals or disaggregating organizational content). The machine learning tool, along with demographic segmentation, helps gain insights from data trends. Knowing where questions are rising, for example, can help to identify information voids and opportunities to target public health advice and information materials.

Although digital social media analysis platforms are an important part of understanding community perceptions and concerns, an integrated analysis of infodemic insights, including the combination of offline and web-based sources to triangulate data, is needed [12]. In addition to expanding the data sets in EARS to include more diverse data, there will always be limitations in infodemic insights data sources from digital platforms and web-based public conversations, making integrated analysis a vital and necessary step in the generation of infodemic insights and recommendations for action. Although the EARS dashboards can provide rich data and "signal" to analysts, human interpretation is needed to contextualize and translate this into actionable insights.

Limitations

This study has some limitations. The machine learning model assumes mutual exclusion for categories and only enables mono-label predictions. This means that we predicted only one category for each text. Conversations about COVID-19 may contain threads that cross several categories, and the inability to multitag data across categories may result in data meaning being lost from analysis within categories that would be relevant. A principal path of improvement would be to extend the model to multilabel classification.

As noted in previous research, the analysis of web-based conversations in academic and marketing sector analytics continues to have overreliance on Twitter data, which has limited representativity even if we can make demographics identification to limit the biases. This is also the case for the EARS platform. The platform will continue to gradually add more data sources, which can expand its coverage and representativity. In addition, comparing with non–text-based sources can help increase representativity, as well as provide a more rounded view of the data. Integrating fact-checker and misinformation data sets could yield more insight into the drivers and effects of the infodemic in the digital data sources. EARS provides a public health–relevant tool for the analysis of digital and web-based data sources in several languages and therefore improves the analysis of data from internet platforms. However, insights that are gleaned from EARS must still be integrated with intelligence from other data sources that cover communities, behaviors, users, and information seeking to inform infodemic management strategies.

The categorization rate of 37% means that most data identified as COVID-19 related is uncategorized. This limitation is addressed by a regular review of uncategorized data and updating of seed words. Although geo-located data are required for country-level analysis, this limits the amount of data that can be included. Gender data were currently disaggregated by men and women only. This is a limitation in interpreting the views of those who were not identified as men or women. The platform currently does not segment by age, which is another planned addition.

Conclusions

The WHO EARS platform described here has been specifically developed to address the changing needs of public health analysts during the COVID-19 pandemic. The platform was designed to allow scalability and iterations, and new countries and languages have been added. The platform's digital architecture has overcome many of the challenges inherent in social media analysis; however, more work is still required. An integrated analysis of other data sources, including offline sources, is still needed, and the implementation of multilabel classification will further aid analysts. Although much of the categorization is automated, human analysts are still needed to contextualize and triangulate the data to create actionable insights.

The application of a public health taxonomy and AI technology to a user-friendly social listening platform, accessible directly by analysts, is a significant step toward a better understanding of global narratives. The scalability and rounds of review and iteration mean that the platform will continue to evolve throughout the pandemic and respond to user needed. As new global emergencies emerge, a key challenge will be remaining agile enough to pivot as needed and incorporate emerging trends.

The WHO EARS platform has applied novel analytic approaches to improve the generation of infodemic intelligence and, therefore, strengthen the evidence base for infodemic management. As the platform matures, there is an increased opportunity for countries to adopt it, or similar tools, in infodemic insight analysis.

Acknowledgments

The authors acknowledge many users from the World Health Organization (WHO) infodemic manager community who have provided useful feedback and testing of the Early Artificial Intelligence–Supported Response With Social Listening (EARS) platform since its launch.

Disclaimer

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.

Conflicts of Interest

AG, LK, AL, and HW are employees of a global digital intelligence company that provided platform development services to the World Health Organization (WHO) as part of a contractual agreement. BKW, TN, BY, AI, MD, CS, RSS, KR, DS, SB, and TDP are consultants or staff of the WHO or the Pan American Health Organization. The authors alone are responsible for the views expressed in this paper and they do not necessarily represent the decisions, policies, or views of the organization with which they are affiliated.

Multimedia Appendix 1
Early Artificial Intelligence–Supported Response With Social Listening (EARS) COVID-19 taxonomy categories and definition.
[DOCX File , 26 KB - infodemiology_v3i1e47317_app1.docx]

Multimedia Appendix 2
Pilot countries.
[DOCX File , 20 KB - infodemiology_v3i1e47317_app2.docx]

Multimedia Appendix 3
COVID-19 keywords and queries to collect social media posts for each country.
[DOCX File , 28 KB - infodemiology_v3i1e47317_app3.docx]

Multimedia Appendix 4
English and Spanish Keywords for each individual category.
[DOCX File , 56 KB - infodemiology_v3i1e47317_app4.docx]

Multimedia Appendix 5
Precision and recall metrics for algorithm versus Boolean methods for all categories.
[DOCX File , 38 KB - infodemiology_v3i1e47317_app5.docx]

Multimedia Appendix 6
Number of velocities alerts by category from December 2020 to February 2022.
[DOCX File , 35 KB - infodemiology_v3i1e47317_app6.docx]

Multimedia Appendix 7
Number of velocities alerts by category and question filter from December 2020 to February 2022.
[DOCX File , 35 KB - infodemiology_v3i1e47317_app7.docx]

References

1. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
2. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. J Med Internet Res 2022 Aug 04;24(8):e37367 [FREE Full text] [doi: 10.2196/37367] [Medline: 35816685]
3. Islam MS, Sarkar T, Khan SH, Mostofa Kamal AH, Hasan SM, Kabir A, et al. COVID-19-related infodemic and its impact on public health: a global social media analysis. Am J Trop Med Hyg 2020 Oct;103(4):1621-1629 [FREE Full text] [doi: 10.4269/ajtmh.20-0812] [Medline: 32783794]

4. Pickles K, Cvejic E, Nickel B, Copp T, Bonner C, Leask J, et al. COVID-19 misinformation trends in Australia: prospective longitudinal national survey. J Med Internet Res 2021 Jan 07;23(1):e23805 [FREE Full text] [doi: 10.2196/23805] [Medline: 33302250]
5. WHO policy brief: COVID-19 infodemic management. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-Policy_Brief-Infodemic-2022.1 [accessed 2023-06-30]
6. Calleja N, AbdAllah A, Abad N, Ahmed N, Albarracin D, Altieri E, et al. A public health research agenda for managing infodemics: methods and results of the first WHO infodemiology conference. JMIR Infodemiology 2021 Sep 15;1(1):e30979 [FREE Full text] [doi: 10.2196/30979] [Medline: 34604708]
7. Mehmet M, Heffernan T, Algie J, Forouhandeh B. Harnessing social listening to explore consumer cognitive bias: implications for upstream social marketing. J Soc Mark 2021 Jul 27;11(4):575-596 [FREE Full text] [doi: 10.1108/jsocm-03-2021-0067]
8. Agenda of 4th infodemic management conference. Advances in social listening for public health. World Health Organization. Geneva: WHO; 2021. URL: https://www.who.int/news-room/events/detail/2021/05/04/default-calendar/4th-virtual-who-infodemic-management-conference-advances-in-social-listening-for-public-health [accessed 2023-06-30]
9. Purnat TD, Nguyen T, Ishizumi A, Yau B, White B, Cecchini S, et al. Delivering actionable infodemic insights and recommendations for the COVID-19 pandemic response. Wkly Epidemiol Rec 2022 Jul 08;97(27):313-324 [FREE Full text] [doi: 10.1093/eurpub/ckac129.645]
10. Rubinelli S, Purnat TD, Wilhelm E, Traicoff D, Namageyo-Funa A, Thomson A, et al. WHO competency framework for health authorities and institutions to manage infodemics: its development and features. Hum Resour Health 2022 May 07;20(1):35 [FREE Full text] [doi: 10.1186/s12960-022-00733-0] [Medline: 35525924]
11. Gruzd A, De Domenico M, Sacco PL, Briand S. Studying the COVID-19 infodemic at scale. Big Data Soc 2021 Jun 10;8(1):1-6 [FREE Full text] [doi: 10.1177/20539517211021115]
12. 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]
13. Purnat TD, Vacca P, Burzo S, Zecchin T, Wright A, Briand S, et al. WHO digital intelligence analysis for tracking narratives and information voids in the COVID-19 infodemic. Stud Health Technol Inform 2021 May 27;281:989-993. [doi: 10.3233/SHTI210326] [Medline: 34042821]
14. Public health taxonomy for social listening on monkeypox conversations. World Health Organization. 2022 Aug 17. URL: https://www.who.int/publications/m/item/public-health-taxonomy-for-social-listening-on-monkeypox-conversations [accessed 2023-06-30]
15. Luengo-Oroz M, Hoffmann Pham K, Bullock J, Kirkpatrick R, Luccioni A, Rubel S, et al. Artificial intelligence cooperation to support the global response to COVID-19. Nat Mach Intell 2020 May 22;2(6):295-297 [FREE Full text] [doi: 10.1038/s42256-020-0184-3]
16. Newberry C, Macready H. What is social listening, why it matters + 14 tools to help. Hootsuite Inc. 2021. URL: https://blog.hootsuite.com/social-listening-business/ [accessed 2023-06-30]
17. Syed Abdul S, Ramaswamy M, Fernandez-Luque L, John O, Pitti T, Parashar B. The pandemic, infodemic, and people's resilience in India: viewpoint. JMIR Public Health Surveill 2021 Dec 08;7(12):e31645 [FREE Full text] [doi: 10.2196/31645] [Medline: 34787574]
18. Beliga S, Martinčić-Ipšić S, Matešić M, Petrijevčanin Vuksanović I, Meštrović A. Infoveillance of the Croatian online media during the COVID-19 pandemic: one-year longitudinal study using natural language processing. JMIR Public Health Surveill 2021 Dec 24;7(12):e31540 [FREE Full text] [doi: 10.2196/31540] [Medline: 34739388]
19. Masngut N, Mohamad E. Association between public opinion and Malaysian government communication strategies about the COVID-19 crisis: content analysis of image repair strategies in social media. J Med Internet Res 2021 Aug 04;23(8):e28074 [FREE Full text] [doi: 10.2196/28074] [Medline: 34156967]
20. Townsend L, Wallace C. Social media research: a guide to ethics. University of Aberdeen, Scotland. 2016. URL: https://www.gla.ac.uk/media/Media_487729_smxx.pdf [accessed 2023-06-30]
21. Graells-Garrido E, Baeza-Yates R, Lalmas M. Representativeness of abortion legislation debate on Twitter: a case study in Argentina and Chile. In: Proceedings of the 2020 Web Conference. 2020 Presented at: WWW '20; April 20-24, 2020; Taipei, Taiwan p. 765-774 URL: https://dl.acm.org/doi/10.1145/3366424.3383561 [doi: 10.1145/3366424.3383561]
22. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv. Preprint posted online October 11, 2018. [FREE Full text] [doi: 10.48550/arxiv.1810.04805]
23. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. Transformers: state-of-the-art natural language processing. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. 2020 Presented at: EMNLP '20; November 16-20, 2020; Virtual Event p. 38-45 URL: https://aclanthology.org/2020.emnlp-demos.6.pdf [doi: 10.18653/v1/2020.emnlp-demos.6]
24. Sun C, Qiu X, Xu Y, Huang X. How to fine-tune BERT for text classification? In: Proceedings of the 18th China National Conference on Chinese Computational Linguistics. 2019 Presented at: CCL '19; October 18–20, 2019; Kunming, China p. 194-206 URL: https://link.springer.com/chapter/10.1007/978-3-030-32381-3_16 [doi: 10.1007/978-3-030-32381-3_16]

25. Li B, Si X, Lyu MR, King IK, Chang EY. Question identification on Twitter. In: Proceedings of the 20th ACM international conference on Information and knowledge management. 2011 Presented at: CIKM '11; October 24-28, 2011; Glasgow, Scotland p. 2477-2480 URL: https://dl.acm.org/doi/10.1145/2063576.2063996 [doi: 10.1145/2063576.2063996]
26. Purnat TD, Wilson H, Nguyen T, Briand S. EARS - a WHO platform for AI-supported real-time online social listening of COVID-19 conversations. Stud Health Technol Inform 2021 May 27;281:1009-1010. [doi: 10.3233/SHTI210330] [Medline: 34042825]
27. WHO EARS platform. World Health Organization. URL: https://www.who-ears.com [accessed 2023-06-30]
28. Subramanya A, Bilmes J. Semi-supervised learning with measure propagation. J Mach Learn Res 2011 Nov;12:3311-3370 [FREE Full text] [doi: 10.1007/bf00114803]
29. Zhu X, Goldberg AB. Introduction to statistical machine learning. In: Zhu X, Goldberg AB, editors. Introduction to Semi-Supervised Learning. Cham, Switzerland: Springer; 2009:1-8.
30. Xiaojin Z, Zoubin G. Learning from labeled and unlabeled data with label propagation. Carnegie Mellon University. 2002. URL: http://reports-archive.adm.cs.cmu.edu/anon/cald/CMU-CALD-02-107.pdf [accessed 2023-06-30]
31. Benítez-Andrades JA, Alija-Pérez JM, Vidal ME, Pastor-Vargas R, García-Ordás MT. Traditional machine learning models and bidirectional encoder representations from transformer (BERT)-based automatic classification of tweets about eating disorders: algorithm development and validation study. JMIR Med Inform 2022 Feb 24;10(2):e34492 [FREE Full text] [doi: 10.2196/34492] [Medline: 35200156]
32. Wang Z, Hale SA, Adelani DI, Grabowicz PA, Hartman T, Flöck F, et al. Demographic inference and representative population estimates from multilingual social media data. In: Proceedings of the 2019 International World Wide Web Conference. 2019 Presented at: WWW '19; May 13-17, 2019; San Francisco, CA p. 2056-2067 URL: https://dl.acm.org/doi/10.1145/3308558.3313684 [doi: 10.1145/3308558.3313684]
33. Hayes P, Zhang M, Habib R, Burgess J, Yilmaz E, Barber D. Integrated weak learning. arXiv Preprint posted online June 19, 2022. [FREE Full text] [doi: 10.48550/arxiv.2206.09496]
34. Van Rijsbergen CJ. Information Retrieval. 2nd edition. London, UK: Butterworths; 1979.
35. Hotelling H. Analysis of a complex of statistical variables into principal components. J Educ Psychol 1933;24(6):417-441 [FREE Full text] [doi: 10.1037/h0071325]
36. Rattan A, Chilazi S, Georgeac O, Bohnet I. Tackling the underrepresentation of women in media. Harvard Business Review. 2019 Jun 06. URL: https://hbr.org/2019/06/tackling-the-underrepresentation-of-women-in-media [accessed 2023-06-30]
37. Kemp S. Digital 2021: the United Kingdom. Data Reportal. 2021. URL: https://datareportal.com/reports/digital-2021-united-kingdom [accessed 2023-06-30]
38. Kemp S. Digital 2021: Mexico. Data Reportal. 2021. URL: https://datareportal.com/reports/digital-2021-mexico [accessed 2023-06-30]
39. Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv. Preprint posted online October 2, 2019. [FREE Full text] [doi: 10.5260/chara.21.2.8]
40. Sager MA, Kashyap AM, Tamminga M, Ravoori S, Callison-Burch C, Lipoff JB. Identifying and responding to health misinformation on reddit dermatology forums with artificially intelligent bots using natural language processing: design and evaluation study. JMIR Dermatol 2021 Sep 30;4(2):e20975 [FREE Full text] [doi: 10.2196/20975]
41. Marshall C, Lanyi K, Green R, Wilkins GC, Pearson F, Craig D. Using natural language processing to explore mental health insights from UK tweets during the COVID-19 pandemic: infodemiology study. JMIR Infodemiology 2022 Mar 31;2(1):e32449 [FREE Full text] [doi: 10.2196/32449] [Medline: 36406146]
42. Hasan A, Levene M, Weston D, Fromson R, Koslover N, Levene T. Monitoring COVID-19 on social media: development of an end-to-end natural language processing pipeline using a novel triage and diagnosis approach. J Med Internet Res 2022 Feb 28;24(2):e30397 [FREE Full text] [doi: 10.2196/30397] [Medline: 35142636]
43. Engel-Rebitzer E, Stokes DC, Meisel ZF, Purtle J, Doyle R, Buttenheim AM. Partisan differences in legislators' discussion of vaccination on Twitter during the COVID-19 era: natural language processing analysis. JMIR Infodemiology 2022 Feb 18;2(1):e32372 [FREE Full text] [doi: 10.2196/32372] [Medline: 35229075]
44. Monselise M, Chang CH, Ferreira G, Yang R, Yang CC. Topics and sentiments of public concerns regarding COVID-19 vaccines: social media trend analysis. J Med Internet Res 2021 Oct 21;23(10):e30765 [FREE Full text] [doi: 10.2196/30765] [Medline: 34581682]
45. Conceptual zero draft for the consideration of the Intergovernmental Negotiating Body at its third meeting. World Health Organization. 2022. URL: https://apps.who.int/gb/inb/pdf_files/inb3/A_INB3_3-en.pdf [accessed 2023-06-30]
46. Article-by-Article Compilation of Proposed Amendments to the International Health Regulations (2005) submitted in accordance with decision WHA75(9) (2022). World Health Organization. 2022. URL: https://apps.who.int/gb/wgihr/pdf_files/wgihr1/WGIHR_Compilation-en.pdf [accessed 2023-06-30]

Abbreviations

AI: artificial intelligence
BERT: bidirectional encoder representations from transformers

EARS: Early Artificial Intelligence–Supported Response With Social Listening
FP: false positive
NLP: natural language processing
TP: true positive
WHO: World Health Organization

Edited by T Mackey; submitted 17.03.23; peer-reviewed by M Monselise, R Pastor-Vargas; comments to author 14.06.23; revised version received 22.06.23; accepted 26.06.23; published 21.08.23.

Please cite as:
White BK, Gombert A, Nguyen T, Yau B, Ishizumi A, Kirchner L, León A, Wilson H, Jaramillo-Gutierrez G, Cerquides J, D'Agostino M, Salvi C, Sreenath RS, Rambaud K, Samhouri D, Briand S, Purnat TD
Using Machine Learning Technology (Early Artificial Intelligence–Supported Response With Social Listening Platform) to Enhance Digital Social Understanding for the COVID-19 Infodemic: Development and Implementation Study
JMIR Infodemiology 2023;3:e47317
URL: https://infodemiology.jmir.org/2023/1/e47317
doi:10.2196/47317
PMID:37422854

Original Paper

Detecting Tweets Containing Cannabidiol-Related COVID-19 Misinformation Using Transformer Language Models and Warning Letters From Food and Drug Administration: Content Analysis and Identification

Jason Turner[1*], MS; Mehmed Kantardzic[1*], PhD; Rachel Vickers-Smith[2*], MPH, PhD; Andrew G Brown[3*], PhD

[1]Data Mining Lab, Department of Computer Science and Engineering, J B Speed School of Engineering, University of Louisville, Louisville, KY, United States

[2]Department of Epidemiology and Environmental Health, College of Public Health, University of Kentucky, Lexington, KY, United States

[3]Department of Criminology and Criminal Justice, Northern Arizona University, Tempe, AZ, United States

[*]all authors contributed equally

Corresponding Author:
Jason Turner, MS
Data Mining Lab
Department of Computer Science and Engineering
Speed School of Engineering, University of Louisville
Louisville, KY, 40292
United States
Phone: 1 859 302 0189
Email: jason.turner@louisville.edu

Abstract

Background: COVID-19 has introduced yet another opportunity to web-based sellers of loosely regulated substances, such as cannabidiol (CBD), to promote sales under false pretenses of curing the disease. Therefore, it has become necessary to innovate ways to identify such instances of misinformation.

Objective: We sought to identify COVID-19 misinformation as it relates to the sales or promotion of CBD and used transformer-based language models to identify tweets semantically similar to quotes taken from known instances of misinformation. In this case, the known misinformation was the publicly available Warning Letters from Food and Drug Administration (FDA).

Methods: We collected tweets using CBD- and COVID-19–related terms. Using a previously trained model, we extracted the tweets indicating commercialization and sales of CBD and annotated those containing COVID-19 misinformation according to the FDA definitions. We encoded the collection of tweets and misinformation quotes into sentence vectors and then calculated the cosine similarity between each quote and each tweet. This allowed us to establish a threshold to identify tweets that were making false claims regarding CBD and COVID-19 while minimizing the instances of false positives.

Results: We demonstrated that by using quotes taken from Warning Letters issued by FDA to perpetrators of similar misinformation, we can identify semantically similar tweets that also contain misinformation. This was accomplished by identifying a cosine distance threshold between the sentence vectors of the Warning Letters and tweets.

Conclusions: This research shows that commercial CBD or COVID-19 misinformation can potentially be identified and curbed using transformer-based language models and known prior instances of misinformation. Our approach functions without the need for labeled data, potentially reducing the time at which misinformation can be identified. Our approach shows promise in that it is easily adapted to identify other forms of misinformation related to loosely regulated substances.

(JMIR Infodemiology 2023;3:e38390) doi:10.2196/38390

KEYWORDS

transformer; misinformation; deep learning; COVID-19; infodemic; pandemic; language model; health information; social media; Twitter; content analysis; cannabidiol; sentence vector; infodemiology

Introduction

Background

The Food and Drug Administration (FDA) defines medical “misinformation” as the information that misleadingly represents a product as able to mitigate, prevent, treat, diagnose, or cure a condition or disease, such as COVID-19 [1]. Misinformation is prevalent on social networking platforms and is often seen in product advertising. Since early 2020, the COVID-19 pandemic has offered a new medical condition for advertisers and sellers to exploit, who seek to profit from the crisis at the expense of public health [2].

Although some misleading posts regarding COVID-19 have been about the virus’s origins or the effectiveness and safety of masks and vaccines [3-5], others have focused on false information about alternative products in treating or preventing COVID-19. For example, since the onset of the pandemic, some cannabidiol (CBD) sellers have claimed that CBD can prevent or cure COVID-19 (Figure 1). Preclinical studies have suggested that CBD could be effective in treating respiratory conditions and might confer cardioprotective, nephroprotective, hepatoprotective, neuroprotective, and anticonvulsant benefits [6]. Meanwhile, other evidence suggests that CBD could decrease one’s ability to fight infections, countering its potential clinical use as an anti-inflammatory agent [7]. Because CBD has not been thoroughly tested for efficacy and safety in treating COVID-19, its benefits are mostly considered unsubstantiated [8]. However, there is plentiful web-based marketing for CBD, touting these unverified benefits. This is not unique to CBD and COVID-19 and occurs with other loosely regulated substances (eg, kratom and herbal supplements) as well as other medical conditions such as Alzheimer disease, cancer, and autism. Although the consumption of CBD is typically well tolerated, misinformation about its effectiveness in treating COVID-19 has become widespread and poses a danger to public health [9,10]. Therefore, it is essential that this type of content be efficiently identified so that its potential harmful effects on public health can be minimized.

Figure 1. Example of a tweet promoting cannabidiol (CBD) as a prevention or treatment for COVID-19.

Misinformation has been shown to spread faster and further than accurate information on social media, and Twitter serves as an example of a social media site where misinformation about smoking products, drugs, vaccines, and diseases is abound [11,12]. Specifically regarding COVID-19, for example, in March 2020, a daily average of 46,000 misleading or inaccurate news posts appeared on Twitter in Italy alone [13]. In Iran, a rumor originating on social networks claimed that ingesting neat alcohol could cure COVID-19, resulting in hundreds of alcohol poisoning deaths [14]. Furthermore, influential Twitter users such as former President Donald J Trump and celebrity Joe Rogan suggested taking the antimalarial medications chloroquine or hydroxychloroquine to treat COVID-19, even though there has never been scientific evidence to support these claims [15-17]. Although the United States FDA warned against improper consumption of these substances in July 2020, there were still dozens of documented deaths and poisonings associated with their use, including at least one person who ingested fish tank cleaner containing chloroquine after being influenced by misinformation on Twitter [18-20]. Not only has misinformation dissuaded consumers from seeking effective treatments but also it has encouraged the use of dangerous and unfounded *treatments*.

Prior research has shown some success in using supervised and unsupervised machine learning techniques to detect and explore COVID-19 misinformation on social media platforms such as

Twitter [21-24] and YouTube [25]. Although some studies used an annotated data approach [21,24,26], it was noted that the rigor and costs of this process could preclude its widespread utility [21]. One study used an unsupervised topic model to examine inaccurate information about vaping CBD and COVID-19 [27], which is a more cost-effective technique than annotation, but is a less explicit way to identify misinformation than supervised approaches. Deep learning [23,28] and transformer language models [22,23,25,29] are the most commonly used techniques, likely because of their efficiency and highly advanced ability to interpret and understand natural language, which is key to examining content on social media.

This study leverages the expertise of the FDA, the regulatory body of the safety and efficacy of numerous products intended for human or animal use in the United States, to define misinformation regarding CBD and COVID-19 using Warning Letters as a gold standard. When the FDA becomes aware that a company has violated FDA regulations, they often issue Warning Letters to the company that outline the nature of the violation (eg, problems with claims about a product and incorrect directions for use), corrective action, and a timeframe [30]. The FDA then follows up to verify whether the company has completed the corrective action, and if not, the FDA may enact regulatory actions such as seizure or civil penalties [31]. These Warning Letters are made available to the public on the FDA's website [32]. By using verbatim quotes taken from relevant FDA Warning Letters and transformer language models, we propose an efficient and consistent approach to identify tweets that contain false claims regarding CBD and COVID-19.

Objectives

This study has two primary objectives: (1) to examine misinformation concerning CBD and COVID-19 disseminated by Twitter advertisers and (2) to propose a framework that helps identify CBD- and COVID-19–related misinformation in tweets more efficiently and can also be easily modified to detect misinformation in advertisements about other substances and conditions.

Traditional text classification requires labeled data for model training and testing, which can be both time consuming and expensive. Transformer-based language models have shown promising results in various semantic textual similarity tasks, including in the medical domain [33].

Methods

Ethics Approval

This study leverages publicly available data and was registered as approved by the University of Louisville Institutional Review Board (approval protocol 20.1122).

Collecting and Annotating Tweets

We used the Snscrape Python package to collect English language tweets from the Twitter website from January 1, 2020, to April 28, 2021, by searching for tweets using the following keywords: *CBD or cannabidiol*, and *COVID-19, COVID, corona* [34]. Although this method does not provide full access to Twitter's past data, it does afford the ability to collect thousands of historical tweets retrospectively after an unexpected event has occurred, such as the pandemic. Using this method of historical tweet searching, we were able to collect 37,526 tweets over a 484-day period.

To extract the commercial CBD tweets referencing COVID-19, we used a model previously developed by our group [35]. The model was trained on an earlier collection of CBD tweets to identify those that reflected commercial sales, promotion, and marketing of CBD. Applying the commercial CBD tweet classifier to the historical tweet collection resulted in 4937 tweets that were classified as commercial CBD tweets referencing COVID-19. We evaluated the performance of this model that identified commercial CBD tweets from the CBD or COVID-19 collection of tweets by annotating a random sample of classified 250 commercial CBD or COVID-19 tweets and 250 noncommercial CBD or COVID-19 tweets. We observed an improvement in the commercial tweet classification performance, with precision, recall, accuracy, and F_1-scores of 0.95 (Table 1).

Table 1. Performance of 2019 commercial cannabidiol (CBD) classifier in differentiating tweets from 2020 to 2021 (N=500).[a]

	Precision	Recall	F_1-score	Random sample
Noncommercial CBD	0.95	0.95	0.95	249
Commercial CBD	0.95	0.95	0.95	251

[a]Accuracy of the model: 0.95.

The United States FDA has warned CBD advertisers and sellers about promoting CBD as a treatment or cure for several conditions, including COVID-19. These Warning Letters inform the advertiser that CBD is not an approved treatment or prevention for the condition that the advertiser mentions, warn the advertiser that further action will be taken if this style of misleading advertising continues, and include the exact misleading quote taken from the advertisement. This list of letters continues to grow as the FDA becomes aware of more advertisers making false claims about the medical benefits of CBD. We drew directly from the FDA Warning Letters about CBD and COVID-19 to understand the types of statements that the FDA had flagged as misinformation.

We used the statements from these FDA Warning Letters to identify and label the CBD tweets containing COVID-19–related misinformation. Instead of using our annotated set of tweets to train and test a tweet classification model, we used them to establish the relationship of text similarity between the misinformation tweets and the statements in the FDA Warning Letters via the transformer language model. The transformer architecture was introduced to address some of the shortcomings of the recurrent neural networks in tasks such as language translation [36].

Because the FDA's CBD or COVID-19 Warning Letters did not include CBD-infused hand sanitizers, we included the Warning Letters sent to advertisers of nonalcohol and essential oil sanitizer products that had made false claims as guidance in our misinformation annotation process [37-43]. Figure 2 is an example of one of the Warning Letters issued by the FDA about misinformation surrounding the use of CBD for treating COVID-19.

The extracted commercial CBD or COVID-19 tweets were annotated for misinformation (yes or no), according to the FDA's definition of misinformation, by 3 university-trained, nonmedical, professional annotators; discrepancies were decided by a majority vote to determine the overall label of the tweet. Before labeling the tweets, we reviewed several FDA Warning Letters with the annotators so that the annotators were familiar with what the FDA considered a misleading statement. Because tweets are relatively short in length, if a tweet contained any misinformation related to CBD and COVID-19, the entire tweet was considered positive for misinformation. There were no noticeable discrepancies among the annotators, and they had an intercoder agreement of 91%.

The Warning Letters on the FDA website provided some of the quotes that sellers used in their advertisements that the FDA deemed misleading. Some examples of these letters are listed in Textbox 1. Along with the FDA's definition of misinformation, these quotes provided guidelines for annotating the collection of tweets as to whether they contained misinformation. Textbox 2 displays selected examples of the nonmisinformation and misinformation tweets encountered during the annotation process.

Figure 2. Example Warning Letter taken from the Food and Drug Administration (FDA) website.

WARNING LETTER

For Our Vets LLC dba Patriot Supreme

MARCS-CMS 611043 — OCTOBER 16, 2020

Share | Tweet | LinkedIn | Email | Print

More Warning Letters

Warning Letters

About Warning and Close-Out Letters

Product: Drugs

Recipient:
Justin Elenburg
For Our Vets LLC dba Patriot Supreme
1910 S. Stapley Dr. #221
Mesa, AZ 85204
United States

Info@patriotsupreme.com

Issuing Office:
Center for Drug Evaluation and Research | CDER
United States

Federal Trade Commission

WARNING LETTER

Date: October 16, 2020

RE: Unapproved and Misbranded Products Related to Coronavirus Disease 2019 (COVID-19)

This is to advise you that the United States Food and Drug Administration (FDA) and the Federal Trade Commission (FTC) reviewed your website at the Internet address https://patriotsupreme.com on October 5, 2020, and October 8, 2020, respectively. We also reviewed your social media websites at https://www.facebook.com/patriotsupreme/ and https://www.instagram.com/patriotsupreme/, where you direct consumers to your website, https://patriotsupreme.com, to purchase your products. The FDA has observed that your website offers CBD products for sale in the United States and that these products are intended to mitigate, prevent, treat, diagnose, or cure COVID-19[1] in people. Based on our review, these products are unapproved new drugs sold in violation of section 505(a) of the Federal Food, Drug, and Cosmetic Act (FD&C Act), 21 U.S.C. § 355(a). Furthermore, these products are misbranded drugs under section 502 of the FD&C Act, 21 U.S.C. § 352. The introduction or delivery for introduction of these products into interstate commerce is prohibited under sections 301(a) and (d) of the FD&C Act, 21 U.S.C. § 331(a) and (d).

There is currently a global outbreak of respiratory disease caused by a novel coronavirus that has been named "severe acute respiratory syndrome coronavirus 2" (SARS-CoV-2). The disease caused by the virus has been named "Coronavirus Disease 2019" (COVID-19). On January 31, 2020, the Department of Health and Human Services (HHS) issued a declaration of a public health emergency related to COVID-19 and mobilized the Operating Divisions of HHS.[2] In addition, on March 13, 2020, the President declared a national emergency in response to COVID-19.[3] Therefore, FDA is taking urgent measures to protect consumers from certain products that, without approval or authorization by FDA, claim to mitigate, prevent, treat, diagnose, or cure COVID-19 in people. As described below, you sell products that are intended to mitigate, prevent, treat, diagnose, or cure COVID-19 in people. We request that you take immediate action to cease the sale of such unapproved and unauthorized products for the mitigation, prevention, treatment, diagnosis, or cure of COVID-19.

Content current as of:
10/19/2020

Regulated Product(s)
Drugs

Some examples of the claims on your websites that establish the intended use of your products and misleadingly represent them as safe and/or effective for the treatment or prevention of COVID-19 include:

- **"CBD AND COVID-19 – CAN CBD HELP COVID-19 LUNG INFLAMMATION?"**
  - "[t]here has been an increased interest in CBD and Covid-19 to treat lung problems and symptoms (mental or physical) associated with the coronavirus."
  - **"CBD May Play a Role in Helping Lung Symptoms"**
    - "Cannabis contains several cannabinoids that have anti-inflammatory properties. Specifically, CBD is the most likely possibility for treating COVID-19 related lung inflammation."
  - **"Using CBD to Alleviate Inflammation"**
    - "CBD is available without a prescription. It is already being used to treat serious medical problems . . . where *pain and/or inflammation are a major factor*. This is why CBD has piqued the interest of the medical world as a significant aid to reduce inflammation for the COVID-19 lung inflammation.

Acute respiratory distress syndrome (ARDS) is a type of respiratory failure characterized by rapid onset of widespread inflammation in the lungs, rapid breathing and the inability to sustain adequate oxygen levels to the body and brain. Shortness of breath or difficulty breathing are some of the early signs of COVID-19 . . ."

  - **"CBD and Covid-19 / Why Do Researchers Believe CBD Can Help?"**
    - "Firstly, the research performed to date has shown that CBD can reduce a number of pro-inflammatory cytokines (numerous different types of substances, such as interferon, interleukin, and growth factors, which are secreted by certain cells of the immune system and have an effect on other cells) including IL-6, the one reduced by other drugs being studied for COVID-19. CBD was also shown to reduce interleukin (IL)-2, IL-1α and β, interferon gamma, inducible protein-10, monocyte chemoattractant protein-1, macrophage inflammatory protein-1α, and tumor necrosis factor-α – all of which are associated with the pathology of severe cases of COVID-19. In addition to reducing these pro-inflammatory cytokines, CBD has also been shown to increase the production of interferons, a type of signaling protein that activates immune cells and prevents viruses from replicating."
    - https://patriotsupreme.com/cbd-covid-19-lung-inflammation]
- **"CBD INFLAMMATION COVID / A PLAUSIBLE METHOD?"**
  - **"CBD Inflammation Covid"**
    - "It has been determined that CBD may be beneficial in curing inflammation of the lungs caused by covid-19 (CBD inflammation covid). In this article we shall discuss it in detail."
  - **"How is CBD inflammation Covid a plausible method?"**
    - "ARDS, which is 'Acute Respiratory Distress Syndrome' is one of the severe symptoms of the covid-19. In this syndrome, the corona virus causes the inflammation of the lungs, which in turn may lead to severe difficulty in breathing, pain or perhaps even coughing."
  - **"How is inflammation cured?"**
    - "Talking about CBD inflammation Covid. Coronavirus causes cytokine production in the body which is an active agent for inflammation. CBD may play a role in reducing the rate of cytokine production, which will give the lungs of a patient enough time to recover, from the damages. As a result, the oxygen levels in the body will spike up again to ideal levels and the lung condition will get back to normal."

Textbox 1. Five example quotes from Warning Letters issued by Food and Drug Administration (FDA) related to cannabidiol (CBD) and COVID-19 misinformation.

- “Firstly, the research performed to date has shown that CBD can reduce a number of pro-inflammatory cytokines (numerous different types of substances, such as interferon, interleukin, and growth factors, which are secreted by certain cells of the immune system and have an effect on other cells) including IL-6, the one reduced by other drugs being studied for COVID-19. CBD was also shown to reduce interleukin (IL)-2, IL-1α and β, interferon gamma, inducible protein-10, monocyte chemoattractant protein-1, macrophage inflammatory protein-1α, and tumor necrosis factor-α – all of which are associated with the pathology of severe cases of COVID-19. In addition to reducing these pro-inflammatory cytokines, CBD has also been shown to increase the production of interferons, a type of signaling protein that activates immune cells and prevents viruses from replicating”
- “There has been an increased interest in CBD and Covid-19 to treat lung problems and symptoms (mental or physical) associated with the coronavirus”
- “CBD oil may help to prevent getting infected by strengthening your immune system. It has also been proven to offer relief to some of the symptoms”
- “By using CBD oil, you can keep inflammation at bay, retain a healthy or even higher than average white blood cell count, stay calm and relaxed (which is best for a strong immune system), and prevent catching a virus or infection beforehand”
- “Is CBD an Anti-Viral Agent for Coronavirus, Influenza, MERS, and Sars Plus Key Antiviral Supplements?”

Textbox 2. Examples of annotated commercial cannabidiol (CBD) or COVID-19 tweets.

- Misinformation
  - “#CBD is readily available for anyone who want to build-up their immune to help guard against the #coronavirus. It’s your responsibility to protect &; take care of yourself, not the government. Order Now”
  - “I’ve ordered a 4 month supply of #CBD to help fortify my immune system &; guard against the #CoronaVirusUpdates #COVID19. What have you done to protect yourself? Order Now”
  - “#COVID19 attacks the inside of body/lungs which are internal so topical solutions will not mitigate what's happening inside your body/lungs. Ask me about #CBD. #cbdoil <MASKED-URL>”
  - “#CBDL Could Double On Product Launch News. CBD Hand Sanitizer Could Help Stop Spread Of Covid-19. [Read Now] LINK #USA #Stocks #Bonds #Equities #Gold #Silver #Bitcoin #CryptoCurrency #Investing #Trading #Options #1Author #USStocks #StockMarket”
- Nonmisinformation
  - “Could CBD offer treatment options for COVID-19? Read more at <MASKED-URL> Global Go does not endorse the use of any product for medicinal purposes. Please consult with a physician before using any such products. #CBDtrials #hempresearch #hempnews #covid19”
  - “First day of the week... First day of the month of June! Would you like to try something new? #Covid19 #SanFrancisco #SanFranciscobayArea” #helpingthecommunity #Realestate #HartFordproperties #CBD Source:
  - “Online sales for cannabidiol (#CBD) products continue to #flourish despite in-store slowdowns amid the COVID-19 #pandemic. #CSPDailyNews”
  - “New post (Iowa Down To One Medical CBD Manufacturer Due To COVID-19 Pandemic) has been published on Buy Premium CBD and CBG Products | 100% Natural Cannabinol Store | Buy CBD Oils, Gummies, Topicals, Pet CBD and more.”

Misinformation Search

The transformer follows an encoder-decoder structure wherein the encoder converts the text input into a vector representation. There is one vector per word in the sequence; the value of each vector representation of each word is partially based on the nearby terms surrounding the word, which provides the context. The decoder portion of the transformer architecture is similar to the encoder but can convert a vector into a sequence of text. The original Bidirectional Encoder Representations from Transformers (BERT) model was introduced by Devlin et al [44] at Google in 2018. The BERT model is trained on the natural language processing tasks of masked language modeling on English Wikipedia and the Brown Corpus texts and it predicts the missing word in a sentence and performs next-sentence prediction. Masked language modeling allows the BERT model to understand the relationships between words, whereas next-sentence prediction allows the BERT model to understand the relationships between sentences.

In this study, we used a transformer language model to encode each of the commercial CBD tweets that we collected. Specifically, we used the Sentence-T5 model because it is a state-of-the-art language model that has outperformed other models in semantic textual similarity [45]. We also computed the encoding for each statement containing misinformation taken from the FDA website into vectors of size 768. We then calculated the cosine similarity (Equation 1) between the encoding vectors of tweets (A) and the encoding vectors of the misinformation quotes taken from the FDA Warning Letters (B).

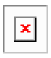
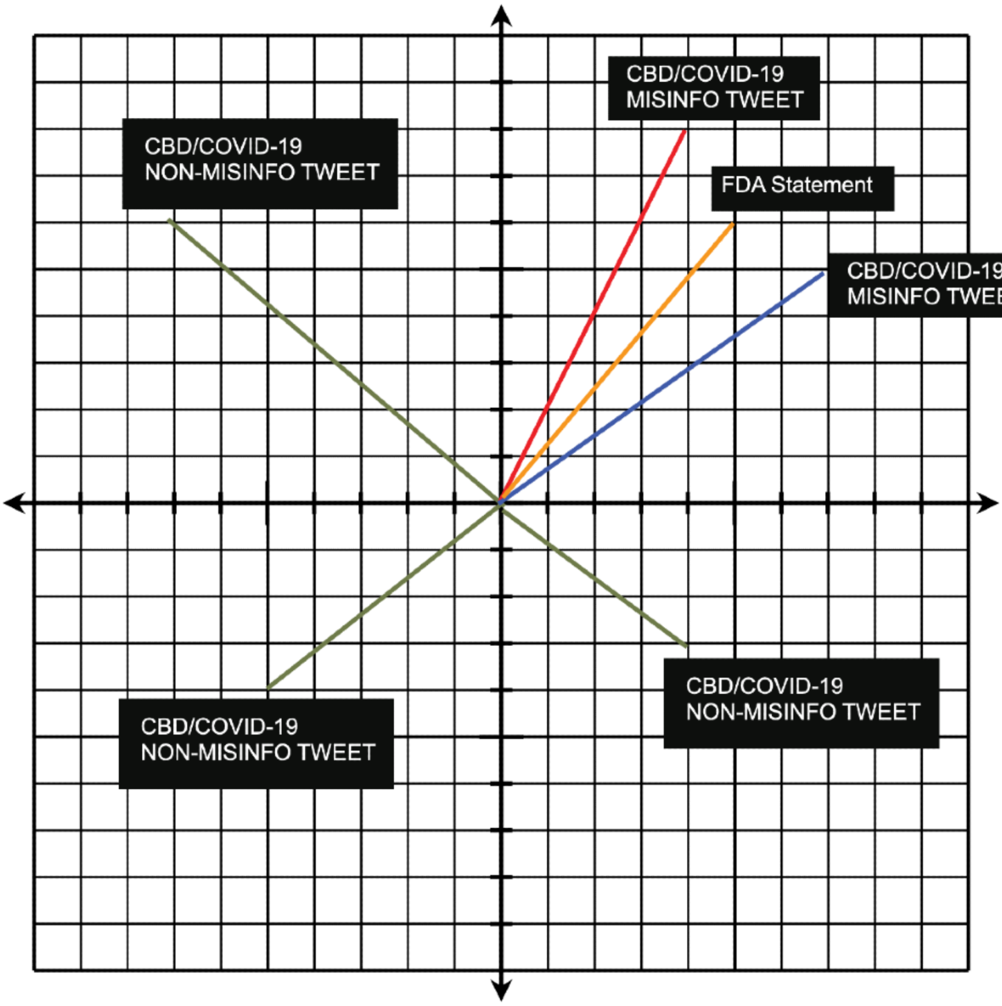


Figure 3 provides a theoretical illustration of how we isolated tweets that made false claims using the quotes taken from the FDA Warning Letters to find contextual similarities using cosine similarity. This example shows that when the tweets and quotes from Warning Letters are encoded into vectors, tweets containing CBD or COVID-19 misinformation will be closer in cosine distance to the FDA quotes than tweets without similar false claims about CBD and COVID-19. Using cosine similarity as the distance of similarity, we expected that the shortest cosine distances would contain more misinformation, that is, tweets that were contextually similar to the FDA's samples. Conversely, we expected the tweets with the longest cosine distances to contain less misinformation. Using these points, we identified a threshold at which we could confidently identify sets of tweets that mostly contained misinformation because of their semantic similarity to an established example of misinformation. The optimal threshold should be the point at which the maximum number of tweets deemed as misinformation is captured while minimizing the number of false positives (ie, tweets that do not contain misinformation) being captured.

Figure 3. Representation of contextual similarity between cannabidiol (CBD) tweets and quotes from Food and Drug Administration (FDA) Warning Letters.



Results

Analysis of Misinformation

After annotating the tweets for misinformation, we observed that approximately 19% (938/4937) of the tweets contained misinformation related to both CBD and COVID-19. Table 2 shows some of the terms most frequently associated with CBD or COVID-19 misinformation. Aside from CBD-related terms, terms such as "help," "boost," "health," "virus," "sanitizer," and "immune system" were among the most frequently occurring terms. We observed an increase in CBD or COVID-19 conversations beginning in February 2020. Misinformation related to CBD and COVID-19 peaked in March 2020; although it appeared to taper down, it did not stop.

Table 2. Top n-grams in the cannabidiol or COVID-19 misinformation tweets.

Term	Frequency
immune	257
cbdoil	244
system	187
immune system	181
products	157
help	152
hemp	139
health	129
virus	127
new	127
cbd cbdoil	125
natural	119
wellness	114
boost	112
oil	99
hand	96
sanitizer	92
cbdproducts	88
cbd products	85
hand sanitizer	83
use	81
immunity	78

Detecting Misinformation

Using 27 misleading quotes taken from the FDA Warning Letters and converting them into vector form and then converting each tweet into vector form, we calculated the cosine similarity. Using the cosine distance, we then counted the number of tweets that were labeled as misinformation compared with the tweets that were considered nonmisinformation. We observed that the nearest tweets indeed contained misinformation, whereas the most distant tweets did not (Table 3).

Table 3. Measuring the cosine distance between sentence vectors of statement 0 and the tweets.

Tweet	Position	Misinformation		Cosine distance
		Yes	No	
“With the growing concern of the COVID-19 virus we understand the importance of boosting the immune system. CBD[a] is a natural way to do that. The Healing Leaf wants to help make CBD more available and lessen costs for those interested. Please come in or call and place your orders!” <MASKED-URL>	Most similar	✓		0.069577
“CBD may reduce cytokine storm and inflammation in COVID-19\n” <MASKED-URL>	Second most similar	✓		0.077852
“They snuck a shipping ban on vape products into the last covid relief bill. All vape products. CBD, nicotine, delta 8, doesn’t matter.”	Second most distant		✓	0.296933
“Final point: When COVID hit, we LOWERED our prices and never put them back up. Funny how reddit and Review haters never use their real names and 9/10 when you find out their real name - they run for the hills or perhaps just back to Mam’s house and their keyboard...”	Most distant		✓	0.297551

[a]CBD: cannabidiol.

We calculated the cosine distance between the sentence vector of each of the 27 statements that we extracted from the FDA Warning Letters, along with each of the tweets, so that we could determine a cosine distance threshold in which we collected the most tweets that contained CBD or COVID-19 misinformation while minimizing false positives (non-CBD or COVID-19 misinformation). Figures 4 and 5 show a consistent observation that as the cosine distance increased, the percentage (recall) and frequency of the tweets, respectively, containing misinformation increased. Figure 6 indicates that as the cosine distance increased, the precision of the tweets containing misinformation decreased. However, if the cosine distance was too small, few to no misinformation tweets were captured, and not all FDA statements performed equally in capturing misinformation tweets.

We also examined the top 5 FDA statements that captured large amounts of misinformation with high precision (Textbox 3). Figures 7-9 are equivalent to Figures 4-6 but include only the statements displayed in Textbox 3. From these figures, we can see that at a cosine distance between 0.10 and 0.13, we were able to capture between 21.9% (206/938) and 28.9% (272/938) of the misinformation tweets with a precision of above 80%. Specifically, statement 8 was able to capture 34.4% (323/938) of tweets at a cosine distance of 0.105, and 84.8% (274/938) of these tweets were labeled as misinformation. The 274 captured tweets represented 29.2% of the 938 misinformation in our data set.

Figure 4. Cosine distance versus proportion of misinformation tweets captured.

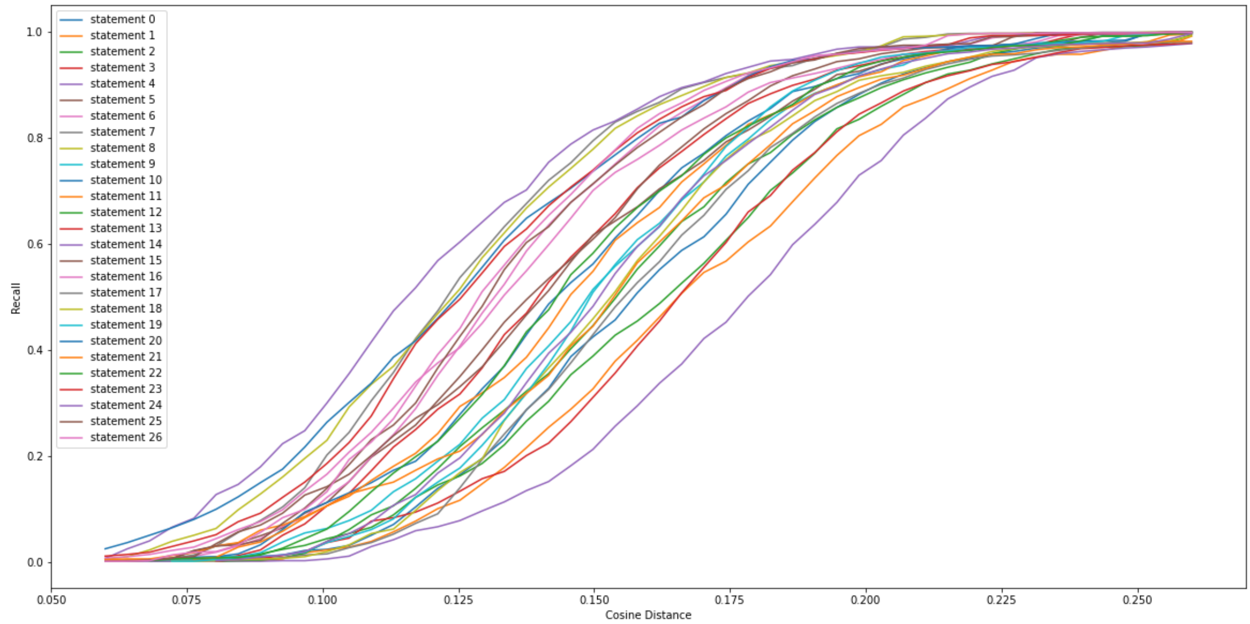


Figure 5. Cosine distance versus number of misinformation tweets captured.

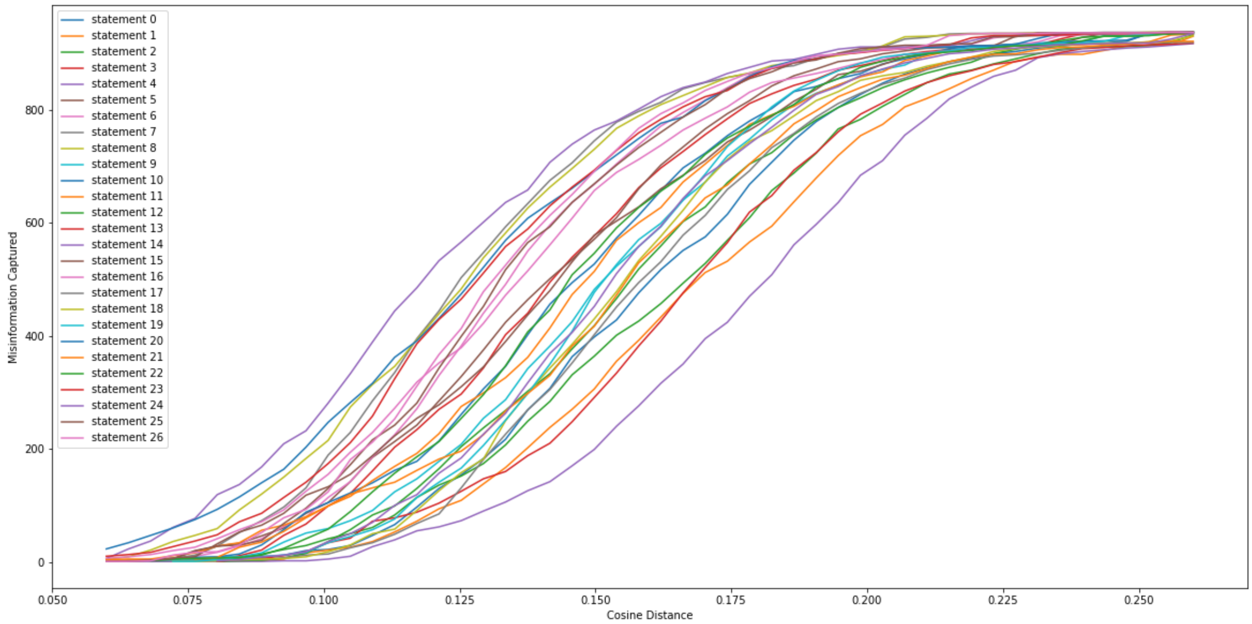
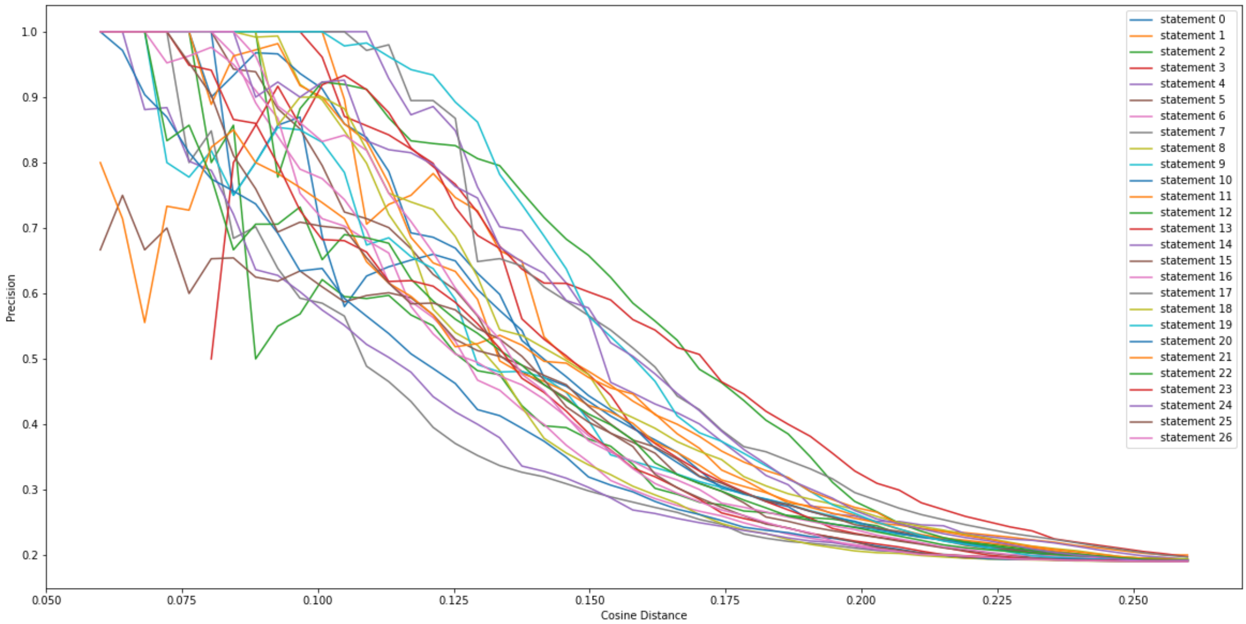


Figure 6. Cosine distance versus misinformation precision of tweets captured.



Textbox 3. Food and Drug Administration (FDA) statements that captured the largest amount of misinformation with the highest amount of precision.

- Statement 3
  - "Firstly, the research performed to date has shown that CBD can reduce a number of pro-inflammatory cytokines (numerous different types of substances, such as interferon, interleukin, and growth factors, which are secreted by certain cells of the immune system and have an effect on other cells) including IL-6, the one reduced by other drugs being studied for COVID-19. CBD was also shown to reduce interleukin (IL)-2, IL-1α and β, interferon gamma, inducible protein-10, monocyte chemoattractant protein-1, macrophage inflammatory protein-1α, and tumor necrosis factor-α – all of which are associated with the pathology of severe cases of COVID-19. In addition to reducing these pro-inflammatory cytokines, CBD has also been shown to increase the production of interferons, a type of signaling protein that activates immune cells and prevents viruses from replicating"
- Statement 8
  - "DML CBD: Immune Boost Pack... ALERT: There is no cure or treatment for COVID19. With this in mind, many doctors claim the best defense is to boost the body's immune system. DML CBD aims to help our customers in an attempt to boost the immune system... WHY TO BUY THE BOOST PACK: Studies suggest that CBD can help fight off inflammation, boost the immune system, and help battle against certain harmful bacteria. Some research suggests it can help suppress the cytokine storm inside the body that can cause great illness and sometimes death... NOTE: The cytokine storm is often triggered in patients with COVID19. Please note there is no proven cure or treatment for COVID19... There has never been a more important time than to boost your immune system. To help our customers get a full CBD experience that aims to boost your immune system, we offer the 'DML CBD Immune Boost' package..."
- Statement 12
  - "What is COVID-19? Coronavirus is referred to as a novel cause for viral pneumonia because it's a virus we haven't seen before and have developed no immunity to... What Happens If You Get Infected and What Can Help? ...Regardless of the shape you're in at this moment, there may be ways you can prepare and protect your body from developing a more severe response to infection. Explore the solutions included in NoronaPak below! [graphic with the following text] 'Selenium, Cannabidiol (CBD), Vitamin-C, Zinc, Vitamin-D, N-Acetylcysteine'... Supplementation with selenium results in changes in the gene expression that is required for protein biosynthesis in lymphocytes, the infection-fighting cells that are crucial to the immune system being able to identify infection and mount an immune response... Selenium is not only important in boosting the immunity of the individual but also to slow the development of more virulent strains of some viral pathogens... CBD may suppress the productions of cytokines in the setting of infection"
- Statement 12
  - "In the wake of the current epidemic, it is now more important than ever to keep your immune system as healthy as can be... Here are 5 key ways to strengthen your immune system during the outbreak...Take supplements such as CBD"
- Statement 26
  - "Crush Corona... While scientists around the world are working 24/7 to develop a COVID-19 vaccine, it will take many more months of testing before it's approved and available. However, there's something you can do right now to strengthen your immune system. Take CBD... CBD can help keep your immune system at the stop of its game... We want everyone to take CBD and take advantage of its potential to help prepare your body to fight a coronavirus infection. So, we're making all of our products more affordable"

Figure 7. Cosine distance versus percentage of misinformation tweets captured by top 5 performing Food and Drug Administration (FDA) statements, which captured large amount of misinformation with high precision.

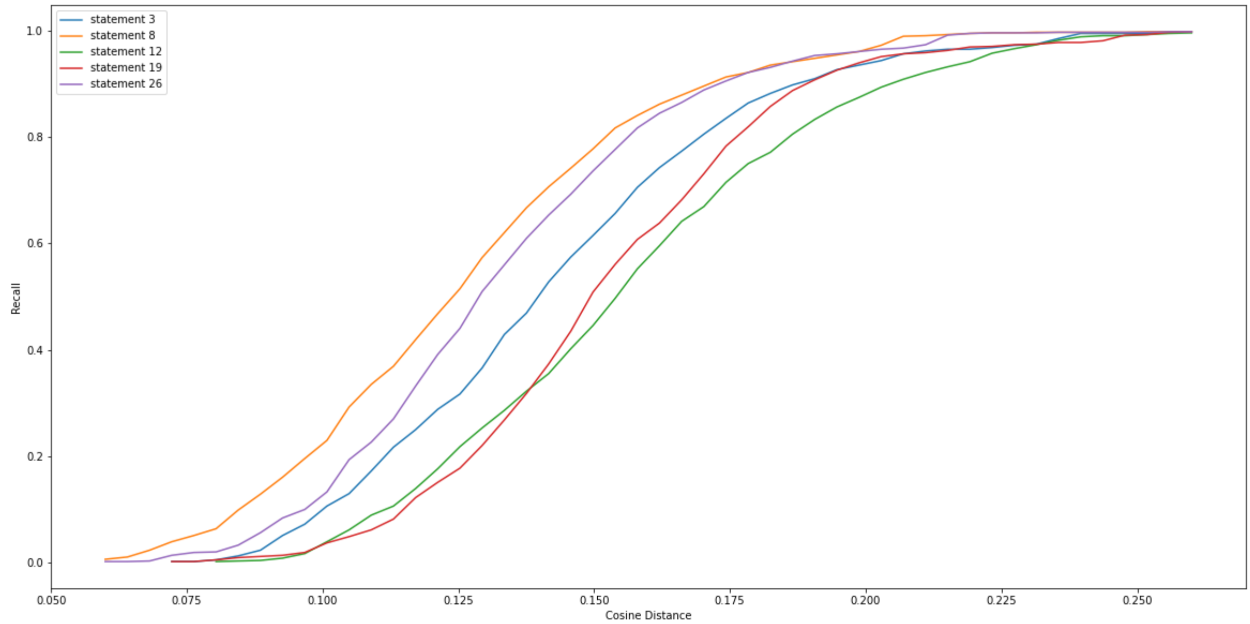


Figure 8. Cosine distance versus number of misinformation tweets captured by top 5 performing Food and Drug Administration (FDA) statements, which captured large amount of misinformation with high precision.

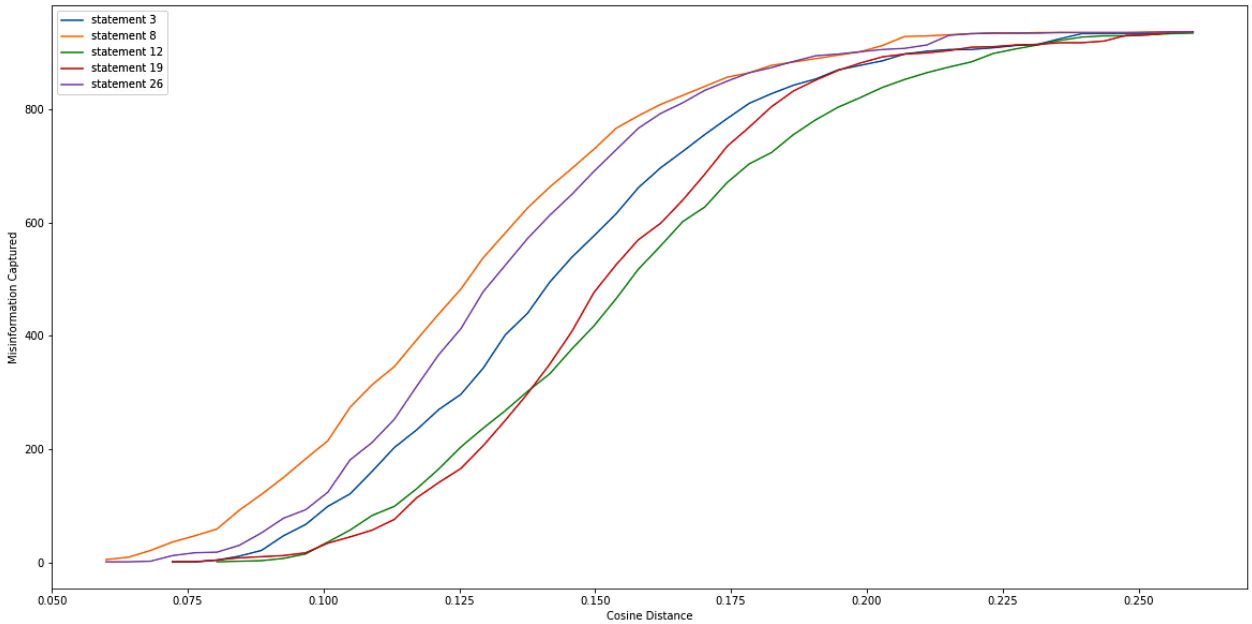
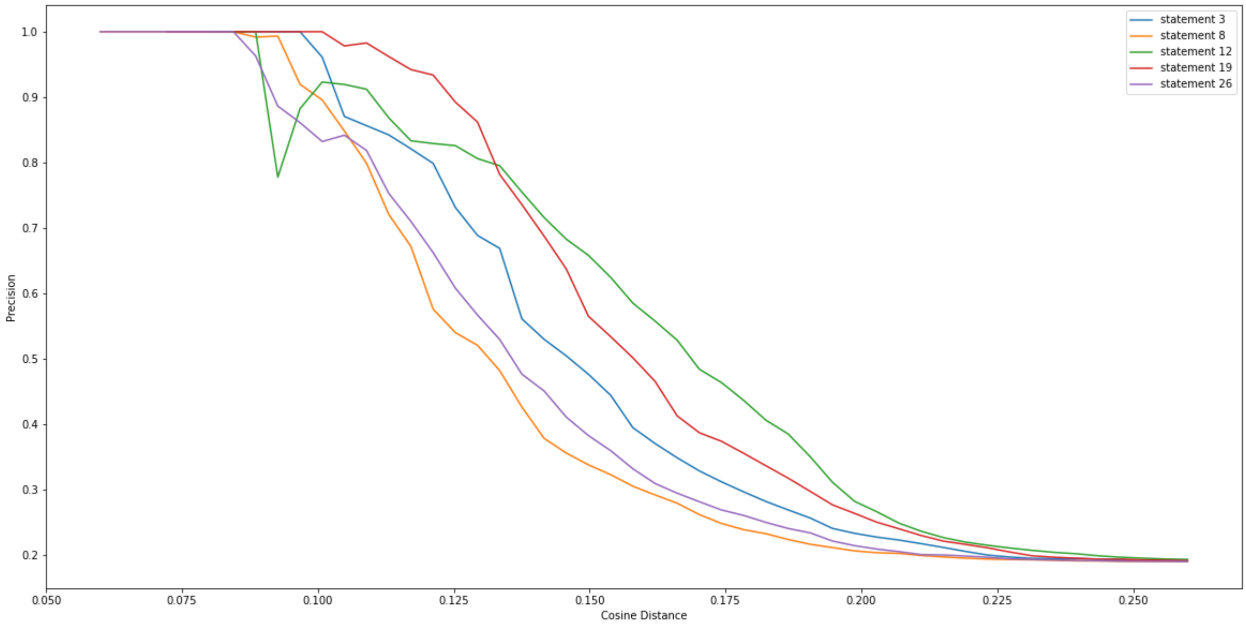


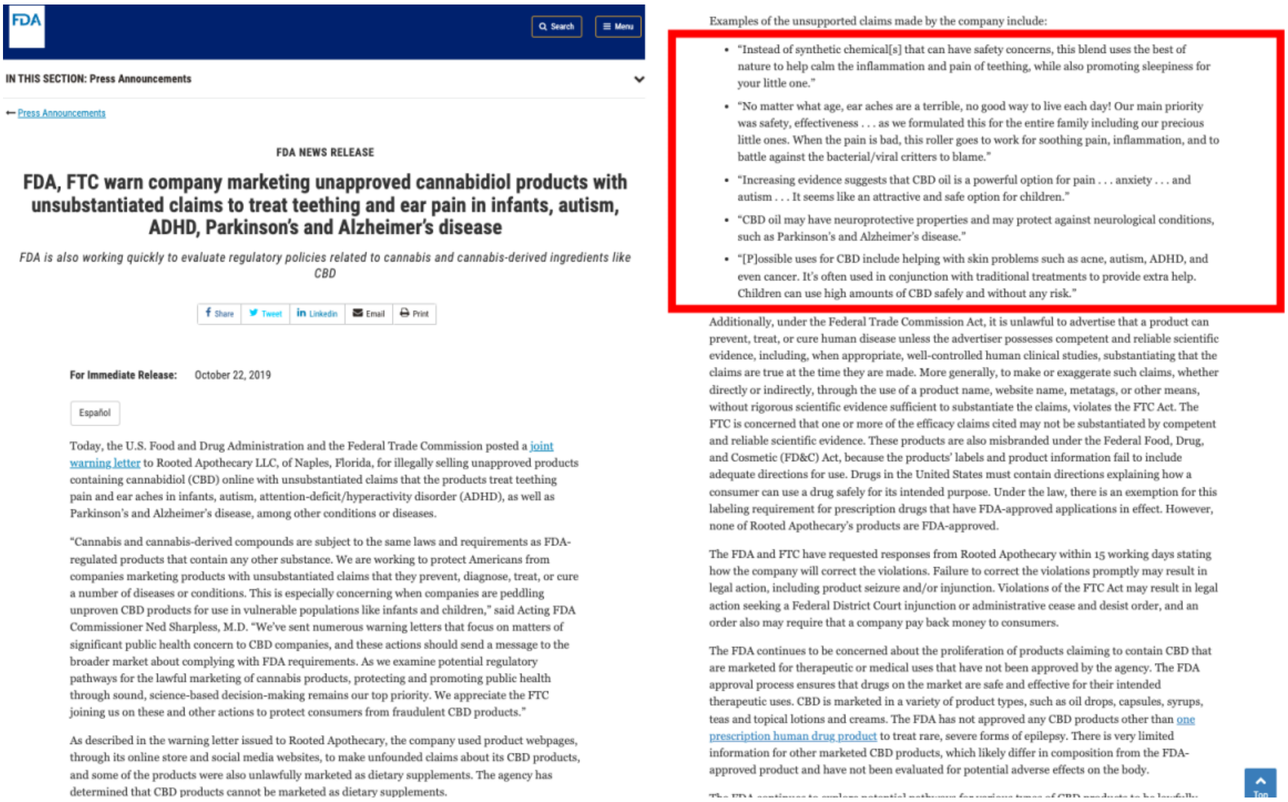
Figure 9. Cosine distance versus misinformation precision of tweets captured.



Application to Other Contexts

To further demonstrate the flexibility of our methods, we applied it to a corpus of tweets collected in 2019 using only the terms "CBD" and "cannabidiol," and misleading quotes from FDA Warning Letters regarding autism and Alzheimer disease (Figure 10). As shown in Table 4, we observed that the tweets most similar to the misinformation samples suggested that CBD could alleviate the symptoms of autism and Alzheimer disease. The most distant tweets did not make false claims about CBD's ability to treat those conditions.

Figure 10. Warning Letter taken from Food and Drug Administration (FDA) website regarding cannabidiol (CBD) as a treat for teething, autism, attention deficit hyperactivity disorder (ADHD), and Alzheimer disease.



FDA

Search Menu

IN THIS SECTION: Press Announcements

← Press Announcements

FDA NEWS RELEASE

FDA, FTC warn company marketing unapproved cannabidiol products with unsubstantiated claims to treat teething and ear pain in infants, autism, ADHD, Parkinson's and Alzheimer's disease

FDA is also working quickly to evaluate regulatory policies related to cannabis and cannabis-derived ingredients like CBD

Share Tweet Linkedin Email Print

For Immediate Release: October 22, 2019

Español

Today, the U.S. Food and Drug Administration and the Federal Trade Commission posted a joint warning letter to Rooted Apothecary LLC, of Naples, Florida, for illegally selling unapproved products containing cannabidiol (CBD) online with unsubstantiated claims that the products treat teething pain and ear aches in infants, autism, attention-deficit/hyperactivity disorder (ADHD), as well as Parkinson's and Alzheimer's disease, among other conditions or diseases.

"Cannabis and cannabis-derived compounds are subject to the same laws and requirements as FDA-regulated products that contain any other substance. We are working to protect Americans from companies marketing products with unsubstantiated claims that they prevent, diagnose, treat, or cure a number of diseases or conditions. This is especially concerning when companies are peddling unproven CBD products for use in vulnerable populations like infants and children," said Acting FDA Commissioner Ned Sharpless, M.D. "We've sent numerous warning letters that focus on matters of significant public health concern to CBD companies, and these actions should send a message to the broader market about complying with FDA requirements. As we examine potential regulatory pathways for the lawful marketing of cannabis products, protecting and promoting public health through sound, science-based decision-making remains our top priority. We appreciate the FTC joining us on these and other actions to protect consumers from fraudulent CBD products."

As described in the warning letter issued to Rooted Apothecary, the company used product webpages, through its online store and social media websites, to make unfounded claims about its CBD products, and some of the products were also unlawfully marketed as dietary supplements. The agency has determined that CBD products cannot be marketed as dietary supplements.

Examples of the unsupported claims made by the company include:

- "Instead of synthetic chemical[s] that can have safety concerns, this blend uses the best of nature to help calm the inflammation and pain of teething, while also promoting sleepiness for your little one."
- "No matter what age, ear aches are a terrible, no good way to live each day! Our main priority was safety, effectiveness . . . as we formulated this for the entire family including our precious little ones. When the pain is bad, this roller goes to work for soothing pain, inflammation, and to battle against the bacterial/viral critters to blame."
- "Increasing evidence suggests that CBD oil is a powerful option for pain . . . anxiety . . . and autism . . . It seems like an attractive and safe option for children."
- "CBD oil may have neuroprotective properties and may protect against neurological conditions, such as Parkinson's and Alzheimer's disease."
- "[P]ossible uses for CBD include helping with skin problems such as acne, autism, ADHD, and even cancer. It's often used in conjunction with traditional treatments to provide extra help. Children can use high amounts of CBD safely and without any risk."

Additionally, under the Federal Trade Commission Act, it is unlawful to advertise that a product can prevent, treat, or cure human disease unless the advertiser possesses competent and reliable scientific evidence, including, when appropriate, well-controlled human clinical studies, substantiating that the claims are true at the time they are made. More generally, to make or exaggerate such claims, whether directly or indirectly, through the use of a product name, website name, metatags, or other means, without rigorous scientific evidence sufficient to substantiate the claims, violates the FTC Act. The FTC is concerned that one or more of the efficacy claims cited may not be substantiated by competent and reliable scientific evidence. These products are also misbranded under the Federal Food, Drug, and Cosmetic (FD&C) Act, because the products' labels and product information fail to include adequate directions for use. Drugs in the United States must contain directions explaining how a consumer can use a drug safely for its intended purpose. Under the law, there is an exemption for this labeling requirement for prescription drugs that have FDA-approved applications in effect. However, none of Rooted Apothecary's products are FDA-approved.

The FDA and FTC have requested responses from Rooted Apothecary within 15 working days stating how the company will correct the violations. Failure to correct the violations promptly may result in legal action, including product seizure and/or injunction. Violations of the FTC Act may result in legal action seeking a Federal District Court injunction or administrative cease and desist order, and an order also may require that a company pay back money to consumers.

The FDA continues to be concerned about the proliferation of products claiming to contain CBD that are marketed for therapeutic or medical uses that have not been approved by the agency. The FDA approval process ensures that drugs on the market are safe and effective for their intended therapeutic uses. CBD is marketed in a variety of product types, such as oil drops, capsules, syrups, teas and topical lotions and creams. The FDA has not approved any CBD products other than one prescription human drug product to treat rare, severe forms of epilepsy. There is very limited information for other marketed CBD products, which likely differ in composition from the FDA-approved product and have not been evaluated for potential adverse effects on the body.

Top

Table 4. Most similar and most distant (in cosine similarity) tweets to misinformation quotes regarding Alzheimer disease.

FDA[a] misinformation quote	Most similar tweet	Most distant tweet
"CBD[b] oil may have neuroprotective properties and may protect against neurological conditions, such as Parkinson's and Alzheimer's disease"	@<MASKED-USER> @<MASKED-USER> @<MASKED-USER> @<MASKED-USER> @<MASKED-USER> "On the up side, some evidence suggests THC[c] and CBD may be neuroprotective, so there's a rationale for some MMJ for you. Alzheimer's prophylaxis"	@<MASKED-USER> "do not believe he was the co-owner and while the store is definitely shitty, the sign did not say CBD cures autism"
"Possible uses for CBD include helping with skin problems such as acne, autism, ADHD[d], and even cancer. It's often used in conjunction with traditional treatments to provide extra help. Children can use high amounts of CBD safely and without any risk."	"CBD is recommended as a treatment for conditions such as seizures, depression and anxiety, and symptoms such as sleeplessness, inflammation, acne, and pain. It has also proven to be effective in treating autistic children. \nSource: <MASKED-URL> <MASKED-URL>"	"4) Sydney is demonstrating that she understands what effectual customer demands are, which is going straight to the source to demand change. She did not see a phone number she could call on the sign that promoted CBD as a cure-all for autism, so she went in to ask for one"

[a]FDA: Food and Drug Administration.

[b]CBD: cannabidiol.

[c]THC: tetrahydrocannabinol.

[d]ADHD: attention deficit hyperactivity disorder.

Discussion

Principal Findings

COVID-19–related misinformation can have fatal consequences [46]. Those consuming this misinformation may have misconceptions about how the virus is transmitted, disease symptoms, or health effects; they may communicate misinformation to others who subsequently spread it and put themselves and others at risk. Although there have been some preliminary studies on the benefits of cannabis in treating the symptoms of COVID-19, findings are uncertain, and at the time of this writing, cannabis is not an approved treatment by the FDA [47]. Therefore, claiming that CBD products can unequivocally treat or prevent COVID-19 is a federal violation in the United States, where the FDA has made numerous attempts to warn web-based retailers making false claims about their products [1]. Our study demonstrates how FDA Warning Letters can be used with transformer language models to identify tweets containing misinformation that are semantically similar to Warning Letters. Our approach reduces the time, labor, and potential monetary costs of other text classification methods. To our knowledge, this is the first study to use FDA Warning Letters as a foundation for identifying misinformation, particularly as it relates to CBD and COVID-19, in web-based social networks.

Transformer language models, such as the one used herein, are powerful tools that have been used for a number of purposes, such as translation [48,49] and text classification [50,51], and have even been used to generate descriptions of images in text [52]. Because of their ability to understand and summarize natural language, several studies have used these models to identify web-based misinformation, which often includes nuanced language and requires techniques that recognize not only semantic but also contextual similarities. Kumar et al [22] used Twitter to build a multilabel tweet classifier system in their study using a RoBERTa-large transformer language model to identify COVID-19–related misinformation. Their model could identify 4 classes: irrelevant, conspiracy, true information, and false information, and it achieved an F_1-score of 76%. Although this study focused on any class of misinformation that was identified in the FDA Warning Letters, future studies could further classify the letters by type and evaluate the accuracy of our approach.

Although this work was built on only Twitter platform, it is possible that it could also be applied to other social network platforms. Serrano et al [25] used transformer-based language models to identify YouTube videos containing COVID-19–related misinformation via comments posted on the videos. They built a text classifier to identify conspiracy-related content and concluded that YouTube videos containing misinformation were accompanied by user comments with a high percentage of conspiracy-related content [25]. Future studies can assess the performance of our method on other platforms.

In a recent approach to concept drift (changes in data and meaning over time) in Twitter data streams, Bechini et al [53] trained a semantic-based classifier using the BERT language model to examine the change in opinions about vaccines within a corpus of Italian tweets; this model outperformed other strategies, such as retraining the ensemble approaches. This suggests that transformer-based models, such as the one described herein, for identifying commercial tweets can be resilient over time. In addition, given more extensive (eg, "firehose" access) and future access to Twitter data, our misinformation tweet classifier could identify newer and current tweets that were not included in our data set, as well as identify tweets making similar violations in near real time. Furthermore, as previously noted, this technology can be applied to other forms of misinformation that threaten public health and safety.

Our approach to identifying tweets that make false claims about CBD and COVID-19 used quotes extracted from FDA Warning Letters to identify tweets that are semantically and contextually similar based on the cosine distance of sentence vectors. Compared with the approaches that require a large amount of data annotation, this substantially reduces the time required to identify the tweets making false health-related claims and flag them with a high amount of confidence. This is attributed to 2 factors: we used minimal manually annotated data for validation purposes and we used a simple calculation of cosine similarity between tweets and quotes.

Our study not only illustrates the scope of misleading information about CBD and COVID-19 in particular but also demonstrates an efficient and affordable approach to identifying other instances of this widespread problem—an approach that can be used by government entities, social networks, and message board administrators concerned with minimizing false advertising and misinformation and the potential threats they pose.

Limitations

This study had several limitations. First, it was built only on the Twitter social networking platform. In addition, although we did not explicitly acknowledge "bots" on Twitter, this was implicitly addressed during the annotation process where the tweets that appeared to be machine-generated were not considered cases of commercial CBD or misinformation. In addition, this model was trained on a collection of CBD tweets from before the COVID-19 pandemic, that is, before words like "COVID," "COVID-19," "corona," and "coronavirus" would have been associated with "CBD" and "cannabidiol." In this case, we applied a CBD commercial tweet classifier that was trained on tweets that were primarily authored in 2019 to a collection of tweets between 2020 and 2021. Although we did observe satisfactory results in testing this model's extraction of commercial CBD tweets that also mentioned COVID-19, we acknowledge that concept drift is always a potential factor in classifying streaming and should always be considered.

Conclusions

There is a clear and pronounced advertising presence on Twitter of loosely regulated substances touted to treat COVID-19, although this type of self-treatment lacks evidentiary support. Twitter is a medium known for rapid spread of medical misinformation, perhaps especially concerning substances like CBD [54-57]. The COVID-19 pandemic has been yet another opportunity for CBD marketers and sellers to mislead the public about CBD's role in treating or preventing the disease. Illegitimate web-based CBD sellers pose a public threat by spreading misinformation, selling unregulated products, and generally sidestepping regulations. Our approach to addressing this issue identified a semantic relationship between tweets containing false claims about CBD in treating COVID-19 and FDA Warning Letters. Using transformer language models and quotes from FDA Warning Letters to other CBD advertisers, this framework can be easily adapted to find misinformation about other conditions and substances, thereby potentially serving a crucial purpose in benefiting public health.

Acknowledgments

The authors wish to thank Mr Steve Claas at the University of Kentucky College of Public Health Office of Scientific Writing for reviewing this manuscript.

Authors' Contributions

JT, MK, and RV-S conceived and designed the study and cowrote and revised the final manuscript. JT and AGB collected and analyzed the data.

Conflicts of Interest

None declared.

References

1. Fraudulent Coronavirus Disease 2019 (COVID-19) products. U.S. Food & Drug Administration. 2022 Oct 26. URL: https://www.fda.gov/consumers/health-fraud-scams/fraudulent-coronavirus-disease-2019-covid-19-products [accessed 2023-01-02]
2. Lee JJ, Kang K, Wang MP, Zhao SZ, Wong JY, O'Connor S, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. J Med Internet Res 2020 Nov 13;22(11):e22205 [FREE Full text] [doi: 10.2196/22205] [Medline: 33048825]
3. Kearney MD, Chiang SC, Massey PM. The Twitter origins and evolution of the COVID-19 "plandemic" conspiracy theory. HKS Misinfo Review. 2020 Oct 9. URL: https://misinforeview.hks.harvard.edu/article/the-twitter-origins-and-evolution-of-the-covid-19-plandemic-conspiracy-theory/ [accessed 2023-01-02]
4. Ayers JW, Chu B, Zhu Z, Leas EC, Smith DM, Dredze M, et al. Spread of misinformation about face masks and COVID-19 by automated software on facebook. JAMA Intern Med 2021 Sep 01;181(9):1251-1253 [FREE Full text] [doi: 10.1001/jamainternmed.2021.2498] [Medline: 34096988]
5. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. J Med Internet Res 2022 Aug 04;24(8):e37367 [FREE Full text] [doi: 10.2196/37367] [Medline: 35816685]
6. Malinowska B, Baranowska-Kuczko M, Kicman A, Schlicker E. Opportunities, challenges and pitfalls of using cannabidiol as an adjuvant drug in COVID-19. Int J Mol Sci 2021 Feb 17;22(4):1986 [FREE Full text] [doi: 10.3390/ijms22041986] [Medline: 33671463]

7. Brown JD. Cannabidiol as prophylaxis for SARS-CoV-2 and COVID-19? Unfounded claims versus potential risks of medications during the pandemic. Res Social Adm Pharm 2021 Jan;17(1):2053 [FREE Full text] [doi: 10.1016/j.sapharm.2020.03.020] [Medline: 32273254]
8. Khalsa JH, Bunt G, Maggirwar SB, Kottilil S. COVID-19 and Cannabidiol (CBD). J Addict Med 2021;15(5):355-356 [FREE Full text] [doi: 10.1097/ADM.0000000000000771] [Medline: 33323690]
9. Iseger TA, Bossong MG. A systematic review of the antipsychotic properties of cannabidiol in humans. Schizophr Res 2015 Mar;162(1-3):153-161. [doi: 10.1016/j.schres.2015.01.033] [Medline: 25667194]
10. Smolev ET, Rolf L, Zhu E, Buday SK, Brody M, Brogan DM, et al. "Pill pushers and CBD oil"-a thematic analysis of social media interactions about pain after traumatic brachial plexus injury. J Hand Surg Glob Online 2021 Jan;3(1):36-40 [FREE Full text] [doi: 10.1016/j.jhsg.2020.10.005] [Medline: 33537664]
11. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
12. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 09;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
13. Fake news in the time of C-19. Tortoise. 2020 Mar 23. URL: https://www.tortoisemedia.com/2020/03/23/the-infodemic-fake-news-coronavirus/ [accessed 2023-01-02]
14. Trew B. Coronavirus: hundreds dead in Iran from drinking methanol amid fake reports it cures disease. Independent. 2020 Mar 27. URL: https://www.independent.co.uk/news/world/middle-east/iran-coronavirus-methanol-drink-cure-deaths-fake-a9429956.html [accessed 2023-01-02]
15. Shabad R. Twitter removes tweet highlighted by Trump falsely claiming COVID-19 'cure'. NBC News. 2020 Jul 28. URL: https://www.nbcnews.com/politics/donald-trump/twitter-removes-tweet-highlighted-trump-falsely-claiming-covid-cure-n1235075 [accessed 2023-01-02]
16. Ozawa S, Billings J, Sun Y, Yu S, Penley B. COVID-19 treatments sold online without prescription requirements in the United States: cross-sectional study evaluating availability, safety and marketing of medications. J Med Internet Res 2022 Feb 16;24(2):e27704 [FREE Full text] [doi: 10.2196/27704] [Medline: 34662286]
17. Mackey TK, Li J, Purushothaman V, Nali M, Shah N, Bardier C, et al. Big data, natural language processing, and deep learning to detect and characterize illicit COVID-19 product sales: infoveillance study on twitter and Instagram. JMIR Public Health Surveill 2020 Aug 25;6(3):e20794 [FREE Full text] [doi: 10.2196/20794] [Medline: 32750006]
18. Yazdany J, Kim AH. Use of hydroxychloroquine and chloroquine during the COVID-19 pandemic: what every clinician should know. Ann Intern Med 2020 Jun 02;172(11):754-755 [FREE Full text] [doi: 10.7326/M20-1334] [Medline: 32232419]
19. Della Porta A, Bornstein K, Coye A, Montrief T, Long B, Parris MA. Acute chloroquine and hydroxychloroquine toxicity: a review for emergency clinicians. Am J Emerg Med 2020 Oct;38(10):2209-2217 [FREE Full text] [doi: 10.1016/j.ajem.2020.07.030] [Medline: 33071096]
20. 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. U.S. Food & Drug Administration. 2020 Jan 7. URL: https://www.fda.gov/drugs/drug-safety-and-availability/fda-cautions-against-use-hydroxychloroquine-or-chloroquine-covid-19-outside-hospital-setting-or [accessed 2023-01-02]
21. 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]
22. Kumar S, Pranesh RR, Carley KM. A fine-grained analysis of misinformation in COVID-19 tweets. Research Square. 2021 Jun 7. URL: http://paperpile.com/b/4ySh57/T7tx [accessed 2023-01-02]
23. Hayawi K, Shahriar S, Serhani MA, Taleb I, Mathew SS. ANTi-Vax: a novel Twitter dataset for COVID-19 vaccine misinformation detection. Public Health 2022 Feb;203:23-30 [FREE Full text] [doi: 10.1016/j.puhe.2021.11.022] [Medline: 35016072]
24. Schroeder DT, Schaal F, Filkukova P, Pogorelov K, Langguth J. WICO Graph: a labeled dataset of twitter subgraphs based on conspiracy theory and 5G-corona misinformation tweets. In: Proceedings of the 13th International Conference on Agents and Artificial Intelligence (ICAART 2021. Setúbal, Portugal: SCITEPRESS – Science and Technology Publications; 2021.
25. Serrano J, Papakyriakopoulos O, Hegelich S. NLP-based feature extraction for the detection of COVID-19 misinformation videos on youtube. In: Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020. 2020 Presented at: 1st Workshop on NLP for COVID-19 at ACL 2020; Jul, 2020; Virtual URL: https://aclanthology.org/2020.nlpcovid19-acl.17
26. Choudrie J, Banerjee S, Kotecha K, Walambe R, Karende H, Ameta J. Machine learning techniques and older adults processing of online information and misinformation: a covid 19 study. Comput Human Behav 2021 Jun;119:106716 [FREE Full text] [doi: 10.1016/j.chb.2021.106716] [Medline: 34866770]
27. Janmohamed K, Soale A, Forastiere L, Tang W, Sha Y, Demant J, et al. Intersection of the web-based vaping narrative with COVID-19: topic modeling study. J Med Internet Res 2020 Oct 30;22(10):e21743 [FREE Full text] [doi: 10.2196/21743] [Medline: 33001829]
28. 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]

29. Tian L, Zhang X, Wang Y, Liu H. Early detection of rumours on twitter via stance transfer learning. In: Advances in Information Retrieval. Cham: Springer; Apr 8, 2020.
30. About warning and close-out letters. U.S. Food & Drug Administration. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/about-warning-and-close-out-letters [accessed 2023-01-02]
31. Regulatory procedures manual. U.S. Food & Drug Administration. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/compliance-manuals/regulatory-procedures-manual [accessed 2023-01-02]
32. Warning letters. U.S. Food & Drug Administration. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/compliance-actions-and-activities/warning-letters [accessed 2023-01-02]
33. Ormerod M, Martínez Del Rincón J, Devereux B. Predicting semantic similarity between clinical sentence pairs using transformer models: evaluation and representational analysis. JMIR Med Inform 2021 May 26;9(5):e23099 [FREE Full text] [doi: 10.2196/23099] [Medline: 34037527]
34. A social networking service scraper in Python. GitHub. URL: https://github.com/JustAnotherArchivist/snscrape [accessed 2023-01-02]
35. Turner J, Kantardzic M, Vickers-Smith R. Infodemiological examination of personal and commercial tweets about cannabidiol: term and sentiment analysis. J Med Internet Res 2021 Dec 20;23(12):e27307 [FREE Full text] [doi: 10.2196/27307] [Medline: 34932014]
36. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, et al. Attention is all you need. arXiv 2017 [FREE Full text]
37. Warning letter Avazo-healthcare, llc. U.S. Food & Drug Administration. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/avazo-healthcare-llc-610282-11302020 [accessed 2023-01-02]
38. Warning letter for our vets llc dba patriot supreme. U.S. Food & Drug Administration. URL: https://tinyurl.com/ms68eh8v [accessed 2023-01-02]
39. Warning letter Nova Botanix LTD DBA CanaBD. U.S. Food & Drug Administration. URL: https://tinyurl.com/58tae34n [accessed 2023-01-02]
40. Warning letter project 1600 inc. U.S. Food & Drug Administration. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/project-1600-inc-608349-06182020 [accessed 2023-01-02]
41. U.S. Food & Drug Administration. URL: https://tinyurl.com/ymhrbkr9 [accessed 2023-01-02]
42. Warning letter CBD online store. U.S. Food & Drug Administration. URL: https://www.fda.gov/inspections-compliance-enforcement-and-criminal-investigations/warning-letters/cbd-online-store-606587-04072020 [accessed 2023-01-02]
43. Warning letter Sanit Technologies LLC dba Durisan. U.S. Food & Drug Administration. URL: https://tinyurl.com/2p8v8ncx [accessed 2023-01-02]
44. Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv 2019 [FREE Full text]
45. Ni J, Ábrego G, Constant N, Ma J, Hall K, Cer D, et al. Sentence-T5: scalable sentence encoders from pre-trained text-to-text models. arXiv 2021 [FREE Full text]
46. Love JS, Blumenberg A, Horowitz Z. The parallel pandemic: medical misinformation and COVID-19 : primum non nocere. J Gen Intern Med 2020 Aug;35(8):2435-2436 [FREE Full text] [doi: 10.1007/s11606-020-05897-w] [Medline: 32410129]
47. Byrareddy SN, Mohan M. SARS-CoV2 induced respiratory distress: can cannabinoids be added to anti-viral therapies to reduce lung inflammation? Brain Behav Immun 2020 Jul;87:120-121 [FREE Full text] [doi: 10.1016/j.bbi.2020.04.079] [Medline: 32360437]
48. Maruyama T, Yamamoto K. Extremely low resource text simplification with pre-trained transformer language model. In: Proceedings of the 2019 International Conference on Asian Language Processing (IALP). 2019 Presented at: 2019 International Conference on Asian Language Processing (IALP); Nov 15-17, 2019; Shanghai, China URL: http://paperpile.com/b/4ySh57/cYQM [doi: 10.1109/IALP48816.2019.9037650]
49. Voita E, Sennrich R, Titov I. The bottom-up evolution of representations in the transformer: a study with machine translation and language modeling objectives. arXiv 2019 [FREE Full text]
50. Zhang J, Chang WZ, Yu HF, Dhillon IS. Fast multi-resolution transformer fine-tuning for extreme multi-label text classification. arXiv 2021 [FREE Full text]
51. Lu H, Ehwerhemuepha L, Rakovski C. A comparative study on deep learning models for text classification of unstructured medical notes with various levels of class imbalance. BMC Med Res Methodol 2022 Jul 02;22(1):181 [FREE Full text] [doi: 10.1186/s12874-022-01665-y] [Medline: 35780100]
52. Zhou Y, Zhang Y, Hu Z, Wang M. Semi-autoregressive transformer for image captioning. In: Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). 2021 Presented at: 2021 IEEE/CVF

International Conference on Computer Vision Workshops (ICCVW)q; Oct 11-17, 2021; Montreal, BC, Canada URL: http://paperpile.com/b/4ySh57/JW4t [doi: 10.1109/iccvw54120.2021.00350]
53. Bechini A, Bondielli A, Ducange P, Marcelloni F, Renda A. Addressing event-driven concept drift in twitter stream: a stance detection application. IEEE Access 2021;9:77758-77770 [FREE Full text] [doi: 10.1109/ACCESS.2021.3083578]
54. Pang RD, Dormanesh A, Hoang Y, Chu M, Allem J. Twitter posts about cannabis use during pregnancy and postpartum: a content analysis. Subst Use Misuse 2021;56(7):1074-1077. [doi: 10.1080/10826084.2021.1906277] [Medline: 33821757]
55. Walsh-Buhi ER. Social media and cancer misinformation: additional platforms to explore. Am J Public Health 2020 Oct;110(S3):S292-S293. [doi: 10.2105/AJPH.2020.305949] [Medline: 33001721]
56. Cresswell L, Espin-Noboa L, Murphy MS, Ramlawi S, Walker MC, Karsai M, et al. The volume and tone of Twitter posts about cannabis use during pregnancy: protocol for a scoping review. JMIR Res Protoc 2022 Mar 29;11(3):e34421 [FREE Full text] [doi: 10.2196/34421] [Medline: 35348465]
57. Bellon-Harn ML, Ni J, Manchaiah V. Twitter usage about autism spectrum disorder. Autism 2020 Oct;24(7):1805-1816. [doi: 10.1177/1362361320923173] [Medline: 32508126]

Abbreviations

BERT: Bidirectional Encoder Representations from Transformers
CBD: cannabidiol
FDA: Food and Drug Administration

Edited by T Mackey; submitted 30.03.22; peer-reviewed by A Pal, SF Tsao; comments to author 16.05.22; revised version received 07.09.22; accepted 30.11.22; published 23.01.23.

Please cite as:
Turner J, Kantardzic M, Vickers-Smith R, Brown AG
Detecting Tweets Containing Cannabidiol-Related COVID-19 Misinformation Using Transformer Language Models and Warning Letters From Food and Drug Administration: Content Analysis and Identification
JMIR Infodemiology 2023;3:e38390
URL: https://infodemiology.jmir.org/2023/1/e38390
doi:10.2196/38390
PMID:36844029

Original Paper

The Quality, Readability, and Accuracy of the Information on Google About Cannabis and Driving: Quantitative Content Analysis

Maria Josey[1], PhD; Dina Gaid[1], PhD; Lisa D Bishop[1], PharmD; Michael Blackwood[1], BSc; Maisam Najafizada[1], MD, PhD; Jennifer R Donnan[1], PhD, MBA

School of Pharmacy, Memorial University of Newfoundland, St. John's, NL, Canada

Corresponding Author:
Jennifer R Donnan, PhD, MBA
School of Pharmacy, Memorial University of Newfoundland
300 Prince Philip Drive
St. John's, NL, A1B 3V6
Canada
Phone: 1 709 864 3587
Email: jennifer.donnan@mun.ca

Abstract

Background: The public perception of driving under the influence of cannabis (DUIC) is not consistent with current evidence. The internet is an influential source of information available for people to find information about cannabis.

Objective: The purpose of this study was to assess the quality, readability, and accuracy of the information about DUIC found on the internet using the Google Canada search engine.

Methods: A quantitative content analysis of the top Google search web pages was conducted to analyze the information available to the public about DUIC. Google searches were performed using keywords, and the first 20 pages were selected. Web pages or web-based resources were eligible if they had text on cannabis and driving in English. We assessed (1) the quality of information using the Quality Evaluation Scoring Tool (QUEST) and the presence of the Health on the Net (HON) code; (2) the readability of information using the Gunning Fox Index (GFI), Flesch Reading Ease Scale (FRES), Flesch-Kincaid Grade Level (FKGL), and Simple Measure of Gobbledygook (SMOG) scores; and (3) the accuracy of information pertaining to the effects of cannabis consumption, prevalence of DUIC, DUIC effects on driving ability, risk of collision, and detection by law enforcement using an adapted version of the 5Cs website evaluation tool.

Results: A total of 82 web pages were included in the data analysis. The average QUEST score was 17.4 (SD 5.6) out of 28. The average readability scores were 9.7 (SD 2.3) for FKGL, 11.4 (SD 2.9) for GFI, 12.2 (SD 1.9) for SMOG index, and 49.9 (SD 12.3) for FRES. The readability scores demonstrated that 8 (9.8%) to 16 (19.5%) web pages were considered readable by the public. The accuracy results showed that of the web pages that presented information on each key topic, 96% (22/23) of them were accurate about the effects of cannabis consumption; 97% (30/31) were accurate about the prevalence of DUIC; 92% (49/53) were accurate about the DUIC effects on driving ability; 80% (41/51) were accurate about the risk of collision; and 71% (35/49) were accurate about detection by law enforcement.

Conclusions: Health organizations should consider health literacy of the public when creating content to help prevent misinterpretation and perpetuate prevailing misperceptions surrounding DUIC. Delivering high quality, readable, and accurate information in a way that is comprehensible to the public is needed to support informed decision-making.

(JMIR Infodemiology 2023;3:e43001) doi:10.2196/43001

KEYWORDS

cannabis; driving; quality; readability; accuracy; public education; internet; Google search; analysis; accessibility; information; evaluation; tool; data; misinterpretation

Introduction

In October 2018, the use of nonmedical cannabis became legal in Canada [1]. By the end of 2020, approximately 20% of Canadians, aged 15 years and older, reported using cannabis over the previous 3 months [2]. Certain cannabis use behaviors can increase the risk of experiencing harmful effects [3], such as daily use of cannabis, using cannabis products with high tetrahydrocannabinol content, or driving under the influence of cannabis (DUIC) [4]. In Canada, approximately 1-2 out of every 5 cannabis consumers engage in some form of risky behaviors [3], with 4%-12% of all injuries and deaths from motor vehicle accidents being attributed to DUIC [5]. Additionally, 40% of participants in a Canada-wide survey reported riding with a cannabis-impaired driver within the past year [6].

There are mixed perceptions among the general public regarding the true risks associated with cannabis use [7-9]. In particular, the mixed beliefs regarding the risks associated with DUIC are concerning given the potential impact on both the consumer and innocent members of the public. Recent literature reported that perception of risks associated with DUIC is low, with one study reporting that 28% of participants believed there was no increased risk of accidents [6]. Another study reported that of those who participated in DUIC, 43% believed it was not a risky behavior [10]. This highlights the need to ensure cannabis consumers have access to evidence-based information to support informed decision-making [7,11,12].

Although information about cannabis can be retrieved from numerous sources, one study reported that 78% of participants relied on knowledge gained from their own personal experiences, while 39% obtained information from the internet [13]. Cannabis-related Google searches increased by 75% between 2004 and 2016 [14,15]; however, the trustworthiness of information retrieved on the internet is questionable. There have been studies that explored the quality of cannabis labels from products sold on the web [16], the accuracy of cannabis claims on common websites [17], and the quality of cannabis-related information in magazines and newspapers [18,19]. In general, these studies reported that the quality of cannabis-related information were very poor. Among studies that specifically looked at cannabis health claims on the internet, one found only 5% of claims on the health benefits of cannabis aligned with evidence [20]. Other studies reported that web-based information about cannabis use for pain was biased as sources often neglected to discuss potential risks [21] or were just unreliable [22]. This points to variable quality of cannabis-related information available on the internet [20,23-26].

Web-based search trends related to health-related topics provide insight on the public perception or cannabis use, which also reflect the availability of public health resources [27]. Taking into consideration that searches related to cannabis increased by 75% on Google from 2004 to 2016 [14,15], high quality, easily accessible, evidence-based information is needed for individuals to make informed decisions about cannabis use behaviors [28], which is especially important given the prevalent misconceptions about DUIC. However, the quality, readability, and accuracy of information found through the Google search engine on DUIC are still unknown [6]. The purpose of this study was to assess the quality, readability, and accuracy of information about DUIC found on the internet through the Google search engine.

Methods

Study Design

A quantitative content analysis about DUIC was performed on public web pages using the Google Canada search engine.

Eligibility Criteria

To be included, the web page had to (1) have information related to cannabis and driving, (2) be available in English, (3) be accessible with no fee, (4) have text to analyze, and (5) be available at the time of analysis. Web pages were excluded if (1) the page became no longer available during analysis and (2) the web page only contained images.

Data Collection

Web pages were identified through the Google search engine. Google was chosen because it is the dominant search engine in Canada, holding 91.98% of market shares [29], and one study showed that 89.8% of people preferred using Google [30]. A private search through incognito mode was used to avoid the search history from biasing results. Six separate Google searches were performed using the terms outlined in Textbox 1, and the first 20 URLs were collected from each search. Neutral search terms were used to ensure the collected web pages were not biased in one direction. The first 20 URLs were collected, as most people consider no more than the first 20 web pages when performing an internet search [15,31]. For our study, one researcher (SS or MJ) extracted web page addresses with Google Chrome (version 99.0.4844.51) [32]. The search was first completed in October 2021 using the first 4 search terms and then repeated fully in April 2022 after 2 new search terms were added.

Textbox 1. Search terms used to collect web pages for analysis.

Google search terms

- Cannabis AND driving
- Marijuana AND driving
- Pot AND driving
- Weed AND driving
- Driving high
- Driving stoned

Data Analysis

Web pages were organized into categories based on categories used in similar studies that assessed the quality of health-related information on the internet [33]. These categories included digital media, commercial web pages, government organizations, health organizations, nonprofit foundations, peer-reviewed materials, and “other.”

Outcome Measures

Quality of Information

The quality of the information was measured by 2 tools: the Health on the Net (HON) code and the Quality Evaluation Scoring Tool (QUEST).

HON Code

HON is a nonprofit foundation aimed to assess and evaluate the quality of web-based health information [34]. The HON certification is designed so that people of the general public can identify trustworthy sources of information [34] and has been used in previous research to evaluate health-related websites as a beneficial tool that shows the intent of a website to publish high-quality information [35-38].The HON code seeks to promote trustworthy health information for the benefit of internet users [39]. HON code is a voluntary certification used on health websites, indicating that their 8 principles were fulfilled. Those principles relate to the authority, complementarity, confidentiality, attribution, justifiability, transparency, financial disclosure, and advertisement policy of the website content [40]. This certification aims to certify websites that are reliable and of high quality, so it is an easy measure for the general public to quickly determine if the web page is a trustworthy source of health information.

The QUEST Tool

The QUEST tool serves as a standard for assessing the quality of web-based health information that does not rely on users’ subjective judgment [41]. The QUEST tool was chosen as it has been validated and assessed for reliability and provides a numeric score allowing for quantitative analysis [42]. The QUEST tool was validated for both treatment and preventative measures of web-based health care information [41] and has since been used in studies to evaluate web-based health care information on various topics, including papillomavirus and oropharyngeal cancer, COVID-19, and using electronic cigarettes [33,43,44]. Additionally, this tool is used for a broad range of health topics as opposed to more focused health topics (eg, treatment) [41]. The QUEST tool assesses 7 aspects of the website information and provides a weighted score out of 28 (Table 1) [41]. Three independent researchers collaborated to assess the quality of the web page, while each page was assessed by at least two researchers (SS, MJ, MB), and any discrepancies were discussed. For our study, if an organization took ownership over the text (rather than a specific author), we gave a score of 1, meaning “all other indications of authorship” on the QUEST tool scoring. Additionally, any language that promoted the sale of cannabis (eg, cannabis brand) or directed the reader to a specific location for purchase was given a score of 1 accordingly under the “Conflicts of Interest” section of the QUEST scoring tool. For example, any mention of a specific cannabis dispensary, even if indirectly mentioned through a picture identifying a dispensary, was considered an endorsement, and therefore, had the potential to be biased.

Table 1. Description of the Quality Evaluation Scoring Tool (QUEST) criteria to evaluate the quality of web-based health information [41]. Scores in the individual sections are weighted and summed to generate a total score of up to 28. This tool is reproduced and distributed under the terms of the Creative Commons Attribution 4.0 International License [45].

Characteristics	Score
Authorship	Score x 1
0: No indication of authorship or username	
1: All other indications of authorship	
2: Author's name and qualification clearly stated	
Attribution	Score x 3
0: No sources	
1: Mention of expert source, research, research findings (although with insufficient information to identify the specific studies), links to various sites, advocacy body, or other	
2: Reference to at least one identifiable scientific study, regardless of format (eg, information in text or reference list)	
3: Reference to mainly identifiable scientific studies, regardless of format (in >50% of claims)	
Type of study (for all articles scoring 2 or 3 on "attribution")	Score x 1
0: In vitro, animal models, and editorials	
1: All observational works	
2: Meta-analyses, randomized controlled trials, and clinical studies	
Conflicts of interest	Score x 3
0: Endorsement or promotion of intervention designed to prevent or treat condition (eg, supplements, brain training games, and foods) within the article	
1: Endorsement or promotion of educational products and services (eg, book and care home services)	
2: Unbiased information	
Currency	Score x 1
0: No date present	
1: Article is dated but is 5 years or older	
2: Article is dated within the last 5 years	
Complimentary	Score x 1
0: No support of the patient-physician relationship	
1: Support of the patient-physician relationship	
Tone (includes title)	Score x 3
0: Fully supported—authors fully and unequivocally support the claims; strong vocabulary is used, such as "cure," "guarantee," and "easy"; use of nonconditional verb tenses mostly (eg, "can" and "will"); and no discussion of limitations	
1: Mainly supported (authors mainly support their claims but with more cautious vocabulary, such as "can reduce your risk" or "may help prevent", and no discussion of limitations)	
2: Balanced or cautious support (authors' claims are balanced by caution and include statements of limitations and contrasting findings)	

Readability

Web page content was assessed for readability by the general public using 4 different scales, including Gunning Fox Index (GFI), Flesch Reading Ease Scale (FRES), Flesh-Kincaid Grade Level (FKGL), and the Simple Measure of Gobbledygook (SMOG) scale (Table 2). There are many scales to measure the readability of information [46], but there is no universally accepted measurement of readability. Therefore, the combination of these 4 readability scores (ie, GFI, FRES, FKGL, and SMOG) has been used together to measure the readability of health information [33,47] in this study. Each web page URL was submitted to the Readable [48] web-based scoring tool by one researcher (DG). If the URL was directed to a PDF, the text was manually entered into the web-based generator by copying and pasting the titles and content. Text were excluded from the analysis if they were advertisements, hyperlinks, author names, or references, as these could bias the results [47]. The scores were compared to a value unique to each readability tool that indicated the content was universally readable.

Table 2. Tools used to measure readability, their range of scores, the score correlated to text that is readable by the general public, and the formula used to calculate the score.

Readability tool	Range	Readable by the general public	Formula
GFI[a]	0-20	<8 [49]	
FRES[b]	0-100	>60 [47]	
FKGL[c]	0-18	<8 [47,50]	
SMOG[d]	—[e]	<10 [47]	

[a]GFI: Gunning Fox Index.

[b]FRES: Flesch Reading Ease Scale.

[c]FKGL: Flesch-Kincaid Grade Level.

[d]SMOG: Simple Measure of Gobbledygook.

[e]Not applicable.

Accuracy

The 5Cs Website Evaluation Tool is a structured tool that evaluates websites using 36 questions, grouped into the following 5 accuracy criteria: credibility, currency, content, construction, and clarity [51]. Since the construction, credibility, currency, and content of websites included in this study were assessed with the quality and readability tools, we only applied the content criteria.

The tool asks if the information on the website is evidence based and represents information from published journals and books [51]. To complete this assessment, current evidence from peer-reviewed journals was gathered, as they pertain to 5 key topics related to cannabis and driving (Multimedia Appendix 1). These topics include (1) the effects of cannabis consumption, (2) the prevalence of DUIC, (3) the effects of cannabis on driving performance, (4) risk of collision after using cannabis, and (5) the detection of cannabis-impaired drivers by law enforcement.

Each web page was assessed for the content across the 5 key categories. For each topic, the web page content was rated as accurate, not accurate, mixed accuracy (ie, some statements were accurate and some were not, or information was not aligning with the literature), or information not present. Each web page was rated independently by 2 researchers (MJ and DG); discrepancies were discussed and resolved. Web pages categorized as peer-reviewed (ie, peer-reviewed journal articles) were not included in the accuracy analysis, as peer-reviewed literature was used to create the evidence-based summary used in the content assessment. This approach has been used by others conducting similar content analyses [21].

Statistical Analysis

Descriptive statistics was performed with the mean (μ), standard deviation (σ), and total sample size (n) being reported for the average QUEST score of all web pages and by category. To assess correlations between QUEST scores and readability scores (ie, GFI, FRES, FKGL, and SMOG), a Pearson 2-tailed test was performed [33]. To assess the QUEST score for the presence of the HON code, an unpaired 1-tailed *t* test was performed, testing if the HON code was present on web pages with higher QUEST scores [33].

Ethics Approval

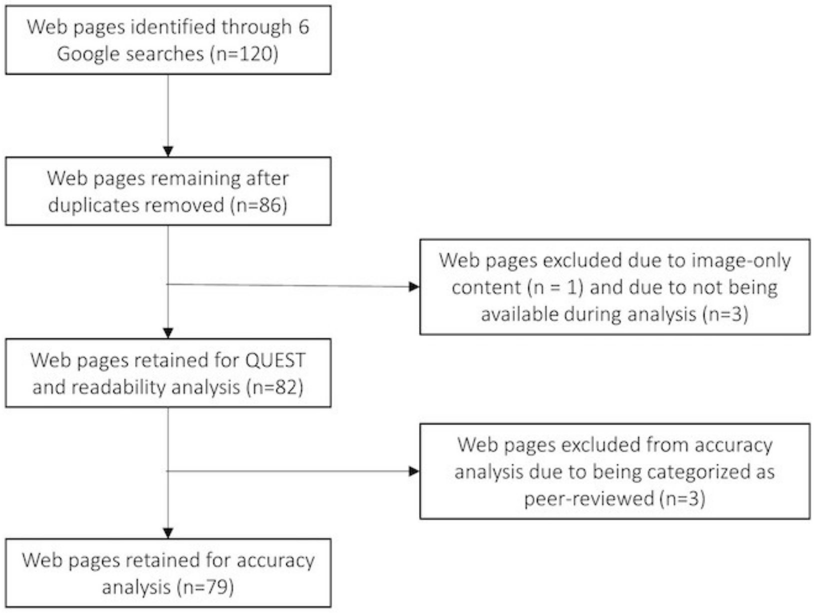
This study was exempted from ethical approval because it does not involve human participants.

Results

Overview

A total of 120 web pages were identified for analysis (Multimedia Appendix 2). Of the 120 web pages, 34 were removed as duplicate web pages, and 4 were removed as they did not meet the eligibility criteria, leaving 82 web pages included in the study (Figure 1). Of these, 40% (33/82) of web pages were categorized as digital media, 20% (16/82) as commercial web pages, 13% (11/82) as government organizations, 12% (10/82) as health organizations, 10% (8/82) as nonprofit foundations, 4% (3/82) as peer-reviewed content, and 1% (1/82) as "other." Multimedia Appendix 3 presents the web pages included in the data analysis.

Figure 1. Included and excluded web pages for quantitative analysis of quality, readability, and accuracy. QUEST: quality evaluation scoring tool.



Quality

The range of the QUEST scores was between 7 and 27, with the average Quest score being 17.4 (SD 5.6) out of a total of 28 (Table 3). Average QUEST scores by category showed that the peer-reviewed category had the highest quality with a score of 26.3 out of 28, and government web pages scored the lowest at 10.0/28.

The HON code was only present on 4 (5%) web pages, and they were found in the categories labelled as commercial (n=2), nonprofit (n=1), or health organization (n=1). There was no significant difference (*P*=.2) between the presence of a HON code on a website and the QUEST score without a HON code. Multimedia Appendix 4 presents the data from the full quality evaluation for each web page.

Table 3. Quality Evaluation Scoring Tool (QUEST) scores by category.

Category	QUEST score		
	Mean (μ)	SD (σ)	Total, n
Peer-reviewed content	26.3	0.6	3
Health organizations	20.5	5.6	10
Other	N/A[a]	N/A	1
Digital media	19.5	3.5	33
Commercial	15.7	6.0	16
Nonprofit foundations	14.9	4.2	8
Government	10.0	2.5	11
Total	18.1	5.6	82

[a]N/A: not applicable.

Readability

The average readability scores were 9.7 (SD 2.3) for FKGL, 11.4 (SD 2.8) for GFI, 12.2 (SD 1.9) for SMOG index, and 49.9 (SD 12.3) for FRES. Assessing the readability scores for all web pages in relation to the universal readability score, 19.5% (16/82) of the web pages were universally readable by the FKGL score (score <8 considered universally readable), 16% (13/82) by the FRES score (score >60 considered universally readable), 11.1% (9/82) by the SMOG index (score <10 considered universally readable), and 9.8% (8/82) by the GFI score (score <8 considered universally readable). None of the web pages in the peer-reviewed or other categories were considered universally readable by any readability scoring tool (Table 4). Multimedia Appendix 5 presents the readability scores for each web page.

Table 4. Web pages by category that were considered universally readable.

Category	Web pages, n (%)			
	FKGL[a]	GFI[b]	SMOG[c]	FRES[d]
Commercial (n=16)	3 (19)	2 (13)	1 (6)	4 (25)
Digital media (n=33)	4 (12)	2 (6)	3 (9)	3 (9)
Government (n=11)	4 (36)	3 (27)	3 (27)	4 (36)
Other (n=1)	0 (0)	0 (0)	0 (0)	0 (0)
Nonprofit foundations (n=8)	2 (25)	0 (0)	1 (13)	1 (13)
Peer-reviewed material (n=3)	0 (0)	0 (0)	0 (0)	0 (0)
Health organization (n=10)	3 (30)	1 (10)	1 (10)	1 (10)

[a]FKGL: Flesh-Kincaid Grade Level.
[b]GFI: Gunning Fox Index.
[c]SMOG: Simple Measure of Gobbledygook.
[d]FRES: Flesch Reading Ease Scale.

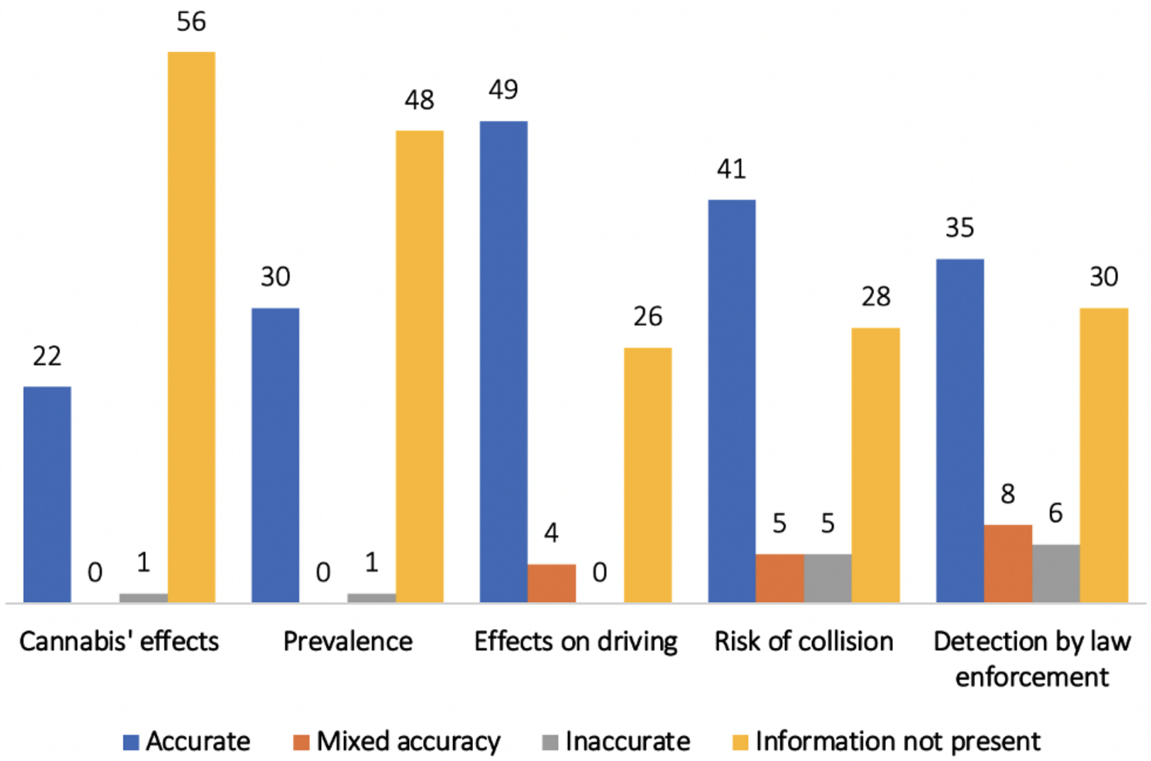
Correlation Between Quality and Readability

A Pearson 2-tailed test showed a significant positive correlation between the QUEST score and the FKGL (r=0.41; P<.001), GFI (r=0.28; P=.01), and SMOG (r=0.34; P=.002) readability scores. A negative correlation was found between the QUEST score and the FRES score (r=–0.40; P<.001).

Accuracy

Of the 79 web pages that were eligible to be reviewed for accuracy, 23 web pages discussed information related to the timing of the effects from cannabis consumption; 31 web pages were related to the prevalence of DUIC; 53 were related to the effects of cannabis impairment on driving ability; 51 were related to the risk of collision; and 49 had information related to detection by law (Figure 2). From those, 96% (22/23) had accurate information on the effects from cannabis consumption; 97% (30/31) of the web pages had accurate information about the prevalence of DUIC; 92% (49/53) of the web pages presented accurate information on the effects of cannabis impairment on driving ability; 80% (41/51) of the web pages had accurate information on the risk of collision; and 71% (35/49) of the web pages presented accurate information on detection by law. Sample excerpts from web pages and accuracy categorization are included in Multimedia Appendix 6.

Figure 2. Accuracy ratings of web page content. Content accuracy of the 5 key topics about driving under the influence of cannabis are represented with colors.



Discussion

Principal Findings

This study assessed the quality, readability, and accuracy of information about DUIC found on the internet using the Google search engine. Our findings showed that peer-reviewed papers had the highest quality of information; however, these web pages were not considered universally readable. The difficulty with comprehension may lead to misinterpretation and inaccurate expectations [52,53].

Surprisingly, our research indicated that government web pages were rated as having the lowest quality, contrary to the general perception that government sources would contain high-quality information. This low rating was likely attributed to the tone of the presented text, given much of the information on government websites was one-sided, used strong words such as "will," and did not discuss limitations of the information presented. This low quality could be due to the fact that government websites often presented information on laws and regulations and did not provide references to other information. This is unfortunate, as government web pages are typically viewed as an accurate source of information as indicated by various academic guides for evaluating information sources [54,55].

Readability for the public was problematic for most pages, with less than 20% of all web pages considered readable based on the FKGL, GFI, SMOG, and FRES readability tools. The majority of the content was written at a higher level of reading, which would often be used in academic settings or postsecondary education. The paucity of web pages written at levels that were considered universally readable was consistent with other health information topics on the internet (eg, general surgical procedures [56] and total joint arthroplasty [57]), suggesting that this could be a wider issue than solely information on cannabis [33,47,58]. Kruger et al [58] also suggests that significant efforts are still needed to provide accurate cannabis-related information on the internet for the health and safety of individuals and society [58].

The readability could be contributing to the misperceptions and behaviors; however, further studies assessing the interpretation of high-quality information with low readability scores could be beneficial. Associations and health advocacy groups should consider the health literacy of the public [59] when creating content to educate the public on DUIC. In addition, a more active form of education for the public could be beneficial as opposed to the passive information presented on a web page.

Our research shows that 80% of the information available about DUIC and its risks for accidents was accurate. However, although most information on DUIC was accurate, it was the lack of complete information that was most concerning. Of the 79 web pages that were analyzed for information about DUIC, 48% (n=38) either had no information on the risk of collision or had mixed or inaccurate information. Misperceptions surrounding cannabis particularly do not recognize the increased risk of accidents associated with DUIC, which highlights the need for comprehensive and accurate information [6,10], as many people turn to the internet to find information about cannabis [8,60] and about health in general [61].

Contrary to our finding that many web pages generally presented accurate information regarding DUIC, Lau et al [20] found that around 80% of the internet claims were inaccurate when investigating the information related to cannabis health benefits. This may suggest that the evidence regarding DUIC is less debated compared to suggested health benefits of cannabis; still, DUIC behaviors persist despite the presence of easily accessible accurate information [5,6]. Studies have shown that both adolescents and adults have a low risk perception of cannabis [62] and feel they are in control of their driving after cannabis consumption [63]. This is problematic given the evidence that cannabis can significantly impair motor coordination, judgment, and reaction time [64,65], increasing the risk of motor vehicle accidents [66].

Limitations

This study has a number of limitations. Although we have used what appears to be the most appropriate tools to evaluate web page information, there are no best practices for conducting this type of research. Second, the QUEST tool does not have a target quality score or a threshold of acceptable quality, and therefore, we can only make relative comparisons with the web pages included in this study. Finally, we made the assumption that peer-reviewed content was accurate and excluded those sites from the accuracy assessment. However, there is no guarantee that all peer-reviewed materials are fully accurate. Fortunately, only 3 web pages fell into this category, so this would have minimal impact on the overall analysis.

Conclusions

Most of the identified web pages on Google Canada search engine provided accurate information about DUIC; however, the information was incomplete, the readability was generally low, and the quality of information varied depending on the source. Health organizations should consider health literacy of the public when creating content to help prevent misinterpretation and perpetuate prevailing misperceptions surrounding DUIC. Delivering high-quality, readable, and accurate information in a way that is comprehensible to the public is needed to support informed decision-making.

Acknowledgments

Sandra Schuhmacher, BSc, PharmD Student, has applied the search strategy and extracted the Health on the Net (HON) code for each web page. The authors received financial support for conduct of the research from Canadian Institutes of Health Research (Grant No. RN407334 - 429120) and the Canadian Centre of Substance Use and Addiction for the Partnerships for Cannabis Policy (Grant No. RN407334 - 429120) inclusive of this research.

Authors' Contributions

All authors have contributed to the design of this study. MJ carried out the search strategy; MJ and MB applied the QUEST tool. DG rated the readability for the web pages. MJ and DG assessed the accuracy of the web pages. All authors contributed to the analysis of the findings. MJ wrote the initial manuscript, and DG helped to draft the manuscript. JD, LDB, and MN conceptualized the research, reviewed discrepancies, developed the accuracy tool, and reviewed the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material.
[PDF File (Adobe PDF File), 158 KB - infodemiology_v3i1e43001_app1.pdf]

Multimedia Appendix 2
Web pages collected form each Google page.
[PDF File (Adobe PDF File), 152 KB - infodemiology_v3i1e43001_app2.pdf]

Multimedia Appendix 3
Web pages included in the data analysis.
[PDF File (Adobe PDF File), 137 KB - infodemiology_v3i1e43001_app3.pdf]

Multimedia Appendix 4
The quality assessment of the included web pages.
[PDF File (Adobe PDF File), 264 KB - infodemiology_v3i1e43001_app4.pdf]

Multimedia Appendix 5
The readability scores of the included web pages.
[PDF File (Adobe PDF File), 104 KB - infodemiology_v3i1e43001_app5.pdf]

Multimedia Appendix 6
Examples of quotations used for the accuracy assessment.
[PDF File (Adobe PDF File), 95 KB - infodemiology_v3i1e43001_app6.pdf]

References

1. Taking stock of progress: cannabis legalization and regulation in Canada. Government of Canada. 2022. URL: https://tinyurl.com/bdfnn6kv [accessed 2023-02-03]
2. Rotermann M. Health report - Looking back from 2020, how cannabis use and related behaviours changed in Canada. Statistics Canada. 2021 Apr 21. URL: https://www150.statcan.gc.ca/n1/pub/82-003-x/2021004/article/00001-eng.htm [accessed 2021-07-30]
3. Lee C, Lee A, Goodman S, Hammond D, Fischer B. The Lower-Risk Cannabis Use Guidelines' (LRCUG) recommendations: how are Canadian cannabis users complying? Prev Med Rep 2020 Dec;20:101187 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101187] [Medline: 33083205]
4. Fischer B, Russell C, Sabioni P, van den Brink W, Le Foll B, Hall W, et al. Lower-risk cannabis use guidelines: a comprehensive update of evidence and recommendations. Am J Public Health 2017 Aug;107(8):e1-e12 [FREE Full text] [doi: 10.2105/AJPH.2017.303818] [Medline: 28644037]
5. Fischer B, Imtiaz S, Rudzinski K, Rehm J. Crude estimates of cannabis-attributable mortality and morbidity in Canada-implications for public health focused intervention priorities. J Public Health (Oxf) 2016 Mar 28;38(1):183-188 [FREE Full text] [doi: 10.1093/pubmed/fdv005] [Medline: 25630540]
6. Erin Goodman S, Leos-Toro C, Hammond D. Risk perceptions of cannabis- vs. alcohol-impaired driving among Canadian young people. Drugs (Abingdon Engl) 2019 May 16;27(3):205-212. [doi: 10.1080/09687637.2019.1611738]
7. Turna J, Balodis I, Van Ameringen M, Busse JW, MacKillop J. Attitudes and beliefs toward cannabis before recreational legalization: a cross-sectional study of community adults in Ontario. Cannabis Cannabinoid Res 2022 Aug 01;7(4):526-536. [doi: 10.1089/can.2019.0088] [Medline: 33998851]
8. Kruger DJ, Kruger JS, Collins RL. Cannabis enthusiasts' knowledge of medical treatment effectiveness and increased risks from cannabis use. Am J Health Promot 2020 May 09;34(4):436-439. [doi: 10.1177/0890117119899218] [Medline: 31916839]

9. Choi NG, DiNitto DM, Marti CN. Older marijuana users' marijuana risk perceptions: associations with marijuana use patterns and marijuana and other substance use disorders. Cambridge University Press. 2017. URL: https://tinyurl.com/2dwsysh8 [accessed 2023-02-03]
10. McDonald A, Hamilton H, Wickens C, Watson T, Elton-Marshall T, Wardell J, et al. Driving under the influence of cannabis risk perceptions and behaviour: a population-based study in Ontario, Canada. Prev Med 2021 Dec;153:106793 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106793] [Medline: 34517043]
11. Kosa K, Giombi K, Rains C, Cates S. Consumer use and understanding of labelling information on edible marijuana products sold for recreational use in the states of Colorado and Washington. Int J Drug Policy 2017 May;43:57-66 [FREE Full text] [doi: 10.1016/j.drugpo.2017.01.006] [Medline: 28222305]
12. Hammond D. Communicating THC levels and 'dose' to consumers: Implications for product labelling and packaging of cannabis products in regulated markets. Int J Drug Policy 2021 May;91:102509 [FREE Full text] [doi: 10.1016/j.drugpo.2019.07.004] [Medline: 31351756]
13. Kruger JS, Kruger D, Collins RL. Knowledge and practice of harm reduction strategies among people who report frequent cannabis use. Health Promot Pract 2021 Jan 22;22(1):24-30. [doi: 10.1177/1524839920923999] [Medline: 32443954]
14. Lubin G. Americans are searching for weed more than ever, according to Google. Business Insider. 2016. URL: https://tinyurl.com/ycr2znh6 [accessed 2023-02-03]
15. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004 Oct;7(5):497-510. [doi: 10.1089/cpb.2004.7.497] [Medline: 15667044]
16. Bonn-Miller MO, Loflin MJE, Thomas BF, Marcu JP, Hyke T, Vandrey R. Labeling accuracy of cannabidiol extracts sold online. JAMA 2017 Nov 07;318(17):1708-1709 [FREE Full text] [doi: 10.1001/jama.2017.11909] [Medline: 29114823]
17. Boatwright KD, Sperry ML. Accuracy of medical marijuana claims made by popular websites. J Pharm Pract 2020 Aug 30;33(4):457-464. [doi: 10.1177/0897190018818907] [Medline: 30595085]
18. Montané E, Duran M, Capellà D, Figueras A. Scientific drug information in newspapers: sensationalism and low quality. The example of therapeutic use of cannabinoids. Eur J Clin Pharmacol 2005 Jul 28;61(5-6):475-477. [doi: 10.1007/s00228-005-0916-7] [Medline: 15983825]
19. Halvorson RT, Stewart CC, Thakur A, Glantz SA. Scientific quality of health-related articles in specialty cannabis and general newspapers in San Francisco. J Health Commun 2018 Oct 25;23(12):993-998 [FREE Full text] [doi: 10.1080/10810730.2018.1534906] [Medline: 30358488]
20. Lau N, Gerson M, Korenstein D, Keyhani S. Internet claims on the health benefits of cannabis use. J Gen Intern Med 2021 Nov 19;36(11):3611-3614 [FREE Full text] [doi: 10.1007/s11606-020-06421-w] [Medline: 33742301]
21. Ng JY, Dzisiak DA, Saini JB. Cannabis for pain: a cross-sectional survey of the patient information quality on the Internet. J Cannabis Res 2021 Aug 16;3(1):36 [FREE Full text] [doi: 10.1186/s42238-021-00093-x] [Medline: 34399853]
22. Premkumar A, Almeida BA, Lopez J, Pean CA, Nwachukwu BU, Sculco PK. The quality of online resources available to patients regarding cannabidiol for symptomatic relief of hip or knee arthritis is poor. J Am Acad Orthop Surg Glob Res Rev 2021 Jan 22;5(1):1-7 [FREE Full text] [doi: 10.5435/JAAOSGlobal-D-20-00241] [Medline: 33497073]
23. Lamy F, Daniulaityte R, Zatreh M, Nahhas R, Sheth A, Martins S, et al. "You got to love rosin: Solventless dabs, pure, clean, natural medicine." Exploring Twitter data on emerging trends in Rosin Tech marijuana concentrates. Drug Alcohol Depend 2018 Mar 01;183:248-252 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.10.039] [Medline: 29306816]
24. Merten JW, Gordon BT, King JL, Pappas C. Cannabidiol (CBD): perspectives from Pinterest. Subst Use Misuse 2020 Jul 25;55(13):2213-2220. [doi: 10.1080/10826084.2020.1797808] [Medline: 32715862]
25. Ramo D, Popova L, Grana R, Zhao S, Chavez K. Cannabis mobile apps: a content analysis. JMIR Mhealth Uhealth 2015 Aug 12;3(3):e81 [FREE Full text] [doi: 10.2196/mhealth.4405] [Medline: 26268634]
26. Ouellette L, Cearley M, Judge B, Riley B, Jones J. Cooking with cannabis: The rapid spread of (mis)information on YouTube. Am J Emerg Med 2018 Jul;36(7):1300-1301. [doi: 10.1016/j.ajem.2017.10.060] [Medline: 29096918]
27. Edberg M. Essential Readings in Health Behavior: Theory and Practice. Sudbury, MA: Jones & Bartlett Learning; 2010:321.
28. Swift W, Copeland J, Lenton S. Cannabis and harm reduction. Drug Alcohol Rev Internet 2000;19(1):101-112 [FREE Full text] [doi: 10.1080/09595230096200]
29. Search engine market share Canada. Statcounter. URL: https://gs.statcounter.com/search-engine-market-share/all/canada [accessed 2021-07-18]
30. Schultheiß S, Lewandowski D. Misplaced trust? The relationship between trust, ability to identify commercially influenced results and search engine preference. J Inf Sci 2021 May 14. [doi: 10.1177/01655515211014157]
31. King A. Website Optimization. Sebastopol, CA: O'Reilly Media Inc; Jul 2008.
32. Google. URL: https://www.google.ca/ [accessed 2023-03-30]
33. Schwarzbach HL, Mady LJ, Kaffenberger TM, Duvvuri U, Jabbour N. Quality and readability assessment of websites on human papillomavirus and oropharyngeal cancer. Laryngoscope 2021 Jan 13;131(1):87-94. [doi: 10.1002/lary.28670] [Medline: 32282087]
34. Safdari R, Gholamzadeh M, Saeedi S, Tanhapour M, Rezayi S. An evaluation of the quality of COVID-19 websites in terms of HON principles and using DISCERN tool. Health Info Libr J 2022 Aug 10 [FREE Full text] [doi: 10.1111/hir.12454] [Medline: 35949046]

35. St John A, Carlisle K, Kligman M, Kavic SM. What's Nissen on the net? the quality of information regarding Nissen fundoplication on the internet. Surg Endosc 2022 Jul 29;36(7):5198-5206. [doi: 10.1007/s00464-021-08895-z] [Medline: 34845552]
36. Flint M, Inglis G, Hill A, Mair M, Hatrick S, Tacchi MJ, et al. A comparative study of strategies for identifying credible sources of mental health information online: Can clinical services deliver a youth-specific internet prescription? Early Interv Psychiatry 2022 Jun 02;16(6):643-650. [doi: 10.1111/eip.13209] [Medline: 34474508]
37. Jasem Z, AlMeraj Z, Alhuwail D. Evaluating breast cancer websites targeting Arabic speakers: empirical investigation of popularity, availability, accessibility, readability, and quality. BMC Med Inform Decis Mak 2022 May 09;22(1):126 [FREE Full text] [doi: 10.1186/s12911-022-01868-9] [Medline: 35534816]
38. Smekal M, Gil S, Donald M, Beanlands H, Straus S, Herrington G, et al. Content and quality of websites for patients with chronic kidney disease: an environmental scan. Can J Kidney Health Dis 2019 Jul 30;6:2054358119863091 [FREE Full text] [doi: 10.1177/2054358119863091] [Medline: 31391944]
39. Boyer C, Baujard V, Geissbuhler A. Evolution of health web certification through the HONcode experience. Stud Health Technol Inform 2011;169:53-57 [FREE Full text] [doi: 10.4414/smi.26.00233]
40. HONcode certification. Health On the Net. URL: https://www.hon.ch/en/certification.html [accessed 2023-02-03]
41. Robillard JM, Jun JH, Lai J, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. BMC Med Inform Decis Mak 2018 Oct 19;18(1):87 [FREE Full text] [doi: 10.1186/s12911-018-0668-9] [Medline: 30340488]
42. Lenaerts G, Bekkering G, Goossens M, De Coninck L, Delvaux N, Cordyn S, et al. Tools to assess the trustworthiness of evidence-based point-of-care information for health care professionals: systematic review. J Med Internet Res 2020 Jan 17;22(1):e15415 [FREE Full text] [doi: 10.2196/15415] [Medline: 31951213]
43. Chen T, Gentry S, Qiu D, Deng Y, Notley C, Cheng G, et al. Online information on electronic cigarettes: comparative study of relevant websites from Baidu and google search engines. J Med Internet Res 2020 Jan 24;22(1):e14725 [FREE Full text] [doi: 10.2196/14725] [Medline: 32012069]
44. Bachu V, Mahjoub H, Holler A, Crihalmeanu T, Bachu D, Ayyaswami V, et al. Assessing COVID-19 health information on Google using the Quality Evaluation Scoring Tool (QUEST): cross-sectional and readability analysis. JMIR Form Res 2022 Mar 11;6(2):e32443 [FREE Full text] [doi: 10.2196/32443] [Medline: 34995206]
45. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2023-04-25]
46. Readability Formulas. URL: https://readabilityformulas.com/ [accessed 2023-02-03]
47. Worrall AP, Connolly MJ, O'Neill A, O'Doherty M, Thornton KP, McNally C, et al. Readability of online COVID-19 health information: a comparison between four English speaking countries. BMC Public Health 2020 Nov 13;20(1):1635 [FREE Full text] [doi: 10.1186/s12889-020-09710-5] [Medline: 33183297]
48. Readable. URL: https://app.readable.com/text/ [accessed 2023-02-03]
49. The Gunning Fog Index. Readable. URL: https://readable.com/readability/gunning-fog-index/ [accessed 2023-02-03]
50. Flesch Reading Ease and the Flesch Kincaid Grade Level. Readable. URL: https://readable.com/readability/flesch-reading-ease-flesch-kincaid-grade-level/ [accessed 2022-02-03]
51. Roberts L. Health information and the Internet: The 5 Cs website evaluation tool. Br J Nurs 2010 Mar 12;19(5):322-325. [doi: 10.12968/bjon.2010.19.5.47075] [Medline: 20335904]
52. Alshaikh L, Shimozono Y, Dankert JF, Ubillus H, Kennedy JG. Evaluation of the quality and readability of online sources on the diagnosis and management of osteochondral lesions of the ankle. Cartilage 2021 Dec 10;13(1_suppl):1422S-1428S [FREE Full text] [doi: 10.1177/19476035211021910] [Medline: 34109846]
53. Hungerford D. Internet access produces misinformed patients: managing the confusion. Orthopedics 2009 Oct;32(9):658-660 [FREE Full text] [doi: 10.3928/01477447-20090728-04] [Medline: 19751023]
54. Government information: evaluating sources/how to read a scholarly article. LibGuides: Wells College. URL: https://libguides.wells.edu/c.php?g=239219&p=1590822 [accessed 2023-02-03]
55. Using sources: evaluating sources and avoiding plagiarism. LibGuides: Northern Virginia Community College. URL: https://libguides.nvcc.edu/c.php?g=361391&p=2440252 [accessed 2023-02-03]
56. Ramli R, Jambor MA, Kong CY. Dr Google - assessing the reliability and readability of information on general surgical procedures found via search engines. ANZ J Surg 2023 Mar 30;93(3):590-596. [doi: 10.1111/ans.18289] [Medline: 36716246]
57. Karimi A, Shah A, Hecht C, Burkhart R, Acuña AJ, Kamath A. Readability of online patient education materials for total joint arthroplasty: a systematic review. J Arthroplasty 2023 Jan 27:A [FREE Full text] [doi: 10.1016/j.arth.2023.01.032] [Medline: 36716898]
58. Kruger DJ, Moffet IM, Seluk LC, Zammit LA. A content analysis of internet information sources on medical cannabis. J Cannabis Res 2020 Sep 18;2(1):29 [FREE Full text] [doi: 10.1186/s42238-020-00041-1] [Medline: 33526127]
59. Gordon N, Crouch E. Digital information technology use and patient preferences for internet-based health education modalities: cross-sectional survey study of middle-aged and older adults with chronic health conditions. JMIR Aging 2019 Apr 04;2(1):e12243 [FREE Full text] [doi: 10.2196/12243] [Medline: 31518291]

60. Ishida JH, Zhang AJ, Steigerwald S, Cohen BE, Vali M, Keyhani S. Sources of information and beliefs about the health effects of marijuana. J Gen Intern Med 2020 Jan 21;35(1):153-159 [FREE Full text] [doi: 10.1007/s11606-019-05335-6] [Medline: 31637640]
61. AlGhamdi K, Moussa N. Internet use by the public to search for health-related information. Int J Med Inform 2012 Jun;81(6):363-373 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.12.004] [Medline: 22217800]
62. Azofeifa A, Mattson M, Schauer G, McAfee T, Grant A, Lyerla R. National estimates of marijuana use and related indicators - national survey on drug use and health, United States, 2002-2014. MMWR Surveill Summ 2016 Oct 02;65(11):1-28 [FREE Full text] [doi: 10.15585/mmwr.ss6511a1] [Medline: 27584586]
63. Berg CJ, Daniel CN, Vu M, Li J, Martin K, Le L. Marijuana use and driving under the influence among young adults: a socioecological perspective on risk factors. Subst Use Misuse 2018 Mar 23;53(3):370-380 [FREE Full text] [doi: 10.1080/10826084.2017.1327979] [Medline: 28777692]
64. Lenné MG, Dietze P, Triggs T, Walmsley S, Murphy B, Redman J. The effects of cannabis and alcohol on simulated arterial driving: Influences of driving experience and task demand. Accid Anal Prev 2010 May;42(3):859-866 [FREE Full text] [doi: 10.1016/j.aap.2009.04.021] [Medline: 20380913]
65. Hartman R, Huestis M. Cannabis effects on driving skills. Clin Chem 2013 Mar;59(3):478-492 [FREE Full text] [doi: 10.1373/clinchem.2012.194381] [Medline: 23220273]
66. Li M, Brady JE, DiMaggio CJ, Lusardi AR, Tzong KY, Li G. Marijuana use and motor vehicle crashes. Epidemiol Rev 2012 Oct 04;34(1):65-72 [FREE Full text] [doi: 10.1093/epirev/mxr017] [Medline: 21976636]

Abbreviations

DUIC: driving under the influence of cannabis
FKGL: Flesch-Kincaid Grade Level
FRES: Flesch Reading Ease Scale
GFI: Gunning Fox Index
HON: Health on the Net
QUEST: quality evaluation scoring tool
SMOG: Simple Measure of Gobbledygook

Edited by C Baur; submitted 27.09.22; peer-reviewed by C Zelaya, F Lamy; comments to author 12.12.22; revised version received 06.02.23; accepted 25.03.23; published 02.05.23.

Please cite as:
Josey M, Gaid D, Bishop LD, Blackwood M, Najafizada M, Donnan JR
The Quality, Readability, and Accuracy of the Information on Google About Cannabis and Driving: Quantitative Content Analysis
JMIR Infodemiology 2023;3:e43001
URL: https://infodemiology.jmir.org/2023/1/e43001
doi:10.2196/43001
PMID:

Original Paper

Lessons Learned From Interdisciplinary Efforts to Combat COVID-19 Misinformation: Development of Agile Integrative Methods From Behavioral Science, Data Science, and Implementation Science

Sahiti Myneni[1*], PhD; Paula Cuccaro[2,3*], PhD; Sarah Montgomery[2,3], BJ; Vivek Pakanati[4], BS, MPH; Jinni Tang[2,3], MPH; Tavleen Singh[1], MS; Olivia Dominguez[2,3], MPH; Trevor Cohen[5], MBChB, PhD; Belinda Reininger[6], DrPH; Lara S Savas[2,3], PhD; Maria E Fernandez[2,3], PhD

[1]School of Biomedical Informatics, The University of Texas Health Science Center, Houston, TX, United States

[2]Department of Health Promotion & Behavioral Sciences, School of Public Health, The University of Texas Health Science Center, Houston, TX, United States

[3]Center for Health Promotion and Prevention Research, School of Public Health, The University of Texas Health Science Center, Houston, TX, United States

[4]The University of Texas Health Science Center, Tyler, TX, United States

[5]Department of Biomedical Informatics and Medical Education, The University of Washington, Seattle, WA, United States

[6]School of Public Health, Brownsville Regional Campus, The University of Texas Health Science Center, Brownsville, TX, United States

[*]these authors contributed equally

Corresponding Author:
Tavleen Singh, MS
School of Biomedical Informatics
The University of Texas Health Science Center
7000 Fannin St.
Houston, TX, 77030
United States
Phone: 1 713 500 3900
Email: tavleen.kaur.ranjit.singh@uth.tmc.edu

Abstract

Background: Despite increasing awareness about and advances in addressing social media misinformation, the free flow of false COVID-19 information has continued, affecting individuals' preventive behaviors, including masking, testing, and vaccine uptake.

Objective: In this paper, we describe our multidisciplinary efforts with a specific focus on methods to (1) gather community needs, (2) develop interventions, and (3) conduct large-scale agile and rapid community assessments to examine and combat COVID-19 misinformation.

Methods: We used the Intervention Mapping framework to perform community needs assessment and develop theory-informed interventions. To supplement these rapid and responsive efforts through large-scale online social listening, we developed a novel methodological framework, comprising qualitative inquiry, computational methods, and quantitative network models to analyze publicly available social media data sets to model content-specific misinformation dynamics and guide content tailoring efforts. As part of community needs assessment, we conducted 11 semistructured interviews, 4 listening sessions, and 3 focus groups with community scientists. Further, we used our data repository with 416,927 COVID-19 social media posts to gather information diffusion patterns through digital channels.

Results: Our results from community needs assessment revealed the complex intertwining of personal, cultural, and social influences of misinformation on individual behaviors and engagement. Our social media interventions resulted in limited community engagement and indicated the need for consumer advocacy and influencer recruitment. The linking of theoretical constructs underlying health behaviors to COVID-19–related social media interactions through semantic and syntactic features using our computational models has revealed frequent interaction typologies in factual and misleading COVID-19 posts and indicated significant differences in network metrics such as degree. The performance of our deep learning classifiers was reasonable, with an F-measure of 0.80 for speech acts and 0.81 for behavior constructs.

Conclusions: Our study highlights the strengths of community-based field studies and emphasizes the utility of large-scale social media data sets in enabling rapid intervention tailoring to adapt grassroots community interventions to thwart misinformation seeding and spread among minority communities. Implications for consumer advocacy, data governance, and industry incentives are discussed for the sustainable role of social media solutions in public health.

(JMIR Infodemiology 2023;3:e40156) doi:10.2196/40156

KEYWORDS

COVID-19; misinformation; social media; health belief model; deep learning; community engagement

Introduction

Exposure to COVID-19 health misinformation has emerged as a global risk factor for human health and wellness [1]. Expanding mobile connectivity and access to digital media allows for the dissemination of both evidence-based and unvetted resources online in this increasingly connected information environment. COVID-19 is the first global pandemic during this social media era, revealing several key shifts in health information consumption by the general public that challenge traditional knowledge and remediation pathways to combat health misinformation [2-5]. Studies show that (1) health consumers are no longer passive readers, but active contributors of misinformation seeding and spread; (2) such contributions can be unintentional and stem from anywhere in the world, affecting the public's perceptions, behaviors, and potential COVID-19–related risks; (3) contamination with other information verticals, including politics, global monetization of media corporations, and inconsistent public health responses around the globe, can multiply mistrust in scientific institutions; and (4) increasing reliance on artificial intelligence and automated content recommendation algorithms confine people to opinion bubbles and echo chambers, with little human moderation, ultimately resulting in polarized social circles that make misinformation easy to proliferate [6-9].

Another striking observation is that COVID-19 misinformation traveled much faster than truth [10-14]. According to a recent report, 20% of COVID-19 misinformation comes from high-profile accounts (celebrities, politicians, and talk radio personalities) and 80% comes from the general public, with the former capturing much higher engagement rates (69% compared to 31%) [15]. Containing misinformation spread is further complicated by rapidly changing public health recommendations that follow emerging COVID-19 research. For instance, mask wearing recommendations have changed throughout the pandemic, resulting in confusion and misinformation associated with mask-wearing behavior [16] as well as mistrust among the population about the validity of recommendations [17].

Research has shown that exposure to COVID-19 misinformation is associated with age, education, and income levels of an individual [18]. Previous studies have also identified main themes related to the spread of COVID-19 misinformation in social media and how it fluctuated with time, for example, there were false stories about the source of the virus in the beginning of the pandemic, followed by false information spread about home remedies, etc [19]. Misinformation about COVID-19 can lead to increased risk of exposure and susceptibility to the virus (eg, promoting vaccine hesitancy), thus affecting the global course of the pandemic. To this end, emerging research suggests that misinformation modeling and management should be considered a critical component of public health campaigns and interventions [20] because of the various dynamics involving information exposure, human behavior, and disease spread. Current tools that automate misinformation detection are prone to algorithmic bias and offer little or no context for individuals to engage in self-reflection and recalibration of their health beliefs, attitudes, and latent heuristics, bringing into question the credibility, equity, and cultural appropriateness of such tools [21-23].

In this paper, our aim is to describe our interdisciplinary efforts to combat COVID-19 misinformation, which include needs assessment, misinformation modeling, and intervention development and dissemination. For the purpose of this work, we used mixed methods community needs assessment and leveraged recent advances in social computing and data science. These methods enable us to conduct large-scale online social listening and gain granular understanding of community needs. Further, these methods allow dissemination of evidence-based information in online settings and at-risk communities in the field to promote COVID-19 testing and vaccination for general and minority populations. In subsequent sections of this paper, we describe how the methods and results of our community needs assessment, integration of behavioral theory, social computing techniques, and social network analysis contributed to COVID-19–related knowledge discovery and interventions. This article aims to help public health researchers, social marketing teams, implementation scientists, disease prevention and health disparity experts, informaticians, and social media technologists expand their understanding of qualitative methods and data science tools, and highlight the missed opportunities in appropriately leveraging these resources for public health and wellness during the COVID-19 pandemic.

Methods

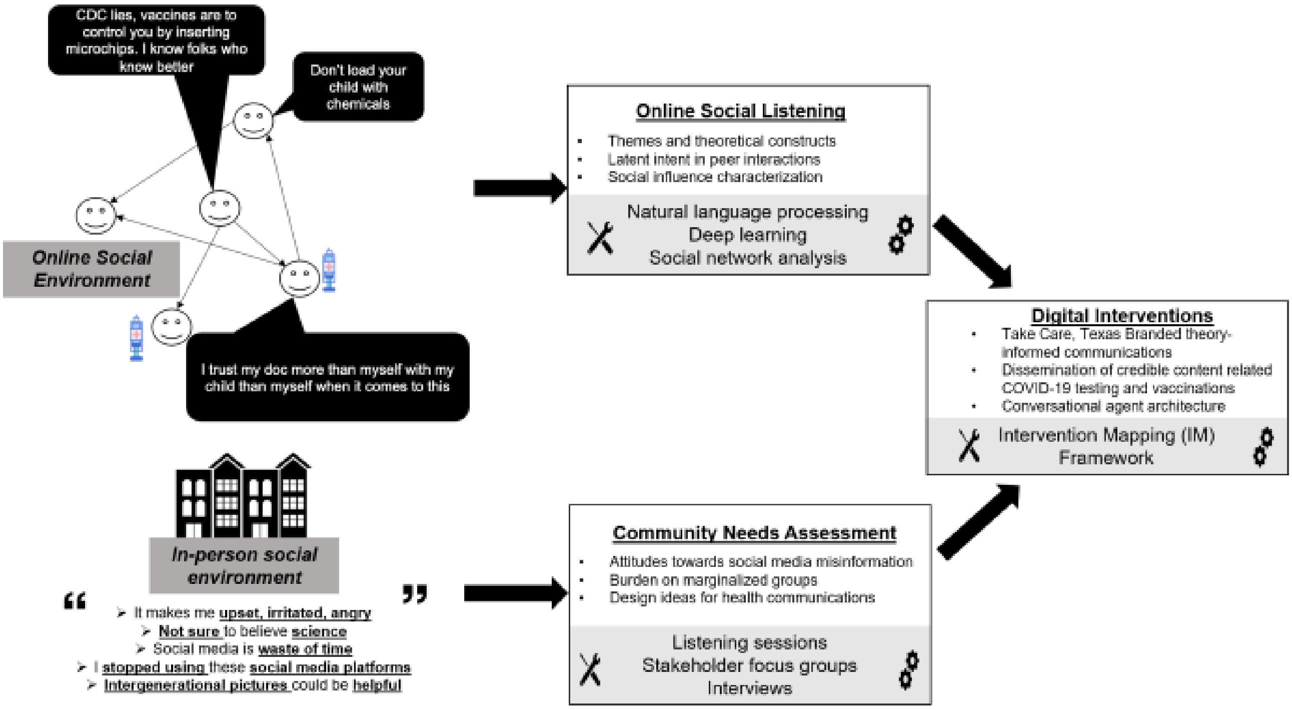
Intervention Mapping

We used the Intervention Mapping (IM) framework [24] that offers a systematic approach to intervention development and adaptation. IM is designed to develop multilevel interventions, such as the one described, in that it considers not only the behavior (COVID-19 testing and vaccination) but also the interpersonal environment (social marketing to promote

COVID-19–related protective behaviors). The IM process comprises the following 6 steps: (1) conducting a needs and assets assessment to create a logical model of the problem for stating intervention goals; (2) flipping the logical model of the problem into the logic model of change by identifying the behavioral and environmental outcomes for the intervention; (3) designing the intervention with theory- and evidence-based change methods; (4) developing the intervention products with a process of pretesting, refining, and producing intervention materials; (5) implementing the intervention plan by identifying potential program users and program performance objectives; and (6) developing indicators and measures for intervention evaluation. In this paper, we focus on IM steps 1, 2, and 4 given the methodological focus, and outlining our activities for rigorous implementation, process evaluation, and effects evaluation is outside the scope of this effort.

Figure 1 illustrates the multilevel nature of our methodology to identify and combat COVID-19 misinformation as described in the sections below.

Figure 1. Overall research methodology. CDC: Centers for Disease Control and Prevention.



Community Needs Assessment

We collected information from multiple sources as part of our community needs assessment. We conducted virtual interviews with participants (n=11) to understand attitudes, beliefs, and knowledge about COVID-19 testing and vaccines in the context of usability evaluation of digital learning environments (eg, chatbots and existing social media) to address misinformation. We recruited participants using direct contact and social media advertising, and they were provided with a description of the study and an informed consent form. Once the consent form was signed, a team of 2 researchers conducted interviews. On average, each interview lasted 32 minutes and was digitally recorded. Once transcribed, all interviews were analyzed with the methods of directed content analysis [25] using Dedoose software (SocioCultural Research Consultants). Each participant who participated in the interview received US $25 compensation. From April to July 2021, local nonprofit agencies trained by the civic engagement group hosted and facilitated 7 listening sessions with nearly 70 community members in areas identified as heavily impacted by COVID-19. The sessions prioritized the experiences of those living in predominantly Black and Latinx neighborhoods, those in refugee and immigrant communities, those in low-income households, young adults, and those whose primary language is Spanish. Participants were asked about motivators, hesitations, structural barriers, rumors, and misinformation pertaining to COVID-19 vaccines. Each participant who attended a listening session received US $50 compensation for participating. The civic organization shared their analysis to inform our intervention development. We subsequently gathered input from stakeholders in 3 meetings with community scientists (around 12-15 participants in each session) from May 2021 to February 2022. The community scientist program is part of the National Institutes of Health–funded Center for Clinical and Translational Sciences to provide feedback from community members trained to understand scientific reasoning about aspects of the research process. Our sessions focused on the applicability and cultural appropriateness of our existing COVID-19 education and intervention materials. Our focus on minority populations was limited to qualitative inquiry, that is, listening sessions and interviews. Our social computing methods described in the next section capture the views of the general population. It is important to note that minority participants from our community needs assessment mentioned the use of specific social media platforms (eg, Facebook, Twitter, and YouTube) where they routinely were exposed to COVID-19 misinformation. Based on this insight, we conducted deeper secondary analysis of online social discourse to examine and portray the sociobehavioral mechanisms underlying misinformation spread and social resistance pathways.

Ethical Considerations

Our virtual interviews were deemed exempt by the institutional review board at the University of Texas Health Science Center at Houston (HSC-SBMI-18-1003). Neither community listening sessions nor Community Scientist sessions underwent review from an institutional review board. The information we present here is with the consent of the organizations involved.

Online Social Listening

We used 2 distinct online social discourse data sets for this analysis. Using a public COVID-19 tweet-ID repository [26], we retrieved tweets published from January 2020 to January 2021. Tweets were hydrated using Twitter's application programming interface (API) and the Twarc package [27], resulting in a total of 416,927 English-language tweets. We only used the original tweets (ie, excluding retweets and quotes) in our analysis. From these, a subset of 1400 tweets was randomly selected for further qualitative analysis as described below. To calculate the interrater reliability, a subset of 100 tweets was initially coded by 2 researchers, and any disagreements were mutually resolved via discussions between the 2 researchers to determine the appropriate label before proceeding with additional coding.

In addition, we employed the COVID-19 Twitter misinformation data set called CMU-MisCov19 [27] that was created to characterize COVID-19–related information in online social media to ensure robustness in our modeling efforts. This data set consisted of a total of 4573 Twitter IDs annotated for 17 categories, including tweets calling out or correcting misinformation, false public health response, false fact or prevention, true public health response, true prevention, etc [27]. We hydrated the tweets using Twitter's API and the Twarc package [26], resulting in a final data set of 3702 tweets. Of these, a total of 1204 tweets exhibited misinformation resistance, in which the users were specifically calling out or correcting COVID-19 misinformation (ie, the stance taking corrective tweets).

Ethical Considerations

Our social media analysis was reviewed and deemed exempt by our institutional review board at the University of Texas Health Science Center at Houston (HSC-SBMI-15-0697).

Content and Intent Characterization

We coded the tweets using a list of constructs included in health behavior theories, including the Health Belief Model (HBM), Social Cognitive Theory, and Theory of Planned Behavior [28-31]. Examples of those constructs include perceived severity, cues to action, social norms, and self-efficacy [32]. For illustration purposes, in this paper, we present our analysis of HBM-related constructs using our high-throughput social computing methods (ie, methods that can be scaled to large volume data sets obtained from social media platforms). To understand the health beliefs associated with COVID-19 spread and misinformation in online social media platforms, we used a subset of 1400 tweets (7%) selected at random from a filtered set of 20,000 high-impact tweets (depending on their dissemination levels such as likes, retweets, etc) obtained from a total of 416,927 tweets. We analyzed every tweet within it for the manifestation of constructs outlined in the HBM. To understand how online users express their latent intent toward COVID-19 misinformation, we used a modified version of Searle's speech act theory [33] and manually coded tweets for various categories of speech acts (eg, declaratives, stance, and assertion). Identification of speech acts in social media content provides a deeper insight into the interactions among individuals derived from attitudes toward topics and actions conveyed through language [33]. The detailed definitions of speech acts and their examples can be found in a previous report [34]. Tweets not falling into any of the categories were labeled as not applicable (NA). The interrater reliability was 0.81 (Cohen kappa) for HBM labels and 0.84 (Krippendorff alpha) for speech act labels.

Social Influence Characterization

The social influence of the tweets was captured via different dissemination levels based on their audience size and popularity. A tweet's audience size was derived from the follower count, and its popularity was reflected by the number of retweets and likes/favorites, which propagate the tweet to other users [35]. The sum of these quantities indicated the total number of user interactions with each tweet. For the CMU-MisCov19 data set, dissemination levels were assigned based on tweet-level metrics capturing users' interactions with the tweets (in this case, retweets and favorites), and tweets were classified as follows: "high" dissemination level (>11 interactions), "low" dissemination level (1-11 interactions), and "no" dissemination level (0 interactions). There were 527 tweets with high dissemination, 1593 with low dissemination, and 1582 with no dissemination in the data set.

Deep Learning Classification

To capture the population-level insights as our society navigated the course of the pandemic through the use of digital media, we used deep learning methods to scale the extraction of health beliefs and speech acts embedded within Twitter user interactions. Such methods have already been applied by researchers to capture the health beliefs associated with health-related conditions [36]. In this study, we evaluated the performance of the following models for classification of the HBM constructs and speech acts embedded within the data sets: (1) BERTweet [37], (2) BERTweet-Covid19 [37], and (3) ensemble of the 2 models (BERTweet+BERTweet-Covid19). These models are the result of unsupervised pretraining on tweets, providing a model with general linguistic information that can then be used by a classification module appended to it. Using the manually coded data set (n=1400 tweets), we first performed text preprocessing in order to convert the text to lowercase and also remove any hyperlinks from the textual data. We then split the entire data set into 90%, 5%, and 5% sets for training, validation, and testing, respectively. We used a learning rate of 1×10^{-5}. We also computed class weights for the loss function to assign a higher weight to the loss encountered by tweets associated with minor classes (ie, the labels that had a lower prevalence as compared to the labels that had a higher prevalence in the given data set). The model was trained for 20 epochs. We converted the probabilities into actual classes based on the threshold value calculated using the validation set. We used recall, precision, and F1-score to evaluate the classifier's predictions on the held-out test data set. Based on the prevalence of various categories of HBM constructs in the manually coded data set, we initially trained the model to distinguish between HBM applicable tweets (all HBM constructs combined) and nonapplicable tweets. We further trained the model to classify the top 2 prevalent categories of HBM constructs within the HBM applicable tweets. The speech acts model was trained to classify the top 5 prevalent categories of speech acts. We applied these models to classify HBM constructs and speech act categories from the CMU-MisCov19 data set (n=3702).

Social Network Analysis

The CMU-MisCov19 data set was further analyzed using 2-mode network analysis by creating affiliation networks composed of 2 modes (the first mode represented the tweets and the second mode represented the various speech act categories with which the tweets were affiliated). We constructed visual representations of HBM construct-based affiliation networks between tweets and speech acts. We compared the structures and topologies across different networks using various social network metrics such as degree, density, diameter, and average path length. For affiliation networks, degree centrality suggests that an actor (in our case, tweet) is popular because of its membership to certain events (in our case, speech acts), while an event (speech act) is popular based on the size of actors that are part of it [38]. The density of an affiliation network is defined as the number of edges divided by the number of pairs of nodes where only edges between vertex sets are possible [39,40]. The diameter of an affiliation network is the length of the longest path between any pair of actors/events [39,40]. Average path length is defined as the average shortest path between the 2 nodes [39,40]. An open-source network visualization tool, UCINET [41], was used for creating and analyzing these networks.

Intervention Planning

As part of the IM framework described earlier, our intervention plan included multicomponent strategies, such as use of phone navigation, community health workers, and social marketing, including use of social media. We ensured that the social media intervention is informed by existing empirical evidence, behavioral theories, and new evidence from the needs assessment and social listening.

Social Media Intervention

To leverage social media to improve the reach of our COVID-19 health promotion materials, we hired a Houston-based creative agency to brand Take Care, Texas (TCT) social media channels tailored to our 3 project regions. The design agency met with project staff experienced in community health promotion to gain insights into population demographics and regional differences, learn about prevalent COVID-19 attitudes, and identify potential structural and psychosocial barriers to accessing COVID-19 testing. Intervention development efforts to counter misinformation necessarily accounted for disparities in COVID-19 morbidity and mortality rates associated with social determinants unique to our 3 project regions. For instance, efforts were adapted in Northeast Texas to account for low population density and in Cameron County to address a largely Spanish-speaking population. Throughout intervention development, community input, including local listening sessions and community scientists' feedback, helped identify social determinants influencing testing and vaccination behaviors as the situation changed over time. Community scientists also provided feedback on the content and quality of our intervention materials.

In June 2021, our social media team launched 3 social media accounts (ie, Facebook, Instagram, and Twitter) for each region, featuring original content with region-specific design elements and branding (eg, photos that reflect the target population), informational posts without TCT branding, and reposted materials from credible resources. Reposted materials were shared from other health agencies or COVID-19 education efforts including but not limited to the Centers for Disease Control and Prevention, Public Health Media Collaborative, and Vaccine Resource Hub. Posts across the 3 regions overlapped and varied based on local events, demographics, and specific barriers or knowledge gaps identified by community health workers conducting door-to-door outreach. Additionally, community health workers conducted in-person COVID-19 education through door-to-door canvassing, community events with local partners, and COVID-19 testing and vaccination events.

Our regional social media accounts were featured on resource and education materials to direct community members to engage with us online, including door hangers, resource flyers, and maps to COVID-19 testing and vaccination sites. QR codes featured on print materials were also linked to regional social media accounts. In regular meetings with the social marketing team, community health workers shared the most frequently reported barriers influencing residents' COVID-19 preventive behaviors, which were captured in encounter forms during outreach, to better tailor regional print and online materials. For example, multiple door hanger designs with alternative messaging to promote COVID-19 vaccines were developed and distributed, and included a project phone number through which residents could ask community health workers questions about COVID-19–related information or resources. Once community health workers identified the most effective door hanger for a neighborhood, based on the number of calls received, that messaging was prioritized in the next round of door hanger distribution in that community.

Custom content for all regional accounts included advertisements of local events, residents' testimonials, and posts that featured each region's team of community health workers as well as local organizations we partnered with to align and engage with a broader audience. Tailored content on these platforms varied between regions by using images that best reflected each region's priority population, hashtags that applied to the region, and languages spoken in the region. For example, Spanish was the primary language used on all Cameron County social media accounts and most testimonials from that region included Spanish speakers at testing and vaccination locations in the community.

Results

Community Needs Assessment

Our qualitative data collection efforts included 11 interviews with participants, a majority of whom identified as female (8/11, 72%). Themes that emerged from these interviews covered a range of topics, including barriers and facilitators to COVID-19 preventive behaviors (eg, testing, vaccines, and masking), trust in clinical and scientific institutions, the role of technological platforms, the burden of misinformation on general and marginalized populations, and the optimal strategies for risk communication and promoting COVID-19–related health information. All participants mentioned exposure to misleading health information on social media even when they were not actively using it. Distrust in social media as a source of health information was exhibited through descriptions of “algorithmic bias,” “persuasive intent,” and “financial motives,” which in turn created “polarization” and “echo chambers.” These social phenomena were also linked to cognitive heuristics applied by individuals to (1) ensure credibility of information sources through the “reputation” heuristic and (2) limit misinformation exposure through the “self-confirmation” heuristic, and they are aligned with existing literature [42].

Findings from the interviews revealed emotional consequences of misinformation exposure. For example, 82% (9/12) of participants mentioned being “baffled,” “upset,” and “angry” by misinformation, while the rest mentioned “being indifferent” and “losing hope.” Participants linked high levels of perceived confidence in misinformation detection and low levels of perceived vulnerability to academic training, their ability to apply cognitive heuristics, such as reputation (to assess source credibility), and their life experiences associated with age (Table 1). Participants often linked vulnerability to the emotional distress caused by challenging circumstances such as the diagnosis of a chronic or terminal condition such as cancer. Participants emphasized the need to be self-reliant, rather than dependent on platform-based misinformation flagging or third-party fact checking tools, given issues with the time-sensitive, evolving, and emotional nature of information-seeking patterns.

Data from the listening sessions elicited similar themes, with fear of the vaccine emerging as a barrier to vaccination. Most fears stemmed from misinformation circulated on social media concerning side effects, the content of the vaccine, and the rushed approval of the vaccine. Another common barrier was mistrust in the government and public officials recommending the vaccine. As seen in the examples in Table 1, participants exhibited awareness, helplessness, and confusion stemming from the pervasive misinformation they encountered in online social media, which they attributed to the financial motives and algorithmic shortcomings of these platforms. Participants mentioned that there is a possibility of echo chambers and polarization in the name of targeted advertising and personalization for extended digital engagement, which they experienced firsthand when using social media.

Community scientists also highlighted the unjust burden misinformation puts on vulnerable populations by widening the disparities that already affect health outcomes and quality of life. Community scientist sessions and stakeholder interviews resulted in common themes related to fear, side effects, worsened outcomes, and tracking devices, enabling us to identify information sharing mechanisms to combat misinformation. One suggested tactic was to provide a frequently asked questions (FAQ) post with a QR code attached with expertise and insight from physicians and other health care professionals.

Table 1. Comments and feedback from interviews and listening sessions.

Theme	Comments and feedback
Attitudes toward social media misinformation	• I'd say it affects—misinformation definitely makes me upset. I guess I get kind of irritated and angry. [Younger White female] • I stopped using social media because it is harmful for mental health and wellbeing, a waste of time. [Older South Asian male] • ...not to really just become something to reach a consensus because that's not happening in today's social media. Nobody gets into social media or into a conversation with the intent of coming to a logical conclusion. [Older South Asian male]
Social media as an echo chamber	• But, you know, if people aren't necessarily aware of or accepting of the fact that they are in an echo chamber, then it's much more difficult. [Younger Hispanic male] • My Facebook generally does tend to be in an echo chamber. I think it's something that, that I'm aware of. [Older South Asian female]
Individual vulnerability and confidence to detect misinformation	• Or people who are not really IT guys. They look at something and they get influenced by this. They think the technology is helping validate what they already know, no it is reading you, been there done that. [Older Asian male] • I'm going to say moderately confident. Sometimes, I have to tell myself, if it's too good, if it looks too good to be true, it's just too good to be true. Check more, google more [Younger Hispanic female]
Burden of misinformation on marginalized groups	• The burden of misinformation is definitely higher in low-income and minority women who may not have access to care or information resources that are credible. [Older Black female] • No one can get accurate information that is sure shot, how can people with no insurance trust anything they see. It is difficult, no information is better than wrong information. I don't know, never so unsure my life. [Older Asian male]
Misinformation example	• There are many myths that you don't know if they are true or lies. … supposedly, there's a chip that people want to put in to see where you are located. And …, they were supposedly putting a chip on the skin with the vaccine. [Younger Spanish-speaking adult] • Vaccines can cause blood clots, one of my close buddies passed away, I am already at higher risk, can't risk more. [Older Hispanic female]
Lack of trust in science and clinical institutions	• Mistrust in the U.S. health system, including the Tuskegee experiments and concerns around population control. [Younger Black male] • When you keep changing words, when so much drama in news, difficult to trust. [Older South Asian male]
Role of community leaders	• I learn from trusted messengers, especially my faith leader. [Older Black female] • Never know whom to trust, I see no leadership, just vacuum top to bottom. [Younger White male]
Design ideas for health communication	• This is like highway driving, having warning signs might help. [Older South Asian male] • One of my friends just doesn't want to learn, not sure how we can make her see. Someone she trusts and can't block like she does in Facebook. Information in the face she can't ignore. [Younger Hispanic male]

Online Social Listening

The most common health belief was perceived severity (369/1400, 26.4%), showing that most of the messages reflected individuals' concerns about the growing severity of the virus (eg, "Coronavirus infections predicted to grow exponentially; first death outside China; outbreak becomes political"). Perceived barriers (239/1400, 17.1%) were also prevalent, reflecting how individuals experienced barriers to performing COVID-19–related prevention behaviors, such as mask-wearing, quarantining, and testing (eg, "I think CNN needs to focus on figuring out how to correct the day-to-day operations and start telling the truth about irrelevance of COVID-19 testing!"). Other health beliefs were expressed in the form of perceived susceptibility (83/1400, 5.9%; eg, "With that healthy asymptomatic person and contract a more severe form of COVID19, then who is responsible for me getting sick?") and perceived benefits (37/1400, 2.6%; eg, "Can't break out the champagne yet, but efforts to avoid a surge have been working, in some jurisdictions"). In terms of speech acts, the most prevalent speech act was assertion (445/1400, 31.8%) where individuals expressed their beliefs about the spread of the virus (eg, "If the UK doesn't go on lockdown virtually now we are going to be in the same position as Italy in another week or two"), followed by declaratives (373/1400, 26.6%) about objective information related to COVID-19 (eg, "The CDC is now performing entry health screening on all passengers with direct and connecting flights from Wuhan, China"). Directives (300/1400, 21.4%) in the form of advice about what precautions one should take were also common (eg, "We need to take things seriously, … and follow the advice of the medical professionals"). Tweets also posed questions (204/1400, 14.6%) regarding concerns about COVID-19 using Twitter (eg, "Can you contract COVID-19 from a mosquito?").

Table 2 shows the F1 scores of deep learning models on the various classification tasks. For HBM construct classification, the ensemble model achieved the highest F1 score for every category and for the overall macro average (0.81) of the model in comparison to the BERTweet model or the BERTweet-Covid19 model. The BERTweet model achieved the highest F1 score for every category of speech acts and for the overall macro average (0.80) of the model in comparison to the BERTweet-Covid19 model or the ensemble model.

For illustration purposes, we compared the 2-mode affiliation networks of tweets and speech act expressions for stance taking misinformation correction tweets, false information tweets (ie, tweets annotated for the labels fake cure, fake treatment, false public health response, and false fact or prevention), and true information tweets (ie, tweets annotated for the labels true public health response and true prevention) within the two HBM constructs (ie, perceived barriers and perceived severity) in Figure 2. In these networks, the tweets' nodes were colored based on their dissemination levels, with blue nodes representing the different speech act categories (Figure 2).

Table 3 shows the network metrics calculated for the various 2-mode networks. Within the perceived barriers health belief, the stance taking misinformation corrective tweets network had the highest average path length, indicating that the efficiency of information transfer expressed using a certain category of speech act was low as compared to the other 2 networks. Given the density and path length of the false information network, the circulation of false information pertaining to health-related barriers about COVID-19 expressed via speech act categories was much faster, whereas the corrective information about health-related barriers regarding COVID-19 took longer to travel within the network as per our data set. In the stance taking misinformation corrective tweets network, the high dissemination tweets had a higher prevalence of assertion and declarative speech acts within the perceived barriers health belief, whereas the high dissemination tweets had a higher prevalence of declarative speech acts within the perceived severity health belief. Thus, misinformation correction strategies should focus on integrating declaratives (eg, provide objective information) within their messages to have a greater impact on the online community of users who are exposed to misinformation.

Within the perceived barriers HBM construct networks, the most commonly used speech acts within the stance taking misinformation correction tweets network were assertion (degree=0.635) and declaratives (degree=0.246). The most commonly used speech acts within the false information tweets network were assertion (degree=0.575) and declaratives (degree=0.275). The most commonly used speech acts within the noncorrective tweets network were assertion (degree=0.423) and declaratives (degree=0.423). Within the perceived severity HBM construct networks, the most commonly used speech acts within the stance taking misinformation correction tweets network were assertion (degree=0.446) and declaratives (degree=0.339). The most commonly used speech acts within the false information tweets network were assertion (degree=0.339) and declaratives (degree=0.279). The most commonly used speech acts within the noncorrective tweets network were declaratives (degree=0.510) and assertion (degree=0.203). Even though the degrees for speech act assertion and declaratives were higher in all 3 networks, the higher values within the stance taking corrective network indicated that tweets were more prominently expressing those speech acts compared to the other 2 networks. ANOVA revealed that there was a statistically significant difference in the degree centrality of tweet nodes between the 3 dissemination levels (P=.006) for perceived severity HBM construct-based networks.

Table 2. Evaluation (F1 scores) of deep learning models for the classification of Health Belief Model constructs and speech acts.

Variable	BERTweet	BERTweet-Covid19	Ensemble (BERTweet+BERTweet-Covid19)
Health Belief Model constructs			
Per class performance			
Perceived barriers	0.64	0.64	0.77
Perceived severity	0.76	0.80	0.83
Overall model performance			
Macro average	0.71	0.74	0.81
Speech acts			
Per class performance			
Assertion	0.80	0.78	0.82
Declaratives	0.92	0.83	0.89
Directive	0.80	0.73	0.67
Question	0.83	0.88	0.83
Statement	0.67	0.59	0.69
Overall model performance			
Macro average	0.80	0.76	0.78

Figure 2. Two-mode affiliation networks for (A) misinformation correction tweets, (B) false information tweets, and (C) true information tweets within the 2 Health Belief Model constructs.

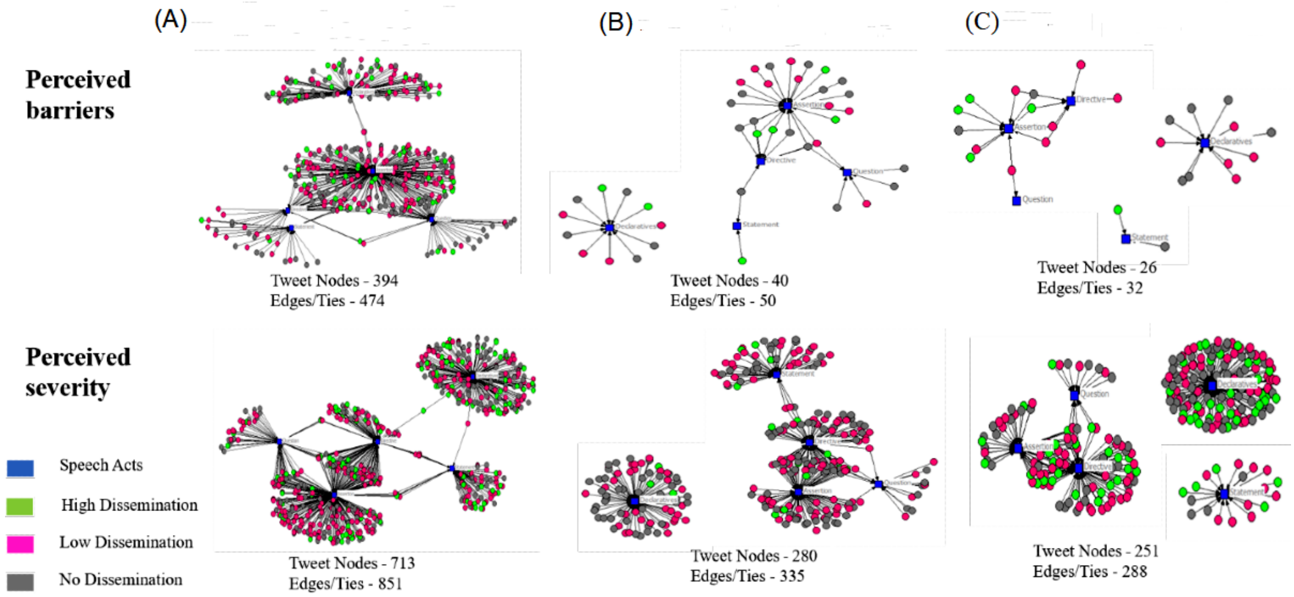


Table 3. Metrics for various Health Belief Model construct-based 2-mode affiliation networks.

Construct and tweet network type	Average path length	Diameter	Density
Perceived barriers			
Stance taking misinformation corrective	3.158	6	0.241
False information	2.635	6	0.250
True information	2.095	5	0.246
Perceived severity			
Stance taking misinformation corrective	3.769	6	0.239
False information	3.185	6	0.239
True information	2.216	4	0.229

Outputs From IM Framework Application

We identified several behavioral science constructs and determinants associated with individuals' intentions to test for COVID-19, intentions to receive COVID-19 vaccinations, attitudes toward testing for COVID-19, and attitudes toward receiving COVID-19 vaccinations, including constructs from the HBM and the Theory of Planned Behavior [32]. Table 4 provides a list of the key constructs and determinants that we used to develop TCT social media content.

Table 4. Behavioral theory constructs and example social media content to promote COVID-19 testing and vaccination behaviors.

Model, COVID-19–related behavior, and construct	Definition	Example messaging	Implementation (social media post examples)
Health Belief Model [32]			
COVID-19 testing and vaccination			
Perceived susceptibility	Belief in the likelihood of getting an illness or disease.	"The Omicron variant is spreading in our community. Anyone can get it."	
COVID-19 vaccination			
Perceived severity	Belief in the severity of an illness or disease.	"Unvaccinated COVID-19 patients have a greater risk of hospitalization and death."	
Perceived barriers	Belief in the obstacles that impede performing the behavior of interest.	"Getting the vaccine is easy. Many drug stores have walk-in appointments."	
Perceived benefits	Belief in the effectiveness of the behavior of interest.	"COVID-19 vaccines are free, safe, and effective."	
COVID-19 testing			
Self-efficacy	Confidence in the ability to perform the behavior of interest.	"It's easy to place your order for free at-home COVID-19 tests at covid.gov. I did it in 5 minutes."	
Theory of Planned Behavior [32]			
COVID-19 testing			
Attitudes	Favorable or unfavorable evaluation of the behavior of interest.	"Regular testing can help give your family peace of mind."	
COVID-19 vaccination			
Subjective norm	Belief about whether important others (parents, partners, or doctors) approve or disapprove of the behavior of interest.	"I vaccinate because I want to keep my children safe. It's the same reason that they wear helmets and seatbelts. #WhyIVaccinate" "I got vaccinated because everyone in my family thought it was an important thing to do. #WhyIVaccinate"	

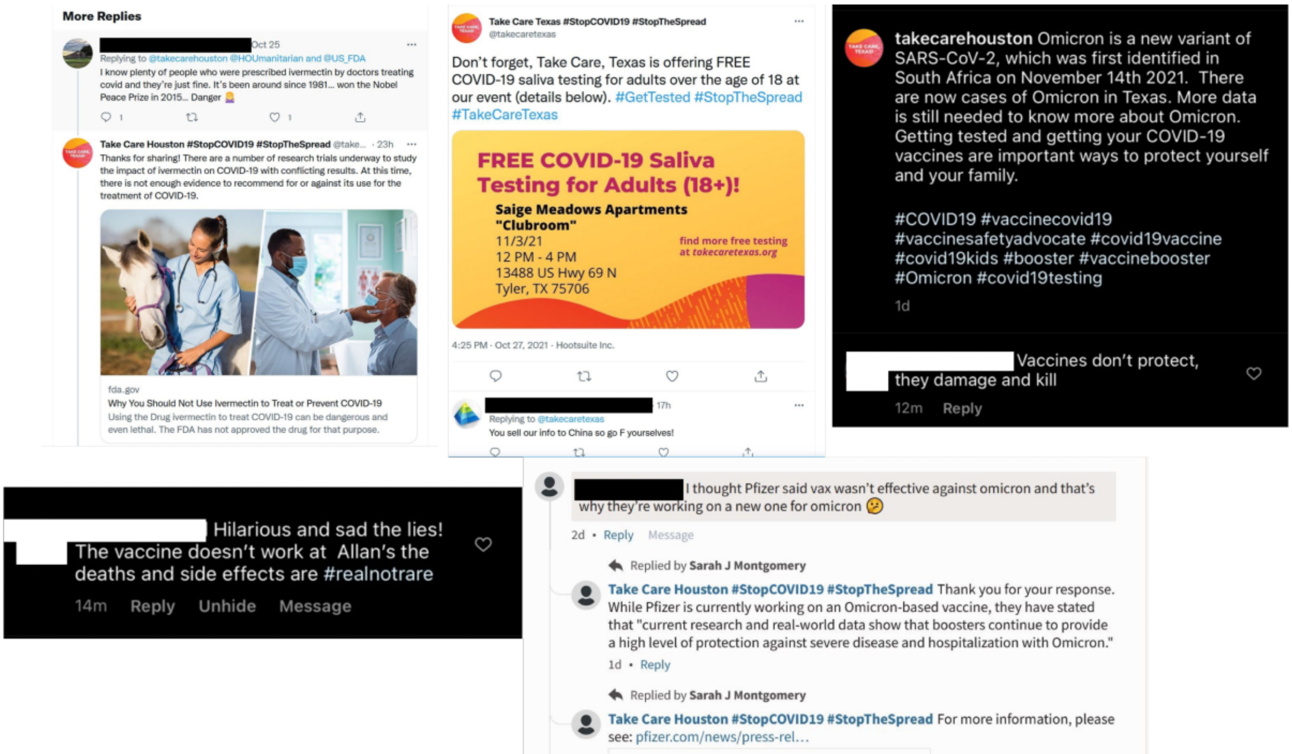
Social Media Interventions and Dissemination

We describe here the cumulative number of posts as well as the engagement metrics for each TCT account in each project region from June 2021 through March 2022, with marked regional differences in engagement by platform. On Facebook, Northeast Texas had the greatest number of posts (n=257), but Cameron County had the largest numbers of fans who liked and followed the account (n=246 and 250, respectively). Cameron County also had the greatest numbers of shared posts and video views (n=38 and 81, respectively), while the Harris County account had the greatest number of reactions (n=116). On Twitter, Northeast Texas had the most tweets (n=332), but Harris County had the greatest number of followers (n=19). While Northeast Texas had the greatest number of replies (n=57), Harris County had the greatest numbers of likes and retweets (n=105 and 95, respectively). On Instagram, Northeast Texas had the most posts (n=233). The Northeast Texas account also had the most community engagement with 440 likes and 2 comments, although the Harris County account had the most followers (n=111). Overall community engagement was determined by the numbers of likes and comments that each post or tweet received. Examples of posts/tweets with high amounts of community engagement for all of the accounts across the 3 regions include informational posts on the virus and its variants (Delta and Omicron), testing and vaccination/booster guidelines (when to test, which vaccine/booster to get, etc), masking guidelines, and community member testimonials.

However, not all community engagement was positive (Figure 3). Some posts, especially those responding to myths, received comments containing misinformation. Examples of such posts include a post debunking the use of ivermectin as a treatment for COVID-19, a post about the Omicron variant, a post about vaccination, a post advertising a testing event in Northeast Texas, and a post debunking the myth that COVID-19 vaccines change a person's DNA. One of the following 3 actions was taken in response to these comments: (1) We attempted to educate the user who made the comment by providing information from reputable sources; this action was taken for comments on a post debunking a myth about the COVID-19 vaccine and a post debunking the use of ivermectin as a treatment for COVID-19; (2) We ignored the user; this action was taken for comments responding to a post about vaccination and a post about the Omicron variant; and (3) We deleted the comment and blocked the user from replying to other posts; this action was taken for comments responding to a post about a testing event in Northeast Texas, as the comment included offensive language.

The analysis of social media posts provided us with insights into the latent needs of the users via the expression of various categories of speech acts based on their health beliefs toward COVID-19 misinformation. Such insights can be translated to design the architecture of just-in-time adaptive interventions, such as chatbots, thus ensuring such virtual interventions are theory-informed and data-driven for efficiently combating COVID-19 misinformation.

Figure 3. Sample bullying and misinformation posts on Take Care, Texas social media channels.



Discussion

Principal Findings

Table 5 presents an overview of the key findings and contributions from the study. Our results from community needs assessment indicated that participants from minority communities mentioned routine social media use and misinformation exposure, and emphasized the increased burden of misinformation among vulnerable populations. These findings enabled us to reposition online social media as a crucial data resource to understand COVID-19 misinformation dynamics not just among the general population, but also among minority groups. This finding is supportive of existing research that indicates no significant racial/ethnic disparities in social media use [43]. The analysis of needs assessment data and social media data sets can provide us with an opportunity to extend the application of human communication models to design informative and credible public health messages for risk communication with the potential to reduce public health burden and inform policy regulations. Such analytical insights can be translated for developing consumer education tools to improve the health literacy of community members such that there is a higher likelihood of individuals recognizing and resisting misinformation. So far, few misinformation mitigation interventions have effectively packaged theory and data insights simultaneously [44-48]. These data-driven theory-guided approaches are important during life-threatening public health events such as the COVID-19 pandemic.

Our results from social media listening and context-aware network analysis portray distinct network topologies and properties, and their dependencies on the semantics and syntactics of human communication [49]. The communication content, intent, structure, and framing together define the persuasive aspects of a social media information facet [50]. Using these tools in the engineering of risk communication can enable us to channel information efficiently, allowing us to target individuals and populations through personal and social contexts using social media.

As evident in our dissemination work, our efforts to promote COVID-19 testing and vaccination reach have been underwhelming. As part of our future efforts, we will apply our online social listening techniques to TCT data to examine content-based and network-driven facilitators and barriers to community engagement. Recruitment of advertising agencies and micro- and macro-influencers has become the new norm for implementation science in health promotion. The budgetary overhead of advertising costs to achieve adequate reach in social media has been concerning, especially for rural public health programs with limited resources, which can be remediated to some extent through the provision of advertising credits (as is done with cloud service credits for National Institutes of Health–funded projects) and reconfiguration of ranking algorithms that promote public health posts from verified scientific sources and nonprofit organizations, providing clear and direct-to-consumer digital engagement pathways for science in social media [51,52]. It is essential to facilitate community investments and educational offerings for social media optimization of outreach and prevention activities with a specialized focus on health marketing techniques, media literacy, and peer modeling. Specifically, prioritizing social media advocacy roles in public health and nonprofit organizations is imperative. One way to support this important function is through curriculum offerings from social media corporations with training on platform-specific optimization processes and financial monetization pathways for sponsored and nonsponsored posts by influencers in social media.

Millions of posts and accounts have been deleted and suspended across multiple social media platforms to combat misinformation spread on social media [53]. To support open science and agile science endeavors in this pertinent topic, we suggest online social media platforms release deidentified batches of social media posts so transdisciplinary theories, data analysis techniques, and innovative social marketing and data science methodologies can be developed and applied to understand the sociotechnical dimensions of misinformation spread [54]. Linking with location-sensing attributes can map misinformation spread to identify communities with the most dissemination. Such modeling efforts can enable us to optimize the positive supportive facets of online social media while examining its unintended consequences of polarization and social chambers [55]. Constant dialogue with community members and scanning of social media are essential to stay relevant and aware of misinformation [56]. Our work has done this, although this paper reports the data from a single time period. The resources and time to acheive sustainability of such efforts are extensive, whereas intentional and unintentional diffusion of misinformation appear to be much less demanding.

Our theoretical contributions are as follows. We built a foundational step to characterize the dynamics of health beliefs and intentions in online social media platforms to develop resilient information dissemination approaches that are resistant to misinformation spread [49,50]. We extended and applied behavioral constructs, which were developed in the era of face-to-face communication, to digital interactions on social media [57]. Our practical contributions include the development of end-to-end full stack social marketing materials using the IM framework, which could be repurposed for other health domains and community settings. Further, our findings from social media analysis can help us understand the mechanisms of social influence and content-specific patterns, which will help us engineer better content development and diffusion strategies.

Table 5. Key findings and contributions from this study.

Methodology	Key findings
Community needs assessment	Several themes regarding COVID-19 misinformation were highlighted in this study, including barriers and facilitators to preventive behaviors, the role of technology, etc. We also identified that social media is a primary resource of information and source of misinformation among minority communities.
Online social listening	This study provided population-level insights and patterns underlying behavioral constructs, communication attributes, and online social ties. Specific network structures and content forms were found to be efficient vehicles for misinformation resistance and true information dissemination. Significant differences were found between different expressions and content areas using online social listening, which can guide the development of impactful risk communication messages and expert conversational agents harnessing naturalistic conversational attributes as expressed in online social media.
Social marketing	Rigorous community needs assessment and theory integration allowed us to curate and develop a portfolio of social marketing materials, including social media postings. Despite evidence-based communication methods, community engagement and traction in online social platforms were limited.

Limitations

Our work is not without limitations. The interviews and listening sessions were conducted after May 2021, while the Twitter data collection covered the time period from January 2020 to January 2021. However, our analysis of Twitter data allowed us to capture key mechanisms underlying peer interactions when discussing COVID-19–related health behaviors. Results from our analysis have both mechanistic and topical findings (eg, how do we structure an interventional message vs what is the trending misinformation topic in a given time frame?). While we used our findings from social listening to inform interview questions and focus group guides, agile integration would further enhance the impact of our methods. Further, we included only the top 5 most prevalent speech act classes to train our deep learning model for the classification of the corrective misinformation Twitter data set, which led to the omission of the remaining speech act classes. In addition, while we only illustrated HBM constructs in the paper for online social listening, we have been using an integrated model with multiple behavior theories in our ongoing work. Not all the tweets were retrieved when the data sets were hydrated because of the retrospective organizational review policies, such as deleted tweets or user account suspensions. Some inherent challenges in using Twitter as the data source include restrictions on the number of requests/calls that can be made to the Twitter API, which leads to increased costs in obtaining the entirety of the data, fictitious or fake accounts, and privacy and ethical issues. Our sample size in participant interviews was limited, although additional data were collected in the form of listening sessions and focus groups. The developed interventions, which include social marketing materials, preliminary chatbot architecture, and social media content, still need to be evaluated for their effects toward combating misinformation.

Conclusions

We described a series of community-based studies and a large-scale observational study to examine and intervene regarding COVID-19 misinformation in our society and its vulnerable populations. While the internet and social media have democratized information, providing tools to marginalized groups to combat misinformation (ie, digital wildfires) is crucial [58]. Our study highlights the potential of online social listening to develop impactful risk communication strategies to combat misinformation spread and seeding on social media platforms. The feasibility and integration of data-intensive and grassroot-focused social engineering can provide a wealth of tools to ensure our communities are aware of and empowered with the skills of social media literacy to stay alert of cognitive blind spots (ie, heuristics) inherent in human reasoning with information environments [59,60]. For us to achieve the scalability and sustainability of social media operations in public health, it is important for industry corporations to provide the government and nonprofit organizations with technical resources and financial incentives along with algorithmic retraining to upvote evidence-based content for public wellness.

Acknowledgments

Research reported in this publication was supported by the National Library of Medicine (award numbers 1R01LM012974-01A1 and 3R01LM012974-02S) and the National Center for Advancing Translational Sciences of the National Institutes of Health (award numbers 3UL1TR003167-03S2, 3UL1TR003167-04S2, and 1U01TR004355-01) and the Centers for Disease Control and Prevention Research Center (award numbers 48DP006408-01-01 and 6U48DP006408-01-01). We thank the efforts of our extended team members, including the research teams, media teams, and community health workers from our project sites; Lisa Mitchell-Bennett; and Halo Agency LLC. In addition, we are thankful for the community partners Houston in Action for their support and commitment towards heralding the true spirit of science and public health. 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 Centers for Disease Control and Prevention.

Conflicts of Interest

None declared.

References

1. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020 Oct 14;7(10):201199 [FREE Full text] [doi: 10.1098/rsos.201199] [Medline: 33204475]
2. 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 Mar 13;12(3):e7255 [FREE Full text] [doi: 10.7759/cureus.7255] [Medline: 32292669]
3. Singh L, Bansal S, Bode L, Budak C, Chi G, Kawintiranon K, et al. A first look at COVID-19 information and misinformation sharing on Twitter. ArXiv 2020 Mar 31:13907v1. [Medline: 32550244]
4. Mian A, Khan S. Coronavirus: the spread of misinformation. BMC Med 2020 Mar 18;18(1):89 [FREE Full text] [doi: 10.1186/s12916-020-01556-3] [Medline: 32188445]
5. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
6. Bursztyn L, Rao A, Roth C, Yanagizawa-Drott D. Misinformation during a pandemic. University of Chicago, Becker Friedman Institute. 2020 Sep 01. URL: https://bfi.uchicago.edu/working-paper/2020-44/ [accessed 2022-12-12]
7. Krause NM, Freiling I, Beets B, Brossard D. Fact-checking as risk communication: the multi-layered risk of misinformation in times of COVID-19. Journal of Risk Research 2020 Apr 22;23(7-8):1052-1059. [doi: 10.1080/13669877.2020.1756385]
8. Rosenberg H, Syed S, Rezaie S. The Twitter pandemic: The critical role of Twitter in the dissemination of medical information and misinformation during the COVID-19 pandemic. CJEM 2020 Jul 06;22(4):418-421 [FREE Full text] [doi: 10.1017/cem.2020.361] [Medline: 32248871]
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 May 01;53(2):335-342. [doi: 10.1017/S0008423920000396]
10. Tasnim S, Hossain M, Mazumder H. Impact of Rumors and Misinformation on COVID-19 in Social Media. J Prev Med Public Health 2020 May;53(3):171-174 [FREE Full text] [doi: 10.3961/jpmph.20.094] [Medline: 32498140]

11. Larson HJ. The biggest pandemic risk? Viral misinformation. Nature 2018 Oct 16;562(7727):309-309. [doi: 10.1038/d41586-018-07034-4] [Medline: 30327527]
12. Merchant RM, Asch DA. Protecting the Value of Medical Science in the Age of Social Media and "Fake News". JAMA 2018 Dec 18;320(23):2415-2416. [doi: 10.1001/jama.2018.18416] [Medline: 30453315]
13. Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? BMJ 2014 Oct 14;349(oct14 5):g6178-g6178. [doi: 10.1136/bmj.g6178] [Medline: 25315514]
14. 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]
15. Simon F, Howard P, Nielsen R. Types, sources, and claims of COVID-19 misinformation. Reuters Institute. 2020. URL: https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation [accessed 2022-12-12]
16. Hébert-Dufresne L, Scarpino SV, Young J. Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement. Nat Phys 2020 Apr 24;16(4):426-431 [FREE Full text] [doi: 10.1038/s41567-020-0791-2] [Medline: 34221104]
17. van der Meer TGLA, Jin Y. Seeking Formula for Misinformation Treatment in Public Health Crises: The Effects of Corrective Information Type and Source. Health Commun 2020 May 14;35(5):560-575. [doi: 10.1080/10410236.2019.1573295] [Medline: 30761917]
18. Lee JJ, Kang K, Wang MP, Zhao SZ, Wong JYH, O'Connor S, et al. Associations Between COVID-19 Misinformation Exposure and Belief With COVID-19 Knowledge and Preventive Behaviors: Cross-Sectional Online Study. J Med Internet Res 2020 Nov 13;22(11):e22205 [FREE Full text] [doi: 10.2196/22205] [Medline: 33048825]
19. 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 Apr 13;2(1):e33827 [FREE Full text] [doi: 10.2196/33827]
20. Hou R, Perez-Rosas V, Loeb S, Mihalcea R. Towards Automatic Detection of Misinformation in Online Medical Videos. In: ICMI '19: 2019 International Conference on Multimodal Interaction. 2019 Presented at: International Conference on Multimodal Interaction; October 14-18, 2019; Suzhou, China p. 235-243. [doi: 10.1145/3340555.3353763]
21. Fell L, Gibson A, Bruza P, Hoyte P. Human Information Interaction and the Cognitive Predicting Theory of Trust. In: CHIIR '20: Proceedings of the 2020 Conference on Human Information Interaction and Retrieval. 2020 Presented at: Conference on Human Information Interaction and Retrieval; March 14-18, 2020; Vancouver, BC, Canada p. 145-152. [doi: 10.1145/3343413.3377981]
22. Lewandowsky S, Ecker UKH, Cook J. Beyond misinformation: Understanding and coping with the “post-truth” era. Journal of Applied Research in Memory and Cognition 2017 Dec;6(4):353-369. [doi: 10.1016/j.jarmac.2017.07.008]
23. Chen E, Lerman K, Ferrara E. Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. JMIR Public Health Surveill 2020 May 29;6(2):e19273 [FREE Full text] [doi: 10.2196/19273] [Medline: 32427106]
24. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory- and evidence-based health education programs. Health Educ Behav 1998 Oct 01;25(5):545-563. [doi: 10.1177/109019819802500502] [Medline: 9768376]
25. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
26. DocNow/twarc. GitHub. URL: https://github.com/DocNow/twarc [accessed 2021-01-06]
27. Memon SA, Carley KM. Characterizing COVID-19 Misinformation Communities Using a Novel Twitter Dataset. arXiv 2020 Sep 19:1-9 [FREE Full text]
28. Rosenstock IM. Why People Use Health Services. The Milbank Memorial Fund Quarterly 1966 Jul;44(3):94-127. [doi: 10.2307/3348967]
29. Bandura A. Social cognitive theory: an agentic perspective. Annu Rev Psychol 2001 Feb;52(1):1-26. [doi: 10.1146/annurev.psych.52.1.1] [Medline: 11148297]
30. Bandura A, Walters RH. Social Learning Theory. Englewood Cliffs, NJ: Prentice-Hall; 1977.
31. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
32. Glanz K, Rimer BK, Viswanath K. Health Behavior: Theory, Research, and Practice. Hoboken, NJ: John Wiley & Sons; 2008.
33. Searle JR, Kiefer F, Bierwisch M. Speech Act Theory and Pragmatics. Berlin, Germany: Springer Dordrecht; 1980.
34. Singh T, Olivares S, Myneni S. Latent Linguistic Motifs in Social Media Postings Resisting COVID-19 Misinformation. Stud Health Technol Inform 2022 Jun 06;290:557-561. [doi: 10.3233/SHTI220139] [Medline: 35673078]
35. Kupavskii A, Umnov A, Gusev G, Serdyukov P. Predicting the Audience Size of a Tweet. ICWSM 2021 Aug 03;7(1):693-696. [doi: 10.1609/icwsm.v7i1.14454]

36. Sesagiri Raamkumar A, Tan SG, Wee HL. Use of Health Belief Model-Based Deep Learning Classifiers for COVID-19 Social Media Content to Examine Public Perceptions of Physical Distancing: Model Development and Case Study. JMIR Public Health Surveill 2020 Jul 14;6(3):e20493 [FREE Full text] [doi: 10.2196/20493] [Medline: 32540840]
37. Nguyen DQ, Vu T, Nguyen AT. BERTweet: A pre-trained language model for English Tweets. arXiv 2020 Oct 05:1-6 [FREE Full text] [doi: 10.18653/v1/2020.emnlp-demos.2]
38. Faust K. Centrality in affiliation networks. Social Networks 1997 Apr;19(2):157-191 [FREE Full text] [doi: 10.1016/S0378-8733(96)00300-0]
39. Borgatti SP. 2-Mode Concepts in Social Network Analysis. Analytic Technologies. URL: http://www.analytictech.com/borgatti/papers/2modeconcepts.pdf [accessed 2022-12-12]
40. Latapy M, Magnien C, Vecchio ND. Basic notions for the analysis of large two-mode networks. Social Networks 2008 Jan;30(1):31-48 [FREE Full text] [doi: 10.1016/j.socnet.2007.04.006]
41. Borgatti SP, Everett MG, Freeman LC. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies; 2002.
42. Maggio LA, Krakow M, Moorhead LL. 'There were some clues': a qualitative study of heuristics used by parents of adolescents to make credibility judgements of online health news articles citing research. BMJ Open 2020 Aug 26;10(8):e039692 [FREE Full text] [doi: 10.1136/bmjopen-2020-039692] [Medline: 32847924]
43. Chou WS, 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]
44. Tay L, Hurlstone M, Kurz T, Ecker U. A comparison of prebunking and debunking interventions for implied versus explicit misinformation. Br J Psychol 2022 Aug;113(3):591-607. [doi: 10.1111/bjop.12551] [Medline: 34967004]
45. Haughey MM, Povolo M, Starbird K. Bridging Contextual and Methodological Gaps on the “Misinformation Beat”: Insights from Journalist-Researcher Collaborations at Speed. In: CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI Conference on Human Factors in Computing Systems; April 29, 2022-May 5, 2022; New Orleans, LA. [doi: 10.1145/3491102.3517503]
46. Aswani R, Kar AK, Ilavarasan PV. Experience: Managing Misinformation in Social Media—Insights for Policymakers from Twitter Analytics. Journal of Data and Information Quality 2020;12(1):1-18. [doi: 10.1145/3341107]
47. Banerjee D, Sathyanarayana Rao T. Psychology of misinformation and the media: Insights from the COVID-19 pandemic. Indian J Soc Psychiatry 2020;36(5):131. [doi: 10.4103/ijsp.ijsp_112_20]
48. Pasquetto IV, Swire-Thompson B, Amazeen MA, Benevenuto F, Brashier NM, Bond RM, et al. Tackling misinformation: What researchers could do with social media data. Harvard Kennedy School Misinformation Review 2020;1(8):49. [doi: 10.37016/mr-2020-49]
49. Singh T, Wang J, Myneni S. Revealing Intention In Health-related Peer Interactions: Implications For Optimizing Patient Engagement In Self-health Management. AMIA Annu Symp Proc 2020;2020:1120-1129 [FREE Full text] [Medline: 33936488]
50. Olivares S, Myneni S. Unpacking Misinfodemic During a Global Health Crisis: A Qualitative Inquiry of Psychosocial Characteristics in Social Media Interactions. Stud Health Technol Inform 2022 Jun 06;290:962-966. [doi: 10.3233/SHTI220223] [Medline: 35673162]
51. Egyedi A, O'Connor MJ, Martínez-Romero M, Willrett D, Hardi J, Graybeal J, et al. Using semantic technologies to enhance metadata submissions to public repositories in biomedicine. CEUR Workshop Proceedings. URL: https://ceur-ws.org/Vol-2275/paper1.pdf [accessed 2022-12-12]
52. Stein LD, Knoppers BM, Campbell P, Getz G, Korbel JO. Data analysis: Create a cloud commons. Nature 2015 Jul 09;523(7559):149-151. [doi: 10.1038/523149a] [Medline: 26156357]
53. Meeks L. Tweeted, deleted: theoretical, methodological, and ethical considerations for examining politicians' deleted tweets. Information, Communication & Society 2016 Nov 16;21(1):1-13. [doi: 10.1080/1369118x.2016.1257041]
54. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun 26;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
55. Terren L, Borge-Bravo R. Echo Chambers on Social Media: A Systematic Review of the Literature. Review of Communication Research 2021;9:99-118. [doi: 10.12840/issn.2255-4165.028]
56. 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]
57. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar 24;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
58. The Global Risks Report 2018. World Economic Forum. URL: https://www3.weforum.org/docs/WEF_GRR18_Report.pdf [accessed 2022-12-12]

59. Marsh EJ, Yang BW. Believing Things That Are Not True: A Cognitive Science Perspective on Misinformation. In: Southwell BG, Thorson EA, Sheble L, editors. Misinformation and Mass Audiences. Austin, TX: University of Texas Press; 2021:15-34.
60. Karduni A. Human-Misinformation interaction: Understanding the interdisciplinary approach needed to computationally combat false information. arXiv. 2019. URL: https://arxiv.org/abs/1903.07136 [accessed 2022-12-11]

Abbreviations

API: application programming interface
HBM: Health Belief Model
IM: Intervention Mapping
TCT: Take Care, Texas

Edited by T Mackey; submitted 08.06.22; peer-reviewed by A Pal, KC Yang, C Golz; comments to author 19.07.22; revised version received 01.10.22; accepted 07.12.22; published 03.02.23.

Please cite as:
Myneni S, Cuccaro P, Montgomery S, Pakanati V, Tang J, Singh T, Dominguez O, Cohen T, Reininger B, Savas LS, Fernandez ME
Lessons Learned From Interdisciplinary Efforts to Combat COVID-19 Misinformation: Development of Agile Integrative Methods From Behavioral Science, Data Science, and Implementation Science
JMIR Infodemiology 2023;3:e40156
URL: https://infodemiology.jmir.org/2023/1/e40156
doi:10.2196/40156
PMID:37113378

Original Paper

The Early Detection of Fraudulent COVID-19 Products From Twitter Chatter: Data Set and Baseline Approach Using Anomaly Detection

Abeed Sarker[1], PhD; Sahithi Lakamana[1], MSc; Ruqi Liao[2], BS; Aamir Abbas[3], MD, MS; Yuan-Chi Yang[1], PhD; Mohammed Al-Garadi[1], PhD

[1]Department of Biomedical Informatics, School of Medicine, Emory University, Atlanta, GA, United States

[2]H Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, United States

[3]Heinz College of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, PA, United States

Corresponding Author:
Abeed Sarker, PhD
Department of Biomedical Informatics
School of Medicine
Emory University
101 Woodruff Circle
Suite 4101
Atlanta, GA, 30030
United States
Phone: 1 6024746203
Email: abeed@dbmi.emory.edu

Abstract

Background: Social media has served as a lucrative platform for spreading misinformation and for promoting fraudulent products for the treatment, testing, and prevention of COVID-19. This has resulted in the issuance of many warning letters by the US Food and Drug Administration (FDA). While social media continues to serve as the primary platform for the promotion of such fraudulent products, it also presents the opportunity to identify these products early by using effective social media mining methods.

Objective: Our objectives were to (1) create a data set of fraudulent COVID-19 products that can be used for future research and (2) propose a method using data from Twitter for automatically detecting heavily promoted COVID-19 products early.

Methods: We created a data set from FDA-issued warnings during the early months of the COVID-19 pandemic. We used natural language processing and time-series anomaly detection methods for automatically detecting fraudulent COVID-19 products early from Twitter. Our approach is based on the intuition that increases in the popularity of fraudulent products lead to corresponding anomalous increases in the volume of chatter regarding them. We compared the anomaly signal generation date for each product with the corresponding FDA letter issuance date. We also performed a brief manual analysis of chatter associated with 2 products to characterize their contents.

Results: FDA warning issue dates ranged from March 6, 2020, to June 22, 2021, and 44 key phrases representing fraudulent products were included. From 577,872,350 posts made between February 19 and December 31, 2020, which are all publicly available, our unsupervised approach detected 34 out of 44 (77.3%) signals about fraudulent products earlier than the FDA letter issuance dates, and an additional 6 (13.6%) within a week following the corresponding FDA letters. Content analysis revealed *misinformation*, *information*, *political*, and *conspiracy theories* to be prominent topics.

Conclusions: Our proposed method is simple, effective, easy to deploy, and does not require high-performance computing machinery unlike deep neural network–based methods. The method can be easily extended to other types of signal detection from social media data. The data set may be used for future research and the development of more advanced methods.

(JMIR Infodemiology 2023;3:e43694) doi:10.2196/43694

KEYWORDS

coronavirus; COVID-19 drug treatment; social media; infodemiology; public health surveillance; COVID-19; misinformation; natural language processing; neural network; data mining

Introduction

As of September 7, 2021, over 220 million confirmed COVID-19 cases have been reported globally, with over 41 million reported cases in the United States alone [1]. As governments and public health agencies worldwide made efforts to mitigate the impact of the pandemic, one persistent problem has been the opportunistic promotion of fraudulent products claiming to treat, prevent, test, or cure COVID-19 infections. The shortage of resources during the pandemic has allowed companies to exploit the public by selling them falsified products. These products include face masks, hand sanitizers, and test kits. Additionally, misinformation from social media has led to the usage of nonrecommended therapies such as ivermectin, methanol, and herbs and vitamins to prevent and treat COVID-19 infections [2]. Fraudulent products pose a threat to public health by inhibiting prevention and enabling the spread of disease, and by drawing people away from seeking recommended care. Furthermore, there have been numerous reports of adverse health events caused by toxic exposures to fraudulent products that have no scientific evidence supporting their use [3,4]. The Ministry of Health of Iran reported that between February and April 2020, there were 5011 patients with methanol poisoning and 505 confirmed deaths due to misinformation that methanol can neutralize COVID-19 [5].

In response to the emergence of many fraudulent products, the US Food and Drug Administration (FDA) has issued warning letters [6]. These warning letters are typically issued after the products become popular and many people have already been exposed to them. Between March and July 2020, approximately 3139 warning letters were released. Of those, 98 (3.14%) pertained to COVID-19–related products [7]. Since it is not possible to advertise fraudulent products on television or via reliable news sources, social media platforms have been exploited for the mass promotion of such products. In fact, promotional content regarding such products over social networks, such as Twitter, is only a subset of the misinformation spread through these platforms, which has been referred to as an *infodemic* [8,9]. The fraudulent products are often promoted directly via the social media accounts (eg, Twitter and Facebook) of the entities profiting from their sales, and, if the promotions gain traction, information about them are circulated by other social media users. It is estimated that from 2020 to 2021, there was a US $500 million consumer loss due to fraudulent products being sold [2]. Consequently, information regarding the products spread through social networks in analogous patterns as other types of misinformation, including those related to COVID-19 [10]. There is, thus, the need to develop toxicovigilance tools that can automatically identify potentially fraudulent COVID-19 products early and generate alerts. While social networks provide fertile grounds for the proliferation of misinformation about fraudulent products, they also provide opportunities for responding to diverse challenges posed by the pandemic, and one potential utility of social media is the automated real-time surveillance of fraudulent COVID-19 products.

In this paper, we demonstrate that chatter about fraudulent products on Twitter, if curated systematically via natural language processing (NLP) and data-centric methods, can provide detectable early signals. We used publicly available streaming data from the Twitter COVID-19 application programming interface (API), which was specifically created by the company to aid COVID-19–related research [11]. Specifically, using Twitter data, we show that social media–based surveillance can detect many fraudulent products early, relative to the FDA warning issuance dates. Our approach to detecting fraudulent products is based on a simple intuition—that products that gain popularity among Twitter users, following their successful promotion, will exhibit increases in their mentions in COVID-19–related chatter. These abrupt increases in the frequency of mentions are likely to be detectable through time-series anomaly detection methods. It is also likely that products that gain *relatively* higher popularity will exhibit anomalous increases of *relatively* higher magnitudes in their mentions among all COVID-19–related Twitter chatter. We present our findings in the following section and detail our methods at the end of the article.

Methods

Ethical Considerations

This study was reviewed by Emory University's institutional review board, which determined on June 11, 2020, that it was exempt from further review (category 4), since only publicly available data were included (STUDY00000711).

Data Collection

We collected data using the COVID-19 streaming API of Twitter [11]. This API was made available by Twitter specifically for supporting COVID-19–related research, and it does not impose throughput limitations or daily or monthly quotas. Consequently, we were able to collect all tweets that mentioned COVID-19–related keywords and phrases (eg, *coronavirus*, *covid19*, and *covid*) [11]. We collected data from February 19 to December 31, 2020. Streaming data were stored in real time in a *mongodb* database hosted on the Google Cloud platform. The collection of data was continuous with only minor down times that were necessary for system modifications or updates.

Product Detection

The list of products and entities were manually collected from the FDA website [6]. The products included were advertised as treatments or cures, tests, or preventative measures for COVID-19. We curated a comprehensive list of entity names, products, FDA letter dates, persons who owned the entities or the products, websites, and social media profiles (if any). We curated this information for a total of 183 letters issued by the FDA. Each warning letter was manually reviewed. From these, we manually curated a set of product names or entity names that were potentially used for promotion over social media. If the same product was mentioned in multiple letters, we only included the first mention of the product or entity and the corresponding date, excluding the later ones. We also manually curated keywords and phrases that were likely to be used to refer to the products or entities on Twitter. The full list of products and entities and their earliest letter dates is provided in Table 1.

Since product and entity names are often misspelled by social media subscribers, keyword-based searches typically miss large numbers of posts that contain misspelled versions of the names. To increase the sensitivity of our searches, we applied NLP to increase the number of keywords we searched for that were associated with each product or entity. Specifically, we generated potential spelling variants or misspellings of the products and entities using a previously developed data-centric tool [12]. The variant generation tool uses a combination of semantic and lexical similarity measures to automatically identify common misspellings and spelling variants of terms or phrases, including multiword expressions. Our past work revealed that such lexical expansion strategies are capable of significantly increasing retrieval or detection rates from Twitter, particularly for medical terms (eg, names of medications) that are often difficult to spell [13]. Examples of product names extracted from the warning letters and their automatically generated lexical variants are shown in Table 2. We included all products or entities and their spelling variants that had at least 10 mentions in our collected data. We excluded key phrases that were mentioned less than 10 times because such low occurrences indicated that the corresponding products or entities were either not promoted over Twitter or never actually gained popularity on the platform. We enumerated the mentions of each product or entity, including their spelling variants, from the entire collected data set. Counts of spelling variants were grouped with the original products or entities. Daily counts were normalized by the total number of posts collected on the same days. The daily relative frequencies were represented as the number of mentions per 1000 tweets.

Table 1. Key phrases included in this study along with their types and the date of the first letter mentioning each of them.

Number	Key phrase	Type	First detected letter date
1	Antimicrobial solution	Treatment	November 2, 2020
2	Aromatherapy	Treatment	March 6, 2020
3	Ayurvedic products	Treatment	April 13, 2020
4	Bee products	Treatment	October 23, 2020
5	Berberine	Treatment	October 23, 2020
6	Betterfly	Treatment	September 1, 2020
7	Bioflavonoids	Treatment	October 23, 2020
8	Biomagnetism	Treatment	August 19, 2020
9	Chlorine dioxide	Treatment	April 8, 2020
10	Cod liver oil	Treatment	May 25, 2020
11	Colloidal silver	Treatment	March 6, 2020
12	Colostrum	Treatment	May 26, 2020
13	Corona-cure	Treatment	March 26, 2020
14	Covid-19 rapid test kit	Test kit	June 10, 2020
15	Curativa	Treatment	June 25, 2020
16	Elderberry syrup	Treatment	November 10, 2020
17	Elderberry tincture	Treatment	March 6, 2020
18	Essential oil	Treatment	March 6, 2020
19	Eupatorium perfoliatum	Treatment	March 6, 2020
20	Grapefruit seed extract	Treatment	May 26, 2020
21	Hypochlorous acid	Treatment	November 2, 2020
22	Iodine products	Treatment	June 10, 2020
23	Kratom	Treatment	May 15, 2020
24	Magnetic therapy	Treatment	August 19, 2020
25	Methylene blue	Treatment	May 29, 2020
26	Nad+	Treatment	May 6, 2020
27	Nephron pharmaceuticals	Treatment	May 22, 2020
28	Niacin product	Treatment	September 1, 2020
29	Novabay	Entity	November 2, 2020
30	Oracare	Treatment	November 18, 2020
31	Pro breath	Treatment	November 18, 2020
32	Quercetin	Treatment	June 15, 2020
33	Salt therapy	Treatment	March 30, 2020
34	Santiste	Entity	April 27, 2020
35	Super C	Treatment	April 21, 2020
36	Superblue silver immune gargle	Treatment	April 9, 2020
37	Supersilver whitening toothpaste	Treatment	April 9, 2020
38	Traditional Chinese medicine	Treatment	May 8, 2020
39	Transdermal patch/defendTM patch	Treatment	April 27, 2020
40	Umbilical cord blood	Treatment	June 4, 2020
41	Vapore	Treatment	July 30, 2020
42	Vidacord	Treatment	June 4, 2020

Number	Key phrase	Type	First detected letter date
43	Vivify	Entity	March 6, 2020
44	Xosomes	Treatment	June 4, 2020

Table 2. Examples of fraudulent product names extracted from the US Food and Drug Administration's warning letters and their automatically generated lexical variants.

Product	Spelling variants
Chlorine dioxide	*chlorinedioxide*, *chloride dioxide*, *chorine dioxide*, *clorine dioxide*, and *clorinedioxide*
Fortify humic beverage concentrate	*fortify humic beverage concentrates* and *fortify humic beverage cocentrate*
Electrify fulvic beverage concentrate	*electrify fulvic beverage cocentrate*, *electrify fulvic beverage concetrate*, and *electrify fulvic beverage concentrates*
Supersilver whitening toothpaste	*supersilver whitening toothpast*, *supersilver whitening toothpastes*, and *supersilver whitening tooth paste*
Superblue fluoride free toothpaste	*superblue fluoride free tooth paste*, *superblue fluoride free toothpastes*, and *superblue fluoride free toothpast*
Prefense hand sanitizers	*prefense handsanitzers*, *prefense hand sanitizes*, *prefense hand sanitiers*, *prefense hand andsanitizers*, *prefense hand*, *prefense hand handsantizers*, *prefense hand handsanitzers*, *prefense handsantizer*, *prefense handsanitizers*, *prefense*, *prefense hand santitizers*, *prefense handsanitisers*, and *prefense handsanitzer*
Covid-19 cough syrup	*covid 19 cough syrups*, *covid 19 coughsyrup*, *covid 19 cough syrup*, and *covid 19 cough coughsyrup*
nCov19 spike protein	*ncov19 spike spike protein*, *ncov19 spike spikeproteins*, *ncov19 spike protei*, *ncov19 spikey proteins*, *ncov19 spike spikeprotein*, *ncov19 spikeprotien*, *ncov19 spike proteins*, *ncov19 spike spikey proteins*, *ncov19 spikeprotein*, *ncov19 spikeproteins*, and *ncov19 spike spikeprotien*

Detecting Anomalies

We applied a 14-day moving average filter to construct a smooth line representing the daily mention frequencies, and anomalies or outliers were detected relative to this moving average line. For each day, the *residual* for SD calculation was computed by subtracting the 14-day moving average from the relative frequency per 1000 tweets on that day. For a given day (n), the SD for the day (σ_n), is computed progressively, given as follows:

where x_i is the relative frequency for day i and μ_i is the 14-day moving average on day i. Thus, the SD computed for a given day includes all the data points starting from day 1. The SD for the first day (February 19, 2020) for any product is by definition 0. This may potentially give the anomaly detection approach an unfair advantage by increasing the sensitivity of detection in the early days easier. Therefore, we artificially added a nonzero SD on day 1, computed as:

where x_1 is the product mention frequency on day 1, X_1 is the total number of tweets collected on day 1, *std*() is the SD function, and $\mu_{1..4}$, are the moving averages over the first 4 days. The value is divided by k to adjust for the $\boldsymbol{k} \times$ *std*()function that is applied to compute the boundaries beyond which a data point would be considered an outlier (***k***=3 in our experiments). This artificial initial bias that we added, therefore, decreases the chances of our approach to detect outliers early on in the time lines and makes the task of detecting anomalies slightly more difficult, particularly for products that have a relatively low number of mentions. For some products, for example, there are many days with 0 mentions early on in their time lines, but the added bias causes the progressive SD to be nonzero. For 3 products with letter issue dates in March 2020, this added bias caused the method to miss early outliers that are detectable without adding the bias. Specific details are provided in Multimedia Appendix 1. The minimum value for daily relative frequency was set at 0.001 (ie, ***k*** × 0.001 served as the minimum threshold for outlier detection).

The chosen window size (14) and SD (3), for which we report results in this paper, were relatively conservative choices for signal detection. We also performed experiments with multiple window sizes (7, 10, and 14) and SD thresholds (2, 2.5, and 3) to study how the anomaly detection performance varied on the basis of these parameters. Slight variations in window sizes and SD did not impact overall performance.

Evaluation

Data points that had a distance of more than 3 SDs from the moving average were considered outliers (ie, signals). For each key phrase, the date of the first outlier was compared with the FDA letter issuance date to determine if the signal was detected earlier, within 1 week, or later than the FDA letter issuance date. System percentage accuracy was computed using the formula: *#early/#total*. For products that were mentioned in multiple letters, our approach was only considered successful in early detection if the outlier was detected prior to the first mention date. Thus, the reported system performance is actually likely to be lower than that in practice.

Content Analysis

To obtain an idea about the contents of the Twitter posts associated with the fraudulent products, we performed a brief, manual content analysis of 400 posts associated with 2 products (200 each). The 2 products chosen—*chlorine dioxide* and *quercetin*—had over 10,000 posts in the data set each and were among the top 5 most frequently mentioned. We performed random sampling to select the posts for manual review. Two authors manually reviewed the posts and identified possible categories for the posts. Following the first round of coding, the categories were collapsed into broader topics. Finally, we computed the distributions of these topics among the manually categorized posts.

Results

The issue dates of the letters ranged from March 6, 2020, to June 22, 2021. Through manual review of each letter, we identified 221 potential keywords or phrases that were either associated with the products (eg, product names) or the entities selling them. From this set, we excluded key phrases collected after the year 2020. Some products were promoted by different entities at different times, causing them to be repeated in the warning letters. Since our primary objective was to assess the possibility of early detection, we excluded repeated key phrases, retaining only their first occurrences (n=56). Furthermore, since our focus was to detect products that gained popularity via promotion on Twitter, we excluded key phrases that were mentioned less than 10 times, including their lexical variants (n=12). In total, 44 key phrases met all the inclusion criteria. Table 1 presents all 44 keywords, their types (ie, product or entity), and the FDA letter issuance dates. The full curated data along with additional information is available in Multimedia Appendix 2.

We included a total of 577,872,350 COVID-19–related tweets in our analysis, which were collected from February 19 to December 31, 2020. We computed the daily counts of the key phrases (along with their spelling variants, if any). Increases in the number of key phrase mentions that were higher than 3 SDs from the 14-day moving average of mentions were flagged as potential "signals." In total, 43 out of the 44 key phrases showed anomalous increases in their mentions at some point of time within our collected data. For 34 out of the 44 (77.3%) key phrases, signals of anomalous increases in chatter were detectable prior to the FDA letter issuance dates. An additional 6 (13.6%) key phrases had anomalous increases within 7 days of the FDA letter issuance dates. Figure 1 presents the daily relative frequencies for 6 sample products or entities from our data set, their 3-SD ranges, and the moving averages. The top 4 panels in the figure represent products or entities for which anomalies were detected prior to the FDA letter issue dates and the bottom 2 panels (highlighted in beige in Figure 1) represent those for which anomalies were not detected prior to the letter issue dates. A larger figure with all 44 products or entities are provided in Multimedia Appendix 3. The daily counts for all 44 key phrases are provided in tabular format in Multimedia Appendix 4.

Table 3 presents the distribution of the topics in terms of percentage for the 2 products identified via manual analysis. We discovered 4 prominent topics—*misinformation*, *information*, *conspiracy theories*, and *political*. Posts that could not be categorized as any of these were labeled as *other*. Misinformation included the spreading of information that these products cure or treat COVID-19. They also consisted of marketing and promotion of these products. Some posts claimed that the user took these products to successfully recover from COVID-19. Particularly for quercetin, many posts encouraged the consumption of multiple dietary supplements such as zinc and vitamin C alongside quercetin. Some of the posts also shared unverified news articles that claimed high efficacy of these products against COVID-19. Many posts shared reliable information and news that countered the unverified claims. Posts belonging to the *information* category also mentioned the FDA letters that we discussed in this paper. We also came across a number of posts that were spreading conspiracy theories, which included false claims about the vaccine or suggestions that the government was intentionally suppressing information about the efficacy of these products. Posts that were categorized as *political* included those that tagged politicians, commented on statements made by politicians, or discussed political mandates. Note that while the proportion of misinformation appears higher for quercetin, many posts that mentioned it were simply speculations about its effectiveness in preventing COVID-19, and the posts often referred to or recommended other forms of protection as well, such as masking. For consistency, we grouped such speculations and advice as misinformation.

Figure 1. Daily relative frequencies for 6 sample products or entities from our data set, their 3-SD ranges, and the moving averages. FDA: US Food and Drug Administration.

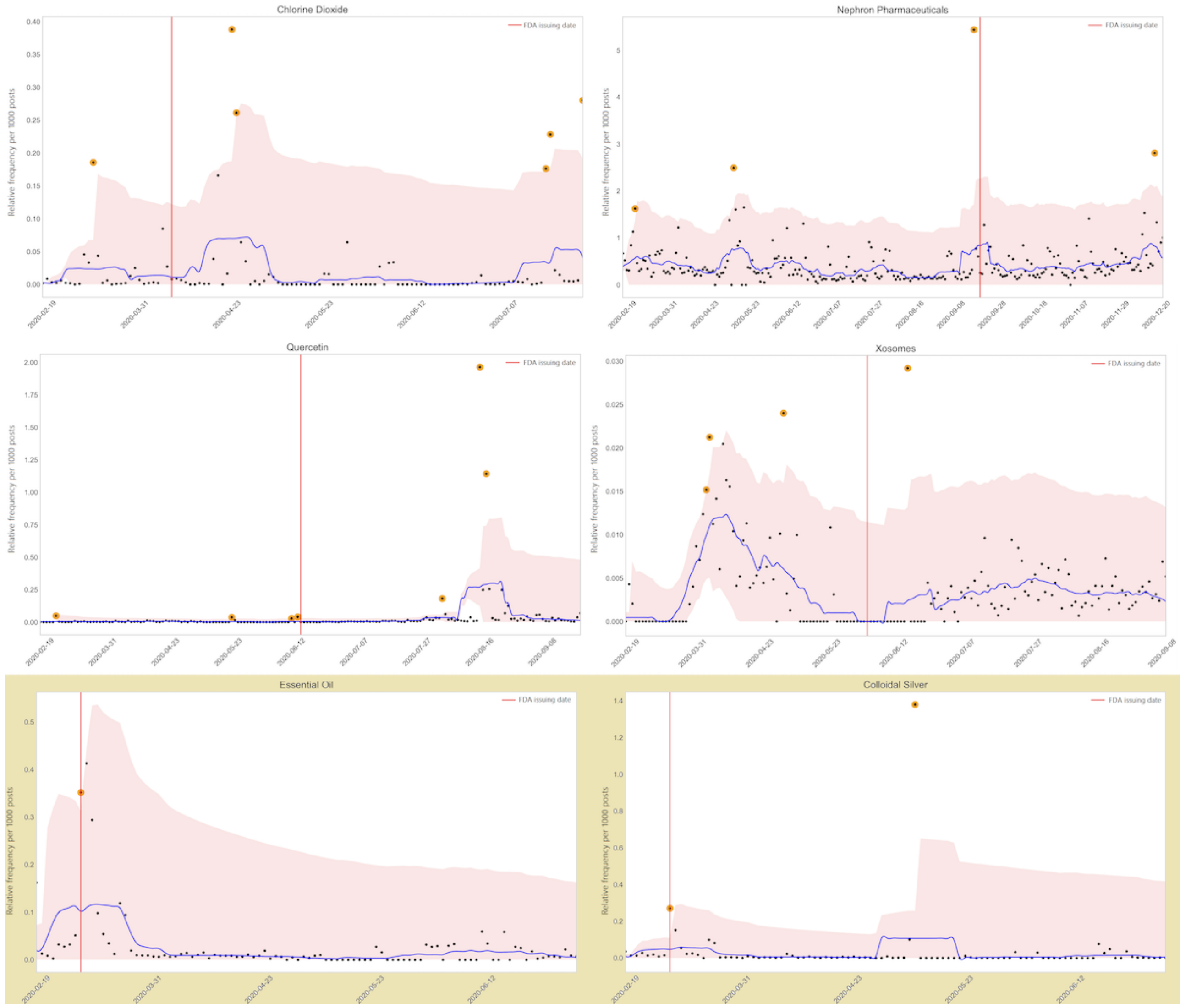


Table 3. Distribution of topics in the manually reviewed posts about chlorine dioxide and quercetin.

Topics	Chlorine dioxide, %	Quercetin, %
Misinformation, marketing, or promotion	37.0	58.0
Information or news	32.5	30.5
Conspiracy theories	13.5	4.0
Political	7.0	2.5
Other	10.0	4.0

Discussion

Principal Results

The primary finding of this study is that our proposed approach allows for anomaly detection in Twitter chatter that is typically associated with a fraudulent product or entity. This method, combined with further in-depth content analysis, can potentially enable us to detect fraudulent products early—as they start getting popular—from Twitter. Since social media serves as a platform for promoting such fraudulent products, increases in their popularity are also likely to cause increases in their web-based mentions. This phenomenon potentially makes it possible to detect fraudulent products that are rising in popularity to the point that renders them a public health concern. Thus, while social media plays an important role in the spread of information about fraudulent products, and misinformation in general, it may also serve as a potential resource for the surveillance of such information. While other information sources are often laggy, social media provides the opportunity to conduct surveillance in close to real time. While our approach is relatively simple, it is very effective in detecting fraudulent products that rise in popularity. Determining the contents of the chatter and the specific dangers that may arise from the content requires further analysis, which is beyond the scope of this study, and we intend for such analyses to be carried out in future work. There is also the potential of developing more advanced and effective methods for detecting such fraudulent products.

In addition to our approach, the data set curated from publicly available FDA reports can help drive future research in this space. To the best of our knowledge, there is no such data set that has been curated and is available for research. Thus, the data set itself can be of high utility to the research community. Importantly, the data set can serve as an important resource for the development of methods to detect misinformation in general from social media data.

Limitations

There are several potential limitations of the proposed approach. First, it requires data that are not rate-limited (eg, data from the standard Twitter streaming API). Anomalous increases may not be detectable from rate-limited streams, since large increases in volume are likely to be dampened by the APIs. For real-time detection of fraudulent product candidates, deployment needs to be performed on streaming data, although it is also possible to periodically run the anomaly detection scripts on stored, static data. Second, we were only able to calculate the percentage of early detection within our given sample, and based on the current data, we were unable to realistically estimate CIs for the percentage values reported. Third, the anomaly detection approach relies on characteristic abrupt increases in chatter volumes about a given topic. It is possible that some fraudulent products may gain popularity gradually, causing the normalized counts to never exceed the SD threshold. In such cases, varying the window size (eg, using 7-day moving averages) and lowering the SD thresholds may improve the detection capability of the method. However, lowering the SD threshold is also likely to result in larger numbers of false positives—an aspect that we did not take into account in this study. We believe that not taking false positives into account in this study is justifiable, since in practical settings, all signals associated with noun phrases would be reviewed by experts; hence, it is perhaps better if the method is biased in favor of *recall* (ie, more true and false positives) rather than *precision.*

We also do not address the detection of candidate fraudulent substances in this study. Several mechanisms can be used for detecting candidates including, but not limited to, named entity recognition (likely to be high precision but low recall), simple part-of-speech tagging to identify noun phrases (high recall and low precision), and topic modeling methods that identify possible topics from texts (low recall and high precision). We intend to explore these strategies in future work. Even without this component, we believe our approach is an improvement over past studies that did not take into account the warning letter dates. We also did not conduct in-depth analysis of the content associated with all the included products or entities or the features associated with the accounts that post the information. Both of these are important future research directions. Finally, since the daily counts are normalized by the total number of tweets on the same day, it is possible that large increases in absolute counts of specific key phrases are not detectable due to equal or larger increases in the total volume of posts on the same day.

Comparison With Prior Work

Our work is not the first to explore the utility of social media as a potential source for detecting fraudulent COVID-19 products. In recent studies, unsupervised NLP methods such as topic modeling and supervised methods such as text classification have been proposed for the automatic detection of such products from social media data [14-16]. Others focused more broadly on detecting misinformation using social media or internet-based data [17,18]. However, these studies did not take into account the time factor. Typically, once the FDA issues a warning about a fraudulent product, there is a rise in chatter regarding the product, but such rises are driven by media coverage or increased public awareness. We observed this phenomenon for most products included in the study, particularly the ones detected within 1 week of the FDA letter issuance dates. Some recent studies have conducted more in-depth analyses of misinformation associated with specific products or substances that were rumored to be effective against COVID-19. For example, Kim et al [19] fine-tuned transformer-based models to automatically classify misinformation related to garlic. Quinn et al [20] analyzed misinformation related to vitamin D and COVID-19 on YouTube. A larger set of studies has focused on COVID-19–related misinformation on social media, in general, for topics such as, for example, vaccines [21-23]. To the best of our knowledge, our approach is the first to attempt to detect fraudulent treatments early. The proposed approach is also simple and computationally inexpensive as it relies on fundamental characteristics of social media chatter (ie, increases in the volume of chatter about a particular topic resulting from increases in its popularity) and is unsupervised (ie, no training data required).

Conclusions

The emergence of fraudulent products associated with COVID-19 has been a significant problem in the fight against the pandemic. Social media has served as a platform for advertising and promoting fraudulent products. While social media makes it easier for opportunist entities to promote and sell fraudulent products, this resource may also be used to conduct surveillance of fraudulent substances. In this paper, we show that it is possible to detect many fraudulent products potentially early from Twitter data. Our simple approach used a time-series anomaly detection method for detecting anomalous increases in mentions of fraudulent substances in Twitter chatter and obtained promising performance. Future work will focus on deploying the NLP pipeline and improving upon the study limitations.

Acknowledgments

This research was funded by the School of Medicine, Emory University.

Conflicts of Interest

None declared

Multimedia Appendix 1
Additional system performance details.
[DOCX File , 16 KB - infodemiology_v3i1e43694_app1.docx]

Multimedia Appendix 2
Full details about the fraudulent products.
[DOCX File , 45 KB - infodemiology_v3i1e43694_app2.docx]

Multimedia Appendix 3
Daily counts for all 44 products/entities from our dataset, their 3-standard deviation ranges, and the moving averages.
[PNG File , 3605 KB - infodemiology_v3i1e43694_app3.png]

Multimedia Appendix 4
Daily counts for all included key phrases.
[XLSX File (Microsoft Excel File), 60 KB - infodemiology_v3i1e43694_app4.xlsx]

References

1. Trends in Number of COVID-19 Cases and Deaths in the US Reported to CDC, by State/Territory. Centers for Disease Control and Prevention. URL: https://covid.cdc.gov/covid-data-tracker/#trends_totalandratedeathssevendayrate [accessed 2021-02-19]
2. Chavda VP, Sonak SS, Munshi NK, Dhamade PN. Pseudoscience and fraudulent products for COVID-19 management. Environ Sci Pollut Res Int 2022 Sep 14;29(42):62887-62912 [FREE Full text] [doi: 10.1007/s11356-022-21967-4] [Medline: 35836045]
3. Turner L. Preying on public fears and anxieties in a pandemic: businesses selling unproven and unlicensed "Stem Cell Treatments" for COVID-19. Cell Stem Cell 2020 Jun 04;26(6):806-810 [FREE Full text] [doi: 10.1016/j.stem.2020.05.003] [Medline: 32464095]
4. Reihani H, Ghassemi M, Mazer-Amirshahi M, Aljohani B, Pourmand A. Non-evidenced based treatment: An unintended cause of morbidity and mortality related to COVID-19. Am J Emerg Med 2021 Jan;39:221-222 [FREE Full text] [doi: 10.1016/j.ajem.2020.05.001] [Medline: 32402498]
5. Mousavi-Roknabadi RS, Arzhangzadeh M, Safaei-Firouzabadi H, Mousavi-Roknabadi RS, Sharifi M, Fathi N, et al. Methanol poisoning during COVID-19 pandemic; a systematic scoping review. Am J Emerg Med 2022 Feb;52:69-84 [FREE Full text] [doi: 10.1016/j.ajem.2021.11.026] [Medline: 34883289]
6. Fraudulent Coronavirus Disease 2019 (COVID-19) Products. US Food and Drug Administration. URL: https://www.fda.gov/consumers/health-fraud-scams/fraudulent-coronavirus-disease-2019-covid-19-products [accessed 2021-08-09]
7. Bramstedt KA. Unicorn Poo and Blessed Waters: COVID-19 Quackery and FDA Warning Letters. Ther Innov Regul Sci 2021 Jan 01;55(1):239-244 [FREE Full text] [doi: 10.1007/s43441-020-00224-1] [Medline: 33001378]
8. The Lancet Infectious Diseases. The COVID-19 infodemic. Lancet Infect Dis 2020 Aug;20(8):875. [doi: 10.1016/s1473-3099(20)30565-x]
9. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
10. Larson HJ. The biggest pandemic risk? Viral misinformation. Nature 2018 Oct 16;562(7727):309-309. [doi: 10.1038/d41586-018-07034-4] [Medline: 30327527]
11. Chen A. Building a data stream to assist with COVID-19 research. Twitter. 2021 Mar 16. URL: https://blog.twitter.com/engineering/en_us/topics/insights/2021/how-we-built-a-data-stream-to-assist-with-covid-19-research [accessed 2023-03-12]
12. Sarker A. LexExp: a system for automatically expanding concept lexicons for noisy biomedical texts. Bioinformatics 2021 Aug 25;37(16):2499-2501 [FREE Full text] [doi: 10.1093/bioinformatics/btaa995] [Medline: 33244602]
13. Sarker A, Gonzalez-Hernandez G. An unsupervised and customizable misspelling generator for mining noisy health-related text sources. J Biomed Inform 2018 Dec;88:98-107 [FREE Full text] [doi: 10.1016/j.jbi.2018.11.007] [Medline: 30445220]
14. Mackey TK, Li J, Purushothaman V, Nali M, Shah N, Bardier C, et al. Big data, natural language processing, and deep learning to detect and characterize illicit COVID-19 product sales: infoveillance study on Twitter and Instagram. JMIR Public Health Surveill 2020 Aug 25;6(3):e20794 [FREE Full text] [doi: 10.2196/20794] [Medline: 32750006]

15. Fittler A, Adeniye L, Katz Z, Bella R. Effect of infodemic regarding the illegal sale of medications on the internet: evaluation of demand and online availability of ivermectin during the COVID-19 pandemic. Int J Environ Res Public Health 2021 Jul 13;18(14):7475 [FREE Full text] [doi: 10.3390/ijerph18147475] [Medline: 34299920]
16. Madhusudhana K. Text classification models for automatic detection of fake COVID products and news on social media. University of Buffalo. 2021. URL: https://search.proquest.com/openview/3ff261c58cc2cbcc203f7e05843db8d5/1.pdf?pq-origsite=gscholar&cbl=18750&diss=y [accessed 2023-03-06]
17. 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. [doi: 10.1016/s2589-7500(20)30318-6]
18. Shams AB, Hoque Apu E, Rahman A, Sarker Raihan MM, Siddika N, Preo RB, et al. Web Search Engine Misinformation Notifier Extension (SEMiNExt): a machine learning based approach during COVID-19 pandemic. Healthcare (Basel) 2021 Feb 03;9(2) [FREE Full text] [doi: 10.3390/healthcare9020156] [Medline: 33546110]
19. Kim MG, Kim M, Kim JH, Kim K. Fine-tuning BERT models to classify misinformation on garlic and COVID-19 on Twitter. Int J Environ Res Public Health 2022 Apr 22;19(9):5126 [FREE Full text] [doi: 10.3390/ijerph19095126] [Medline: 35564518]
20. Quinn EK, Fenton S, Ford-Sahibzada CA, Harper A, Marcon AR, Caulfield T, et al. COVID-19 and vitamin D misinformation on YouTube: content analysis. JMIR Infodemiology 2022 Mar 14;2(1):e32452 [FREE Full text] [doi: 10.2196/32452] [Medline: 35310014]
21. Gabarron E, Oyeyemi SO, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bull World Health Organ 2021 Mar 19;99(6):455-463A. [doi: 10.2471/blt.20.276782]
22. 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 Nov 17;7(11):e30642 [FREE Full text] [doi: 10.2196/30642] [Medline: 34653016]
23. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. J Med Internet Res 2022 Aug 04;24(8):e37367 [FREE Full text] [doi: 10.2196/37367] [Medline: 35816685]

Abbreviations

API: application programming interface
FDA: US Food and Drug Administration
NLP: natural language processing

Edited by T Mackey; submitted 20.10.22; peer-reviewed by MG Kim, J Lei; comments to author 11.01.23; revised version received 31.01.23; accepted 25.02.23; published 14.03.23.

Please cite as:
Sarker A, Lakamana S, Liao R, Abbas A, Yang YC, Al-Garadi M
The Early Detection of Fraudulent COVID-19 Products From Twitter Chatter: Data Set and Baseline Approach Using Anomaly Detection
JMIR Infodemiology 2023;3:e43694
URL: https://infodemiology.jmir.org/2023/1/e43694
doi:10.2196/43694
PMID:37113382

Original Paper

Advertising Alternative Cancer Treatments and Approaches on Meta Social Media Platforms: Content Analysis

Marco Zenone[1], MSc; Jeremy Snyder[2], PhD; Jean-Christophe Bélisle-Pipon[2], PhD; Timothy Caulfield[3], LLB, LLM; May van Schalkwyk[1], MBBS, MPH; Nason Maani[4], PhD

[1]Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, United Kingdom

[2]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

[3]Health Law Institute, University of Alberta, Edmonton, BC, Canada

[4]Global Health Policy Unit, University of Edinburgh, Edinburgh, United Kingdom

Corresponding Author:
Marco Zenone, MSc
Faculty of Public Health and Policy
London School of Hygiene and Tropical Medicine
Keppel St
London, WC1E 7HT
United Kingdom
Phone: 1 7789086246
Email: marco.zenone@lshtm.ac.uk

Abstract

Background: Alternative cancer treatment is associated with a greater risk of death than cancer patients undergoing conventional treatments. Anecdotal evidence suggests cancer patients view paid advertisements promoting alternative cancer treatment on social media, but the extent and nature of this advertising remain unknown. This context suggests an urgent need to investigate alternative cancer treatment advertising on social media.

Objective: This study aimed to systematically analyze the advertising activities of prominent alternative cancer treatment practitioners on Meta platforms, including Facebook, Instagram, Messenger, and Audience Network. We specifically sought to determine (1) whether paid advertising for alternative cancer treatment occurs on Meta social media platforms, (2) the strategies and messages of alternative cancer providers to reach and appeal to prospective patients, and (3) how the efficacy of alternative treatments is portrayed.

Methods: Between December 6, 2021, and December 12, 2021, we collected active advertisements from alternative cancer clinics using the Meta Ad Library. The information collected included identification number, URL, active/inactive status, dates launched/ran, advertiser page name, and a screenshot (image) or recording (video) of the advertisement. We then conducted a content analysis to determine how alternative cancer providers communicate the claimed benefits of their services and evaluated how they portrayed alternative cancer treatment efficacy.

Results: We identified 310 paid advertisements from 11 alternative cancer clinics on Meta (Facebook, Instagram, or Messenger) marketing alternative treatment approaches, care, and interventions. Alternative cancer providers appealed to prospective patients through eight strategies: (1) advertiser representation as a legitimate medical provider (n=289, 93.2%); (2) appealing to persons with limited treatments options (n=203, 65.5%); (3) client testimonials (n=168, 54.2%); (4) promoting holistic approaches (n=121, 39%); (5) promoting messages of care (n=81, 26.1%); (6) rhetoric related to science and research (n=72, 23.2%); (7) rhetoric pertaining to the latest technology (n=63, 20.3%); and (8) focusing treatment on cancer origins and cause (n=43, 13.9%). Overall, 25.8% (n=80) of advertisements included a direct statement claiming provider treatment can cure cancer or prolong life.

Conclusions: Our results provide evidence alternative cancer providers are using Meta advertising products to market scientifically unsupported cancer treatments. Advertisements regularly referenced “alternative” and “natural” treatment approaches to cancer. Imagery and text content that emulated evidence-based medical providers created the impression that the offered treatments were effective medical options for cancer. Advertisements exploited the hope of patients with terminal and poor prognoses by sharing testimonials of past patients who allegedly were cured or had their lives prolonged. We recommend that Meta introduce a mandatory, human-led authorization process that is not reliant upon artificial intelligence for medical-related advertisers before giving advertising permissions. Further research should focus on the conflict of interest between social media platforms advertising products and public health.

(JMIR Infodemiology 2023;3:e43548) doi:10.2196/43548

KEYWORDS

cancer; advertising; misinformation; false hope; Meta; Facebook; Instagram; Messenger; social media; exploitation; infodemiology; cancer treatment; online health information

Introduction

Social media is both a valuable resource and a challenging arena for cancer patients and their families to navigate. Patients with cancer can find community [1,2], support [3,4], identity [5], and resources [6] across social media groups, pages, and forums. Social media also allows patients with cancer and their families to share updates and appeal for support within their networks [7,8]. Simultaneously, the internet contains widespread misinformation [9-14] about cancer, including its causes, evidence-based cancer treatments, and purported cancer treatments represented as efficacious when little, no, or disproven evidence exists for its use [15,16]. Nonetheless, content and articles with cancer misinformation shared on social media receive more engagement than factual sources [17].

Cancer misinformation reaches patients on social media and may have negative consequences, such as misinformed treatment decisions, worsened clinician-patient dynamics, and damaged caregiver-patient relationships [18,19]. In some cases, cancer misinformation can lead to patients with treatable or early-stage cancers opting out of evidence-based treatments in preference for alternative cancer treatments [20]. In other cases, patients with advanced cancers or limited treatment options may reasonably want to exhaust all options in search of a cure or to prolong life, leading them to try unproven, experimental, or alternative cancer therapies against their medical provider's recommendation [21,22]. Patients who distrust health care, lack health literacy, do not have their informational needs met [23-25], and those with lower educational attainment are the most susceptible to cancer misinformation [26]. Alternative cancer treatment in patients with treatable or terminal cancer is associated with a reduced time to death than in patients with cancer undergoing scientifically supported treatment [27,28].

Compounding the misinformation difficulties faced by patients with cancer, alternative cancer treatment providers are alleged to actively promote unproven, experimental, and potentially harmful treatments [29,30]. Promotion occurs through various mediums and strategies, including websites and social media. Facebook groups, which can support patients with cancer through community and shared experiences, are targeted by posts advertising alternative cancer treatments or products [31]. Providers make unsubstantiated health claims, share disinformation [32], and distort the scientific evidence supporting their services in promotional activities [15]. The marketing of cancer treatments, especially by alternative providers, is harmful in that it provides false hope, utilizes medical resources inappropriately, and disrupts doctor-patient relationships [33]. The US Food and Drug Administration (FDA) regularly issues warnings to companies and services promoting unproven cancer products and treatments. In 2018, the FDA warned 14 alternative providers for "fraudulently claiming to diagnose, treat, or cure cancer," with some selling and promoting their products on Facebook and Instagram [34]. However, warnings typically lead to limited negative consequences for providers.

While it is understood that advertising by alternative cancer providers is a source of harmful misinformation, an important area yet to be explored is how alternative cancer treatment providers utilize paid social media advertising products and tools to market their services. As opposed to other types of nondigital direct-to-consumer and nonpaid social media promotional activities or strategies [35-37] (eg, hosting a Facebook page without paid advertisements), targeted advertisements are uniquely effective at reaching specific groups via tailored messaging with little cost. Social media advertisers can target users in a certain age group, gender, geographic area, and income group, as well as people who demonstrate specific interests [38]. Advertisers can also employ advanced targeting features such as "custom" [39] or "lookalike" [40] audiences for further in-depth advertisement audience targeting. Applying social media–targeted advertising strategies for alternative cancer treatments may enable potential advertisers to target demographic groups fitting their target demographic profile, such as groups at a statistically higher risk of cancer or high-income earners. Targeted advertising may also enable advertisers to focus on demographics with "interests" or social activities demonstrating a higher likelihood of receptivity to their services (eg, "natural products"). Meta banned certain detailed terms (eg, "cancer") to target as interests on January 19, 2022 [41]. However, as an advertiser, it is still possible to target the followers of known proponents of alternative medicine, such as Gwyneth Paltrow [42]. In summary, social media advertisements can reach and track a large, defined audience with little investment and effort.

To prevent the misuse of advertising tools, social media platforms require advertisers' adherence to their platform-specific health-related advertising policies [43]. For example, Meta's advertising policy states that "ads must not contain deceptive, false or misleading claims…that set unrealistic expectations for users." Despite policies against misleading or harmful health advertising, Meta advertising tools promote scientifically unsupported public health messages and unproven products or services. Past research has found that paid Meta advertisements disseminated antivaccine [44] and protobacco content [45]. Patients with cancer have shared anecdotes of how they started to see advertisements for fake cancer cures after their diagnosis [18,29]. As recently as June 2022, paid advertisements for scientifically unsupported cancer treatments were reported on Meta platforms [46]. The current context suggests an immediate need to investigate the extent of alternative and unproven cancer treatment advertisements on Meta social platforms.

In this study, we partially address this need by systematically analyzing the advertising activities of prominent alternative cancer treatment practitioners on Meta platforms, including Facebook, Instagram, Messenger, and Audience Network. We specifically sought to determine (1) whether alternative cancer treatment paid advertising occurs on Meta social media platforms, (2) the strategies and messages alternative cancer providers use to reach and appeal to prospective patients, and (3) how the efficacy of alternative treatments is portrayed. Analyzing the advertising activities of alternative cancer treatment providers serves as a useful case study to examine Meta's advertising infrastructure and its role in the propagation of misinformed cancer treatments.

Methods

Identifying and Retrieving Advertisements

To identify alternative cancer advertisements, we searched the Meta Ad Library [47]—a publicly accessible database of current advertisements running on Facebook, Instagram, Messenger, or Audience Network—by well-known alternative cancer providers to determine if marketing was occurring. We identified prominent alternative providers from a patient directory of nontraditional cancer clinics [48] and treatment destinations identified from a study investigating alternative cancer treatment crowdfunding [22]. The first source, Heal Navigator, is a website specializing in providing information on alternative treatment clinics outside of conventional care options. We chose this source because prospective patients and their families may use similar directories when researching alternative care options. The second source was a research study that investigated the crowdfunding activities of patients with cancer seeking complementary and alternative cancer treatment options. The study developed a list of 110 complementary and alternative cancer treatments, searched each treatment with the term "cancer" on GoFundMe, and subsequently collected the frequency of specific treatments being crowdfunded and the names of each alternative cancer clinic where patients sought to receive treatment. We chose this source because it reflects a novel data source to understand where patients are seeking to receive alternative cancer treatment. We considered "alternative cancer treatments" to include any cancer-specific treatment that is not medically supported, disproven, unproven, experimental, or in an early stage of research outside of a registered clinical trial or provided by an oncology trial unit. We identified 86 clinics to search for evidence of marketing alternative cancer treatments on a Meta social media platform.

Between December 6, 2021, and December 12, 2021, we visited each clinic's unique advertising page daily on the Meta Ad Library and collected active advertisements. The information collected included the advertisement identification number, advertisement URL, date retrieved, active/inactive status, dates launched/run, advertiser page name, and a screenshot (image) or recording (if containing a video) of the advertisement. We collected 383 advertisements from 17 alternative cancer providers. To determine inclusion in the study, the first author (MZ) reviewed each advertisement to determine if the advertisement directly or indirectly offers an alternative, experimental, disproven, or unproven cancer treatment or approach to prospective patients with cancer through a paid Meta product advertisement. TC reviewed 50% of the inclusion decisions to ensure consistency in the inclusion criteria application. In total, we marked 310 advertisements for inclusion.

Ethical Considerations

This study did not require ethics approval because all data collected were publicly available.

Content Analysis

We conducted a content analysis [49] to analyze how alternative cancer providers communicate the benefits of their services through advertisements on Meta platforms. Content analysis has been used to study cancer content on numerous social media platforms [50-53] and is useful to observe, systematically categorize, and quantify communication message strategies and characteristics [54]. Authors MZ, JS, JCBP, and TC independently reviewed between 77 and 78 (25%) advertisements and met to determine pattern observations and identify key thematic frames. The authors developed an initial coding frame, and MZ test coded the advertisements. MZ then consulted with authors NM and MvS for their input into the coding frame. After minor modifications and similar code grouping, MZ coded the advertisements on the mixed methods software analysis program Dedoose (University of California, Los Angeles). We identified 8 advertising strategies. We also coded for the treatments mentioned and evaluated how alternative cancer providers portrayed treatment efficacy. When assessing efficacy representation, we chose to have another author review each statement for application consistency due to potential subjective interpretations of being cured or having one's life prolonged. Author MvS reviewed efficacy statement coding decisions and agreed on 93% of efficacy code applications. Authors MZ and MvS then resolved disagreements through open discussion.

Results

We identified 310 paid advertisements from 11 alternative cancer clinics on Meta (Facebook, Instagram, or Messenger) marketing alternative treatment approaches, care, and interventions. The clinic profiles of those hosting advertisements are summarized in Table 1. The clinics found in our study were in the United States (n=4, 36.4%), Mexico (n=4, 36.4%), Spain (n=2, 18.2%), and Thailand (n=1, 9.1%). Clinics may offer services in multiple locations. An expanded table detailing the treatments offered at each clinic and their treatment provider qualifications according to their websites is available in Multimedia Appendix 1.

Table 1. Clinic profile overview of alternative cancer treatment providers.

Clinic name	Total advertisements, n (%)	Location
Brio-Medical	146 (47.1)	Scottsdale, Arizona
Conners Clinic	44 (14.2)	Lake Elmo, Minnesota
CHIPSA[a] Hospital	34 (11)	Tijuana, Mexico
Verita Life	23 (7.4)	Bangkok, Thailand
Budwig Center	14 (4.5)	Málaga, Spain
Immucura	14 (4.5)	Málaga, Spain
Hope4Cancer Treatment Centers	12 (3.9)	Tijuana, Mexico
Immunity Therapy Center	12 (3.9)	Tijuana, Mexico
Envita Medical Centers	6 (1.9)	Scottsdale, Arizona
Dayspring Cancer Clinic	4 (1.3)	Scottsdale, Arizona
Issels Immuno-Oncology	1 (0.3)	Tijuana, Mexico

[a]CHIPSA: Centro Hospitalario Internacional del Pacifico, SA.

Nearly all (n=289, 93.2%) advertisements featured imagery or text signifying that the provider is a qualified medical expert and may legitimately advise on and administer cancer treatment. Visual cues included images or text mentioning qualified health care providers (eg, doctors, surgeons), reference to interventions (treatment, medications, intravenous administration, therapies), medical imagery and equipment, and labeling the provider location as a "medical treatment center," "clinic," or related terms. Many clinics had staff providers with credentials that were not associated with expertise in primary cancer care or were legally barred from recommending primary cancer treatment. This included naturopaths, chiropractors, and other alternative medicine practitioners. Despite representation as a legitimate medical option, certain providers' websites specify that they do not offer medical advice. Figure 1 displays illustrative examples of clinics presenting themselves as qualified cancer care and treatment providers. Figure 2 depicts an advertisement from Conners Clinic where the primary service provider refers to himself as "Dr" in a medical context giving cancer treatment advice. However, according to the Conners Clinic website [55,56], he practices under a "Pastoral Medical License" and does not offer medical advice.

Figure 1. Advertisements depicting alternative cancer treatment provider is qualified to advise and administer cancer treatment.

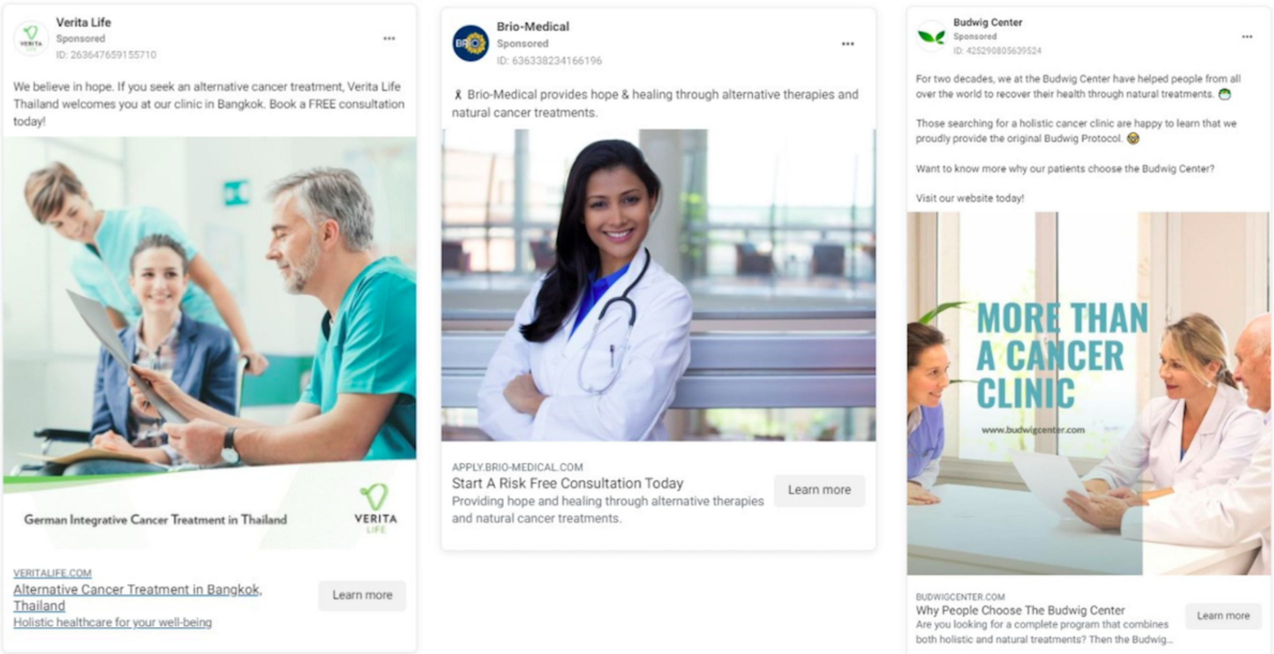
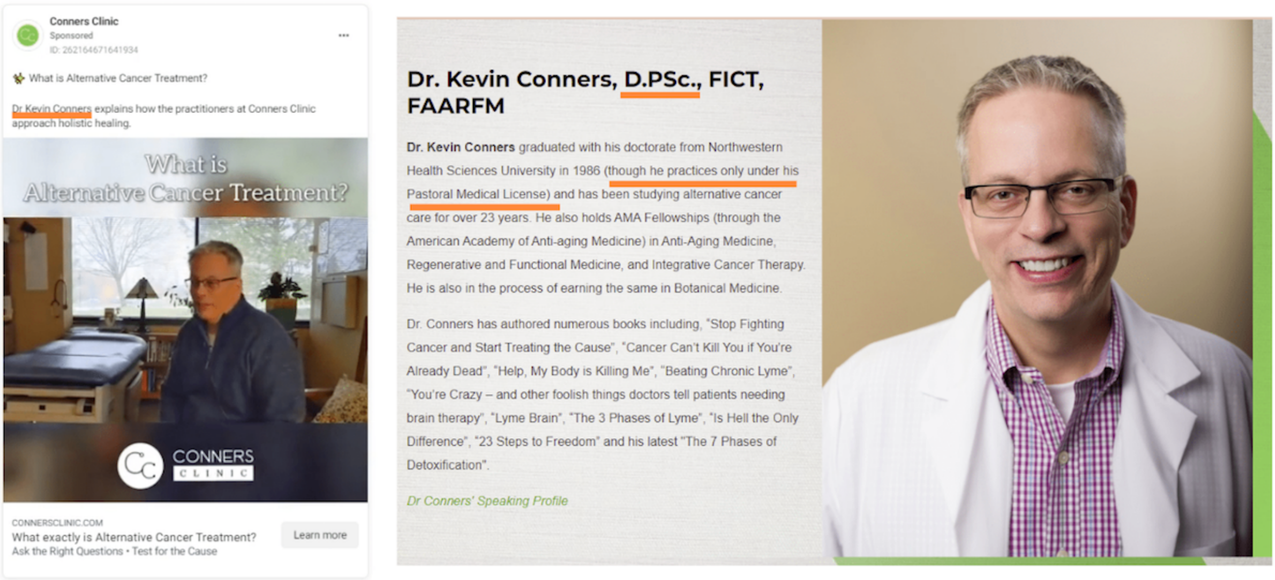


Figure 2. An example of an alternative cancer treatment advertisement depicting that the provider is qualified to advise cancer treatment with a website screenshot of provider qualifications.



In 65.5% (n=203) of the advertisements, providers appealed to persons with limited treatment options due to an advanced, aggressive, or terminal cancer prognosis to not give up seeking treatment because other "effective" options exist. Clinic advertisements invoked skepticism regarding noncurable cancer cases and gave examples of their alleged "success" in treating or curing terminal cancer cases. A CHIPSA (Centro Hospitalario Internacional del Pacifico, SA) hospital advertisement states their client "was told it [their cancer] was noncurative. But nearly 2 years after her initial diagnosis, and treatment at CHIPSA, she is cancer free." Many advertisements invoked direct skepticism toward other health providers labeling a patient's cancer as noncurative. In other cases, clinics offered examples where past clients were allegedly abandoned by their medical teams once their cancer reached an advanced or noncurative form. For example, an advertisement states, "She [patient] was 'dropped' by her doctors, put on hospice, and given only months to live. [Patient] and her husband [names redacted] refused this death sentence and ventured to CHIPSA Hospital in Mexico." Illustrative screenshots are shown in Figure 3.

Across the advertisements, 54.2% (n=168) featured 1 or more people with cancer who received treatment from a provider and spoke about their experience, either about the impact of alternative treatment on their cancer diagnosis or their experience with the advertiser. Figure 4 displays examples of advertisements depicting supposed clients speaking to services received as improving or curing their cancer. Many contain specific references to being cancer-free after receiving treatment with an alternative provider. An advertisement from Envita Medical Centres includes a statement from a person depicted as a patient stating, "I came in here with stage 4 colorectal cancer, [and] I'm leaving cancer free." Another advertisement reads, "My oncologist didn't believe it was possible to cure my cancer. Thanks to Immunity Therapy Center, I proved him wrong!"

Alternative cancer providers marketed holistic approaches to healing in 39% (n=121) of advertisements, including emotional health, addressing trauma, and other factors impacting a person's ability to treat and fight their cancer. Providers emphasized incorporating psychological wellness into their treatments. For example, a Budwig Medical Centre advertisement states, "It is a treatment for your physical body, but it also a treatment for your soul—it is an emotional and psychological treatment."

Approximately 26.1% (n=81) of advertisements featured language conveying care about their patients' well-being, often emphasizing the relationship they want/do build with their patients. For example, an Immunity Therapy Centre advertisement states, "Our knowledgeable and loving team invests time in developing relationships that bless everyone involved." Other promotions highlight apparent vulnerable, caring moments between staff and patients. A Brio-Medical advertisement reads, "Dr Larry was there [while patient crying], and he hugged me, and I knew after that it was going to be great." Last, advertisements emphasize treating clients not just as another case. A Budwig Centre ad states: "You are not a chart or diagnosis—you are an individual who deserves the absolute best care."

Providers sought to support the effectiveness or legitimacy of their treatments or approach by referencing rhetoric or imagery related to science, research, evidence, and well-known science-related organizations or institutions in 23.2% (n=72) of advertisements. Cues for coding included the terms "research-based," "Harvard medical," "NASA," "new research," "Nobel prize," "proven," "published," "scientific evidence," "researched," "scientifically proven," and related terms. Here, providers gave little to no details about the research mentioned and included images of cells or other biological processes with no context (see Figure 5). In many cases, unproven, disproven, or experimental treatments were represented as being supported by research. For example, Brio-Medical states, "Researchers are using vitamin C and oxygen to kill cancer." Advertisements also included misleading statements about the research quality or implications for specific treatments.

In 20.3% (n=63) of the advertisements, providers represented themselves as keeping up to date and offering the latest technological advances in their cancer treatments and approach. Providers used terminology signaling major innovation, including “groundbreaking,” “breakthrough,” “new,” “paradigm shift,” “state of the art,” and related terms. In some cases, the clinic’s latest advanced technology was used as an appeal. For example, an advertisement from Conners Clinic links to the clinic founder’s “groundbreaking book” on treating cancer. In other cases, advertisements mention the facilities as a “state-of-the-art center.”

Messaging stating that the key to treating cancer is understanding why it developed in the first place was observed in 13.9% (n=43) of advertisements. Here, clinics argue that treating cancer requires identifying and removing the reasons leading to cancer development. For example, an advertisement from Brio-Medical states, “Stop fighting cancer and address the cause by asking why your body is sick.” Most often, clinics recommend making certain lifestyle or diet changes to prevent reoccurrence and promote healing. For example, Conners Clinic recommends a 4-pronged treatment for healing cancer that consists of “cause, nutrition, technology, diet, and detoxification.”

Figure 3. Advertisements appealing to persons with limited or no treatment options due to an advanced or terminal cancer prognosis.

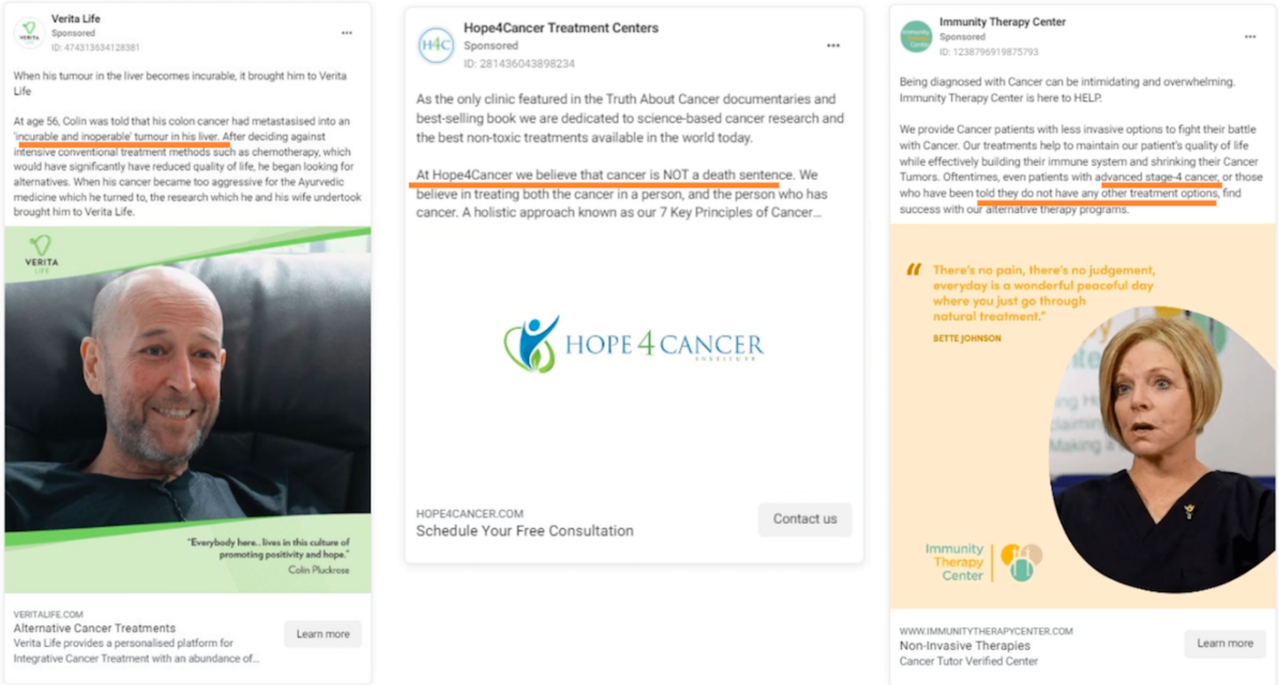


Figure 4. Advertisements featuring testimonials of past clients allegedly cured of cancer.

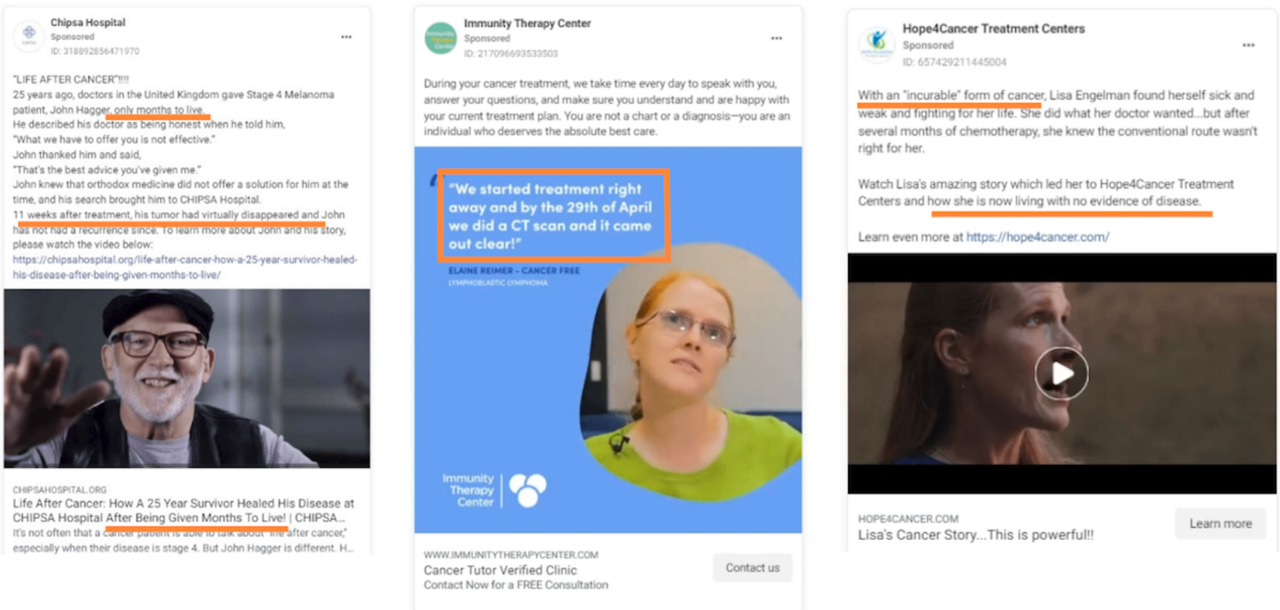
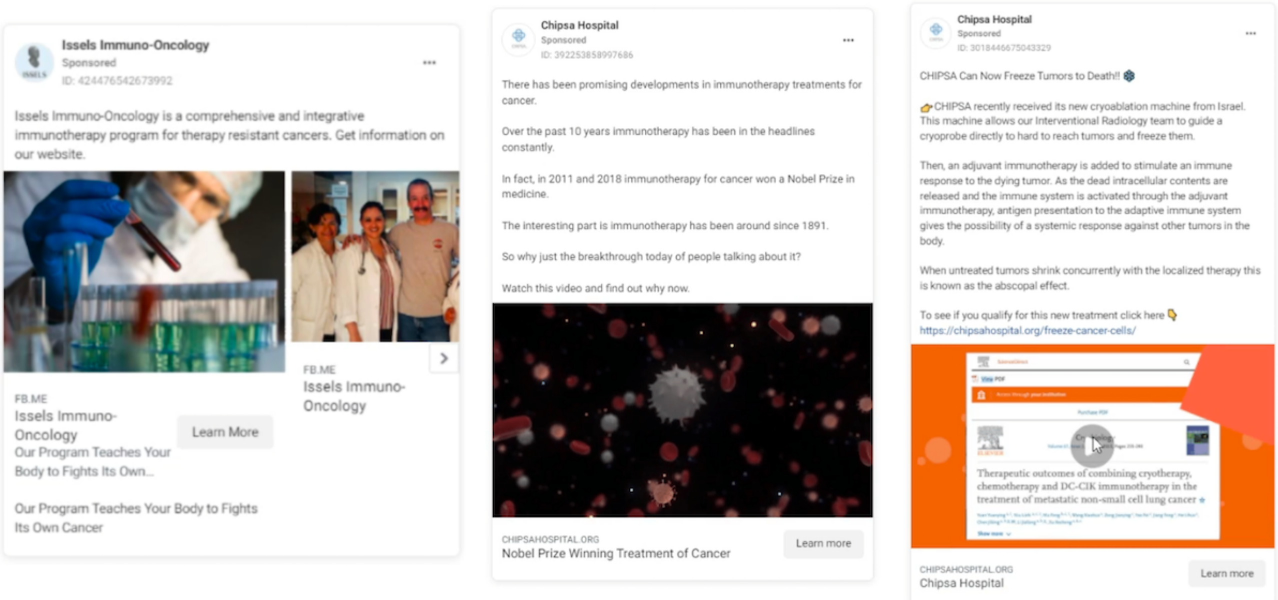


Figure 5. Advertisements referencing provider treatments rooted in science and research.



Across the 310 advertisements analyzed, 25.8% (n=80) included 1 or more direct statements signifying the offered cancer treatments at their facility are effective for curative or life-prolonging purposes or that the treatment offered has cured or prolonged life in patients. Example excerpts of these direct statements are included in Textbox 1; note that these are raw text and have not been edited for grammar. We found 78 cancer treatments mentioned in advertisements (Multimedia Appendix 2). The most mentioned treatments or approaches were alternative (n=191, 61.6%) and natural (n=153, 49.4%). Many clinics do not advertise the full range of treatments they offer.

Textbox 1. Example statements signifying that the offered cancer treatment is effective for curative or life-prolonging purposes, or treatment offered has cured or prolonged life in patients.

Sample excerpts

- “From hospice to healed! CHIPSA saves another cancer patient.”
- “It really was just about the 2-week mark where I really had noticeable improvement in how I felt, and my breast lump started shrinking so that was pretty amazing.”
- “Craig was diagnosed with colorectal cancer and came to the Budwig Center in August 2014 to receive treatment pursue the natural approach. Just a year later, in May 2015, the doctors shared with some good news: his cancer had totally disappeared.”
- “Eight years later: Bailey O’Brien shares how she be terminal melanoma at CHIPSA.”
- “Aaron’s stage IV glioblastoma survivor story.”
- “My oncologist didn’t believe It was possible to cure my cancer, thanks to Immunity Therapy Center I proved him wrong!”
- “11 weeks after treatment, his tumor had virtually disappeared and John has not had a recurrence since.”
- “But nearly two years after her initial diagnosis, and treatment at CHIPSA, Amanda is still alive to share her story, and remarkably, she’s cancer free!”
- “Rebecca’s battle with thyroid cancer led her to seek a more integrative approach. She found Verita Life Thailand. Following treatment at our clinic in Bangkok, today, she is cancer-free.”
- “How Michelle overcame breast cancer with immunotherapy based on dendritic cells: ‘I’ve been getting treatments for about a month and there is no evidence of the tumour whatsoever.’”
- “Envita totally saved my life.”
- “I stayed the full 6 weeks just to get all the good therapies and it took me to a place of being cancer free.”
- I came in here with stage 4 colorectal cancer, I’m leaving cancer free.”
- “Find out like I did yesterday that my tumor is gone.”

Discussion

Principal Findings

Our results provide evidence that alternative cancer providers are using Meta products to advertise alternative cancer treatments to social media users. Advertisements regularly referenced “alternative” and “natural” approaches to cancer treatments. Imagery and text content emulated evidence-based medical providers and created the impression the treatments were legitimate medical options for cancer. Similarly, advertisements exploited the hope [57] of patients with terminal cancer and poor prognoses by sharing testimonials of past patients who allegedly were cured, had their lives prolonged, or had their quality of life improved. Providers framed their services as filling a gap once conventional medicine runs out of treatment options and sought to differentiate themselves from evidence-based medical providers who delivered a terminal diagnosis by undermining the efficacy of their administered cancer treatments (eg, radiation, chemotherapy) and their care and compassion for their patients.

Providers appealed to prospective patients with cancer through “scienceploitation” [58], which occurs “when popular scientific ideas…are used to take advantage of the social capital associated with them and induce consumer interest in products or services” and can “create misunderstandings and/or posits false connections” [59]. Providers shared narratives of their clinics offering breakthrough, advanced, and scientifically supported services outside the traditional medical scope. In other cases, providers conveyed information about promising scientific treatments, such as immunotherapy, but did not contextualize the inability of the clinic to properly manufacture, administer, and monitor such advanced treatments or correctly explain its evidence base [60]. We also identified scientific language and imagery used in an effort to legitimize unproven therapies and approaches. References and imagery of research, science, specific studies, or experiments in advertisements may distort the viewer’s assessment of how medically accepted the ideas are to which the advertisements were referring. This, in turn, leads to an unfounded belief in the likelihood of treatment success and unnecessary financial and time expenditure.

Meta advertising tools enable alternative clinics to promote and at some level target their advertisements to people with cancer. Prior studies demonstrate how established platform features and tools (groups, timelines, sharing posts) are employed by users and providers to purposefully or inadvertently spread cancer misinformation [4,11,17,61-63]. In difference from such studies, we demonstrate an active element in social media platforms spreading and profiting from misleading medical information. Meta platforms approve advertisements [64], provide targeting options, and earn direct revenue from advertisements. When unproven cancer advertising is found, Meta publicly frames the advertisements against their policies, removes the advertisements, and details interventions to minimize or prevent health misinformation [46,65-67]. Despite removal, alternative cancer treatments can still create new advertisements with disproven claims and use targeting tools. Our results suggest that the case-by-case ad removal after media or user reporting [68] and overreliance upon artificial intelligence by Meta have not addressed nor will be able to address the problem.

Currently, Meta requires an authorization process, “written permissions,” or application procedures for select advertisements (ie, prescription drug advertising, addiction treatment, cryptocurrency, social issues, elections, politics, online pharmacies, online gambling and gaming, and dating) [38]. Expanding the authorization processes to all medical advertisements could potentially limit the dissemination of misleading or exploitative medical advertising identified in this paper. Approval processes should not rely on artificial intelligence tools [69] but instead, be coordinated by qualified medical professionals. Regular audits of approved medical advertisers would likely be necessary to ensure compliance. Strong disincentives, such as banning and reporting advertisers who violate legal and platform policies, may also help limit this harmful practice. Cross-border advertising tools and the reach of advertisements create difficulties [70-73] for country-specific regulatory adherence and enforcement, positioning Meta as the only party with the competency and capability to efficiently police advertisements.

In providing public health recommendations for Meta, the power dynamics between public health researchers and social media platforms must be made transparent and discussed extensively. While we believe these aforementioned recommendations would be effective, they are framed and scaled to what national public health systems have the authority to intervene upon and what is likely to be accepted by Meta [74]. Although a growing body of literature provides recommendations for Meta and other social media platforms to improve public health, we argue that it is important to acknowledge that these proposals likely will not be pursued if they adversely impact social media platform interests or business models [75,76]. With the little power public health researchers and practitioners are availed to change social media policies and processes, recommendations to social media businesses such as Meta are created to appeal to the good nature of platforms or make a case that our suggestions are beneficial to their interests. In most other contexts, appealing to or working with a for-profit industry to improve health in ways against their financial interests is not effective [77] and can hurt public health interests [78], even if case-by-case gains are achieved. This context is emblematic of a larger power dynamic in how social media businesses reinforce their political power, acting as both infrastructure and advertiser, thus both judge and interested party [79,80].

Fully acting upon the issue of misleading advertisements requires examining and confronting the conflict of interest between social media business interests and public health [81]. In the case of misleading health advertisements, this is only a single symptom of a larger infrastructure in pursuit of profit [82,83], and it is at odds with public health objectives. Meta, like most social media businesses, relies on advertisements for revenue. Many advertisements hosted by Meta are harmful to public health or cause direct harm, including those promoting health-harming products [84], dis/misinformation [85], hateful speech [86,87], and other content types. Advertising tools allow invasive targeting [88] for products or messages using data that many users may not know are collected [89] or sold. However, the public health response [90], and indeed Meta's response, is to accept this system as a status quo and seek ways to improve it incrementally while not recognizing or acknowledging that the business model itself is harmful [91]. It is important to understand the shared responsibility between advertisers and social media platforms, both of whom benefit greatly from deceptive advertising being relayed to the public. This calls for political courage and the use of effective means to avoid such harmful practices.

Limitations

This study has several limitations. First, the advertisements collected are only a brief snapshot of the advertising of unproven cancer treatments across Meta platforms. The search strategy attempted to identify the most well-known clinics administering unproven medicine; therefore, our results likely undercount the true scale of unproven cancer treatment advertising. The advertisements and clinics identified are also geared toward English-speaking audiences located in North America. Next, we cannot objectively state the testimonial content seen in this study is untrue or that specific cases of cancer were not cured or improved. However, the marketing of curative and life-prolonging testimonials for scientifically unsupported treatment is still dangerous because it provides false hope to patients with advanced or terminal cancer. This study employed a single-coder approach, which may have subjected the data set to the interpretative bias of the coder. However, we took several steps to mitigate this, including cocreating a defined coding frame, test coding, team discussion, and auditing categories with perceived subjectiveness, such as advertising claims of being cured or having life prolonged. Finally, the Meta Ad Library does not provide advertisement viewership data (reach, demographics), advertisement targeting details, conversions, or financial spending information. Thus, we cannot speculate on the viewership impact of the specific advertisements in our sample.

Conclusion

In this study, we found alternative health providers advertise scientifically unsupported cancer treatments and approaches through paid advertising products on Meta platforms. Advertisements contained 8 distinct strategies to appeal to viewers: advertiser representation as a legitimate medical provider, appealing to persons with limited treatments options, client testimonials, promoting holistic approaches, rhetoric related to science and research, rhetoric pertaining to the latest technology, and focusing treatments on cancer origins and cause. Among the advertisements, 25.8% (n=80) included a direct statement claiming that their treatment can cure or prolong life. The dissemination of advertising poses a serious concern to public health, which may spread misinformation, distrust in evidence-based health care, exploitation of vulnerable groups, unnecessary financial expenditure on unproven treatments, and disengagement from evidence-based cancer treatments. This study also illustrates how Meta advertising tools promote unproven medical therapies and the inadequacy of existing deterrents to prevent misleading medical advertisements. We recommend that Meta introduce a mandatory, human-led authorization process for medically related advertisers before receiving advertising permissions. As social media platforms have historically failed to fully act on such recommendations, we also suggest public health policies be enacted to compel social media companies to better monitor and remove problematic advertisements and ban advertising from companies and individuals with a history of spreading misinformation. Further research should consider an enhanced focus on the conflict of interest between social media platforms advertising products and public health and better characterize the nature and scale of the harm caused by such targeted advertisements.

Acknowledgments

The authors would like to thank the Public Health Agency of Canada and the Canadian Institutes of Health Research for their generous support of ScienceUpFirst.

Data Availability

The coded data and media files are available upon reasonable request to the corresponding author.

Authors' Contributions

MZ conceptualized the manuscript. MZ, JS, J-CB-P, and TC contributed to the methodology by reviewing the data source and creating the data collection procedure. MZ collected the data. All authors provided input into formal analysis and result interpretation. MZ wrote the manuscript. All authors edited and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Clinic profile overview of alternative cancer treatment providers.

[DOCX File , 21 KB - infodemiology_v3i1e43548_app1.docx]

Multimedia Appendix 2
Frequency of specific treatments mentioned or displayed in advertisements.
[DOCX File , 15 KB - infodemiology_v3i1e43548_app2.docx]

References

1. Harkin LJ, Beaver K, Dey P, Choong K. Navigating cancer using online communities: a grounded theory of survivor and family experiences. J Cancer Surviv 2017 Dec;11(6):658-669 [FREE Full text] [doi: 10.1007/s11764-017-0616-1] [Medline: 28470506]
2. Hale BJ, Gonzales AL, Richardson M. Vlogging cancer: predictors of social support in YouTube cancer vlogs. Cyberpsychol Behav Soc Netw 2018 Sep;21(9):575-581. [doi: 10.1089/cyber.2018.0176] [Medline: 30132690]
3. Lazard AJ, Collins MKR, Hedrick A, Varma T, Love B, Valle CG, et al. Using social media for peer-to-peer cancer support: interviews with young adults with cancer. JMIR Cancer 2021 Sep 02;7(3):e28234 [FREE Full text] [doi: 10.2196/28234] [Medline: 34473063]
4. Petukhova T, Wilson B, Gadjiko M. Utilization of Facebook for supporteducation by patients with skin cancer. Dermatol Online J 2020 Mar 3;26(3):2. [doi: 10.5070/d3263047973] [Medline: 32609441]
5. Daniels SR, Yang C, Toohey SJ, Willard VW. Perspectives on social media from adolescents and young adults with cancer. J Pediatr Oncol Nurs 2021 Mar 17;38(4):225-232. [doi: 10.1177/1043454221992319] [Medline: 33729902]
6. Pellino G, Simillis C, Qiu S, Rasheed S, Mills S, Warren O, et al. Social media and colorectal cancer: A systematic review of available resources. PLoS One 2017 Aug 23;12(8):e0183031 [FREE Full text] [doi: 10.1371/journal.pone.0183031] [Medline: 28832603]
7. Bloom RD, Reblin M, Chou WS, Beck SL, Wilson A, Ellington L. Online social support for cancer caregivers: alignment between requests and offers on CaringBridge. J Psychosoc Oncol 2021 Aug 24;39(1):118-134. [doi: 10.1080/07347332.2020.1806174] [Medline: 32835649]
8. Snyder J, Zenone M, Crooks V, Schuurman N. What medical crowdfunding campaigns can tell us about local health system gaps and deficiencies: exploratory analysis of British Columbia, Canada. J Med Internet Res 2020 May 22;22(5):e16982 [FREE Full text] [doi: 10.2196/16982] [Medline: 32348269]
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. Warner EL, Waters AR, Cloyes KG, Ellington L, Kirchhoff AC. Young adult cancer caregivers' exposure to cancer misinformation on social media. Cancer 2021 Apr 15;127(8):1318-1324. [doi: 10.1002/cncr.33380] [Medline: 33368215]
11. Wilner T, Holton A. Breast cancer prevention and treatment: misinformation on Pinterest, 2018. Am J Public Health 2020 Oct;110(S3):S300-S304. [doi: 10.2105/ajph.2020.305812]
12. Loeb S, Reines K, Abu-Salha Y, French W, Butaney M, Macaluso JN, et al. Quality of bladder cancer information on YouTube. Eur Urol 2021 Jan;79(1):56-59. [doi: 10.1016/j.eururo.2020.09.014] [Medline: 33010986]
13. Xu AJ, Taylor J, Gao T, Mihalcea R, Perez-Rosas V, Loeb S. TikTok and prostate cancer: misinformation and quality of information using validated questionnaires. BJU Int 2021 Oct 21;128(4):435-437. [doi: 10.1111/bju.15403] [Medline: 33811424]
14. Teplinsky E, Ponce SB, Drake EK, Garcia AM, Loeb S, van Londen G, et al. Online medical misinformation in cancer: distinguishing fact from fiction. JCO Oncology Practice 2022 Aug;18(8):584-589. [doi: 10.1200/op.21.00764]
15. Delgado-López PD, Corrales-García EM. Influence of internet and social media in the promotion of alternative oncology, cancer quackery, and the predatory publishing phenomenon. Cureus 2018 May 13;10(5):e2617 [FREE Full text] [doi: 10.7759/cureus.2617] [Medline: 30027009]
16. Mazzocut M, Truccolo I, Antonini M, Rinaldi F, Omero P, Ferrarin E, et al. Web conversations about complementary and alternative medicines and cancer: content and sentiment analysis. J Med Internet Res 2016;18(6):e120 [FREE Full text] [doi: 10.2196/jmir.5521] [Medline: 27311444]
17. 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 2021 Jul 22. [doi: 10.1093/jnci/djab141] [Medline: 34291289]
18. King A. When hope kills: social media's false promises to cancer patients. Healthy Debate. 2021 Sep 9. URL: https://healthydebate.ca/2021/08/topic/social-media-fake-cancer-treatment/ [accessed 2022-08-07]
19. The challenges of cancer misinformation on social media. National Cancer Institute. 2021 Sep 21. URL: https://www.cancer.gov/news-events/cancer-currents-blog/2021/cancer-misinformation-social-media [accessed 2022-08-15]
20. Zenone M, Snyder J, Caulfield T. Crowdfunding cannabidiol (CBD) for cancer: hype and misinformation on GoFundMe. Am J Public Health 2020 Oct;110(S3):S294-S299. [doi: 10.2105/ajph.2020.305768]
21. Zenone MA, Snyder J, Crooks VA. What are the informational pathways that shape people's use of cannabidiol for medical purposes? J Cannabis Res 2021 May 06;3(1):13 [FREE Full text] [doi: 10.1186/s42238-021-00069-x] [Medline: 33957993]
22. Snyder J, Zenone M, Caulfield T. Crowdfunding for complementary and alternative medicine: What are cancer patients seeking? PLoS One 2020 Nov 20;15(11):e0242048 [FREE Full text] [doi: 10.1371/journal.pone.0242048] [Medline: 33216790]

23. Ginossar T, Shah SFA, West AJ, Bentley JM, Caburnay CA, Kreuter MW, et al. Content, usability, and utilization of plain language in breast cancer mobile phone apps: a systematic analysis. JMIR Mhealth Uhealth 2017 Mar 13;5(3):e20 [FREE Full text] [doi: 10.2196/mhealth.7073] [Medline: 28288954]
24. Wang X, Shi J, Kong H. Online health information seeking: a review and meta-analysis. Health Commun 2021 Sep;36(10):1163-1175. [doi: 10.1080/10410236.2020.1748829] [Medline: 32290679]
25. Zhao Y, Zhang J. Consumer health information seeking in social media: a literature review. Health Info Libr J 2017 Dec;34(4):268-283. [doi: 10.1111/hir.12192] [Medline: 29045011]
26. Scherer LD, McPhetres J, Pennycook G, Kempe A, Allen LA, Knoepke CE, et al. Who is susceptible to online health misinformation? A test of four psychosocial hypotheses. Health Psychol 2021 Apr;40(4):274-284. [doi: 10.1037/hea0000978] [Medline: 33646806]
27. Johnson S, Park H, Gross C, Yu J. Use of alternative medicine for cancer and its impact on survival. J Natl Cancer Inst 2018 Jan 01;110(1):121-124. [doi: 10.1093/jnci/djx145] [Medline: 28922780]
28. Johnson SB, Park HS, Gross CP, Yu JB. Complementary medicine, refusal of conventional cancer therapy, and survival among patients with curable cancers. JAMA Oncol 2018 Oct 01;4(10):1375-1381. [doi: 10.1001/jamaoncol.2018.2487] [Medline: 30027204]
29. King A. I have cancer. Now my Facebook feed is full of 'alternative care' ads. New York Times. 2020 Jul 10. URL: https://www.nytimes.com/2020/07/10/opinion/facebook-cancer-ads.html [accessed 2022-07-26]
30. Ohlheiser A. They turn to Facebook and YouTube to find a cure for cancer and get sucked into a world of bogus medicine. Washington Post. 2019 May 19. URL: https://tinyurl.com/6hs4n9n6 [accessed 2022-07-26]
31. Sissung TM, Figg WD. Facebook groups for alternative treatments for cancer: advertising masquerading as community support. Lancet Oncol 2021 Jan;22(1):25-26. [doi: 10.1016/s1470-2045(20)30457-5]
32. The Lancet Oncology. Oncology, "fake" news, and legal liability. Lancet Oncol 2018 Sep;19(9):1135. [doi: 10.1016/S1470-2045(18)30610-7] [Medline: 30191835]
33. Hlubocky FJ, McFarland DF, Spears PA, Smith L, Patten B, Peppercorn J, et al. Direct-to-consumer advertising for cancer centers and institutes: ethical dilemmas and practical implications. Am Soc Clin Oncol Educ Book 2020 May(40):e207-e217. [doi: 10.1200/edbk_279963]
34. McCarthy M. US targets companies touting unproven cancer cures. BMJ 2017 Apr 28;357:j2116. [doi: 10.1136/bmj.j2116] [Medline: 28455380]
35. Bélisle-Pipon J, Williams-Jones B. Regulating direct-to-consumer drug information: a case study of Eli Lilly's Canadian 40over40 erectile dysfunction campaign. Healthc Policy 2015 May 1;10(4):16-22. [doi: 10.12927/hcpol.2015.24209]
36. Bélisle-Pipon JC, Williams-Jones B. Drug familiarization and therapeutic misconception via direct-to-consumer information. J Bioeth Inq 2015 Jun 12;12(2):259-267. [doi: 10.1007/s11673-015-9634-8] [Medline: 25963773]
37. Pipon JB, Williams-Jones B. Preparing for the arrival of "pink Viagra": strengthening Canadian direct-to-consumer information regulations. Can Med Assoc J 2016 Mar 15;188(5):319-320 [FREE Full text] [doi: 10.1503/cmaj.150705] [Medline: 26668190]
38. Facebook advertising targeting options. Meta. URL: https://en-gb.facebook.com/business/ads/ad-targeting [accessed 2021-05-17]
39. Reach existing customers using Facebook ads. Meta. URL: https://en-gb.facebook.com/business/goals/retargeting [accessed 2022-08-08]
40. How to use Ads Manager to create custom or lookalike audiences. Meta. URL: https://en-gb.facebook.com/business/learn/lessons/using-ads-manager-to-create-custom-lookalike-audiences [accessed 2021-07-28]
41. Removing certain ad targeting options and expanding our ad controls. Meta. URL: https://www.facebook.com/business/news/removing-certain-ad-targeting-options-and-expanding-our-ad-controls [accessed 2022-08-09]
42. Caulfield T. Is Gwyneth Paltrow Wrong About Everything? When Celebrity Culture and Science Clash. New York, NY: Viking; 2015.
43. Facebook Inc profile: summary. Open Secrets. URL: https://www.opensecrets.org/orgs/facebook-inc/summary?cycle=A&id=D000033563 [accessed 2021-05-27]
44. Jamison AM, Broniatowski DA, Dredze M, Wood-Doughty Z, Khan D, Quinn SC. Vaccine-related advertising in the Facebook Ad Archive. Vaccine 2020 Jan 16;38(3):512-520 [FREE Full text] [doi: 10.1016/j.vaccine.2019.10.066] [Medline: 31732327]
45. Majmundar A, Moran M. For or against tobacco control: sponsored tobacco advocacy messages on Facebook and Instagram. Nicotine Tob Res 2021 Oct 07;23(11):1861-1868. [doi: 10.1093/ntr/ntab111] [Medline: 34049405]
46. Leffer L. Meta allowed misleading ads for cancer treatments on Facebook. Gizmodo. 2022 Jun 27. URL: https://gizmodo.com/facebook-meta-misleading-ads-cancer-treatment-1849112333 [accessed 2022-08-08]
47. Ad Library. Meta. URL: https://www.facebook.com/ads/library/?active_status=all&ad_type=political_and_issue_ads&country=CA&media_type=all [accessed 2021-05-09]
48. Heal Navigator. URL: https://healnavigator.com/ [accessed 2021-08-08]
49. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]

50. Thomas J, Zheng R, Prabhu A, Heron D, Beriwal S. Content analysis of posts about cancer on the social media website Reddit. Int J Radiat 2019 Sep;105(1):E461-E462. [doi: 10.1016/j.ijrobp.2019.06.1420]
51. Chen L, Wang P, Ma X, Wang X. Cancer communication and user engagement on Chinese social media: content analysis and topic modeling study. J Med Internet Res 2021 Nov 10;23(11):e26310 [FREE Full text] [doi: 10.2196/26310] [Medline: 34757320]
52. Nastasi A, Bryant T, Canner JK, Dredze M, Camp MS, Nagarajan N. Breast cancer screening and social media: a content analysis of evidence use and guideline opinions on Twitter. J Cancer Educ 2017 Jan 17. [doi: 10.1007/s13187-017-1168-9] [Medline: 28097527]
53. Sunsuwan W, Bejrananda T. A content analysis of YouTube videos related to bladder cancer. PSU Med J 2021 Mar 12;1(1):3-7. [doi: 10.31584/psumj.2021245369]
54. Maier M. Content analysis: advantages and disadvantages. In: The SAGE Encyclopedia of Communication Research Methods. London, UK: SAGE Publications Inc; 2017.
55. Meet your team. Conners Clinic. 2022. URL: https://www.connersclinic.com/meet-your-team/ [accessed 2022-03-12]
56. Conners Clinic. URL: https://www.connersclinic.com/ [accessed 2022-03-11]
57. Snyder J. Exploiting Hope: How the Promise of New Medical Interventions Sustains Us--and Makes Us Vulnerable. Oxford, UK: Oxford University Press; 2020.
58. Caulfield T. Blinded by science. The Walrus. URL: https://thewalrus.ca/blinded-by-science/ [accessed 2022-08-09]
59. Murdoch B, Zarzeczny A, Caulfield T. Exploiting science? A systematic analysis of complementary and alternative medicine clinic websites' marketing of stem cell therapies. BMJ Open 2018 Feb 28;8(2):e019414 [FREE Full text] [doi: 10.1136/bmjopen-2017-019414] [Medline: 29490963]
60. Marth C, Wieser V, Tsibulak I, Zeimet AG. Immunotherapy in ovarian cancer: fake news or the real deal? Int J Gynecol Cancer 2019 Jan 07;29(1):201-211. [doi: 10.1136/ijgc-2018-000011] [Medline: 30640705]
61. Xu AJ, Myrie A, Taylor JI, Matulewicz R, Gao T, Pérez-Rosas V, et al. Instagram and prostate cancer: using validated instruments to assess the quality of information on social media. Prostate Cancer Prostatic Dis 2022 Apr 01;25(4):791-793. [doi: 10.1038/s41391-021-00473-7] [Medline: 34853412]
62. Shi S, Brant A, Sabolch A, Pollom E. False news of a cannabis cancer cure. Cureus 2019 Jan 19;11(1):e3918 [FREE Full text] [doi: 10.7759/cureus.3918] [Medline: 30931189]
63. Massey PM, Kearney MD, Hauer MK, Selvan P, Koku E, Leader AE. Dimensions of misinformation about the HPV vaccine on Instagram: content and network analysis of social media characteristics. J Med Internet Res 2020 Dec 03;22(12):e21451 [FREE Full text] [doi: 10.2196/21451] [Medline: 33270038]
64. How Facebook adverts get approved. Meta. URL: https://en-gb.facebook.com/business/a/ad-review-process [accessed 2022-08-09]
65. Ohlheiser A. Facebook is bombarding cancer patients with ads for unproven treatments. MIT Technology Review. 2022 Jun 27. URL: https://www.technologyreview.com/2022/06/27/1054784/facebook-meta-cancer-treatment-ads-misinformation/ [accessed 2022-08-09]
66. Ohlheiser A. Facebook wants to limit the reach of bogus medical "cures" by treating them like spam. Washington Post. 2019 Jul 2. URL: https://www.washingtonpost.com/technology/2019/07/02/facebook-wants-limit-reach-bogus-medical-cures-by-treating-them-like-spam/ [accessed 2022-08-07]
67. Hernandez D, McMillan R. Facebook, YouTube Overrun with bogus cancer treatment claims. Wall Street Journal. 2019 Jul 02. URL: https://www.wsj.com/articles/facebook-youtube-overrun-with-bogus-cancer-treatment-claims-11562072401 [accessed 2022-08-07]
68. Citron DK. Fix Section 230 and hold tech companies to account. Wired UK. 2021 Jun 05. URL: https://www.wired.co.uk/article/section-230-social-media [accessed 2022-08-09]
69. Hendrix J. FTC issues report on using AI to address online harms. Tech Policy Press. 2022 Jun 17. URL: https://techpolicy.press/ftc-issues-report-on-using-ai-to-address-online-harms/ [accessed 2022-08-10]
70. Thomas K. Key opinion leaders supercharged by the internet: paid doctor and patient influencers on social media. BMJ 2019 May 31;365:l2336. [doi: 10.1136/bmj.l2336] [Medline: 31152003]
71. Mills A. The law applicable to cross-border defamation on social media: whose law governs free speech in 'Facebookistan'? J Media Law 2015 Jul 25;7(1):1-35. [doi: 10.1080/17577632.2015.1055942]
72. Vassallo A, Jones A, Freeman B. Social media: frenemy of public health? Public Health Nutr 2021 Jun 24;25(1):61-64. [doi: 10.1017/s136898002100269x]
73. Reducing the harm from alcohol – by regulating cross-border alcohol marketing, advertising and promotion: a technical report. World Health Organization. 2022 May 10. URL: https://www.who.int/publications-detail-redirect/9789240046504 [accessed 2022-05-18]
74. Kenworthy N, Storeng K, Zenone M. The digital technology sector and health systems. In: The Commercial Determinants of Health. Oxford: Oxford University Press; 2022.
75. Pelley S. Whistleblower: Facebook is misleading the public on progress against hate speech, violence, misinformation. CBS News. 2021 Oct 04. URL: https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/ [accessed 2022-08-10]

76. Perrigo B, Bergengruen V. Why Facebook employees “deprioritized” a misinformation fix. Time. 2021 Nov 10. URL: https://time.com/6116354/facebook-employees-deprioritized-misinformation/ [accessed 2022-08-10]
77. Erzse A, Karim SA, Foley L, Hofman KJ. A realist review of voluntary actions by the food and beverage industry and implications for public health and policy in low- and middle-income countries. Nat Food 2022 Aug 05;3(8):650-663. [doi: 10.1038/s43016-022-00552-5]
78. de Lacy-Vawdon C, Livingstone C. Defining the commercial determinants of health: a systematic review. BMC Public Health 2020 Jun 29;20(1):1022 [FREE Full text] [doi: 10.1186/s12889-020-09126-1] [Medline: 32600398]
79. Helberger N. The political power of platforms: how current attempts to regulate misinformation amplify opinion power. Digit 2020 Jul 07;8(6):842-854. [doi: 10.1080/21670811.2020.1773888]
80. Lacy-Nichols J, Marten R. Power and the commercial determinants of health: ideas for a research agenda. BMJ Glob Health 2021 Feb 16;6(2):e003850 [FREE Full text] [doi: 10.1136/bmjgh-2020-003850] [Medline: 33593758]
81. Zenone M, Kenworthy N, Maani N. The social media industry as a commercial determinant of health. Int J Health Policy Manag 2022 Apr 27. [doi: 10.34172/ijhpm.2022.6840] [Medline: 35490262]
82. Dwoskin E, Newmyer T, Mahtani S. The case against Mark Zuckerberg: insiders say Facebook's CEO chose growth over safety. Washington Post. 2021 Oct 25. URL: https://www.washingtonpost.com/technology/2021/10/25/mark-zuckerberg-facebook-whistleblower/ [accessed 2022-08-08]
83. Ortutay B. People or profit? Facebook papers show deep conflict within. AP News. 2021 Oct 25. URL: https://apnews.com/article/the-facebook-papers-whistleblower-misinfo-trafficking-64f11ccae637cdfb7a89e049c5095dca [accessed 2022-08-10]
84. Donaldson SI, Dormanesh A, Perez C, Majmundar A, Allem J. Association between exposure to tobacco content on social media and tobacco use: a systematic review and meta-analysis. JAMA Pediatr 2022 Sep 01;176(9):878-885. [doi: 10.1001/jamapediatrics.2022.2223] [Medline: 35816331]
85. Zenone M, Kenworthy N. Pre-emption strategies to block taxes on sugar-sweetened beverages: a framing analysis of Facebook advertising in support of Washington state initiative-1634. Glob Public Health 2022 Sep 19;17(9):1854-1867 [FREE Full text] [doi: 10.1080/17441692.2021.1977971] [Medline: 34542004]
86. Levin S. Facebook allowed advertisers to target "Jew haters". The Guardian. 2017 Sep 15. URL: https://www.theguardian.com/technology/2017/sep/14/facebook-advertising-jew-hater-antisemitism [accessed 2022-08-10]
87. Nix N. Facebook bans hate speech but still makes money from white supremacists. Washington Post. 2022 Aug 10. URL: https://www.washingtonpost.com/technology/2022/08/10/facebook-white-supremacy-ads/ [accessed 2022-08-10]
88. González-Cabañas J, Cuevas, Cuevas R, López-Fernández J, García D. Unique on Facebook: formulation and evidence of (nano)targeting individual users with non-PII data. In: Proceedings of the 21st ACM Internet Measurement Conference. In; 2021 Presented at: 21st ACM Internet Measurement Conference; November 2-4; New York, NY p. 464-479. [doi: 10.1145/3487552.3487861]
89. Padyab A, Päivärinta T, Ståhlbröst A, Bergvall-Kåreborn B. Awareness of indirect information disclosure on social network sites. Soc Media Soc 2019 May 19;5(2):205630511882419. [doi: 10.1177/2056305118824199]
90. Zenone M, Kenworthy N, Barbic S. The paradoxical relationship between health promotion and the social media industry. Health Promot Pract 2021 Dec 29:15248399211064640 [FREE Full text] [doi: 10.1177/15248399211064640] [Medline: 34963368]
91. van Schalkwyk MCI, Zenone M, Maani N, Petticrew M, McKee M. Back to our roots or sowing new seeds: thinking anew on the paradigms of health, harm and disease. J Public Health (Oxf) 2022 Nov 21;44(Suppl 1):i28-i33. [doi: 10.1093/pubmed/fdac093] [Medline: 36465052]

Abbreviations

CHIPSA: Centro Hospitalario Internacional del Pacifico, SA
FDA: Food and Drug Administration

Edited by T Mackey; submitted 17.10.22; peer-reviewed by B Rein, T Ginossar; comments to author 15.11.22; revised version received 13.01.23; accepted 27.03.23; published 31.05.23.

Please cite as:
Zenone M, Snyder J, Bélisle-Pipon JC, Caulfield T, van Schalkwyk M, Maani N
Advertising Alternative Cancer Treatments and Approaches on Meta Social Media Platforms: Content Analysis
JMIR Infodemiology 2023;3:e43548
URL: https://infodemiology.jmir.org/2023/1/e43548
doi:10.2196/43548
PMID:37256649

Original Paper

The Role of Social Media in Health Misinformation and Disinformation During the COVID-19 Pandemic: Bibliometric Analysis

Funmi Adebesin[1], BSc, BSc (Hons), MSc, PhD, PGCHE; Hanlie Smuts[1], BSc, BSc (Hons), MSc, PhD; Tendani Mawela[1], BCom, HDip, MBA, PhD, PGCHE; George Maramba[1], BSc, BSc (Hons), MSc, PhD; Marie Hattingh[1], BMil, BMil (Hons), MMil, PhD

Department of Informatics, University of Pretoria, Pretoria, South Africa

Corresponding Author:
Funmi Adebesin, BSc, BSc (Hons), MSc, PhD, PGCHE
Department of Informatics
University of Pretoria
Information Technology Building, Level 5
Lynwood Road
Pretoria, 0028
South Africa
Phone: 27 12 420 5667 ext 5667
Email: funmi.adebesin@up.ac.za

Abstract

Background: The use of social media platforms to seek information continues to increase. Social media platforms can be used to disseminate important information to people worldwide instantaneously. However, their viral nature also makes it easy to share misinformation, disinformation, unverified information, and fake news. The unprecedented reliance on social media platforms to seek information during the COVID-19 pandemic was accompanied by increased incidents of misinformation and disinformation. Consequently, there was an increase in the number of scientific publications related to the role of social media in disseminating health misinformation and disinformation at the height of the COVID-19 pandemic. Health misinformation and disinformation, especially in periods of global public health disasters, can lead to the erosion of trust in policy makers at best and fatal consequences at worst.

Objective: This paper reports a bibliometric analysis aimed at investigating the evolution of research publications related to the role of social media as a driver of health misinformation and disinformation since the start of the COVID-19 pandemic. Additionally, this study aimed to identify the top trending keywords, niche topics, authors, and publishers for publishing papers related to the current research, as well as the global collaboration between authors on topics related to the role of social media in health misinformation and disinformation since the start of the COVID-19 pandemic.

Methods: The Scopus database was accessed on June 8, 2023, using a combination of Medical Subject Heading and author-defined terms to create the following search phrases that targeted the title, abstract, and keyword fields: ("Health*" OR "Medical") AND ("Misinformation" OR "Disinformation" OR "Fake News") AND ("Social media" OR "Twitter" OR "Facebook" OR "YouTube" OR "WhatsApp" OR "Instagram" OR "TikTok") AND ("Pandemic*" OR "Corona*" OR "Covid*"). A total of 943 research papers published between 2020 and June 2023 were analyzed using Microsoft Excel (Microsoft Corporation), VOSviewer (Centre for Science and Technology Studies, Leiden University), and the *Biblioshiny* package in Bibliometrix (K-Synth Srl) for RStudio (Posit, PBC).

Results: The highest number of publications was from 2022 (387/943, 41%). Most publications (725/943, 76.9%) were articles. *JMIR* published the most research papers (54/943, 5.7%). Authors from the United States collaborated the most, with 311 coauthored research papers. The keywords "Covid-19," "social media," and "misinformation" were the top 3 trending keywords, whereas "learning systems," "learning models," and "learning algorithms" were revealed as the niche topics on the role of social media in health misinformation and disinformation during the COVID-19 outbreak.

Conclusions: Collaborations between authors can increase their productivity and citation counts. Niche topics such as "learning systems," "learning models," and "learning algorithms" could be exploited by researchers in future studies to analyze the influence of social media on health misinformation and disinformation during periods of global public health emergencies.

(JMIR Infodemiology 2023;3:e48620) doi:10.2196/48620

KEYWORDS

bibliometric analysis; COVID-19; fake news; health disinformation; health misinformation; social media

Introduction

Background

The use of social media platforms, including Facebook, Twitter, YouTube, WhatsApp, and Instagram, has revolutionized how we seek information and connect [1]. As of January 2023, there were 4.76 billion social media users worldwide, accounting for 59.4% of the world's population [2]. Therefore, it was no surprise that many people worldwide relied heavily on social media to seek information about the coronavirus at the height of the COVID-19 pandemic, when governments instituted national lockdowns and restricted movement to contain the spread of the virus [3]. According to Naeem et al [4], global social media use surged from 20% to 87% at the height of the COVID-19 pandemic.

Governments and policy makers worldwide turned to social media to provide relevant information about COVID-19 to their citizens [5,6]. For example, the Macao government (China) used its official Facebook page to disseminate information about COVID-19 to its citizens. This included the government's plans to combat the virus, dissemination of public health messages aimed at changing behavior, postings aimed at debunking fake news about the virus, and updates about the virus [6]. Health professionals also used social media platforms to rapidly disseminate care guidelines to health workers in different countries, with the guidelines translated into local languages [7].

The viral nature of social media platforms makes them suited to the rapid transmission of credible information, fake news, and unverified information [8]. Although government officials were using social media platforms to disseminate information about COVID-19, peddlers of misinformation and disinformation were exploiting social media to distribute false and unverified information about the virus [4,9].

Misinformation, disinformation, and fake news are not new [10-12] and can be traced back to the Roman Empire [11]. However, the advent of social media, made popular by Web 2.0 technology, means that content creation is no longer restricted to traditional news media. Users are now empowered to become content creators [13], making the widespread distribution of misinformation, disinformation, and fake news a huge challenge for society at large [12].

According to Wardle and Derakhshan [12], misinformation and disinformation are forms of information disorder that involve the dissemination of nonfactual information. The distinguishing feature between the 2 terms is intention. Misinformation entails the distribution of false information without the intention to cause harm. This definition can be expanded to encompass health misinformation, which involves making health-related claims that are not based on scientific evidence without intending to cause harm [14]. In contrast, the motive for disinformation is to deliberately cause harm by sharing false information. A third form of information disorder identified by Wardle and Derakhshan [12] is malinformation, which is the sharing of factual information outside its original context with the intention to cause harm.

Wardle and Derakhshan [12] identified 7 types of mis- and disinformation that lie on a continuum scale that loosely measures the intention to deceive. *Satire or parody* are not intended to cause harm but could be mistaken by some audiences as facts. *Misleading contents* entail the deceptive use of information to manipulate a situation or an issue. *Imposter contents* are disguised to look like or mimic genuine contents, whereas *fabricated contents* are based on nonfactual information with the intention to cause harm. *Content manipulation*, in contrast, involves changing information, images, or videos with the intention to deceive the recipients. In *false content*, factual information is shared with nonfactual contextual information. Finally, *false connection* is when headlines, visuals, or captions are not aligned with the associated content [12].

In contrast, fake news can be described as fabricated content disguised to look like real news [10,15]. According to Tandoc et al [16], fake news can manifest in 6 different forms. *Satire* entails inducing humor based on factual topical issues or critiquing those in power. *Parody* is similar to satire in that it is also intended to induce humor. However, parody is different in that it uses nonfactual or fictitious information to induce humor. *Fabrication* is another manifestation of fake news in which content based on false information is disguised as real news with the intent to misinform the consumers of such content. Fake news can also take the form of *manipulation* of original photos or videos to create incorrect or misleading narratives. Another manifestation of fake news is when *advertisement and public relations* content produced by public relations professionals is included in real news content for financial gain. Although the advertisements may be based on facts, their mashup with real news content could mislead consumers into believing that the entire content is from a real news agency. Finally, fake news can also manifest in the form of *propaganda*, in which contents that may or may not be based on facts are created by a celebrity or public figure with the intention to influence public perception about a particular topic or issue [16].

Despite the overlap between the different manifestations of fake news, misinformation, and disinformation, authors such as Wardle and Derakhshan [12] have refrained from using the term *fake news* because of the complexity of misinformation and disinformation and the propensity of public figures and politicians to misuse *fake news* to refer to critical news items that they do not agree with.

The spread of misinformation, disinformation, and fake news, especially during pandemics, can create panic among citizens and erode their trust in governments and policy makers [17], with the potential for fatal consequences [18]. For example, Soltaninejad [19] reported that misinformation on social media played a role in many Iranian citizens consuming large volumes of methanol in early 2020 as they believed that this would protect them from being infected by the virus. This led to >2000 people being admitted to hospitals across the country with methanol poisoning and 264 subsequent fatalities. In a similar study, Naeem et al [4] analyzed the sources of 1225 fake news stories between January 2020 and April 2020. The authors found that social media platforms accounted for 50% of distributed fake news and identified three main categories of misinformation about the coronavirus: (1) false claims, in which incorrect information about the mode of transmission of COVID-19, the cure for the virus (eg, that consuming large amounts of methanol can cure COVID-19), and the promotion of false prejudices about people of specific ethnic groups being responsible for COVID-19 were being disseminated; (2) the spread of conspiracy theories driven by the pronouncements of some world leaders and public figures about the origin of the virus (eg, that the virus is spread through 5G towers); and (3) the spread of pseudoscientific remedies and treatments purported to be capable of preventing the virus or curing people infected with COVID-19 (eg, that steam inhalation can cure COVID-19).

Another study by Fieselmann et al [20] found that vaccine hesitancy among German citizens could be attributed to the spread of misinformation on Twitter about the benefits of the COVID-19 vaccination. Similarly, the findings by Crouse and Dupuis [18] suggest that people who believe in misinformation and conspiracy theories about COVID-19 have a higher probability of refusing vaccination.

The unprecedented reliance on social media platforms to seek information about the virus [21,22] was accompanied by the dissemination of information about the virus at an alarming rate from both credible and unreliable sources on social media platforms [23]. Social media platforms made it easier to spread misinformation, disinformation, and fake news worldwide [18]. There was an overabundance of factual and nonfactual information [18]. As a consequence, the World Health Organization issued a warning about an impending infodemic and the risk it posed to the global efforts to combat the pandemic [24]. An infodemic can be defined as an overabundance of factual and incorrect information, especially during a pandemic, which could lead to panic and risky behaviors among people [25]. More specifically, Gisondi et al [26] defined the COVID-19 infodemic as “the overwhelming amount of complex and often contradictory information available about COVID-19, inclusive of substantial fake news about the origins of the virus, treatment options unsupported by rigorous clinical data, and baseless claims regarding adverse effects of lifesaving vaccines.” The authors further state that the false narratives about COVID-19 were sometimes propagated by authoritative institutions or public figures who exploited the trust of the general populace to influence their views about, and behaviors toward the virus, with the potential for negative health outcomes.

Several authors [18,26-28] have identified social media platforms as an important driver of the COVID-19 infodemic. Given the huge number of active social media users worldwide [2], it has become much easier to rapidly disseminate information (factual and nonfactual) worldwide. Although humans were generally responsible for sharing and resharing misinformation and disinformation about COVID-19 [18,29], automated web-based accounts, called social bots, have also been identified as superspreaders of the COVID-19 infodemic [28,29]. For example, social bots were responsible for the rapid propagation of COVID-19 misinformation, disinformation, and fake news, including conspiracy theories about the spread of the virus through 5G towers or weakening of the immune system, thus making people more susceptible to the virus, and that Bill Gates was the creator of the virus to obtain financial benefits from the pandemic [30]. In addition to the COVID-19 infodemic, scientific publications related to the virus also surged significantly during the pandemic, with a reduction in the publication of studies that were not related to COVID-19 [31]. For example, a Google Scholar search using the phrase “social media as a driver of misinformation about Covid-19” filtered between 2020 and 2023 yielded 24,500 search results. Given the apparent increase in the number of publications related to COVID-19, it is important to investigate scientific productivity related to the role of social media in health misinformation and disinformation since the start of the COVID-19 pandemic.

Objectives

The rapid pace at which studies related to the role of social media platforms in driving health misinformation and disinformation about the coronavirus were published at the height of the pandemic makes a bibliometric analysis a suitable research method for this study. Although authors such as Yeung et al [32] have conducted a bibliometric analysis of medical and health-related misinformation on social media, this study is different in that it is specifically focused on the bibliometric analysis of publications related to the role of social media in health misinformation and disinformation during the COVID-19 pandemic. Consequently, this paper reports the results of a bibliometric analysis aimed at answering the following research questions: (1) How have research publications related to the role of social media as a driver of health misinformation and disinformation evolved since the start of the COVID-19 pandemic? (2) What are the trending keywords and niche topics in publications related to the role of social media as a driver of health misinformation and disinformation since the start of the COVID-19 pandemic? (3) Who are the most influential authors of documents and the most influential publishers of studies related to the role of social media as a driver of health misinformation and disinformation since the start of the COVID-19 pandemic? and (4) Who has made up the collaborative networks of authors of publications related to the role of social media as a driver of health misinformation and disinformation since the start of the COVID-19 pandemic?

Studies that report on COVID-19–related health misinformation and disinformation through other media sources such as web-based news and print media are not the focus of this paper.

The remaining sections of the paper are structured as follows. The research design is presented in the *Methods* section. This is followed by the presentation of the results and a discussion of the bibliometric analysis in the *Results* and *Discussion* sections, respectively. Thereafter, the study’s limitations and conclusions are presented in the *Limitations* and *Conclusions* sections, respectively.

Methods

Overview

This study used the bibliometric analysis method to investigate the research outputs related to the role of social media as a driver of health misinformation and disinformation since the start of the COVID-19 pandemic. A bibliometric analysis is an objective method of analyzing and synthesizing extant literature to measure the productivity of science, scientists, or scientific activities in a particular field or about a particular topic [33].

Source Selection Process

Data for the bibliometric analysis were retrieved from the Scopus database on June 8, 2023. Scopus was the database of choice because of its collection of a wide range of abstracts and sources from different disciplines. To ensure that relevant sources were retrieved, we used a combination of Medical Subject Heading and author-defined terms to create the following search phrases, targeting the title, abstract, and keyword fields: (“Health*” OR “Medical”) AND (“Misinformation” OR “Disinformation” OR “Fake News”) AND (“Social media” OR “Twitter” OR “Facebook” OR “YouTube” OR “WhatsApp” OR “Instagram” OR “TikTok”) AND (“Pandemic*” OR “Corona*” OR “Covid*”).

Source Screening Process

A total of 1570 search results were retrieved. There were no duplicate documents. The sources were then filtered by limiting them to documents that had been published between 2020 and 2023, resulting in the exclusion of 0.7% (11/1570) of the documents. Thereafter, the sources were filtered by limiting them to documents that were published in English. This resulted in the elimination of 3.01% (47/1559) of the documents. An additional 12.43% (188/1512) of the documents (notes, commentaries, opinions, and letters) were excluded during the screening stage, thereby retaining 1324 documents. The remaining 1324 documents were then screened for eligibility by checking the document types and reading the titles and abstracts. This step resulted in the exclusion of sources with an unspecified document type (26/1324, 1.96%) and 9.14% (121/1324) of the documents, which had no abstract. An additional 17.67% (234/1324) of the documents, which were not relevant to the study (eg, dissemination of fake news about monkeypox, antiviral drugs for smallpox, and social media’s role in panic buying during the pandemic), were excluded. The eligibility screening step resulted in a total of 943 documents being retained as the final set of sources included in the bibliometric analysis. The screening process of the data sources is illustrated in Figure 1, whereas Table 1 summarizes the main information about the sources included in the bibliometric analysis. A copy of the 943 sources included in the bibliometric analysis is provided in Multimedia Appendix 1.

We used Microsoft Excel (Microsoft Corporation) to analyze the number of publications per year, the annual publication growth rate, and document type. VOSviewer (version 1.6.19; Centre for Science and Technology Studies, Leiden University) [34] was used to analyze the keyword co-occurrences and authors’ collaboration networks. Finally, we used the *Biblioshiny* package in Bibliometrix (K-Synth Srl) for RStudio (version 2023.03.0+386 “Cherry Blossom” release; Posit, PBC) [33] to analyze the authors’ number of publications, the most active publishers, the most cited documents, and a thematic map of the topics related to the role of social media in health misinformation and disinformation since the start of the COVID-19 pandemic.

Figure 1. Source screening process.

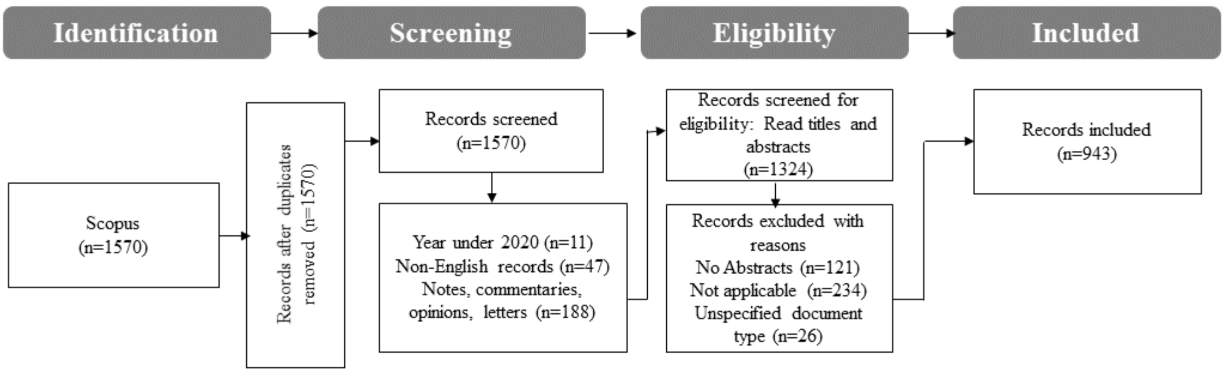


Table 1. Main information on sources included in the bibliometric analysis generated using Biblioshiny.

Category and description	Values
Time span	2020-2023
References	44,612
Authors, n	
Authors	4038
Authors of single-authored documents	88
Author collaboration	
Single-authored documents, n	91
Coauthors per document, mean	4.77
International coauthorships (%)	28.31
Document contents, n	
Keywords Plus (ID)	2990
Author keywords (DE)	1777
Documents	
Total documents, n	943
Age of document, mean	1.42
Citations per document, mean	17.4

Ethical Considerations

We did not apply for ethics clearance for this study because the University of Pretoria only requires ethics clearance for studies that include humans or animals. The bibliometric analysis reported in this study did not include human or animal subjects.

Results

Overview

As stated in the *Methods* section, the bibliometric analysis was conducted with the aid of Microsoft Excel, VOSviewer (version 1.6.19) [34], and the *Biblioshiny* package in Bibliometrix for RStudio (version 2023.03.0+386 "Cherry Blossom" release) [33]. *Biblioshiny* is "a shiny app" that provides a web interface for Bibliometrix and can be used by researchers with no programming experience [35]. The results of the bibliometric analysis are presented in the following subsections.

Annual Publication Growth

A citation analysis of the 943 research papers was conducted to understand how research related to the role of social media in driving health misinformation and disinformation about COVID-19 evolved from 2020 to 2023.

Basic publication growth analysis was conducted in Microsoft Excel using the following annual publication growth rate formula: annual publication growth rate = (number of publications in the year under consideration / number of publications in the preceding year − 1) × 100.

As illustrated in Table 2, a total of 113 documents were published during the first year under consideration (2020). In 2021, there were 306 publications, representing a massive 171% publication growth rate. The number of publications increased to 387 in 2022, representing a 26.4% annual publication growth rate. In 2023, there were only 137 publications as of June 8, 2023, representing a 64.4% decline compared with the preceding year. The decline in the annual publication growth rate can be attributed to the fact that the sources for the bibliometric analysis were extracted on June 8, 2023.

The sources were also analyzed according to their document types. Most of the documents (725/943, 76.9%) were classified as articles, followed by conference papers (107/943, 11.3%). There were 8.3% (78/943) reviews and 3.5% (33/943) book chapters.

The documents were also analyzed from the perspective of the number of publications per author. A total of 4038 authors contributed to the 943 sources that were included in the bibliometric analysis (Table 1). As shown in Table 3, a combined total of 5.3% (50/943) of the documents were published by the top 10 authors. In total, 0.7% (7/943) of the documents were published by the author in the topmost position, whereas 0.5% (5/943) of the documents were published by each of the authors in positions 2 to 8. The authors in positions 9 and 10 published 0.4% (4/943) of the documents each.

Finally, the sources were analyzed according to the publishers to determine the most active ones. The 943 documents were published by 489 unique publishers. Analysis through *Biblioshiny* showed that most of the publishers published only 1 document each. *JMIR*, consisting of *JMIR* (54/943, 5.7%), *JMIR Infodemiology* (22/943, 2.3%), *JMIR Formative Research* (14/943, 1.5%), and *JMIR Public Health Surveillance* (13/943, 1.4%), accounted for 10.9% (103/943) of the publications from the top 10 publishers (Table 4). There were 4.8% (45/943) of publications from the *International Journal of Environmental Research and Public Health*, 2.9% (27/943) from *PLOS ONE*, and 2.4% (23/943) from the *Vaccines* journals. The *Frontiers in Public Health* journal had 1.7% (16/943) of publications, the *Human Vaccines & Immunotherapeutics* journal published 1.5% (14/943), and the *BMC Public Health* journal published 1.4% (13/943).

The annual publication, depicted in Figure 2, is only for the top 5 publishers. This is to ensure a meaningful visualization. The line graph corroborates the number of publications by the top 10 publishers shown in Table 4, confirming *JMIR* as the top publisher with 12 documents at the height of the COVID-19 pandemic in 2020. In 2021, the number of publications through *JMIR* increased to 17 (41.7% growth) before peaking at 20 (17.6% growth) in 2022, with only 5 (75% decline) as of June 8, 2023, when the data set was retrieved. The *International Journal of Environmental Research and Public Health* published 6 documents in 2020. This increased to 20 (233% growth) in 2021 before dipping to 15 (25% decline) and 4 (73.3% decline) in 2022 and 2023, respectively. The *PLOS ONE* journal published 4 papers in 2020, a total of 10 (150% growth) in 2021, and 11 (10% growth) in 2022. There were only 2 (81.8% decline) publications from *PLOS ONE* as of June 8, 2023, when the data set was retrieved. There were no publications from the *Vaccines* journal in 2020. This journal published 7 papers in 2021, which increased to 12 (71.4% growth) in 2022. There were only 4 (66.7% decline) publications from the *Vaccines* journal as of June 8, 2023, when the data set was retrieved. In 2020, there were no publications from *JMIR Infodemiology*. The lack of publications by *JMIR Infodemiology* in 2020 can be explained by the fact that the journal was launched in the middle of 2021. Only 1 paper was published in this journal in 2021, also because the journal was launched in the middle of this year. There was a sharp increase to 15 publications (1400% growth) in 2022. There were only 6 (60% decline) publications from *JMIR Infodemiology* as of June 8, 2023, when the data set was retrieved.

Table 2. Number of publications per year (the year 2023 represents data up until June 2023).

Year	Publications (n=943), n (%)
2020	113 (11.97)
2021	306 (32.45)
2022	387 (41.04)
2023	137 (14.53)
Total	943 (100)

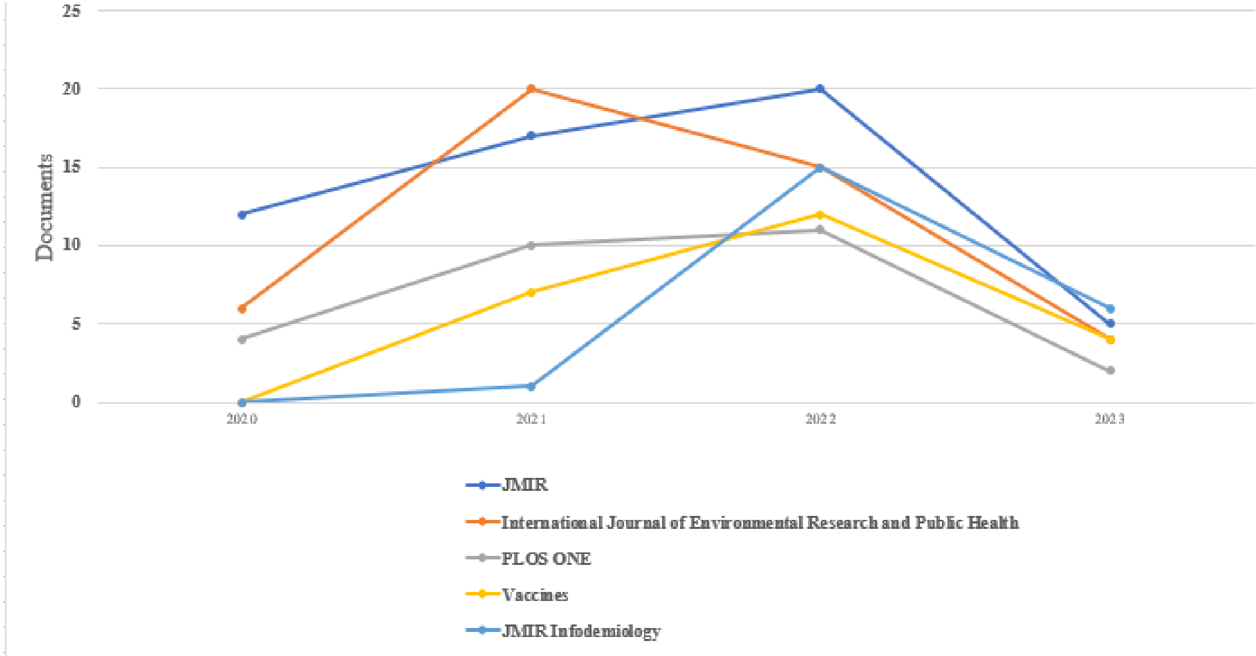
Table 3. Number of publications by the top 10 authors.

Author	Publications (n=943), n (%)
Ahmed	7 (0.7)
Da San Martino	5 (0.5)
Li	5 (0.5)
Liu	5 (0.5)
Luo	5 (0.5)
Nakov	5 (0.5)
Shaban-Nejad	5 (0.5)
Wang	5 (0.5)
Briand	4 (0.4)
Chakraborty	4 (0.4)
Total	50 (5.3)

Table 4. Number of publications by the top 10 publishers.

Journal	Publications (n=943), n (%)
JMIR	54 (5.7)
International Journal of Environmental Research and Public Health	45 (4.8)
PLOS ONE	27 (2.9)
Vaccines	23 (2.4)
JMIR Infodemiology	22 (2.3)
Frontiers in Public Health	16 (1.7)
Human Vaccines & Immunotherapeutics	14 (1.5)
JMIR Formative Research	14 (1.5)
BMC Public Health	13 (1.4)
JMIR Public Health and Surveillance	13 (1.4)
Total	241 (25.56)

Figure 2. Annual publications per publisher (the year 2023 represents data up until June 2023).



Co-Word Analysis

A co-word analysis of the 943 research papers was conducted using VOSviewer to determine the trending keywords in publications related to the role of social media in driving health misinformation and disinformation since the start of the COVID-19 pandemic.

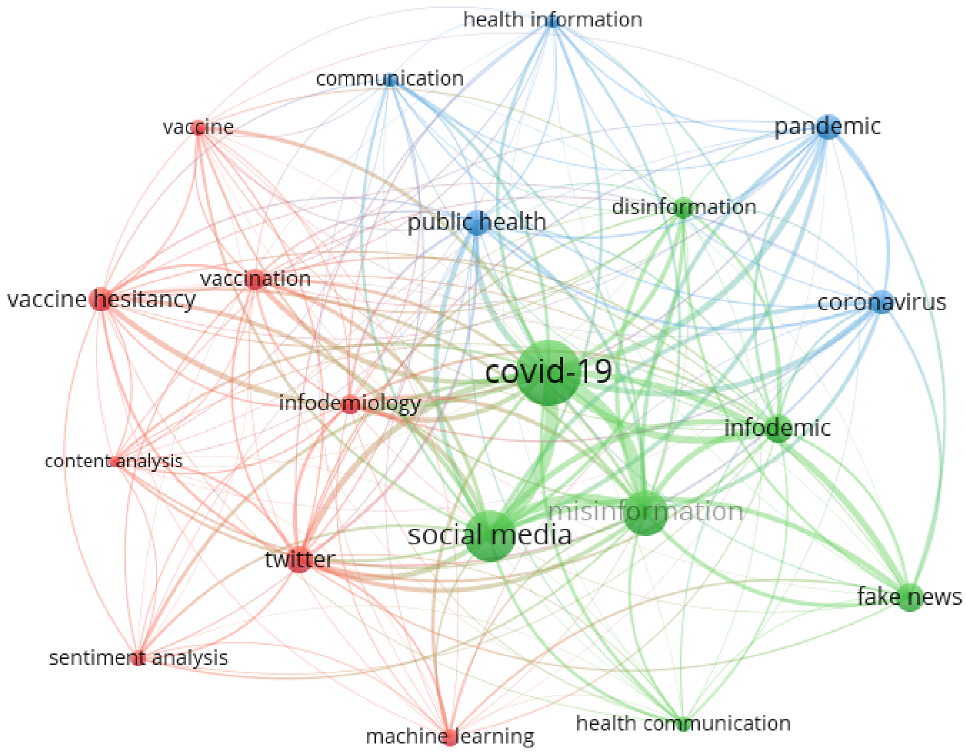
Co-word analysis is a bibliometric analysis technique used to analyze the co-occurrence of keywords in texts and map the strength of the relationships between the keywords [36]. The relationships between the keywords are expressed in terms of the number of occurrences and the total link strength (TLS) of the keywords [34].

VOSviewer detected 1779 keywords (as opposed to the 1777 detected by *Biblioshiny*) from the 943 documents using the full counting method. Although VOSviewer detected 107 keywords that met the threshold of 5 keywords as the minimum number of occurrences, the number of keywords to be selected was restricted to 20 to ensure the generation of a meaningful visualization of keywords. As illustrated in Table 5, “Covid-19,” “social media,” and “misinformation” were the top 3 keywords, with “Covid-19” having the highest number of occurrences at 566 (TLS=1247), followed by “social media” with 342 occurrences (TLS=956) and “misinformation” with 277 occurrences (TLS=797). It should be noted that the higher the value of the TLS for a keyword, the stronger the link [34] and the higher the ranking assigned to a keyword. Figure 3 illustrates the top 20 authors’ keyword networks. The size of each circle in the network is an indication of the number of keyword occurrences. The connections (links) between keywords indicate the strength of their relationships with one another.

Table 5. Top 20 keywords according to total link strength (TLS) using VOSviewer.

Keyword rank	Keyword	Number of occurrences	TLS
1	“Covid-19”	566	1247
2	“Social media”	342	956
3	“Misinformation”	277	797
4	“Infodemic”	114	385
5	“Twitter”	96	362
6	“Public health”	89	283
7	“Pandemic”	86	275
8	“Coronavirus”	84	272
9	“Fake news”	111	268
10	“Infodemiology”	52	262
11	“Vaccine hesitancy”	82	226
12	“Vaccination”	62	210
13	“Disinformation”	65	202
14	“Health information”	31	128
15	“Vaccine”	35	120
16	“Communication”	29	118
17	“Content analysis”	21	107
18	“Sentiment analysis”	37	106
19	“Health communication”	37	105
20	“Machine learning”	38	97

Figure 3. VOSviewer visualization of networks of the top 20 authors’ keywords related to the role of social media in health misinformation and disinformation during the COVID-19 pandemic.



Coauthorship and Collaboration by Country

To understand the collaboration networks of authors from different countries in publications related to the role of social media in driving health misinformation and disinformation since the start of the COVID-19 pandemic, the country of affiliation of the coauthored documents was analyzed using VOSviewer (version 1.6.19).

Of the 111 countries detected by VOSviewer, 51 (45.9%) met the minimum threshold of 5 documents from a country with at least 1 citation from that country. Table 6 illustrates the top 20 coauthorships among countries. As shown in Table 6, authors from the United States had the highest number of collaborations, with 311 coauthored documents. The collaborations between the United States and other countries yielded the highest number of citations (n=5388), with a TLS of 141. The higher the TLS, the stronger the collaboration between authors [34]. South Africa and Nigeria were the only African countries in the top 20 coauthorship list. The 2 countries were ranked 10th (coauthorship=26; TLS=32) and 17th (coauthorship=26; TLS=26), respectively, in their collaborations with other countries.

Figure 4 depicts the collaboration networks among the top 20 countries on topics related to the role of social media on health misinformation and disinformation during the COVID-19 pandemic. The collaboration networks between the United States and other countries, as well as between South Africa and other countries, are accentuated in Figures 5 and 6, respectively.

Table 6. Top 20 coauthorships by country using VOSviewer.

Country rank	Country	Number of documents	Total citations, n	TLS[a]
1	United States	311	5388	141
2	United Kingdom	94	2350	99
3	Canada	66	2021	64
4	China	61	987	54
5	Australia	54	1505	44
6	India	70	788	40
7	Pakistan	24	713	34
8	Hong Kong	21	463	32
9	Italy	42	1011	32
10	South Africa	26	86	32
11	Germany	30	278	31
12	Spain	50	1012	30
13	Singapore	18	232	28
14	South Korea	18	128	28
15	Switzerland	17	406	28
16	Saudi Arabia	35	376	27
17	Nigeria	26	137	26
18	France	19	190	23
19	Belgium	7	301	18
20	United Arab Emirates	20	198	17

[a]TLS: total link strength.

Figure 4. VOSviewer visualization of the top 20 countries' collaboration networks on publications related to the role of social media in health misinformation and disinformation during the COVID-19 pandemic.

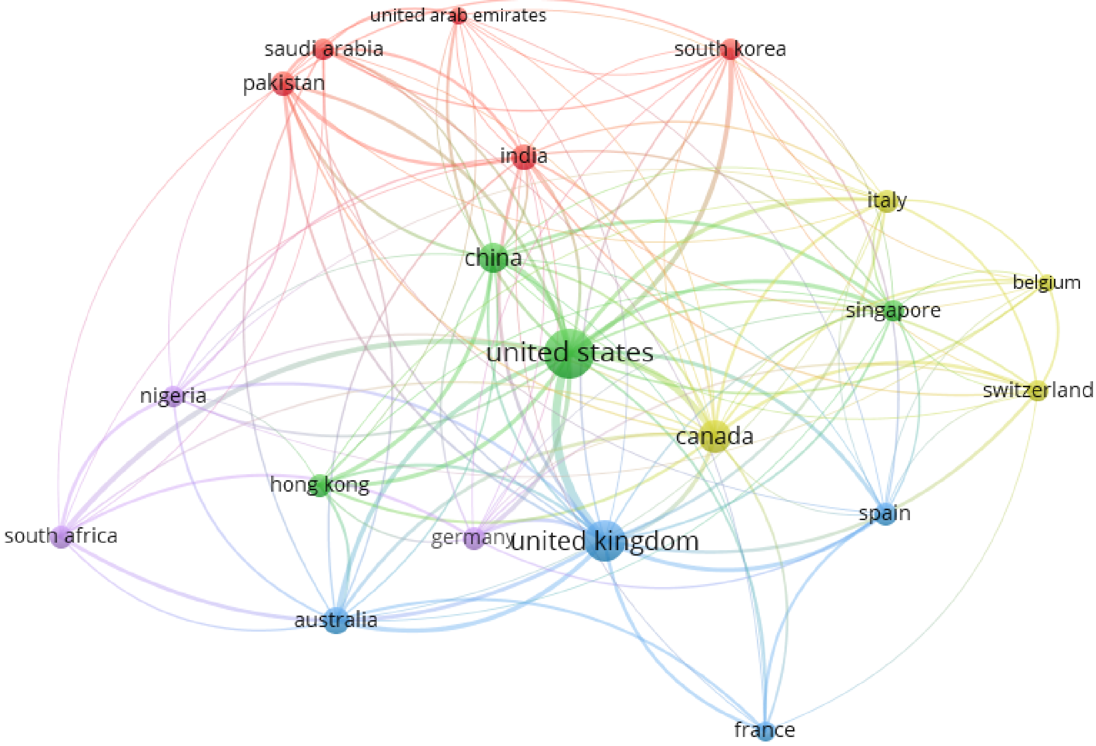


Figure 5. VOSviewer visualization accentuating the collaboration networks between the United States and other countries on publications related to the role of social media in health misinformation and disinformation during the COVID-19 pandemic.

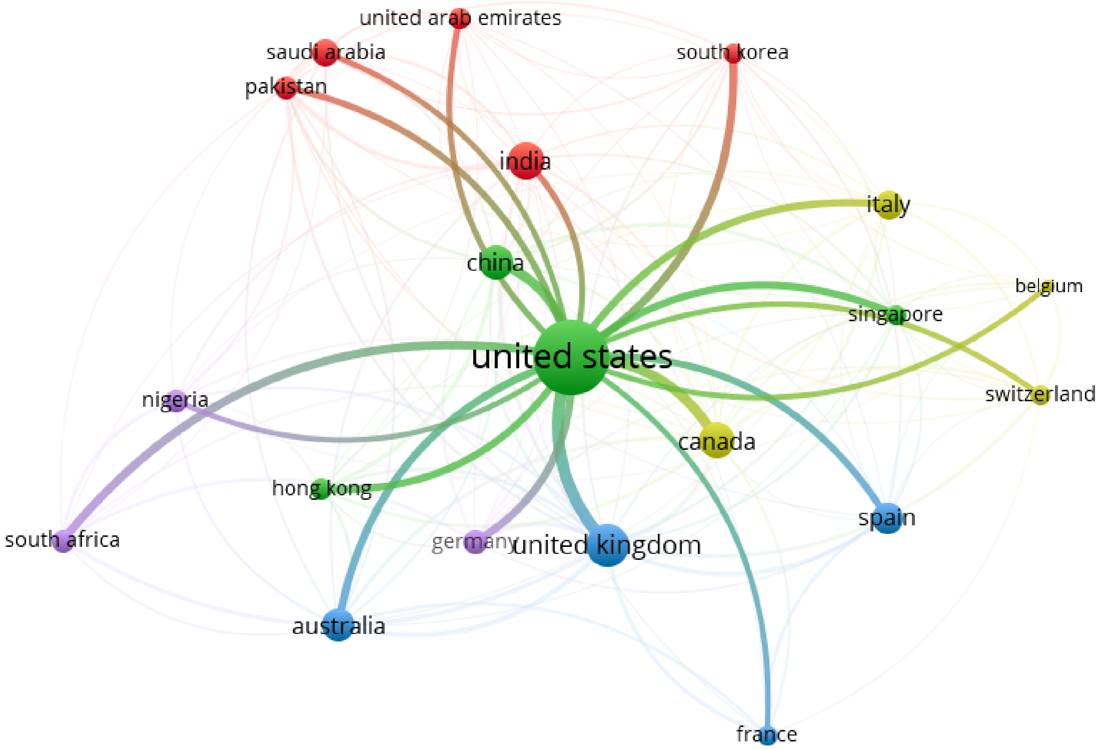
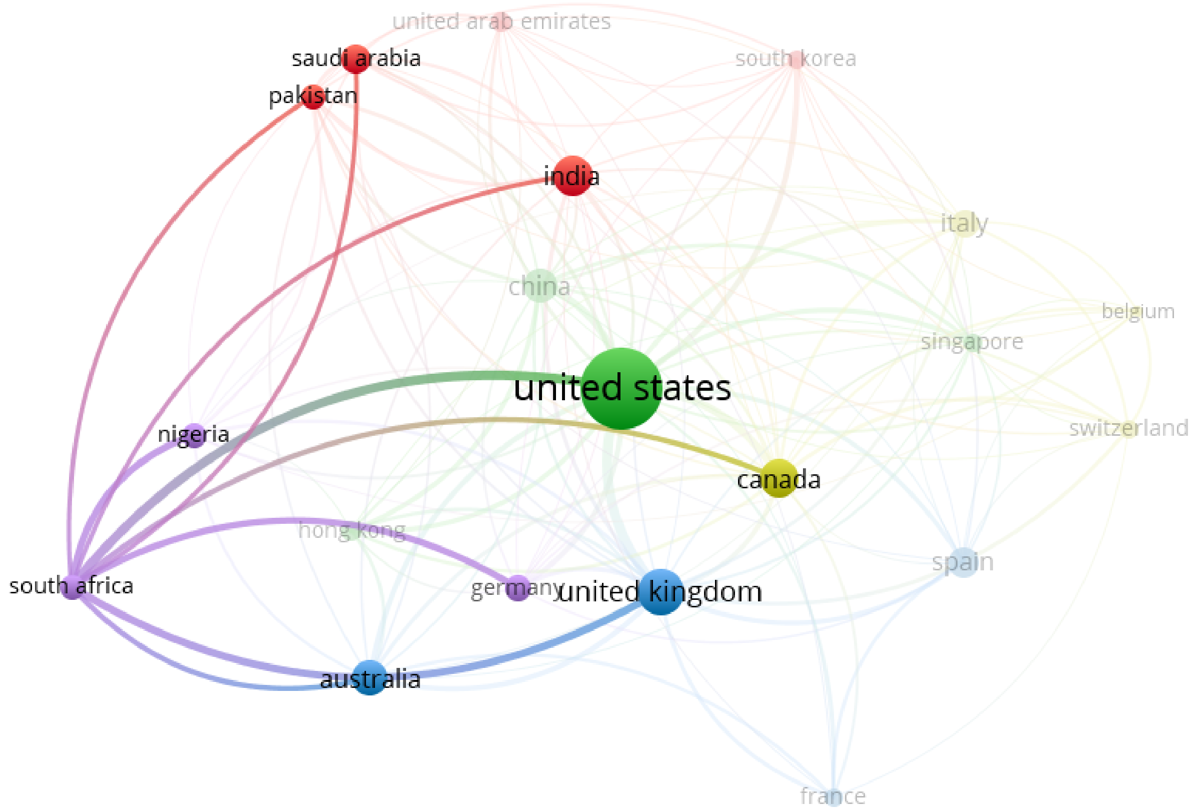


Figure 6. VOSviewer visualization accentuating the collaboration networks between South Africa and other countries on publications related to the role of social media in health misinformation and disinformation during the COVID-19 pandemic.



Most Cited Documents

To determine the authors with the most impactful publications, the top 20 most cited publications related to the role of social media in driving health misinformation and disinformation since the start of the COVID-19 pandemic were analyzed.

One of the indicators of a scientific publication's impact is the number of citations of an author's published research. The influence of the authors of the 943 research papers included in the bibliometric analysis was measured using the global citation feature in *Biblioshiny* and the impact factor of the journals in which the research papers were published. All the top 20 most cited papers were coauthored by at least 2 authors. The paper with the highest number of citations (n=522) was written by Puri et al [37], followed by the paper by Romer and Jamieson [38] in the second position with 491 citations and the paper by Islam et al [39] with 461 citations in the third position. All the top 20 most cited papers were published in journals with impact factors >3, with 3.7 being the lowest impact factor and 30.8 being the highest (Table 7).

Table 7. Top 20 most cited papers.

Study, year	Title	Total citations (Scopus), n	Journal	Journal impact factor[a]
Puri et al [37], 2020	Social Media and Vaccine Hesitancy: New Updates for the Era of COVID-19 and Globalized Infectious Diseases	522	Human Vaccines & Immunotherapeutics	4.52
Romer and Jamieson [38], 2020	Conspiracy Theories as Barriers to Controlling the Spread of COVID-19 in the U.S.	491	Social Science & Medicine	5.4
Islam et al [39], 2020	COVID-19-Related Infodemic and its Impact on Public Health: A Global Social Media Analysis	461	American Journal of Tropical Medicine and Hygiene	3.7
Abd-Alrazaq et al [40], 2020	Top Concerns of Tweeters During the COVID-19 Pandemic: A Surveillance Study	442	JMIR	7.08
Ahmed et al [41], 2020	COVID-19 and the 5G Conspiracy Theory: Social Network Analysis of Twitter Data	385	JMIR	7.08
Tasnim et al [42], 2020	Impact of Rumors and Misinformation on COVID-19 in Social Media	380	Journal of Preventive Medicine & Public Health	3.97
Allington et al [43], 2021	Health-Protective Behaviour, Social Media Usage and Conspiracy Belief During the COVID-19 Public Health Emergency	350	Psychological Medicine	6.03
Ahmad and Murad [44], 2020	The Impact of Social Media on Panic During the COVID-19 Pandemic in Iraqi Kurdistan: Online Questionnaire Study	348	JMIR	7.08
Hua and Shaw [45], 2020	Corona Virus (Covid-19) “Infodemic” and Emerging Issues Through a Data Lens: The Case of China	315	IJERPH[b]	4.61
Li et al [21], 2020	YouTube as a Source of Information on COVID-19: A Pandemic of Misinformation?	315	BMJ Global Health	8.05
Tsao et al [46], 2021	What Social Media Told Us in the Time of COVID-19: a Scoping Review	279	The Lancet Digital Health	30.8
Laato et al [47], 2020	What Drives Unverified Information Sharing and Cyberchondria During the COVID-19 Pandemic?	250	European Journal of Information Systems	9.01
Tangcharoensathien et al [48], 2020	Framework for Managing the COVID-19 Infodemic: Methods and Results of an Online, Crowdsourced WHO Technical Consultation	247	JMIR	7.08
van der Linden et al [49], 2020	Inoculating Against Fake News About COVID-19	219	Frontiers in Psychology	3.9
Gallotti et al [50], 2020	Assessing the Risks of “Infodemics” in Response to COVID-19 Epidemics	205	Nature Human Behaviour	24
Suarez-Lledo and Alvarez-Galvez [14], 2021	Prevalence of Health Misinformation on Social Media: Systematic Review	201	JMIR	7.08
Islam et al [51], 2021	COVID-19 Vaccine Rumors and Conspiracy Theories: The Need for Cognitive Inoculation Against Misinformation to Improve Vaccine Adherence	199	PLOS ONE	3.75
Malecki et al [52], 2021	Crisis Communication and Public Perception of COVID-19 Risk in the Era of Social Media	198	Clinical Infectious Diseases	11.8
Jennings et al [53], 2021	Lack of Trust, Conspiracy Beliefs, and Social Media Use Predict COVID-19 Vaccine Hesitancy	173	Vaccines	7.8
Barua et al [54], 2020	Effects of Misinformation on COVID-19 Individual Responses and Recommendations for Resilience of Disastrous Consequences of Misinformation	166	Progress in Disaster Science	6.3

[a]2022 journal impact factor.

[b]IJERPH: International Journal of Environmental Research and Public Health.

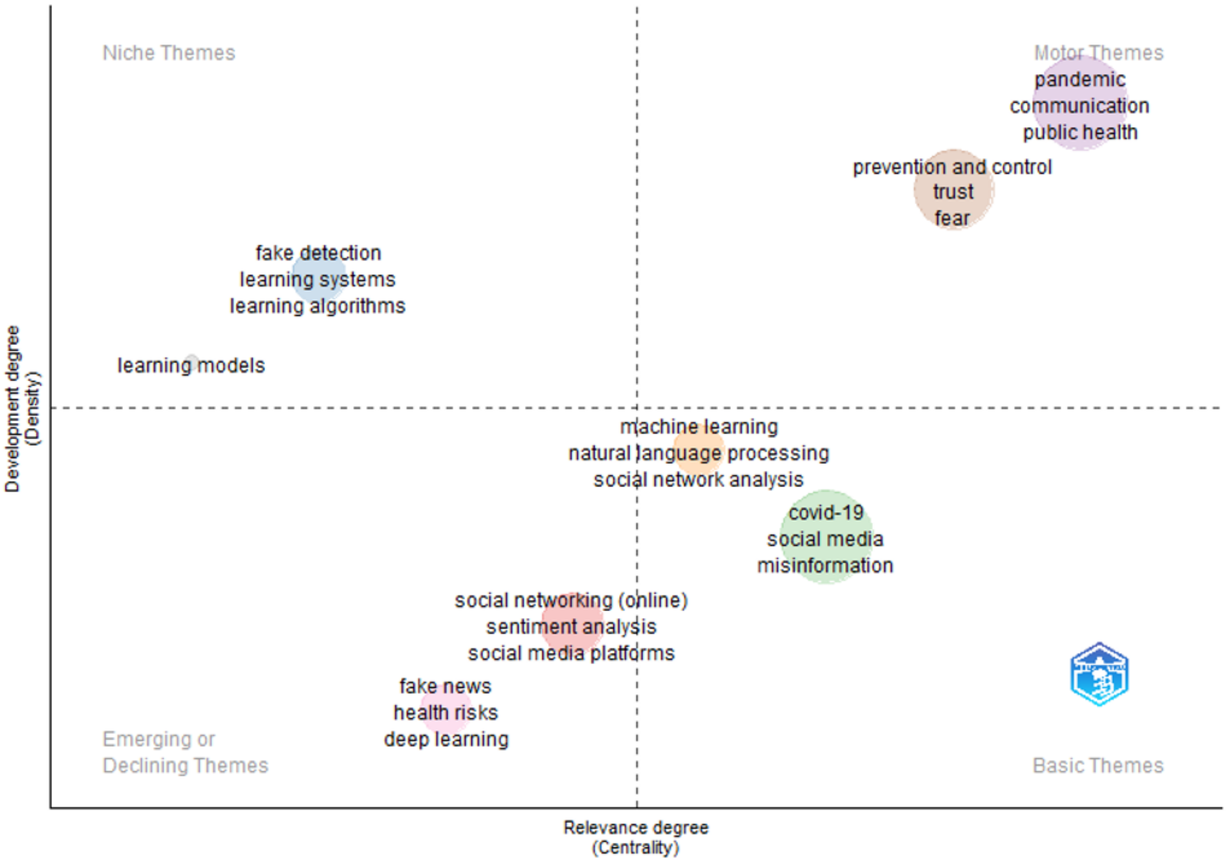
Thematic Map of Keywords

Finally, the *Biblioshiny* package in Bibliometrix for RStudio was used to generate the thematic map illustrated in Figure 7. The thematic map provides a 2D view of the degree of development of the themes, which is computed by *Biblioshiny* based on the external links (centrality) and internal strengths (density) of the keywords in the documents that were analyzed [55,56]. The *Biblioshiny* package has been used by many authors, including Wilczewski and Alon [55], Di Cosmo et al

[56], and Bretas and Alon [57], to derive a conceptual thematic map for a specific research topic. We used the Keyword Plus feature in *Biblioshiny* based on 250 units of keywords, with a minimum cluster frequency of 5 per 1000 document units and 3 labels per cluster. To ensure a meaningful visualization, basic keywords such as "human," "female," "male," "article," and "aged" were removed from the set of keywords that were used by *Biblioshiny* to generate the conceptual thematic map. The following words were also merged with their synonyms: (1) "Coronavirus disease 2019," "coronavirus infection," "coronavirus infections," "coronaviruses," and "sars-cov-2" were merged with "Covid-19"; (2) "Covid-19 vaccines," "vaccination," "sars-cov-2 vaccine," "vaccine," and "vaccines" were merged with "covid-19 vaccine"; and (3) "Pandemics" was merged with "pandemic."

As illustrated in Figure 7, there are 4 quadrants on the thematic map [55,56] (Textbox 1).

Figure 7. Biblioshiny visualization of thematic mapping of keywords related to the role of social media in health misinformation and disinformation during the COVID-19 pandemic.



Textbox 1. The 4 quadrants on the thematic map.

- Topics in the motor themes (the upper right quadrant) have high density and centrality. Topics in this quadrant are seen as the mainstream topics in a specific field of research.
- Topics in the niche themes (the upper left quadrant) have high density but low centrality. This implies that topics in this quadrant are seen as specialized topics in a field of research.
- Topics in the basic themes (the lower right quadrant) have low density but high centrality. This implies that topics in this quadrant are general themes in a field of research.
- Topics in the emerging or declining themes (the lower left quadrant) have low density and centrality. This implies that topics in this quadrant are weakly developed themes in a field of research.

As shown in Figure 7, there are 2 clusters of *motor themes*. The first is "pandemic," "communication," and "public health," whereas the second is "prevention and control," "trust," and "fear."

There are 2 clusters of *basic themes*. The first cluster is "machine learning," "natural language processing," and "social network analysis." This could imply that an increasing number of publications included in the bibliometric analysis used these techniques to detect health misinformation and disinformation on social media platforms during the COVID-19 pandemic. As expected, "Covid-19," "social media," and "misinformation" were the dominant topics in the second cluster of the *basic themes* quadrant. This is not a surprise given that these topics are the focus of this bibliometric analysis.

The *niche themes* quadrant also consists of 2 clusters. The first is “fake detection,” “learning systems,” and “learning algorithms,” whereas the second cluster has only 1 item, “learning models.” As stated earlier, topics that appear in the *niche themes* quadrant are generally seen as specialized topics in a research area [55,56]. This could imply that these topics and techniques were not widely used by the authors of the documents included in the bibliometric analysis.

Finally, the *emerging or declining themes* quadrant also has 2 clusters of themes. The first is “social networking (online),” “sentiment analysis,” and “social media platforms.” The second is “fake news,” “health risks,” and “deep learning.”

Discussion

Principal Findings

The bibliometric analysis of studies related to the role of social media in health misinformation and disinformation since the start of the COVID-19 pandemic revealed a high level of publications by authors from countries worldwide, with 4038 authors involved in the 943 documents analyzed. However, most of the publications were by authors from high-income countries. The study also showed a high level of global collaboration, with the highest level between authors from the United States and those from other countries. The study showed that the higher the number of documents published through collaborations, the higher the total number of citations of the documents. This implies that a high level of collaboration can increase the productivity of researchers and the citation counts of their publications. As reported by Abramo et al [58] and Ceballos et al [59], collaboration between researchers has a positive influence on their productivity.

Keywords such as “Covid-19,” “social media,” “misinformation,” and “infodemic” featured prominently in the 943 documents that were analyzed. The prominence of these keywords was no surprise given the focus of the bibliometric analysis. The study revealed important keyword themes that are related to the role of social media in health misinformation and disinformation since the outbreak of COVID-19. Basic topics that included “Covid-19,” “social media,” and “misinformation” were revealed. The analysis also showed that techniques such as “machine learning,” “social network analysis,” and “natural language processing” were used by researchers to detect health misinformation and disinformation on social media platforms during the COVID-19 pandemic. In addition, niche topics such as “learning systems,” “learning models,” and “learning algorithms” were revealed through the *Biblioshiny* thematic mapping of keywords. According to Wilczewski and Alon [55] and Di Cosmo et al [56], terms or topics that appear in the *niche themes* quadrant of *Biblioshiny*’s thematic map are seen as specialized topics in a research area.

The bibliometric analysis showed that the publication by Puri et al [37], “Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases,” was the document with the highest number of citations (n=522). This was followed by the publication by Romer and Jamieson [38] titled “Conspiracy theories as barriers to controlling the spread of COVID-19 in the US,” cited 491 times. The paper by Islam et al [39], titled “COVID-19-related infodemic and its impact on public health: A global social media analysis,” was the third most cited with 461 citations. These top 3 most cited papers were published by authors from high-income countries (Canada, the United States, and the United Kingdom), Asia, and the Middle East. In addition, *JMIR* as an umbrella journal had the highest number of publications.

The research reported in this paper is similar to that of Yeung et al [32] in that both studies used the bibliometric analysis method. These 2 studies are also related to social media and health misinformation and disinformation. However, this study is different in that it is specifically concerned with the analysis of research papers that focused on social media and health misinformation and disinformation during the COVID-19 pandemic. In addition, the *Biblioshiny* tool that we used is a methodological contribution to additional tools that can be used by researchers when conducting bibliometric analysis.

Although the worst of the COVID-19 pandemic appears to be behind us, similar outbreaks could reoccur. For example, an outbreak of monkeypox was reported in May 2022 [60]. The findings of this bibliometric analysis can be leveraged by other researchers through studies that investigate how niche topic areas (“learning systems,” “learning algorithms,” and “learning models”) can be used to analyze the influence of social media on health misinformation and disinformation during periods of global public health emergencies.

Limitations

The results of the bibliometric analysis reported in this paper are based on the documents retrieved from the Scopus database on June 8, 2023. As such, the validity of the results is only based on the sources that were analyzed. Other relevant studies that were not available through the Scopus database could have been excluded. Hence, the results are not necessarily a full representation of studies related to the role of social media in health misinformation and disinformation since the outbreak of the COVID-19 pandemic.

Conclusions

This bibliometric analysis aimed to understand scientific productivity on topics related to the role of social media in health misinformation and disinformation since the start of the COVID-19 pandemic. A total of 943 publications between 2020 and June 8, 2023, were included in the bibliometric analysis. This study revealed a number of key findings. Our analysis showed that the highest number of publications related to the role of social media in health misinformation and disinformation since the start of the COVID-19 pandemic was 387 in 2022. We also found that *JMIR* was the most productive publisher with a combined 10.9% (103/943) of the publications. The results of our analysis also showed that “Covid-19,” “social media,” and “misinformation” were the top 3 keywords, with 566, 342, and 277 occurrences, respectively. This was corroborated by the thematic mapping of keywords through *Biblioshiny*, which identified “Covid-19,” “social media,” and “misinformation” as one of the clusters of topics of the basic themes related to the role of social media in health misinformation and disinformation during the COVID-19 outbreak. In addition, the terms “learning systems,” “learning models,” and “learning algorithms” were revealed as niche topics on the role of social media in health misinformation and disinformation during the COVID-19 outbreak.

Finally, the study revealed that authors from the United States had the highest number of collaborations with authors from different parts of the world in publishing research papers related to the role of social media in health misinformation and disinformation during the COVID-19 outbreak, whereas the research paper by Puri et al [37] titled “Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases” had the highest number of citations at 522.

Future Research

Niche topics offer opportunities for researchers in new areas that can be exploited to analyze the influence of social media on health misinformation and disinformation during periods of global public health emergencies.

Acknowledgments

The research reported in this paper was supported by the South African Department of Science and Innovation and the South African Medical Research Council under the Brazil, Russia, India, China, and South Africa Joint Application Form 2020/33. The content and findings reported or illustrated herein are the sole deduction, views, and responsibility of the researchers and do not reflect the official position and sentiments of the funders.

Authors' Contributions

FA contributed to the conceptualization, method, results, and proofreading of the paper. HS contributed to the overall quality assurance, conclusion, and proofreading of the paper. TM contributed to the introduction and proofreading of the paper. GM contributed to the discussion and proofreading of the paper. MH contributed to the abstract, limitations, and proofreading of the paper. All authors contributed to the data acquisition. All the authors gave consent for the paper to be submitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data source.
[XLSX File (Microsoft Excel File), 5058 KB - infodemiology_v3i1e48620_app1.xlsx]

References

1. Kross E, Verduyn P, Sheppes G, Costello CK, Jonides J, Ybarra O. Social media and well-being: pitfalls, progress, and next steps. Trends Cogn Sci 2021 Jan;25(1):55-66 [FREE Full text] [doi: 10.1016/j.tics.2020.10.005] [Medline: 33187873]
2. Kemp S. Digital 2023: global overview report. DataReportal. 2023. URL: https://datareportal.com/reports/digital-2023-global-overview-report [accessed 2023-04-14]
3. 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]
4. Naeem SB, Bhatti R, Khan A. An exploration of how fake news is taking over social media and putting public health at risk. Health Info Libr J 2021 Jun;38(2):143-149 [FREE Full text] [doi: 10.1111/hir.12320] [Medline: 32657000]
5. Pang PC, Cai Q, Jiang W, Chan KS. Engagement of government social media on Facebook during the COVID-19 pandemic in Macao. Int J Environ Res Public Health 2021 Mar 28;18(7):3508 [FREE Full text] [doi: 10.3390/ijerph18073508] [Medline: 33800621]
6. Górska A, Dobija D, Grossi G, Staniszewska Z. Getting through COVID-19 together: understanding local governments' social media communication. Cities 2022 Feb;121:103453 [FREE Full text] [doi: 10.1016/j.cities.2021.103453] [Medline: 34566232]
7. Chan AK, Nickson CP, Rudolph JW, Lee A, Joynt GM. Social media for rapid knowledge dissemination: early experience from the COVID-19 pandemic. Anaesthesia 2020 Dec;75(12):1579-1582 [FREE Full text] [doi: 10.1111/anae.15057] [Medline: 32227594]
8. Chen S, Xiao L, Kumar A. Spread of misinformation on social media: what contributes to it and how to combat it. Comput Human Behav 2023 Apr;141:107643 [FREE Full text] [doi: 10.1016/j.chb.2022.107643]
9. Scardigno R, Paparella A, D’Errico F. Faking and conspiring about COVID-19: a discursive approach. Qual Rep 2023 May 01;28(1):49-68 [FREE Full text] [doi: 10.46743/2160-3715/2023.5660]
10. Lazer DM, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, et al. The science of fake news. Science 2018 Mar 09;359(6380):1094-1096. [doi: 10.1126/science.aao2998] [Medline: 29590025]

11. Posetti J, Matthews A. A short guide to the history of 'fake news' and disinformation. International Center for Journalists. 2018. URL: https://en.wikisource.org/wiki/A_short_guide_to_the_history_of_%27fake_news%27_and_disinformation [accessed 2023-07-12]
12. Wardle C, Derakhshan H. Information disorder: toward an interdisciplinary framework for research and policy making. Council of Europe. 2017. URL: https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html [accessed 2023-07-12]
13. Ghani NA, Hamid S, Targio Hashem IA, Ahmed E. Social media big data analytics: a survey. Comput Human Behav 2019 Dec;101:417-428 [FREE Full text] [doi: 10.1016/j.chb.2018.08.039]
14. 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]
15. Tandoc ECJ. The facts of fake news: a research review. Sociol Compass 2019 Jul 25;13(9):e12724 [FREE Full text] [doi: 10.1111/soc4.12724]
16. Tandoc ECJ, Lim ZW, Ling R. Defining “fake news”. Digit Journal 2018;6(2):137-153 [FREE Full text] [doi: 10.1080/21670811.2017.1360143]
17. Sylvia Chou W, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. Am J Public Health 2020 Oct;110(S3):S273-S275 [FREE Full text] [doi: 10.2105/ajph.2020.305905]
18. Crouse L, Dupuis M. A dangerous infodemic: an examination of the impact social media misinformation has on COVID-19 vaccination status. In: Proceedings of the 23rd Annual Conference on Information Technology Education. 2022 Presented at: SIGITE '22; September 21-24, 2022; Chicago, IL p. 37-43 URL: https://dl.acm.org/doi/abs/10.1145/3537674.3554754 [doi: 10.1145/3537674.3554754]
19. Soltaninejad K. Methanol mass poisoning outbreak: a consequence of COVID-19 pandemic and misleading messages on social media. Int J Occup Environ Med 2020 Jul 10;11(3):148-150 [FREE Full text] [doi: 10.34172/ijoem.2020.1983] [Medline: 32225177]
20. Fieselmann J, Annac K, Erdsiek F, Yilmaz-Aslan Y, Brzoska P. What are the reasons for refusing a COVID-19 vaccine? A qualitative analysis of social media in Germany. BMC Public Health 2022 Apr 27;22(1):846-848 [FREE Full text] [doi: 10.1186/s12889-022-13265-y] [Medline: 35484619]
21. 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 May;5(5):e002604 [FREE Full text] [doi: 10.1136/bmjgh-2020-002604] [Medline: 32409327]
22. Lelisho ME, Pandey D, Alemu BD, Pandey BK, Tareke SA. The negative impact of social media during COVID-19 pandemic. Trends Psychol 2022 May 16;31(1):123-142 [FREE Full text] [doi: 10.1007/s43076-022-00192-5]
23. Jamison AM, Broniatowski DA, Dredze M, Sangraula A, Smith MC, Quinn SC. Not just conspiracy theories: vaccine opponents and proponents add to the COVID-19 'infodemic' on Twitter. Harv Kennedy Sch Misinformation Rev 2020 Sep 9;1:37016 [FREE Full text] [doi: 10.37016/mr-2020-38] [Medline: 34368805]
24. Managing the COVID-19 infodemic: promoting healthy behaviours and mitigating the harm from misinformation and disinformation. World Health Organization. 2020 Sep 23. 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 2023-04-04]
25. Infodemic. World Health Organization. URL: https://www.who.int/health-topics/infodemic#tab=tab_1 [accessed 2023-04-04]
26. 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]
27. Toquero CM. Addressing infodemic through the comprehensive competency framework of media and information literacy. J Public Health (Oxf) 2023 Mar 14;45(1):e132-e133 [FREE Full text] [doi: 10.1093/pubmed/fdab412] [Medline: 35151229]
28. Suarez-Lledo V, Alvarez-Galvez J. Assessing the role of social bots during the COVID-19 pandemic: infodemic, disagreement, and criticism. J Med Internet Res 2022 Aug 25;24(8):e36085 [FREE Full text] [doi: 10.2196/36085] [Medline: 35839385]
29. Weng Z, Lin A. Public opinion manipulation on social media: social network analysis of Twitter bots during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Dec 07;19(24):16376 [FREE Full text] [doi: 10.3390/ijerph192416376] [Medline: 36554258]
30. Himelboim I, Borah P, Lee DK, Lee J, Su Y, Vishnevskaya A, et al. What do 5G networks, Bill Gates, Agenda 21, and QAnon have in common? Sources, distribution, and characteristics. New Media Soc (Forthcoming) 2023 Jan 04 [FREE Full text] [doi: 10.1177/14614448221142800]
31. Riccaboni M, Verginer L. The impact of the COVID-19 pandemic on scientific research in the life sciences. PLoS One 2022 Feb 9;17(2):e0263001 [FREE Full text] [doi: 10.1371/journal.pone.0263001] [Medline: 35139089]
32. Yeung AW, Tosevska A, Klager E, Eibensteiner F, Tsagkaris C, Parvanov ED, et al. Medical and health-related misinformation on social media: bibliometric study of the scientific literature. J Med Internet Res 2022 Jan 25;24(1):e28152 [FREE Full text] [doi: 10.2196/28152] [Medline: 34951864]
33. Aria M, Cuccurullo C. bibliometrix : an R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975 [FREE Full text] [doi: 10.1016/j.joi.2017.08.007]

34. van Eck NJ, Waltman L. VOSviewer manual: manual for VOSviewer version 1.6.19. The Centre for Science and Technology Studies. 2023. URL: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.19.pdf [accessed 2023-03-17]
35. Biblioshiny: the shiny app for no coders. Bibliometrix. 2023. URL: https://www.bibliometrix.org/home/index.php/layout/biblioshiny [accessed 2023-04-10]
36. Leung XY, Sun J, Bai B. Bibliometrics of social media research: a co-citation and co-word analysis. Int J Hosp Manag 2017 Sep;66:35-45 [FREE Full text] [doi: 10.1016/j.ijhm.2017.06.012]
37. Puri N, Coomes EA, Haghbayan H, Gunaratne K. Social media and vaccine hesitancy: new updates for the era of COVID-19 and globalized infectious diseases. Hum Vaccin Immunother 2020 Nov 01;16(11):2586-2593 [FREE Full text] [doi: 10.1080/21645515.2020.1780846] [Medline: 32693678]
38. Romer D, Jamieson KH. Conspiracy theories as barriers to controlling the spread of COVID-19 in the U.S. Soc Sci Med 2020 Oct;263:113356 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113356] [Medline: 32967786]
39. Islam MS, Sarkar T, Khan SH, Mostofa Kamal A, Hasan SM, Kabir A, et al. COVID-19-related infodemic and its impact on public health: a global social media analysis. Am J Trop Med Hyg 2020 Oct;103(4):1621-1629 [FREE Full text] [doi: 10.4269/ajtmh.20-0812] [Medline: 32783794]
40. 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]
41. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 May 06;22(5):e19458 [FREE Full text] [doi: 10.2196/19458] [Medline: 32352383]
42. Tasnim S, Hossain MM, Mazumder H. Impact of rumors and misinformation on COVID-19 in social media. J Prev Med Public Health 2020 May;53(3):171-174 [FREE Full text] [doi: 10.3961/jpmph.20.094] [Medline: 32498140]
43. Allington D, Duffy B, Wessely S, Dhavan N, Rubin J. Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. Psychol Med 2021 Jul;51(10):1763-1769 [FREE Full text] [doi: 10.1017/S003329172000224X] [Medline: 32513320]
44. Ahmad AR, Murad HR. The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. J Med Internet Res 2020 May 19;22(5):e19556 [FREE Full text] [doi: 10.2196/19556] [Medline: 32369026]
45. Hua J, Shaw R. Corona virus (COVID-19) "infodemic" and emerging issues through a data lens: the case of China. Int J Environ Res Public Health 2020 Mar 30;17(7):2309 [FREE Full text] [doi: 10.3390/ijerph17072309] [Medline: 32235433]
46. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. The Lancet Digital Health 2021 Mar;3(3):e175-e194 [FREE Full text] [doi: 10.1016/s2589-7500(20)30315-0]
47. Laato S, Islam AK, Islam MN, Whelan E. What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? Eur J Inf Syst 2020;29(3):288-305 [FREE Full text] [doi: 10.1080/0960085x.2020.1770632]
48. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. J Med Internet Res 2020 Jun 26;22(6):e19659 [FREE Full text] [doi: 10.2196/19659] [Medline: 32558655]
49. van der Linden S, Roozenbeek J, Compton J. Inoculating against fake news about COVID-19. Front Psychol 2020 Oct 23;11:566790 [FREE Full text] [doi: 10.3389/fpsyg.2020.566790] [Medline: 33192844]
50. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat Hum Behav 2020 Dec 29;4(12):1285-1293. [doi: 10.1038/s41562-020-00994-6] [Medline: 33122812]
51. Islam MS, Kamal AM, 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 May 12;16(5):e0251605 [FREE Full text] [doi: 10.1371/journal.pone.0251605] [Medline: 33979412]
52. Malecki KM, Keating JA, Safdar N. Crisis communication and public perception of COVID-19 risk in the era of social media. Clin Infect Dis 2021 Feb 16;72(4):697-702 [FREE Full text] [doi: 10.1093/cid/ciaa758] [Medline: 32544242]
53. Jennings W, Stoker G, Bunting H, Valgarðsson VO, Gaskell J, Devine D, et al. Lack of trust, conspiracy beliefs, and social media use predict COVID-19 vaccine hesitancy. Vaccines (Basel) 2021 Jun 03;9(6):593 [FREE Full text] [doi: 10.3390/vaccines9060593] [Medline: 34204971]
54. Barua Z, Barua S, Aktar S, Kabir N, Li M. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. Prog Disaster Sci 2020 Dec;8:100119 [FREE Full text] [doi: 10.1016/j.pdisas.2020.100119] [Medline: 34173443]
55. Wilczewski M, Alon I. Language and communication in international students' adaptation: a bibliometric and content analysis review. High Educ (Dordr) 2023 Jul 12;85(6):1235-1256 [FREE Full text] [doi: 10.1007/s10734-022-00888-8] [Medline: 35855684]
56. Di Cosmo A, Pinelli C, Scandurra A, Aria M, D'Aniello B. Research trends in Octopus biological studies. Animals (Basel) 2021 Jun 17;11(6):1808 [FREE Full text] [doi: 10.3390/ani11061808] [Medline: 34204419]
57. Bretas VP, Alon I. Franchising research on emerging markets: bibliometric and content analyses. J Bus Res 2021 Sep;133:51-65 [FREE Full text] [doi: 10.1016/j.jbusres.2021.04.067]

58. Abramo G, D'Angelo AC, Murgia G. The relationship among research productivity, research collaboration, and their determinants. J Informetr 2017 Nov;11(4):1016-1030 [FREE Full text] [doi: 10.1016/j.joi.2017.09.007]
59. Ceballos HG, Fangmeyer JJ, Galeano N, Juarez E, Cantu-Ortiz FJ. Impelling research productivity and impact through collaboration: a scientometric case study of knowledge management. Knowl Manag Res Pract 2017;15(3):346-355 [FREE Full text] [doi: 10.1057/s41275-017-0064-8]
60. WHO. Mpox (monkeypox) outbreak. World Health Organization. 2022. URL: https://www.who.int/emergencies/situations/monkeypox-oubreak-2022 [accessed 2023-04-20]

Abbreviations

TLS: total link strength

Edited by T Mackey; submitted 01.05.23; peer-reviewed by AWK Yeung, W Ceron; comments to author 30.05.23; revised version received 27.07.23; accepted 11.08.23; published 20.09.23.

Please cite as:
Adebesin F, Smuts H, Mawela T, Maramba G, Hattingh M
The Role of Social Media in Health Misinformation and Disinformation During the COVID-19 Pandemic: Bibliometric Analysis
JMIR Infodemiology 2023;3:e48620
URL: https://infodemiology.jmir.org/2023/1/e48620
doi:10.2196/48620
PMID:37728981

Original Paper

Content Quality of YouTube Videos About Pain Management After Cesarean Birth: Content Analysis

Natalie A Squires[1], MD; Elizabeth Soyemi[2,3]; Lynn M Yee[3], MPH, MD; Eleanor M Birch[4], MPH, MD; Nevert Badreldin[3], MS, MD

[1]Department of Obstetrics and Gynecology, New York Presbyterian Hospital Weill Cornell Medical Center, New York, NY, United States

[2]Illinois Math and Science Academy, Aurora, IL, United States

[3]Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[4]Department of Emergency Medicine, Madigan Army Medical Center, Joint Base Lewis-McChord, WA, United States

Corresponding Author:
Nevert Badreldin, MS, MD
Division of Maternal-Fetal Medicine
Department of Obstetrics and Gynecology
Northwestern University Feinberg School of Medicine
Suite 05-2149
250 E. Superior Street
Chicago, IL, 60611
United States
Phone: 1 312 472 4685
Fax: 1 312 472 0511
Email: nevert.badreldin@northwestern.edu

Abstract

Background: YouTube is an increasingly common source of health information; however, the reliability and quality of the information are inadequately understood. Several studies have evaluated YouTube as a resource during pregnancy and found the available information to be of poor quality. Given the increasing attention to postpartum health and the importance of promoting safe opioid use after birth, YouTube may be a source of information for birthing individuals. However, little is known about the available information on YouTube regarding postpartum pain.

Objective: The purpose of this study is to systematically evaluate the quality of YouTube videos as an educational resource for postpartum cesarean pain management.

Methods: A systematic search of YouTube videos was conducted on June 25, 2021, using 36 postpartum cesarean pain management–related keywords, which were identified by clinical experts. The search replicated a default YouTube search via a public account. The first 60 results from each keyword search were reviewed, and unique videos were analyzed. An overall content score was developed based on prior literature and expert opinion to evaluate the video's relevance and comprehensiveness. The DISCERN instrument, a validated metric to assess consumer health information, was used to evaluate the reliability of video information. Videos with an overall content score of ≥5 and a DISCERN score of ≥39 were classified as high-quality health education resources. Descriptive analysis and intergroup comparisons by video source and quality were conducted.

Results: Of 73 unique videos, video sources included medical videos (n=36, 49%), followed by personal video blogs (vlogs; n=32, 44%), advertisements (n=3, 4%), and media (n=2, 3%). The average overall content score was 3.6 (SD 2.0) out of 9, and the average DISCERN score was 39.2 (SD 8.1) out of 75, indicating low comprehensiveness and fair information reliability, respectively. High-quality videos (n=22, 30%) most frequently addressed overall content regarding pain duration (22/22, 100%), pain types (20/22, 91%), return-to-activity instructions (19/22, 86%), and nonpharmacologic methods for pain control (19/22, 86%). There were differences in the overall content score (P=.02) by video source but not DISCERN score (P=.45). Personal vlogs had the highest overall content score at 4.0 (SD 2.1), followed by medical videos at 3.3 (SD 2.0). Longer video duration and a greater number of comments and likes were significantly correlated with the overall content score, whereas the number of video comments was inversely correlated with the DISCERN score.

Conclusions: Individuals seeking information from YouTube regarding postpartum cesarean pain management are likely to encounter videos that lack adequate comprehensiveness and reliability. Clinicians should counsel patients to exercise caution when using YouTube as a health information resource.

(JMIR Infodemiology 2023;3:e40802) doi:10.2196/40802

KEYWORDS

health information; internet; YouTube; cesarean section; cesarean; C-section; postpartum; social media; web-based video; maternal; postnatal; pain; systematic search; patient education; information quality; accuracy; credibility; health education; educational video; education resource; health video

Introduction

YouTube is a frequently visited website in the United States and a common source of eHealth information [1,2]. As an alternative to written communication, YouTube provides an opportunity to narrow the health literacy gap if quality health information is presented clearly and comprehensively [3]. Indeed, studies have demonstrated that some patients prefer video over written sources of medical information [4]. However, accessing YouTube for health information remains problematic, as there are few regulations governing the information available.

Recent studies have evaluated YouTube as a source of health information during pregnancy. Chandrasekaran et al [5] evaluated the use of various social media platforms as a resource for the Zika virus. The authors found that while YouTube provided a similar number of informative results when compared to other platforms, it also included a higher number of outdated and misleading results, including hoax messages and conspiracy theories [5]. Similarly, YouTube videos discussing medication use in pregnancy were found to have inconsistent or inadequate safety information [6].

Pain is a significant concern among postpartum individuals [7,8]. Inadequately controlled pain in the early postpartum period increases individuals' risk of experiencing persistent pain, depressive symptoms, and opioid abuse [9,10]. As such, practice guidelines make a strong recommendation for patient education and antenatal counseling regarding postpartum pain management protocols to optimize their recovery [11]. However, the optimal mode and content of this counseling have not been established.

Given the unique challenges of the early postpartum period, individuals may use internet resources to address concerns related to their postcesarean birth pain and recovery. While the growing popularity of YouTube has the possibility to improve access to postpartum care and the postpartum pain experience, there are limited data evaluating the quality of available resources for postpartum pain management. Thus, the purpose of this study was to systematically evaluate the quality of YouTube videos on postpartum cesarean recovery.

Methods

Search Strategy

A systematic search of YouTube videos was conducted using postpartum cesarean pain management–related keywords on June 25, 2021. Search terms were identified by expert consensus and expanded using Google Trends to identify related searches. The final terms included 36 iterations of the search "postpartum cesarean pain" (Multimedia Appendix 1).

To duplicate a public search, the search was performed in incognito mode in a cache-cleared browser, and no registered account was used. Search results were sorted by relevance, which is the default setting for YouTube searches. The first 60 results from each keyword search were collected, and duplicates were omitted. This sort of strategy and sample size were selected based on data showing that 83% of searchers will not view more than three web pages of results [12]. Videos were excluded if the full video was unavailable, was >30 minutes in duration, or was in a language other than English. Video duration was capped based on research showing that web search queries for adults were on average 18 minutes in duration, and thus longer videos are unlikely to be viewed by the general public [13]. The remaining videos were assessed for inclusion by screening the video titles, comments, and channels for terms related to postcesarean pain. If any uncertainty remained, videos less than 10 minutes in duration were watched in their entirety. If videos were longer than 10 minutes, the first 10 minutes were watched, and the reviewer reviewed additional time stamps or sections indicating a shift in content to verify eligibility. This process was designed to mirror that of a traditional systematic review, wherein a sample of the content (ie, abstracts) is initially reviewed to determine relevance prior to the review of the full content. Videos were also excluded if the content was unrelated to cesarean delivery, postpartum pain management, or recovery (ie, if the overall content score was 0, as described below).

Data Extraction

Descriptive characteristics of each video were gathered, including the date posted, video length, number of comments, likes, dislikes, and channel subscribers. Values that accumulate over time were collected within one day (July 12, 2021) by a single reviewer to minimize variability. Video source and presenter characteristics were also collected. Video sources were categorized as personal video blogs (vlogs), medical or hospital videos, advertisements, and media. The source was determined based on the affiliation of the video author and the channel description, when applicable. Videos were labeled as vlogs when the video author had an independent channel describing their personal experience and recommendations. Medical or hospital videos were differentiated by a clear affiliation with a hospital or medical service company. Video bloggers who identified as medical professionals on their independent channels were characterized as personal vloggers. Advertisement videos differed from medical videos in that they clearly described the benefits of a single product in the postpartum period. Media videos included news clips and talk show interviews. Videos were labeled as character videos if a specific, identifiable person presented the information. Presenter characteristics were identified when applicable, and reviewers subjectively identified the presenter's gender, race or ethnicity, and age.

Content Analysis

Two content scores were developed using the expert opinions of maternal-fetal medicine specialists (NB and LMY) in conjunction with American College of Obstetrician and Gynecologists guidelines regarding pain management [14]. Both scores were used to evaluate the video's relevance and comprehensiveness as a health education resource. The first, an "overall content score," included nine topics relevant to postcesarean pain management: (1) pain duration, (2) pain types, (3) when to notify a clinician, (4) activity recommendations, (5) pain medication timing, (6) multimodal pharmacologic methods, (7) nonpharmacologic methods, (8) maternal risks of treatments, and (9) risks to newborns. Second, given growing awareness regarding opioid use in the postpartum period, a second "opioid content score" was used to evaluate the comprehensiveness with regard to opioid use in postpartum pain management. This was scored based on the following nine topics: (1) addressing opioid use, (2) when to use, (3) limitations of use, (4) general maternal risks of treatment, (5) risk of addiction, (6) risks to newborns, (7) length of use, (8) discharge instructions, and (9) disposal of remaining tablets. For each of the content scores, one point was awarded if a topic was mentioned, for a total possible score of 9. Higher content scores indicated greater comprehensiveness in the video. Similar content assessments have been used in prior studies to evaluate YouTube as a health information resource [15,16].

DISCERN Analysis

The DISCERN instrument was used to assess the quality and reliability of the videos as an information source. This tool has been widely used to evaluate web-based sources of health information, including YouTube videos [16-21]. Studies have demonstrated that the DISCERN tool enables both professionals and consumers to distinguish between high- and low-quality sources of health information [19,22]. The DISCERN instrument consists of 15 questions plus an overall quality rating to assess consumer health information on treatment choices. The first 8 questions address the reliability or trustworthiness of a source, followed by 7 questions evaluating whether consumers had access to detailed information regarding their treatment options. Questions are rated on a Likert scale of 1-5, with a score of 1 indicating the criterion was not satisfied and a score of 5 indicating the criterion was fully satisfied. Specific guidelines on the application scoring criteria are provided via the Online Discern Tool [23]. Like prior studies, we report the DISCERN score as a sum of the first 15 questions, and the score was interpreted with established categories describing source reliability: excellent (63-75 points), good (51-62), fair (39-50), poor (28-38), or very poor (≤27) [20,21,24].

Quality Analysis

A combination of the overall content score and the DISCERN score was used to establish video quality as a health education resource that is both comprehensive and reliable. Videos with an overall content score of ≥5 and DISCERN score of ≥39 were classified as high-quality. These criteria were chosen as a DISCERN score of ≥39 indicates at least fair information reliability, and an overall content score of ≥5 indicates that greater than half of the content criteria were met.

Consensus regarding the application of scoring criteria was obtained through a collaborative review of 3 videos among 3 authors (NS, ES, and NB). Subsequently, the application of the scoring criteria was tested via an independent review of 10 videos. The average DISCERN scoring disparity was 0.18 points. Intraclass correlation was 0.76 and interclass correlation was 0.81, indicating good interrater reliability. Areas of discordance were resolved by team discussion. The remaining videos were divided and scored by authors NS or ES. Data were extracted and stored using REDCap (Research Electronic Data Capture; Vanderbilt University) software.

Statistical Analysis

All statistical analysis was performed using Excel software (version 16.56, Microsoft Corp). Interrater agreement was analyzed by intraclass correlation coefficients and a single-factor ANOVA. Video characteristics were analyzed via descriptive statistics. Associations among video source, quality, and descriptive characteristics were evaluated using nonparametric correlations. A *P* value of less than .05 was considered significant.

Ethics Approval

This study does not involve human subject research and was deemed exempt by Northwestern University's institutional review board (reference number: STU00214706).

Results

Video Characteristics

A total of 233 unique videos were identified. Following the application of inclusion and exclusion criteria, 73 videos remained for analysis (Figure 1, Multimedia Appendix 2). Most videos (69/73, 95%) were character videos. Among these, most presenters appeared to be female (63/69, 91%), of reproductive age (56/69, 81%), and non-Hispanic White race or ethnicity (39/69, 57%). Video sources were most commonly medical videos (36/73, 49%), followed by personal vlogs (32/73, 44%), advertisements (3/73, 4%), and media (2/73, 3%; Table 1). The median length of videos was 8.03 minutes and they were uploaded for a median of 1230 days at the time of access (Table 1).

Figure 1. YouTube video selection regarding postcesarean pain management. The figure illustrates a flow diagram of the identification, selection, and exclusion of YouTube videos. The first 60 video titles for 36 unique search terms were collected for a total of 2160 videos. 73 videos were included for the final analysis.

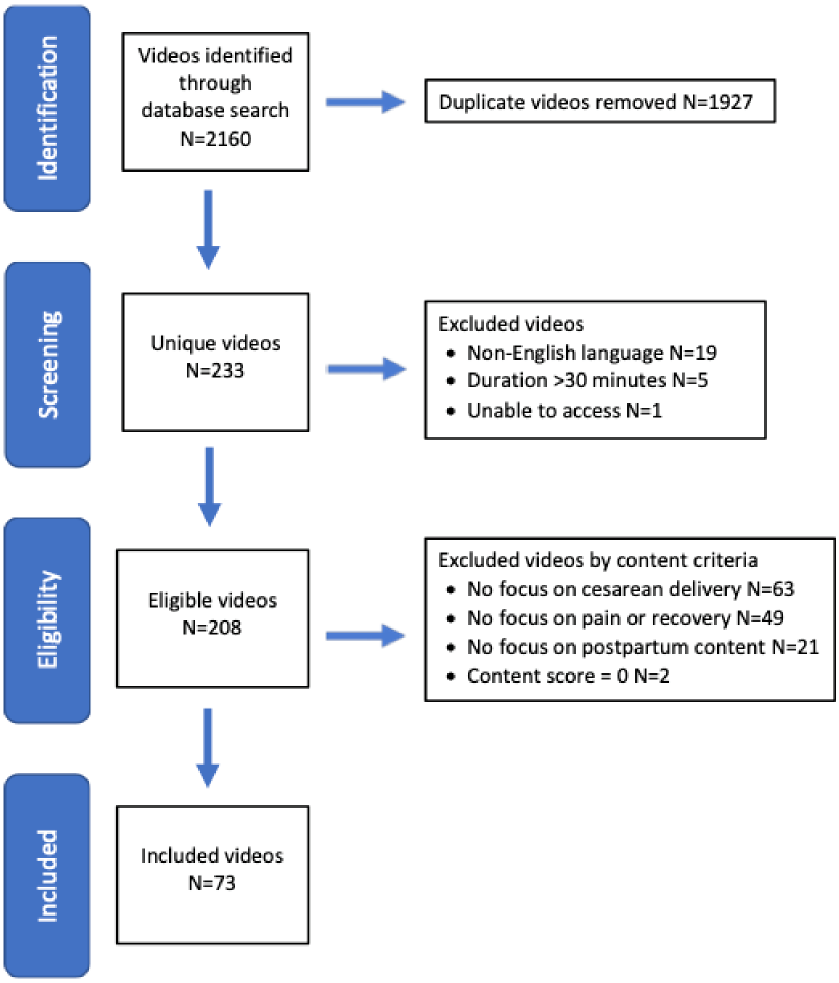


Table 1. Characteristics of videos included in review (N=73).

Characteristics	Values
Character video, n (%)	69 (95)
Presenter age, n (%)	
Adolescent	1 (1)
Reproductive age	58 (84)
Older adult	10 (14)
Presenter race or ethnicity, n (%)[a]	
Black	3 (4)
White	44 (64)
Latinx	1 (1)
Asian	17 (25)
Undetermined	4 (6)
Video source, n (%)	
Medical	36 (49)
Personal vlog	32 (44)
Advertisement	3 (4)
Media	2 (3)
Video characteristics, median (IQR)	
Days since post	817 (551-1853)
Duration (minutes)	7.8 (3.0-13.7)
Views	40614 (6841-82748)
Comments	15 (0-54)
Likes	213 (53-902)
Dislike	11 (3-34)
Channel subscriber number	35100 (9020-177000)

[a]Presenter age and race or ethnicity were subjectively assigned.

Content Analysis

Regarding the overall content score, videos most frequently covered the expected duration of pain (50/73, 68%), different types of pain (44/73, 60%), and return to activity (44/73, 60%), whereas information on when to use medication (14/73, 19%) and risks to the newborn (6/73, 8%) were less frequently included (Figure 2). The mean overall content score was 3.6 (SD 2.0) out of 9. The overall content score significantly differed by the video source (*P*=.02). Personal vlog videos had the highest overall content score at 4.0 (SD 2.1), followed by medical videos at 3.3 (SD 2.0; Table 2).

Most videos (57/73, 78%) did not specifically address opioids and, therefore, had an opioid content score of 0. For those videos that did address opioids (16/73, 22%), videos most often covered maternal risks (9/16, 56%), limitations of opioids (7/16, 44%), and when to use opioids (6/16, 38%). Videos rarely discuss the risk of addiction (2/16, 13%), the recommended duration of use (1/16, 6%), or proper opioid disposal (0/16, 0%). For videos that addressed opioids, the mean opioid content score was 3.1 (SD 1.6) out of 9. There was no difference in opioid content score by video source (*P*=.77; Table 3).

Figure 2. YouTube video content inclusion by topic area. The figure illustrates the overall content score by topic area. The y-axis demonstrates the percentage of total videos covering each of the 9 total topic areas. (A) The percentage of videos covering each topic area from all videos. (B) The percentage of videos by quality designation covering each topic area.

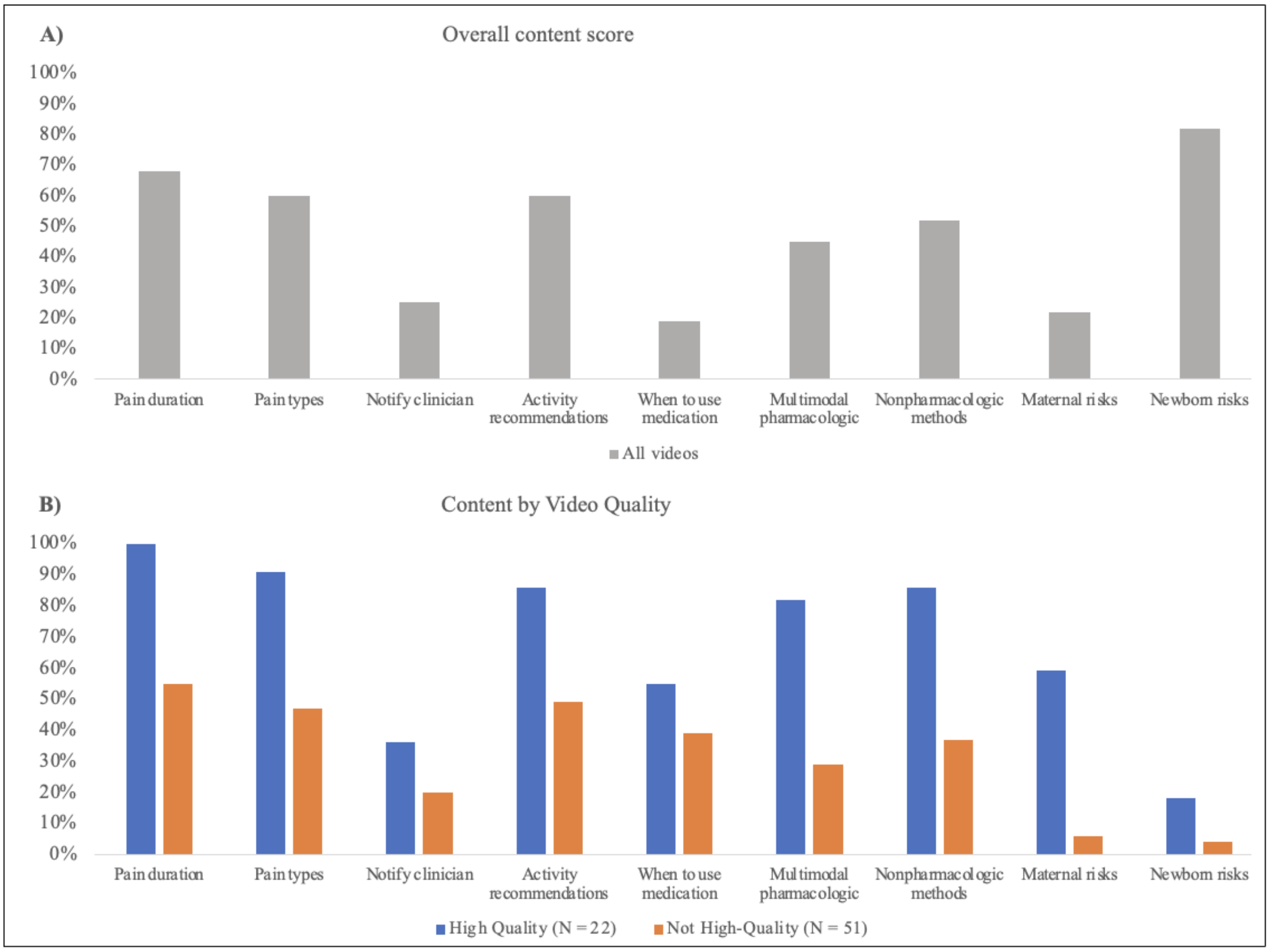


Table 2. Quality of postpartum pain management videos on YouTube by video source.[a,b]

Video source	Overall content score		DISCERN score	
	Mean (SD)	*P* value	Mean (SD)	*P* value
Total (N=73)	3.6 (2.0)	N/A[c]	3.2 (8.1)	N/A
Medical (n=36)	3.3 (2.0)	.02	39.4 (8.7)	.45
Personal vlog (n=32)	4.0 (2.1)	N/A	38.9 (7.5)	N/A
Advertisement (n=3)	2.0 (0)	N/A	35.7 (8.6)	N/A
Media (n=2)	2.5 (0.7)	N/A	46 (5.6)	N/A

[a]The “overall content score” assesses video comprehensiveness related to postcesarean pain and is scored out of a maximum of 9 points.

[b]The DISCERN instrument evaluates the reliability of consumer health information. Higher scores indicate greater reliability. Scores are reported out of a maximum of 75. The following categories were used for score interpretation: excellent (63-75 points), good (51-62), fair (39-50), poor (28-38), and very poor (≤27).

[c]N/A: not applicable.

Table 3. Opioid content score by video source.[a,b]

Video source	Opioid content score	
	Mean (SD)	*P* value
Total (N=16)	3.1 (1.6)	N/A[c]
Medical (n=9)	2.9 (1.7)	.77
Personal vlog (n=5)	3.4 (1.8)	N/A
Advertisement (n=0)	—[d]	N/A
Media (n=2)	3.5 (0.7)	N/A

[a]The "opioid content score" assesses video comprehensiveness related to postcesarean opioid use and is scored out of a maximum of 9.

[b]Videos with an "opioid content score=0" were excluded from the analysis.

[c]N/A: not applicable.

[d]Not available.

DISCERN Analysis

The DISCERN scores ranged from 22 (very poor reliability) to 59 (good reliability), with a mean DISCERN score of 39.2 (SD 8.1), consistent with fair reliability. No videos met the criteria for excellent reliability. The overall DISCERN score did not significantly differ by video source (P=.45; Table 2). Videos received the highest average score for DISCERN question 2, "Does it achieve its aims?" (mean 3.4), and question 3, "Is it relevant?" (mean 3.5). Videos received the lowest score for DISCERN question 4, "Does it provide sources?" (mean 1.8), and question 11, "Does it describe the risks of each treatment?" (mean 1.7; Multimedia Appendix 1).

Of the video characteristics, video duration (r=0.38; P<.01), the number of comments (r=0.30; P<.01), and the number of likes (r=0.32; P<.01) were significantly correlated with the overall content score. The number of comments was inversely correlated with the DISCERN score (r=–0.40; P<.01). No video characteristics were significantly correlated with the opioid content score (Table 4).

Table 4. Association of YouTube video comprehensiveness and reliability with video characteristics.

Video characteristics	Median (IQR)	Overall content score		Opioid content score		DISCERN score	
		Correlation (r)	*P* value	Correlation (r)	*P* value	Correlation (r)	*P* value
Days since post	817 (551-1853)	–0.01	.96	0.01	.95	–0.14	.25
Duration (minutes)	7.8 (3.0-13.7)	0.38	<.001	0.05	.64	<0.1	.94
Views	40,614 (6841-82,748)	0.20	.10	0.01	.94	–0.18	.14
Comments	15 (0-54)	0.30	.009	0.07	.82	–0.40	<.001
Likes	213 (53-902)	0.32	.005	0.03	.99	–0.08	.49
Dislike	11 (3-34)	0.17	.14	<–0.01	.56	–0.19	.10
Channel subscribers	35,100 (9020-177,000)	0.06	.59	–0.08	.49	0.13	.28

Quality Analysis

A minority of videos (22/73, 30%) met the criteria for high quality. High-quality videos most frequently addressed overall content regarding pain duration (22/22, 100%) and pain types (20/22, 91%; Figure 2). High-quality videos infrequently address when to notify a clinician (8/22, 36%), and the risks of treatment to the newborn (4/22, 18%). Like trends for the overall content score, high-quality videos had significantly greater median video duration (13.0 minutes vs 7.6 minutes; P=.03), number of comments (24 vs 6; P=.04), and number of likes (397 vs 159; P=.04; Table 5). High-quality videos received the highest score for DISCERN question 2, "Does it achieve its aims?" (mean 3.9) and question 3, "Is it relevant?" (mean 4.1; Multimedia Appendix 3).

Table 5. YouTube video characteristics by quality designation.[a]

Video characteristics	High quality (N=22), median (IQR)	Not high quality (N=51), media (IQR)	*P* value
Days since post	738 (458-2220)	846 (551-1788)	.54
Duration (minutes)	13.0 (5.5-16.6)	7.6 (2.7-11.7)	.03
Views	53,865 (13,817-96,211)	29,661 (4857-75,572)	.20
Comments	24 (11-77)	6 (0-46)	.04
Likes	397 (210-992)	159 (32-632)	.04
Dislikes	20 (3-37)	10 (1-33)	.28
Subscribers	53,900 (9940-240,000)	32,900 (10,970-126,000)	.85

[a]High-quality videos were defined as videos with DISCERN scores greater than or equal to 39 and covering at least five topics out of 9 on the content score.

Discussion

Principal Results

In this study of the top 73 YouTube videos on postcesarean pain management, average video comprehensiveness was low and reliability was fair. Videos rarely address the full scope of health education topics relevant to preparing patients for their postcesarean pain experience. Interestingly, a greater number of comments and likes was positively correlated with better overall content, although more comments were also associated with poorer reliability according to the DISCERN instrument. These findings suggest greater video comprehensiveness is not necessarily associated with improved video reliability, and vice versa. Furthermore, only a minority of videos met the criteria for a high-quality health education resource, suggesting the information currently available on YouTube for postcesarean individuals has important limitations.

Limitations

Like all web content, YouTube is a dynamic source of information. The search results in this study are limited in that they represent a cross-sectional sample. Additionally, the search strategy using the filter "relevance," the default search setting on YouTube, represents only one filter method available to users. We used 36 different search terms to capture relevant videos; however, a different filter setting or search term may yield different findings. However, the chosen search terms were purposefully specific to established content criteria. The limited sample size may limit the ability to detect relationships between video characteristics and quality. Limitations exist in our screening process, where videos longer than 10 minutes were not watched in their entirety. It is possible that relevant content was missed using this strategy. Furthermore, our evaluation of the presenter's characteristics (age, gender, and race or ethnicity) was limited by the fact that YouTube presenters rarely provide self-identifying information. Though our assessment was subjective, we felt that it was important to note representation, as this may influence viewership. Finally, our assessment of video quality was subjective. While we recognize the possibility that YouTube videos may be purposefully narrow in scope with high reliability, we purposefully defined high-quality videos as those that presented both comprehensive and reliable health information to the public.

Comparison With Prior Work

Studies have evaluated eHealth information on pain management outside of the obstetrical population. In one study, the average DISCERN score for chronic pain websites was 55.9 out of a possible 80 points [20], suggesting that written resources may have a higher level of information reliability. However, there is a growing body of evidence that patient comprehension and satisfaction may improve with video over written resources [25,26]. These findings may be related to the average readability of written content. Despite recommendations that written patient education material be at a sixth grade reading level [27,28], the average readability of websites on chronic pain management was that of a 10th-11th grade student [20]. Another study found that web-based patient education materials across obstetric and gynecologic societies ranged from a 9th to 12th grade reading level [27,29]. Thus, video resources have the opportunity to minimize literacy as a barrier to obtaining reliable health information.

Several studies have evaluated the reliability of YouTube videos as a source of health information during pregnancy. Studies regarding COVID-19 during pregnancy, gestational diabetes, and epidural analgesia for labor pain identified DISCERN scores of low to moderate information reliability [18,30,31]. Lee et al [32] recently studied the content and quality of the most frequently viewed YouTube videos related to cesarean birth. According to their content-quality analysis, medical videos were of greater quality than nonmedical video sources, and videos describing personal experiences scored significantly lower than other video content. These findings are consistent with our data, which found that while personal vlog videos commonly contained more content, they did not necessarily contain more reliable content.

Clinical Implications

Uncontrolled postoperative pain may delay hospital discharge and prolong recovery [33]. For birthing individuals, this presents a barrier to independence and caring for a newborn, highlighting the importance of optimizing pain management following cesarean delivery. Experience with pain management interventions, such as enhanced recovery protocols following cesarean, suggests a significant role for thorough education through counseling and written instructions [17]. Additionally, a meta-analysis of emergency room discharge instructions suggested that correct recall may be highest among those who view video discharge instructions [34]. These data suggest a need to translate evidence-based patient education information into a more accessible video format.

While YouTube provides an opportunity to supplement patient education regarding recovery after cesarean birth, current content, including opioid content, is inadequate. Few videos address safe opioid use in the postoperative, outpatient setting, despite campaigns for judicious use. Opioids are known to be prescribed at high rates following cesarean delivery [35]. Fulfillment of postpartum opioid prescriptions and increasing doses are known to increase the risk of serious opioid-related events following cesarean birth [36]. A recent study found a 25% reduction in opioid use when patients viewed an educational video regarding pain management after cesarean delivery [37]. Our content analyses indicate a need to expand upon current YouTube videos to include information regarding opioid use. Improved video content is required for the public to have access to comprehensive information on postcesarean pain management.

YouTube videos provide an opportunity to share quality information on postpartum pain management with a large audience; however, it is essential that clinicians and patients be aware of the limitations of the available videos. This is particularly relevant, as many patients may not discuss the content of electronic sources with their clinicians. A review examining patterns of electronic health use during pregnancy in an underserved, racially diverse population found that while the majority of patients used electronic health sources, approximately 70% of patients discussed their searches with their clinician [38]. Therefore, clinicians may not have an opportunity to discuss the quality of their findings. Interestingly, videos in our study scored low in promoting shared decision-making according to the DISCERN criteria. Even high-quality videos infrequently mention notifying a clinician of warning signs in the postpartum period. Taken together, this highlights the importance of encouraging patients to discuss web-based health information.

Research Implications

Further research is required to understand how obstetric patients are using YouTube during pregnancy and postpartum. The availability of videos and associated subscribers indicates public interest, but further studies are required to understand the needs of postcesarean individuals as they generate their own YouTube searches. Further work is required to evaluate the information available regarding recovery from vaginal birth as well as pain control in the antepartum, intraoperative, and immediate postoperative periods. This study highlights the need for pain management videos that combine medical expertise with consumer needs. While clinicians should caution patients about the reliability of YouTube videos as a health resource, they should also take an interest in what information their patients are looking for on the internet.

Conclusions

Patients seeking information from YouTube regarding postcesarean pain management are likely to encounter videos that lack adequate comprehensiveness and reliability. YouTube is an easily accessible resource and an increasingly common source of health information; however, clinicians should counsel patients to use caution when using current YouTube videos as a resource in the postpartum period.

Acknowledgments

This study was supported by the “Eunice Kennedy Shriver” National Institute of Child Health and Human Development (R01 HD098178-02S1). This grant is also supported, in part, by the National Institutes of Health’s National Center for Advancing Translational Sciences (grant UL1TR001422). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This abstract was presented at the 2022 Society of Maternal-Fetal Medicine 42nd Annual Meeting (virtual, January 31-February 5, 2022).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Post-Cesarean Pain Management Search Term.
[DOCX File , 15 KB - infodemiology_v3i1e40802_app1.docx]

Multimedia Appendix 2
List of videos (N=73) included in the study.
[XLSX File (Microsoft Excel File), 17 KB - infodemiology_v3i1e40802_app2.xlsx]

Multimedia Appendix 3
Discern Score Summary by Video source.
[DOCX File , 28 KB - infodemiology_v3i1e40802_app3.docx]

References

1. The Moz top 500 websites. Moz. 2022. URL: https://moz.com/top500 [accessed 2023-05-19]
2. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. Health Informatics J 2015;21(3):173-194 [FREE Full text] [doi: 10.1177/1460458213512220] [Medline: 24670899]
3. Yeung DL, Alvarez KS, Quinones ME, Clark CA, Oliver GH, Alvarez CA, et al. Low-health literacy flashcards and mobile video reinforcement to improve medication adherence in patients on oral diabetes, heart failure, and hypertension medications. J Am Pharm Assoc (2003) 2017;57(1):30-37 [FREE Full text] [doi: 10.1016/j.japh.2016.08.012] [Medline: 27816544]
4. Sheele JM, Bhangu J, Wilson A, Mandac E. Patient preference for medical information in the emergency department: post-test survey of a random allocation intervention. J Emerg Nurs 2019;45(5):517-522.e6. [doi: 10.1016/j.jen.2019.05.009] [Medline: 31445628]
5. Chandrasekaran N, Gressick K, Singh V, Kwal J, Cap N, Koru-Sengul T, et al. The utility of social media in providing information on Zika virus. Cureus 2017;9(10):e1792 [FREE Full text] [doi: 10.7759/cureus.1792] [Medline: 29282437]
6. Hansen C, Interrante JD, Ailes EC, Frey MT, Broussard CS, Godoshian VJ, et al. Assessment of YouTube videos as a source of information on medication use in pregnancy. Pharmacoepidemiol Drug Saf 2016;25(1):35-44 [FREE Full text] [doi: 10.1002/pds.3911] [Medline: 26541372]
7. Sultan P, Jensen SE, Taylor J, El-Sayed Y, Carmichael S, Cella D, et al. Proposed domains for assessing postpartum recovery: a concept elicitation study. BJOG 2022;129(1):9-20 [FREE Full text] [doi: 10.1111/1471-0528.16937] [Medline: 34536324]
8. Wexler A, Davoudi A, Weissenbacher D, Choi R, O'Connor K, Cummings H, et al. Pregnancy and health in the age of the internet: a content analysis of online "birth club" forums. PLoS One 2020;15(4):e0230947 [FREE Full text] [doi: 10.1371/journal.pone.0230947] [Medline: 32287266]
9. Bjørnstad J, Ræder J. Post-operative pain after caesarean section. Tidsskr Nor Laegeforen 2020 May 05;140(7) [FREE Full text] [doi: 10.4045/tidsskr.19.0506] [Medline: 32378860]
10. Eisenach JC, Pan PH, Smiley R, Lavand'homme P, Landau R, Houle TT. Severity of acute pain after childbirth, but not type of delivery, predicts persistent pain and postpartum depression. Pain 2008;140(1):87-94 [FREE Full text] [doi: 10.1016/j.pain.2008.07.011] [Medline: 18818022]
11. Wilson RD, Caughey AB, Wood SL, Macones GA, Wrench IJ, Huang J, et al. Guidelines for antenatal and preoperative care in cesarean delivery: enhanced recovery after surgery society recommendations (part 1). Am J Obstet Gynecol 2018;219(6):523.e1-523.e15. [doi: 10.1016/j.ajog.2018.09.015] [Medline: 30240657]
12. Jansen BJ, Spink A. How are we searching the world wide web? A comparison of nine search engine transaction logs. Inf Process Manag 2006;42(1):248-263 [FREE Full text] [doi: 10.1016/j.ipm.2004.10.007]
13. Fang Y, Somasundaram N, Si L, Ko J, Mathur AP. Analysis of an expert search query log. 2011 Presented at: SIGIR '11: The 34th International ACM SIGIR Conference on Research and Development in Information Retrieval; July 24 - 28, 2011; Beijing, China p. 1189-1190. [doi: 10.1145/2009916.2010113]
14. American College of Obstetricians and Gynecologists. ACOG Committee Opinion No. 742 Summary: Postpartum Pain Management. Obstet Gynecol 2018 Jul;132(1):252-253. [doi: 10.1097/AOG.0000000000002711] [Medline: 29939935]
15. Kecojevic A, Basch C, Basch C, Kernan W. Pre-exposure prophylaxis YouTube videos: content evaluation. JMIR Public Health Surveill 2018;4(1):e19 [FREE Full text] [doi: 10.2196/publichealth.7733] [Medline: 29467119]
16. Leong AY, Sanghera R, Jhajj J, Desai N, Jammu BS, Makowsky MJ. Is YouTube useful as a source of health information for adults with type 2 diabetes? A South Asian perspective. Can J Diabetes 2018;42(4):395-403.e4. [doi: 10.1016/j.jcjd.2017.10.056] [Medline: 29282200]
17. Wiese AD, Osmundson SS, Mitchel E, Adgent M, Phillips S, Patrick SW, et al. The risk of serious opioid-related events associated with common opioid prescribing regimens in the postpartum period after cesarean delivery. Am J Obstet Gynecol MFM 2021;3(6):100475 [FREE Full text] [doi: 10.1016/j.ajogmf.2021.100475] [Medline: 34455101]
18. Yuksel B, Cakmak K. Healthcare information on YouTube: pregnancy and COVID-19. Int J Gynaecol Obstet 2020;150(2):189-193 [FREE Full text] [doi: 10.1002/ijgo.13246] [Medline: 32471005]
19. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999;53(2):105-111 [FREE Full text] [doi: 10.1136/jech.53.2.105] [Medline: 10396471]
20. Kaicker J, Borg Debono V, Dang W, Buckley N, Thabane L. Assessment of the quality and variability of health information on chronic pain websites using the DISCERN instrument. BMC Med 2010;8:59 [FREE Full text] [doi: 10.1186/1741-7015-8-59] [Medline: 20939875]
21. Śledzińska P, Bebyn MG, Furtak J. Quality of YouTube videos on meningioma treatment using the DISCERN instrument. World Neurosurg 2021;153:e179-e186. [doi: 10.1016/j.wneu.2021.06.072] [Medline: 34166826]
22. Charnock D, Shepperd S. Learning to DISCERN online: applying an appraisal tool to health websites in a workshop setting. Health Educ Res 2004;19(4):440-446 [FREE Full text] [doi: 10.1093/her/cyg046] [Medline: 15155597]
23. Charnock D, Shepperd S. The DISCERN instrument. Discern Online. British Library and the University of Oxford; 1997. URL: http://www.discern.org.uk/discern_instrument.php [accessed 2023-06-06]
24. Ozsoy-Unubol T, Alanbay-Yagci E. YouTube as a source of information on fibromyalgia. Int J Rheum Dis 2021;24(2):197-202. [doi: 10.1111/1756-185X.14043] [Medline: 33355406]

25. Atzema CL, Austin PC, Wu L, Brzozowski M, Feldman MJ, McDonnell M, et al. Speak fast, use jargon, and don't repeat yourself: a randomized trial assessing the effectiveness of online videos to supplement emergency department discharge instructions. PLoS One 2013;8(11):e77057 [FREE Full text] [doi: 10.1371/journal.pone.0077057] [Medline: 24244272]
26. Bloch SA, Bloch AJ. Using video discharge instructions as an adjunct to standard written instructions improved caregivers' understanding of their child's emergency department visit, plan, and follow-up: a randomized controlled trial. Pediatr Emerg Care 2013;29(6):699-704. [doi: 10.1097/PEC.0b013e3182955480] [Medline: 23714763]
27. Weiss BD. Health Literacy: A Manual for Clinicians. Chicago: American Medical Associaton Foundation and American Medical Association; 2003.
28. Weiss BD. CME questionnaire: health literacy: help your patients understand. In: Health Literacy: A Manual for Clinicians. USA: American Medical Association Foundation and American Medical Association; 2003.
29. Fahimuddin FZ, Sidhu S, Agrawal A. Reading level of online patient education materials from major obstetrics and gynecology societies. Obstet Gynecol 2019;133(5):987-993. [doi: 10.1097/AOG.0000000000003214] [Medline: 30969212]
30. D'Souza RS, D'Souza S, Sharpe EE. YouTube as a source of medical information about epidural analgesia for labor pain. Int J Obstet Anesth 2021;45:133-137. [doi: 10.1016/j.ijoa.2020.11.005] [Medline: 33339713]
31. Birch EM, Leziak K, Jackson J, Dahl E, Niznik CM, Yee LM. Content quality of YouTube videos about gestational diabetes: systematic evaluation. JMIR Diabetes 2022;7(2):e30156 [FREE Full text] [doi: 10.2196/30156] [Medline: 35389355]
32. Lee KN, Joo YJ, Choi SY, Park ST, Lee KY, Kim Y, et al. Content analysis and quality evaluation of cesarean delivery-related videos on YouTube: cross-sectional study. J Med Internet Res 2021;23(7):e24994 [FREE Full text] [doi: 10.2196/24994] [Medline: 34328422]
33. American Society of Anesthesiologists Task Force on Acute Pain Management. Practice guidelines for acute pain management in the perioperative setting: an updated report by the American society of anesthesiologists task force on acute pain management. Anesthesiology 2012;116(2):248-273 [FREE Full text] [doi: 10.1097/ALN.0b013e31823c1030] [Medline: 22227789]
34. Hoek AE, Anker SCP, van Beeck EF, Burdorf A, Rood PPM, Haagsma JA. Patient discharge instructions in the emergency department and their effects on comprehension and recall of discharge instructions: a systematic review and meta-analysis. Ann Emerg Med 2020;75(3):435-444. [doi: 10.1016/j.annemergmed.2019.06.008] [Medline: 31439363]
35. Badreldin N, Grobman WA, Chang KT, Yee LM. Opioid prescribing patterns among postpartum women. Am J Obstet Gynecol 2018;219(1):103.e1-103.e8 [FREE Full text] [doi: 10.1016/j.ajog.2018.04.003] [Medline: 29630887]
36. Mullman L, Hilden P, Goral J, Gwacham N, Tauro C, Spinola K, et al. Improved outcomes with an enhanced recovery approach to cesarean delivery. Obstet Gynecol 2020;136(4):685-691 [FREE Full text] [doi: 10.1097/AOG.0000000000004023] [Medline: 32925620]
37. Mokhtari NB, Saeed H, Kawakita T, Huang JC, Iqbal SN. Educational video on pain management and subsequent opioid use after cesarean delivery: a randomized controlled trial. Obstet Gynecol 2021;138(2):253-259 [FREE Full text] [doi: 10.1097/AOG.0000000000004468] [Medline: 34237764]
38. Narasimhulu DM, Karakash S, Weedon J, Minkoff H. Patterns of internet use by pregnant women, and reliability of pregnancy-related searches. Matern Child Health J 2016;20(12):2502-2509. [doi: 10.1007/s10995-016-2075-0] [Medline: 27456311]

Abbreviations

REDCap: Research Electronic Data Capture
Vlog: video blog

Edited by T Mackey; submitted 29.09.22; peer-reviewed by A Zheluk, D Carvalho; comments to author 08.02.23; revised version received 22.02.23; accepted 12.04.23; published 23.06.23.

Please cite as:
Squires NA, Soyemi E, Yee LM, Birch EM, Badreldin N
Content Quality of YouTube Videos About Pain Management After Cesarean Birth: Content Analysis
JMIR Infodemiology 2023;3:e40802
URL: https://infodemiology.jmir.org/2023/1/e40802
doi:10.2196/40802
PMID:37351938

Original Paper

Open-Source Intelligence for Detection of Radiological Events and Syndromes Following the Invasion of Ukraine in 2022: Observational Study

Haley Stone[1], BSc, MIDI, MIPH; David Heslop[2], BSc (Hons), MBBS, MPH, PhD; Samsung Lim[3], BA, MA, PhD; Ines Sarmiento[1], BSc, BE (Hons); Mohana Kunasekaran[1], BSc, MPH, MIPH, PhD; C Raina MacIntyre[1,4], MBBS (Hons), MAE, PhD

[1]Biosecurity Program, The Kirby Institute, Faculty of Medicine, University of New South Wales, Kensington, Australia

[2]School of Population Health, Faculty of Medicine, University of New South Wales, Sydney, Australia

[3]School of Civil & Environmental Engineering, University of New South Wales, Sydney, Australia

[4]College of Public Service & Community Solutions, Arizona State University, Tempe, AZ, United States

Corresponding Author:
Haley Stone, BSc, MIDI, MIPH
Biosecurity Program, The Kirby Institute, Faculty of Medicine
University of New South Wales
Level 6 Wallace Wurth Building
High Street
Kensington, 2052
Australia
Phone: 61 2 9348 0672
Email: haley.c.stone@protonmail.com

Abstract

Background: On February 25, 2022, Russian forces took control of the Chernobyl power plant after continuous fighting within the Chernobyl exclusion zone. Continual events occurred in the month of March, which raised the risk of potential contamination of previously uncontaminated areas and the potential for impacts on human and environmental health. The disruption of war has caused interruptions to normal preventive activities, and radiation monitoring sensors have been nonfunctional. Open-source intelligence can be informative when formal reporting and data are unavailable.

Objective: This paper aimed to demonstrate the value of open-source intelligence in Ukraine to identify signals of potential radiological events of health significance during the Ukrainian conflict.

Methods: Data were collected from search terminology for radiobiological events and acute radiation syndrome detection between February 1 and March 20, 2022, using 2 open-source intelligence (OSINT) systems, EPIWATCH and Epitweetr.

Results: Both EPIWATCH and Epitweetr identified signals of potential radiobiological events throughout Ukraine, particularly on March 4 in Kyiv, Bucha, and Chernobyl.

Conclusions: Open-source data can provide valuable intelligence and early warning about potential radiation hazards in conditions of war, where formal reporting and mitigation may be lacking, to enable timely emergency and public health responses.

(JMIR Infodemiology 2023;3:e39895) doi:10.2196/39895

KEYWORDS

artificial intelligence; contamination; data source; early warning; emergency response; environmental health; open source; open-source intelligence; OSINT; power plant; public health; radiation; radiobiological events; radiological; sensor; Ukraine

Introduction

On February 24, 2022, the Russian invasion of Ukraine began. On the first day of the invasion, battles between Russian and Ukrainian forces occurred in the vicinity of the Chernobyl power plant [1,2]. Following the invasion of Chernobyl on February 25, 2022, the Ukrainian government reported increased levels of radiation in the air [3]. Unverified reports at the time raised concerns of increased radiation levels in the area, potentially due to the disruption of the soil in highly contaminated areas around the power plant due to the fighting and military vehicles moving over the exclusion zone [1,2]. Armed conflict continued throughout March 2022, with intense fighting in Slavutych, a town nearby that houses workers at the power plant. On March 22, 2022, forest fires broke out within the Chernobyl exclusion zone with the potential for generation of contaminated smoke [4]. On March 31, there were reports of confirmed radiation exposure of Russian soldiers, most likely due to soldiers digging trenches in the soil within the Red Forest area to the west of the nuclear power plant [5]. This highlighted the possibility of contamination of previously uncontaminated areas and the potential for subsequent radiological impacts on human and environmental health.

Epidemic open-source intelligence (OSINT) systems provide new approaches to public health surveillance and are increasingly used for epidemic early warning [6]. Early warning OSINT systems can complement and improve the performance of formal surveillance systems by enabling early detection of serious events or fill a gap when routine surveillance systems fail or are absent. Indicator-based surveillance systems largely require clinicians to link cardinal clinical features of specific diseases with key historical, geographic, and social data, thus recognizing the potential occurrence of disease either in an individual or in populations. This process can easily be undermined by a lack of clinical experience, biological variability of presentations in populations, and most importantly, a delay in the recognition of potential disease due to the time it takes for cardinal features to manifest in patients. By contrast, OSINT systems can provide earlier warning through the analysis of large volumes of unstructured digital data and communications. Such data do not rely on clinical experience or acumen, official health system reporting, or the results of laboratory testing. Through the use of specialized processes and algorithms, early warning of potential outbreaks of diseases in populations can be flagged from unstructured sources such as new articles and social media [7]. EPIWATCH and Epitweetr are examples of such systems [6,8]. While OSINT lacks verification, an early warning can be followed by a formal investigation to verify a signal.

Early warning for radiation effects is time critical, as exposures may result in severe outcomes and affect large populations. Following radiation and radioisotope exposure, acute radiation syndrome (ARS) can manifest as early as hours after exposure, and certain therapies require immediate delivery. During the Ukraine conflict, public health surveillance and health protection programs relevant to radiological exposures have been limited or completely ceased. Therefore, in the context of conflict and degraded public health systems, the use of OSINT to rapidly identify locations where a radiological event may have occurred is important and enables the most efficient and timely allocation of limited health resources to limit the spread and impact of contamination. A key distinction in the Ukrainian conflict, as compared broadly to other conflicts, has been the widespread and continued access to high quality open-source data communications, including social media and news sources, across the broad Ukrainian geography and continued penetration of access within the Ukrainian population. Unstructured data from local news reports, social media, and various open-source channels from the Ukrainian population and occupying forces can be used by systems such as EPIWATCH and Epitweetr to detect signals for health-related events of importance.

This study aims to demonstrate the value of OSINT in Ukraine to identify signals of potential radiological events of health significance during the Ukrainian conflict.

Methods

Data Collection

To determine the potential detection of radiobiological events, data were analyzed between the timeframe of February 1 and March 20, 2022, using both EPIWATCH and Epitweetr systems. EPIWATCH is an artificial intelligence (AI)-driven system that uses both curated information, such as governmental reports, and broader web searches to generate automated early warnings for epidemics worldwide [9-12]. Outbreak signals in EPIWATCH are obtained from reports collected in real time using prespecified search terms applied to open-source data. These can be monitored for deviations from baseline or unusual, newly emerging diseases. The system contains 52 translated languages, together with geographic information system capability. In addition to 2 AI subsystems (natural language processing [NLP] and a prioritization algorithm), the information collected is curated by epidemic analysts. Epitweetr is an R-based open-source data surveillance tool. Epitweetr's data are routinely collected. In order to monitor trends in tweets' geolocation, time, and topic using the Twitter Standard Search API, data are collected by sending queries to the predetermined list of topics and associated keywords. The default topics list consists of 71 unique topics but can be customized to the user's choice [8,13]. EPIWATCH, at the time of this study, did not query Twitter. Epitweetr was used to enhance the data set to include social media coverage, as social media is more likely to pick up early signals for acute radiation syndrome. However, as social media is more vulnerable to manipulation, both systems are needed to validate potential detections or events.

These systems were originally created to detect infectious disease outbreaks but can be rapidly adapted for the detection of radiobiological events. A series of search terms were created by a domain expert on radiation (DH) that were indicative of potential radiobiological events (acute exposure to radioisotopes, contamination by radioisotopes, ARS, and related medical symptoms and signs). The terms were translated into Ukrainian and Russian. The search terms are listed in Table 1 and their definitions are in Multimedia Appendix 1. In addition, since users often do not disclose direct illness on social media but rather discuss symptoms, we individualized each symptom and added variations for analysis for radiation poisoning. Symptomology terms for acute radiation poisoning were also investigated and are described in Table 2. Reports collected from EPIWATCH were obtained through a manual search within the system; they did not undergo machine learning classification and were gathered solely through noncurated broader web searches using Google Alerts (Table 3). Data collected through Epitweetr queried the terms added through the Epitweetr Shiny app interface using Twitter APIs 1 and 2. The tweets gathered from the queries are then aggregated and geolocated, and an

Early Aberration Reporting System (EARS) signal detection algorithm is applied. Each individual tweet is counted as a report and is visualized in the results (Table 3) [8,13]. Data collection occurred after the search timeframe for potential radiobiological events, from March 20 to April 12, 2022. Both systems used the same search terms to investigate the potential for radiobiological events and acute radiation poisoning, which are justified and explained in Multimedia Appendix 1 and Tables 3 and 1. Specific terms related to features of acute radiation exposure (eg, radiation types, the Cherenkov effect, radioisotopes, and initial medical impacts) and terms relating to the short-term effects and immediate medical management of exposure were used.

Table 1. Terms used in the search for radiobiological events in Ukraine by subtopics: event-based terms (n=14), radiological substance–based terms (n=14), medical terms (n=13), and radiation preparedness terms (n=8).

English	Ukrainian	Russian
Event-based terms		
Radiation	Радіація	Радиация
Radiological	Радіологічний	Радиологический
Reactor	реактор	Реактор
Alpha radiation	Альфа-випромінювання	Альфа-излучение
Beta radiation	Бета-випромінювання	Бета-излучение
Gamma radiation	Гамма-випромінювання	Гамма-излучение
Isotope	Ізотоп	Изотоп
Geiger	Гігер	Гигер
Curie	Кюрі	Кюри
Becquerel	Бекерель	Беккерель
Sievert	Зіверт	Зиверт
REM[a]	REM	REM
RAD[b]	RAD	RAD
Cherenkov	Черенков	Черенков
Radiological substance–based terms		
Iodine	йод	Йод
I-131	I-131	I-131
Cesium	цезій	Цезий
Cs-137	Cs-137	Cs-137
Cs-134	Cs-134	Cs-134
Plutonium	плутоній	Плутоний
Strontium	стронцій	Стронций
Sr-90	Sr-90	Sr-90
Americium	америцій	Америций
Am-241	Ам-241	Ам-241
Uranium	уран	Уран
Nuclear fuel	Ядерне паливо	Ядерное топливо
Nuclear waste	Ядерні відходи	Ядерные отходы
Graphite	Графіт	Графит
Medical terms		
Beta burn	Бета-запис / Бета опік	Бета ожог
Desquamation	Десквамація	Десквамация (latin) / Шелушение
Hair loss	Втрата волосся	Выпадение волос
Mucositis	Мукозит	Мукозит
Gastrointestinal syndrome	Шлунково-кишковий синдром	Желудочно-кишечный синдром
Cardiovascular syndrome	Серцево-судинний синдром	Сердечно-сосудистый синдром
Neurological syndrome	Неврологічний синдром	Неврологический синдром

English	Ukrainian	Russian
Melena	Мелена	Мелена
Vomiting	Блювота	Рвота
Lymphopaenia	Лімфопенія	Лимфопения
Bone marrow suppression	Пригнічення кісткового мозку	Подавление костного мозга
Bone marrow transplant	Пересадка кісткового мозку	Пересадка костного мозга
Sepsis	Сепсис	Сепсис
Radiation preparedness terms		
Potassium Iodide	Калій йодид/Йодистий Калій	Йодистый калий
Heavy Metal Chelation	Хелатування важких металів	Хелотирование тяжелых металлов
Calcium DTPA	Кальцій DTPA	Кальцый DTPA
Zinc DTPA	Цинк DTPA	Цынк DTPA
Decontamination	Дезактивація/знезараження	Деконтаминация/Обеззараживание
Prussian Blue	Прусський блакитний/берлінска блакитність	Берлинская лазурь
Granulocyte Monocyte Colony Stimulating Factor	Фактор, що стимулює колонію гранулоцитів моноцитів	Фактор стимулирующий колонию гранулоцитов моноцитовю/Гранулоциты Моноциты Колониестимулирующий фактор
Granulocyte Colony Stimulating Factor	Фактор, стимулюючий колонію гранулоцитів	Фактор стимулирующий колонию гранулоцитов/Гранулоцитарный колониестимулирующий фактор

[a]RAD: radiation absorbed dose.

[b]REM: roentgen equivalent man.

Table 2. Syndromic terms and variants of each term used to search for acute radiation poisoning in Ukraine [14].

Syndromic terms	English	Ukrainian	Russian
Radiation	radiation OR RAD OR radiated OR glowing	випромінювання АБО RAD АБО випромінюваний АБО світиться	излучение ИЛИ RAD ИЛИ излучаемое ИЛИ светящееся
Nausea	nausea OR nauseated	нудота АБО нудота	тошнота ИЛИ тошнота
Vomiting	vomiting OR vomit OR throwup OR puke	блювота АБО блювота АБО блювота АБО блювота	рвота ИЛИ рвота ИЛИ рвота ИЛИ рвота
Headaches	headaches OR headache OR migraine	головні болі АБО головний біль АБО мігрень	головные боли ИЛИ головная боль ИЛИ мигрень
Fatigue	fatigue OR drowsy OR disoriented	втома АБО сонливість АБО дезорієнтація	усталость ИЛИ сонливость ИЛИ дезориентированность
Fever	fever OR feverish OR temperature OR shivering	гарячка АБО лихоманка АБО температура АБО тремтіння	лихорадка ИЛИ лихорадка ИЛИ температура ИЛИ озноб
Rash and fever	skin-reddening OR rash	почервоніння шкіри АБО висип	покраснение кожи ИЛИ сыпь

Table 3. Data output for Epitweetr and EPIWATCH.

System	Output data type	Data analysis before output
Epitweetr [8,13]	Aggregated tweets by search term	Twitter results for individual terms are initially geolocated. The tweets are then aggregated on terms and geolocation. Finally, the Early Aberration Reporting System algorithm is applied to identify if a signal was detected by qualitatively comparing baseline activity to aberrations (2 standard deviations) from the baseline.
EPIWATCH	Web results by search term	Manual search through the EPIWATCH system by term is performed by a human analyst. Each individual web result is deemed as a report. Aggregate by term is performed manually. Data generated are reviewed to identify if a signal was detected by qualitatively comparing the baseline activity to aberrations (2 standard deviations) from the baseline.

Data Analysis

A comprehensive line list was created for both EPIWATCH and Epitweetr. Analysis was completed separately on potential radiological events and acute radiation detection analysis. Data were sorted using MATLAB by date, subtopic, language, and, for Epitweetr, subnational geolocation. Data from Epitweetr were only used if the tweet's geolocation was within Ukraine. Analysis and reporting for this study followed STROBE (Strengthening the Reporting of Observational Studies) guidelines for epidemiological studies [15]. For EPIWATCH's signal reports, for both radiobiological event analysis and acute radiation detection analysis, a mean daily signal count (σ) was established for each subtopic using Stata/IC (Stata Corp). The total daily signal was adjusted for this factor. The signal curve was constructed using the date of the signal and the total adjusted daily signal. The plots were analyzed on the date of peak signal and compared to key dates and events around Ukraine. We used the data to identify if a signal was detected by qualitatively comparing baseline activity over time to aberrations (2 standard deviations) from baseline and in relation to key events around the war. Analysis of Epitweetr included searches for the individual terms in order to identify increased signals within the given time period. In addition, geolocation of the total signal amount and subtopic were performed using descriptive statistics and plotted using ArcGIS Pro (Esri).

Ethical Considerations

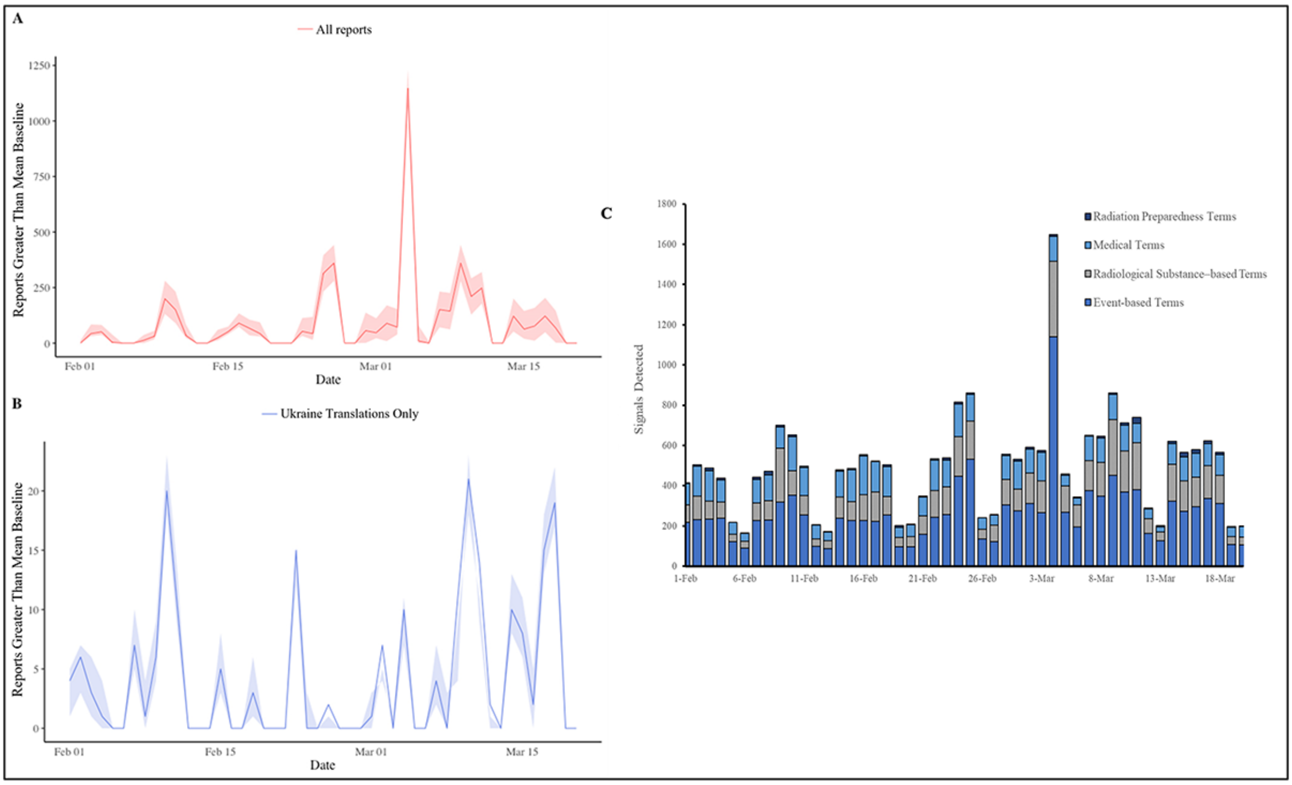
This study only contains open-source data which are publicly available. No individual or identifying data about patients or people were collected. In addition, all the data presented are in aggregate form and have been deidentified before data analyses were completed.

Results

Potential Radiobiological Event Detection

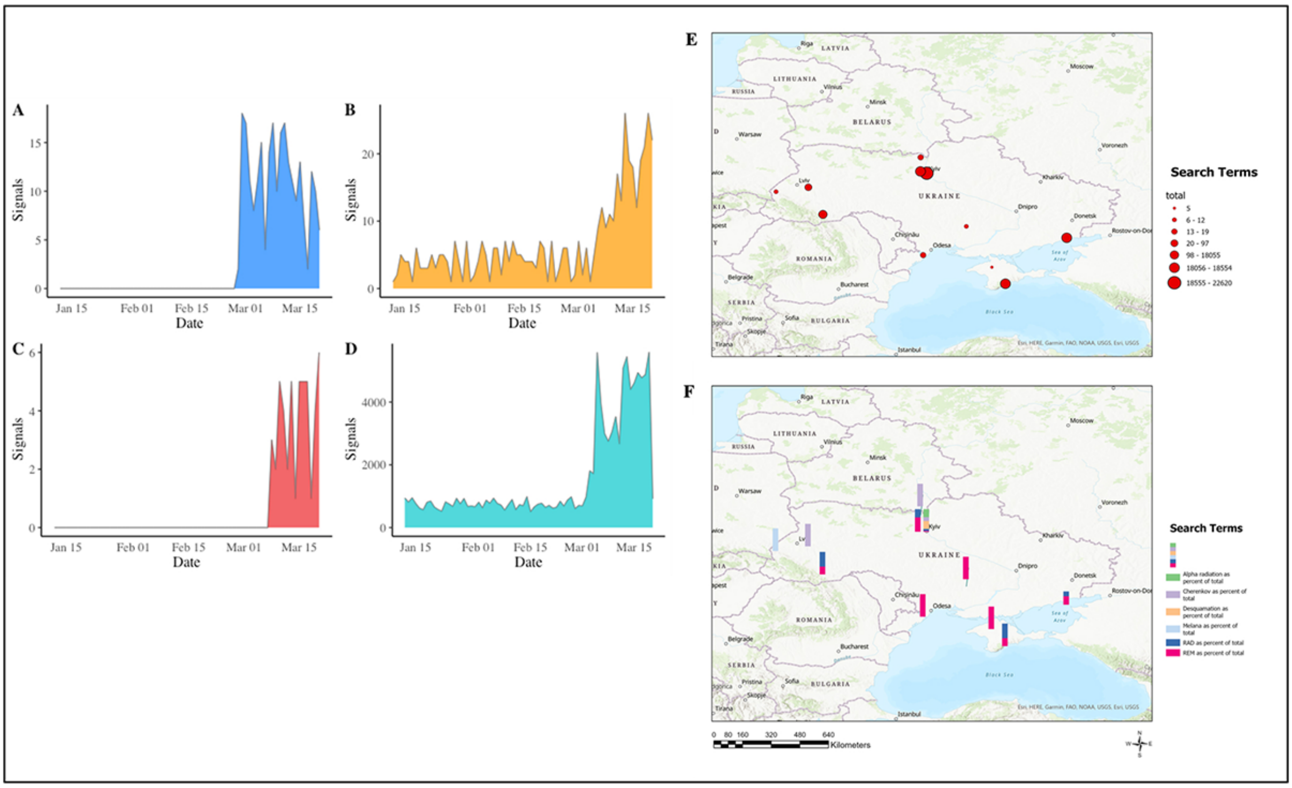
Both systems detected potential radiological events from February 1 to March 20, 2022. Terms used to mine these open-source news were separated into 4 subgroups: event-based surveillance, radiological substance, radiological preparedness, and medical based terms. EPIWATCH overall detected 24,071 reports with a mean of 502 reports per day (σ=36.6; 95% CI 427.8-575.2) with an adjusted peak on March 4 (n=1147) using both English and Ukrainian translations. Of the reports, 5.6% (n=1348) were Ukrainian. Adjusted daily reports for both English and Ukrainian translations found 5 distinct peaks on February 24 and 28 and March 4, 9, and 17 (Figure 1A). Likewise for Ukrainian-only translations, 5 peaks were observed on February 10 and 23 and March 4, 10, and 18 (Figure 1B).

Figure 1. EPIWATCH's adjusted daily signal detection for all languages (A) and Ukrainian (B) were examined in addition to a time series by subtopic (C).



Event-based surveillance terms, described in Multimedia Appendix 1, were detected with a daily mean of 265 (σ=23.9; 95% CI 216.7- 312.8) reports and had 2 peaks from February 24 and March 4 (Figure 2C). For radiological substance–based terms, EPIWATCH detected a mean of 123 reports per day (σ=10.1; 95% CI 102.2- 142.9) and had a peak on March 4 (n=376). For medical-based terms, EPIWATCH detected a mean of 106 reports per day (σ=6.1; 95% CI 94.0-118.3) and had a peak on February 24 (n=162). Lastly, for radiation preparedness terms, EPIWATCH detected a mean of 8 reports per day and had a peak of reports on March 11 (n=29) (Figure 1B). Using exclusively the Ukrainian translations, EPIWATCH detected a mean of 27 reports per day (σ=1.3; 95% CI 24.0-29.4) and a peak observed on March 18 (n=46) for event-based surveillance terms. For radiological substance-based terms, EPIWATCH detected a mean of a report per day (σ=0.3; 95% CI 0.6-1.9) and peaks on March 2, 9, and 11 for radiological substance-based terminology. For medical-based terms, EPIWATCH detected a mean of less than a report per day with single reports found on March 3 and 17. Lastly, for radiation preparedness, no reports were detected with Ukrainian translations.

Figure 2. Epitweetr's signal (tweets) detection for (A) Cherenkov, (B) roentgen equivalent man (REM), (C) alpha radiation, and (D) radiation absorbed dose (RAD). Event and radiological terms by (E) region (n=96,094) and (F) individual term.



For Epitweetr, a total of 4 different search terms were identified to have distinct peaks during the invasion in Ukraine: Cherenkov radiation, which was first reported on February 28 and peaked on February 28 (Figure 2A); REM, which rose from the baseline average of 4 signals a day on March 3 and peaked twice on March 13 and 19 (Figure 2B); alpha radiation, which was first reported on March 8 and peaked on March 20 (Figure 2C); and RAD, which rose from the baseline average of 743 signals per day on March 2 and had 3 peaks on March 4, 12, and 20 (Figure 2D).

The 3 highest regions within Ukraine for the radiological terms were within Kyiv (22,620/96,094, 23.5%), Semidvor'e (18,405/96,094, 19.2%), and Mariupol (18,301/96094, 19%) for Epitweetr (Figure 2E). Additionally, a total of 2 terms reported signals within the Chernobyl area: *Cherenkov* (n=12) and *RAD* (n=4) (Figure 2F).

Potential Acute Radiation Syndrome Detection

EPIWATCH detected 51,248 reports of symptoms related to radiation poisoning throughout the period between February 24 and March 20, 2022 with a mean of 2050 reports per day (σ=131.7; 95 CI 1778.1-2321.8) and an overall peak on February 28 (n=2898). Radiation reports had a mean of 354 (σ=32.6; 95% CI 286.5-421.3) and a noticeable peak on March 4 (n=940), which consisted of 54.4% of all reports on 1 day. Nausea reports had a mean of 38 (σ=2.8; 95% CI 31.8-43.2) and a peak on March 7 (n=66). Vomiting reports had a mean of 44 (σ=2.5; 95% CI 39.3-49.4) and a peak on March 1 (n=65). Headache reports had a mean of 147 (σ=10.1; 95% CI 125.8-167.4) and a peak on March 2 (n=231). Fatigue reports had a mean of 153 (σ=9.8; 95% CI 133.2-173.6) and a peak on February 26 (n=210). Fever reports had a mean of 624 (σ=43.2; 95% CI 534.4-712.84) and a peak on February 28 (n=1001). Skin-reddening reports had a mean of 44 (σ=2.8; 95% CI 37.6-49.3) and a peak on March 1 (n=78). Rash and fever reports had a mean of 5 (σ=0.7; 95% CI 3.1-6.0) and a noticeable peak on March 1 (n=14).

A total of 757 signals were detected with the symptomology related to radiation poisoning from Epitweetr. The 2 regions with the most signals detected were within Bucha (n=287) and Kyiv (n=196) at the time of the search (Figure 3A). Of the 757 detected signals for the symptoms related to radiation poisoning, 27.6% (n=209) signals were for vomiting, 24% (n=182) were for fever, 15.6% (n=118) were for nausea, 15.3% (n=116) were for skin reddening, 10.3% (n=78) were for fatigue, and 7.1% (n=54) were for headaches (Figure 4B).

Figure 3. EPIWATCH reports on syndromes for acute radiation poisoning by detecting reports above the baseline daily mean and by individual syndromes between February 24 and March 20, 2022, (n=51,248).

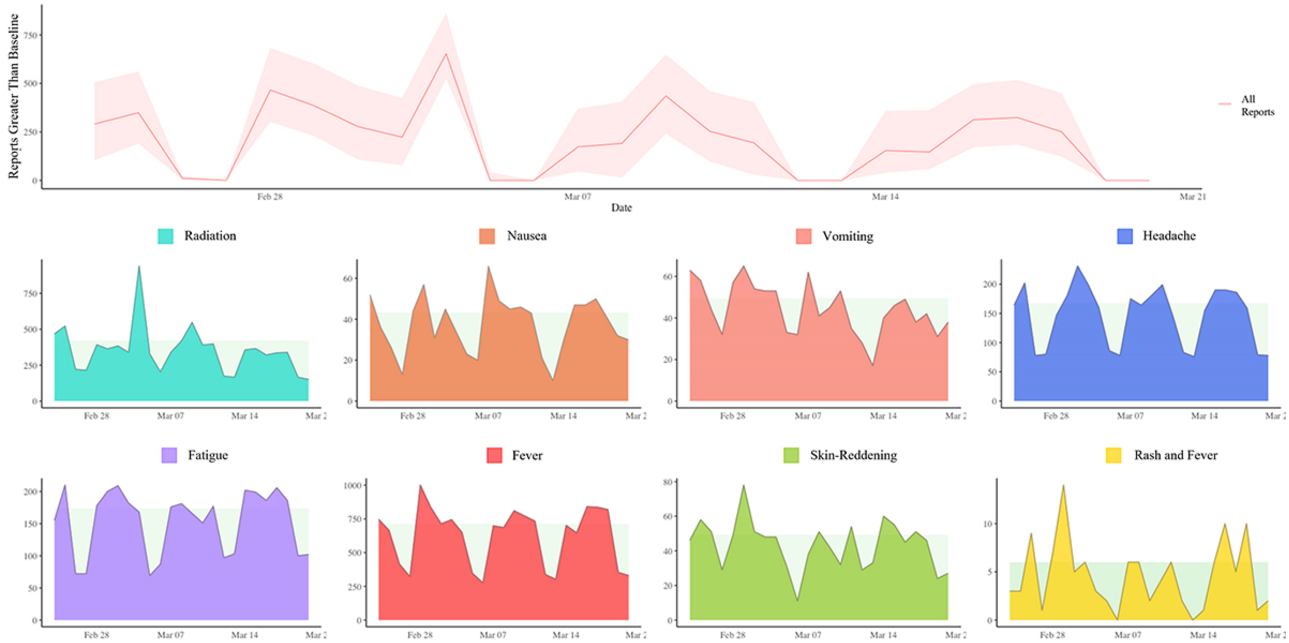
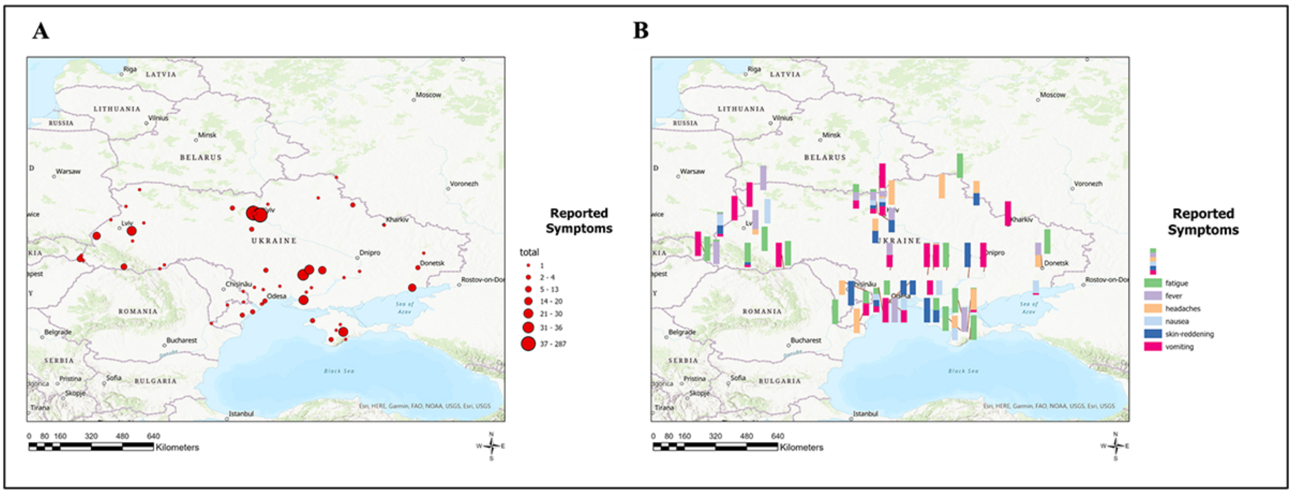


Figure 4. Epitweetr's signals reported by region for both total reports per region (A) and by individual symptom per region (B) (N=757).



Discussion

We have shown that under conditions of war, when routine reporting and monitoring may be disrupted or absent, open-source intelligence from news reports or social media can be used for early warning of potential radiation events. While signals were detected on February 24 with the beginning of the invasion of the Chernobyl plant, both systems detected further signals in March, which could be linked to the rise in radiobiological events, such as radiation exposure during the armed conflict within the exclusion zone. The Russians occupied the Chernobyl plant from February 24 to March 31, 2022 with acute radiation syndrome reported in Russian soldiers on March 31 and one death reported [5,16]. Potential exposure could have been throughout the occupation of Chernobyl and surrounding areas by the Russian soldiers. Additionally, there was global concern about the disruption of Chernobyl and other nuclear sites during the invasion. Geolocation analysis of radiobiological events for Epitweetr found 2 terms within the Chernobyl region: *RAD* and *Cherenkov*. For acute radiation poisoning syndromic analysis, *vomiting* and *headaches* were identified within regions surrounding the Chernobyl exclusion zone in the month of March. Clustering of signals in our syndromic analysis for radiation sickness appeared in or around Kyiv. The results from this study show the usefulness of immediate, timely information, particularly in a war zone where access for investigations might be minimal. This information, obtained rapidly, can complement the formal intelligence systems already in place.

OSINT systems have already been used in Ukraine to aid in the detection of potential war crimes and military movements [17]. We showed that OSINT can detect potential radiation events and can be used in real time for early warning. While not a replacement for validated data, such as radiation measurements, open-source data can provide early intelligence when formal reporting is absent and can provide a trigger for an early investigation or emergency response.

A potential limitation to using open-source data is the possibility of manipulation or interference by third parties through the injection of tweets or news sources to boost sentiment. This can be mitigated through multi-source data fusion, triangulation of data, and correlation within and between NLP and machine learning (ML) identified data and other sources of data. In this study, the use of news-based OSINT allowed the validation of Twitter-based OSINT. EPIWATCH, an AI system, applies a model using contemporary NLP and named entity recognition (NER) algorithms in order to detect unusual spikes or signals in particular topics. In addition to system filtration, human moderation is implemented to verify the authenticity of reports. Epitweetr, unlike EPIWATCH, does not individually filter tweets but instead uses a modified EARS, which is a well-established model developed by the US Centers for Disease Control as a baseline for signal detection. OSINT can result in lexical bias that can lead to overreporting of signals and cause issues establishing signal validity. The bias can be allayed by using specific terminology to decrease irrelevant outside noise. This mitigation was confirmed by the detection of distinct spikes in specific terminology not used in regular vernacular.

A further limitation of this study is the dependence on the quality of the data inputs. The data obtained from EPIWATCH could have delayed reporting time or contain biases. There are also language biases, with a predominant amount of news reports being in English for EPIWATCH, despite searching in Ukrainian, which could indicate events outside the scope of the Ukrainian invasion. However, we did perform an analysis on the total reports from EPIWATCH in addition to solely reports in Ukrainian to detect varying signals, if any, from the 2 languages. Additionally, searching in Ukrainian only began in February 2022, whereas searching in Russian was part of EPIWATCH since 2019. For Epitweetr, the tweets are aggregated and rely on built-in signal detection algorithms to distinguish actual events from “white noise.” In addition, the symptoms of radiation poisoning can be indicative of other diseases rather than radiation poisoning. We, however, attempted to mitigate by clustering and geolocating symptoms, in which all symptoms in our syndromic analysis appeared in or around Kyiv. Lastly, the signal detected using open-source syndromic analysis may not reflect radiation exposure and may be a false positive. However, the purpose of OSINT is to monitor the baseline, detect early warning signals above the baseline, and then formally investigate for confirmation.

Using OSINT systems such as EPIWATCH and Epitweetr, signal detection from war zones can be used in the absence of formal detection methods to help rapidly discover and control public health risks. Several studies have identified social media, particularly Twitter, that can be used to identify particular syndromes [18-20]. The value of these open-source data systems, like Epitweetr and EPIWATCH, is the rapid detection of outbreaks and public health events when surveillance systems are not as robust or have been weakened, such as with the invasion of Ukraine [21,22]. An estimated 50% of the stakeholders in epidemic response report lacking access to timely surveillance data, yet 90% do not use available open-source systems, highlighting the potential to improve the use of OSINT [9].

Both systems identified potential radiobiological events throughout Ukraine, particularly on March 4 in Kyiv, Bucha, and 16 reports within Chernobyl. The risk of a nuclear accident will remain a pressing matter as the conflict continues in Ukraine. While Chernobyl has been returned to the Ukrainian government, the Zaporizhzhia plant, where spent fuel assemblies can be damaged, is still under the control of Russia [23,24]. An accident involving spent fuel assemblies could be equivalent in magnitude to the initial Chernobyl event in 1986 and requires the site to undergo constant preventive activities and monitoring. Additionally, normal preventive activities and radiation monitoring sensors have been nonfunctional during parts of the occupation, specifically in Chernobyl, and do not allow for real-time data to be received at this time [25]. OSINT reports can support governmental classified intelligence sources, gather information where formal surveillance might not be as robust or be hindered during the conflict, and provide this information in real time, which can inform timely government responses to the data presented. The significance of OSINT during the invasion, where formal information is scarce, will be to supplement more formal data sources, provide essential early warning of radiobiological events, and ensure timely emergency and public health responses. Both Epitweetr and EPIWATCH can be rapidly adapted to evolving biosecurity or other acute threats. In addition, EPIWATCH continues with the search terminology presented in this study, which routinely monitors potential radiobiological events. These systems can be used as collaborative tools with many stakeholders as a means of surveillance, both in peacetime and in active war zones.

Acknowledgments

The authors would like to thank Kate Cole for arranging the translation of the Ukrainian terms used in this study. We acknowledge Epitweeter, a program developed and provided as an open source tool by the European Centers for Disease Control. Funding for EPIWATCH comes from the Balvi Filantropic Fund.

Conflicts of Interest

Authors CRM, SL, and DH have been involved in the development of EPIWATCH at the University of New South Wales but do not receive any financial remuneration (such as shares or stock options) from EPIWATCH, which is not a commercial entity.

Multimedia Appendix 1
Terms and associated definitions used in the search for radiobiological events in Ukraine by subtopics.

[DOCX File , 43 KB - infodemiology_v3i1e39895_app1.docx]

References

1. Gill V. Chernobyl: Why radiation levels spiked at nuclear plant. BBC News. 2022. URL: https://www.bbc.com/news/science-environment-60528828 [accessed 2023-06-01]
2. Russian forces trying to seize control of Chernobyl Nuclear Plant in Ukraine, says President Zelensky 2022. ANI News. 2022. URL: https://www.aninews.in/news/world/asia/russian-forces-trying-to-seize-control-of-chernobyl-nuclear-plant-in-ukraine-says-president-zelensky20220224212625/ [accessed 2023-06-01]
3. Polityuk P, Crellin F. Ukraine reports higher chernobyl radiation after Russians capture plant. Reuters. 2022. URL: https://www.reuters.com/world/europe/ukraine-nuclear-agency-reports-higher-chernobyl-radiation-levels-due-heavy-2022-02-25/ [accessed 2023-06-01]
4. Milman O. Forest fires erupt around chernobyl nuclear plant in Ukraine. The Guardian. 2022. URL: https://tinyurl.com/4mm4sfef [accessed 2023-06-01]
5. The first death of a Russian soldier from radiation sickness. UKRNEWS. 2022. URL: https://ukrnews.org.ua/2022/03/the-first-death-of-a-russian-soldier-from-radiation-sickness/ [accessed 2023-06-01]
6. MacIntyre CR, Lim S, Quigley A. Preventing the next pandemic: use of artificial intelligence for epidemic monitoring and alerts. Cell Rep Med 2022;3(12):100867 [FREE Full text] [doi: 10.1016/j.xcrm.2022.100867] [Medline: 36543103]
7. Yan SJ, Chughtai AA, Macintyre CR. Utility and potential of rapid epidemic intelligence from internet-based sources. Int J Infect Dis 2017;63:77-87 [FREE Full text] [doi: 10.1016/j.ijid.2017.07.020] [Medline: 28765076]
8. Espinosa L, Wijermans A, Orchard F, Höhle M, Czernichow T, Coletti P, et al. Epitweetr: early warning of public health threats using Twitter data. Euro Surveill 2022;27(39):2200177 [FREE Full text] [doi: 10.2807/1560-7917.ES.2022.27.39.2200177] [Medline: 36177867]
9. Hii A, Chughtai AA, Housen T, Saketa S, Kunasekaran MP, Sulaiman F, et al. Epidemic intelligence needs of stakeholders in the Asia-Pacific region. Western Pac Surveill Response J 2018;9(4):28-36 [FREE Full text] [doi: 10.5365/wpsar.2018.9.2.009] [Medline: 30766745]
10. Kpozehouen EB, Chen X, Zhu M, Macintyre CR. Using open-source intelligence to detect early signals of COVID-19 in China: descriptive study. JMIR Public Health Surveill 2020;6(3):e18939 [FREE Full text] [doi: 10.2196/18939] [Medline: 32598290]
11. Thamtono Y, Moa A, MacIntyre CR. Using open-source intelligence to identify early signals of COVID-19 in Indonesia. Western Pac Surveill Response J 2021;12(1):40-45 [FREE Full text] [doi: 10.5365/wpsar.2020.11.2.010] [Medline: 34094623]
12. Puca C, Trent M. Using the surveillance tool EpiWATCH to rapidly detect global mumps outbreaks. Global Biosecurity 2020;2(1) [FREE Full text] [doi: 10.31646/gbio.54]
13. Twitter API. Twitter. 2022. URL: https://developer.twitter.com/en/docs/twitter-api [accessed 2023-06-01]
14. Acute radiation syndrome: a fact sheet for clinicians. National Center for Environmental Health. 2018. URL: https://tinyurl.com/4f84k4s2 [accessed 2023-06-01]
15. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596):1453-1457 [FREE Full text] [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
16. Kilner J. Russian soldier dies from radiation poisoning in Chernobyl. Telegraph. 2022. URL: https://www.telegraph.co.uk/world-news/2022/04/01/russian-soldier-dies-radiation-poisoning-chernobyl/ [accessed 2023-06-01]
17. Schwartz L. Amateur open-source researchers went viral unpacking the war in Ukraine. Rest of World. 2022. URL: https://restofworld.org/2022/osint-viral-ukraine/ [accessed 2023-06-01]
18. de Araujo DHM, de Carvalho EA, da Motta CLR, da Silva Borges MR, Gomes JO, de Carvalho PVR. Social networks applied to Zika and H1N1 epidemics: a systematic review. In: Congress of the International Ergonomics Association. 2018 Presented at: Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018); August 26-30, 2018; Florence, Italy p. 679-692. [doi: 10.1007/978-3-319-96089-0_74]
19. Masri S, Jia J, Li C, Zhou G, Lee MC, Yan G, et al. Use of twitter data to improve Zika virus surveillance in the United States during the 2016 epidemic. BMC Public Health 2019;19(1):761 [FREE Full text] [doi: 10.1186/s12889-019-7103-8] [Medline: 31200692]
20. van de Belt TH, van Stockum PT, Engelen LJLPG, Lancee J, Schrijver R, Rodríguez-Baño J, et al. Social media posts and online search behaviour as early-warning system for MRSA outbreaks. Antimicrob Resist Infect Control 2018;7:69 [FREE Full text] [doi: 10.1186/s13756-018-0359-4] [Medline: 29876100]
21. Demographic and Social Statistics 2018. United Nations. 2018. URL: https://unstats.un.org/unsd/demographic-social/products/dyb/ [accessed 2023-06-01]
22. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395(10229):1054-1062 [FREE Full text] [doi: 10.1016/S0140-6736(20)30566-3] [Medline: 32171076]

23. Henly J. Russia steps up attack on Azovstal plant as convoy arrives in Zaporizhzhia. Guardian. 2022. URL: https://www.theguardian.com/world/2022/may/03/russia-steps-up-attack-on-azovstal-plant-as-convoy-arrives-in-zaporizhzhia [accessed 2023-06-01]
24. Brumfiel G, Rizzo M, Le T, Hurt A. Video analysis reveals Russian attack on Ukrainian nuclear plant veered near disaster. National Public Radio. 2022. URL: https://www.npr.org/2022/03/11/1085427380/ukraine-nuclear-power-plant-zaporizhzhia [accessed 2023-06-01]
25. Nuclear monitor loses contact with Chernobyl radiation sensors, raising alarm. Times of Israel. 2022. URL: https://www.timesofisrael.com/nuclear-monitor-loses-contact-with-chernobyl-radiation-sensors-raising-alarm/ [accessed 2023-06-01]

Abbreviations

AI: artificial intelligence
ARS: acute radiation syndrome
EARS: Early Aberration Reporting System
ML: machine learning
NER: named entity recognition
NLP: natural language processing
RAD: radiation absorbed dose
REM: roentgen equivalent man
OSINT: open-source intelligence
STROBE: Strengthening the Reporting of Observational Studies

Edited by T Mackey; submitted 27.05.22; peer-reviewed by M Katz, CL Sánchez Bocanegra; comments to author 11.01.23; revised version received 26.01.23; accepted 11.04.23; published 28.06.23.

Please cite as:
Stone H, Heslop D, Lim S, Sarmiento I, Kunasekaran M, MacIntyre CR
Open-Source Intelligence for Detection of Radiological Events and Syndromes Following the Invasion of Ukraine in 2022: Observational Study
JMIR Infodemiology 2023;3:e39895
URL: https://infodemiology.jmir.org/2023/1/e39895
doi:10.2196/39895
PMID:37379069

Original Paper

Compliance With the US Food and Drug Administration's Guidelines for Health Warning Labels and Engagement in Little Cigar and Cigarillo Content: Computer Vision Analysis of Instagram Posts

Jiaxi Wu[1], MS; Juan Manuel Origgi[2], MS; Lynsie R Ranker[3], PhD; Aruni Bhatnagar[4,5], PhD; Rose Marie Robertson[5,6], MD; Ziming Xuan[3], MA, SCD; Derry Wijaya[2], PhD; Traci Hong[1*], PhD; Jessica L Fetterman[7*], PhD

[1]College of Communication, Boston University, Boston, MA, United States

[2]Department of Computer Science, Boston University, Boston, MA, United States

[3]Department of Community Health Sciences, School of Public Health, Boston University, Boston, MA, United States

[4]Department of Medicine, University of Louisville, Louisville, KY, United States

[5]American Heart Association Tobacco Regulation and Addiction Center, Dallas, TX, United States

[6]Department of Medicine, School of Medicine, Vanderbilt University, Nashville, TN, United States

[7]Evans Department of Medicine and Whitaker Cardiovascular Institute, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Jessica L Fetterman, PhD
Evans Department of Medicine and Whitaker Cardiovascular Institute
Boston University Chobanian & Avedisian School of Medicine
700 Albany Street
W-602A
Boston, MA, 02118
United States
Phone: 1 6173587544
Email: jefetter@bu.edu

Abstract

Background: Health warnings in tobacco advertisements provide health information while also increasing the perceived risks of tobacco use. However, existing federal laws requiring warnings on advertisements for tobacco products do not specify whether the rules apply to social media promotions.

Objective: This study aims to examine the current state of influencer promotions of little cigars and cigarillos (LCCs) on Instagram and the use of health warnings in influencer promotions.

Methods: Instagram influencers were identified as those who were tagged by any of the 3 leading LCC brand Instagram pages between 2018 and 2021. Posts from identified influencers, which mentioned one of the three brands were considered LCC influencer promotions. A novel Warning Label Multi-Layer Image Identification computer vision algorithm was developed to measure the presence and properties of health warnings in a sample of 889 influencer posts. Negative binomial regressions were performed to examine the associations of health warning properties with post engagement (number of likes and comments).

Results: The Warning Label Multi-Layer Image Identification algorithm was 99.3% accurate in detecting the presence of health warnings. Only 8.2% (n=73) of LCC influencer posts included a health warning. Influencer posts that contained health warnings received fewer likes (incidence rate ratio 0.59, P<.001, 95% CI 0.48-0.71) and fewer comments (incidence rate ratio 0.46, P<.001, 95% CI 0.31-0.67).

Conclusions: Health warnings are rarely used by influencers tagged by LCC brands' Instagram accounts. Very few influencer posts met the US Food and Drug Administration's health warning requirement of size and placement for tobacco advertising. The presence of a health warning was associated with lower social media engagement. Our study provides support for the implementation of comparable health warning requirements to social media tobacco promotions. Using an innovative computer vision approach to detect health warning labels in influencer promotions on social media is a novel strategy for monitoring health warning compliance in social media tobacco promotions.

(JMIR Infodemiology 2023;3:e41969) doi:10.2196/41969

KEYWORDS

tobacco; cigar; little cigar; cigarillo; Instagram; social media; influencer promotion; tobacco advertising; health warning; machine learning; computer vision; warning label; health label; health promotion; advertising; advertise; smoking; smoker; algorithm; visualization

Introduction

In 2021, overall 5.2% of middle and high school students in the United States reported ever using cigars, representing 1,400,000 youths who ever experimented with cigar products [1]. Cigars were also the most commonly used combustible tobacco products among US high school students in 2021 [1]. That cigars surpassed cigarettes in becoming the most popular combustible tobacco products among non-Hispanic Black middle and high school students is concerning [1]. Compared to non-Hispanic White youths, non-Hispanic Black youths had greater odds of initiating tobacco product use via cigars [2]. On the contrary, non-Hispanic White youths were more likely to initiate tobacco use through e-cigarettes [2]. Youths who initiated tobacco use via cigars were also more likely to become current tobacco product users of multiple products than youths who initiated tobacco use via e-cigarettes [2]. Importantly, longitudinal research suggests that the use of cigars may be a predictor of marijuana initiation among young college students [3]. Thus, cigar smoking among youths not only presents a critical public health issue but also raises concerns about health equity in tobacco prevention and control.

Tobacco advertising plays an important role in shaping tobacco-related knowledge, attitudes, and behaviors among youths [4]. A substantial body of evidence links advertising exposure with key factors that lead to youth tobacco use, such as curiosity about smoking, tobacco brand awareness, positive attitudes, and intentions to smoke [5-7]. In the United States, the tobacco industry has shifted its marketing efforts to the internet to circumvent restrictions on broadcast media (eg, TV and video), outdoor media (eg, billboards and public transit advertising), and tobacco product placement in entertainment media [8,9]. Specifically, in the United States, the 1998 Master Settlement Agreement prohibited tobacco companies from targeting youths in the advertising, promotion, or marketing of tobacco products [10]. The 2009 Family Smoking Prevention and Tobacco Control Act (Tobacco Control Act) established additional restrictions on youth-targeted tobacco marketing [11]. However, both the Master Settlement Agreement and Tobacco Control Act did not directly address social media–based tobacco marketing. As a result, the tobacco industry shifted its marketing efforts from traditional media forms such as print and billboards to the internet [12,13]. Social media–based advertising is largely an unregulated environment for tobacco companies to reach and engage current and potential customers [14,15].

Tobacco companies promote products on social media mainly through 2 means: brand-owned pages and influencer promotions. An analysis of 112 leading brands of tobacco products found that most brands had pages on at least 2 of the following social media platforms: Instagram, Facebook, Twitter, YouTube, Pinterest, and Tumblr [16]. In addition, tobacco companies promoted products through paid “influencers” who have large social media–based followings [17]. Social media promotions may be a more effective means for influencing tobacco perceptions and use than traditional advertising such as TV and print media that provide no interactive features, as social media gives the audience more options for engagement and interactions with tobacco content [18].

Influencers discussing tobacco products have the potential to affect followers’ attitudes and product use [19]. Followers of tobacco influencers are younger, have lower education, and are more likely to report past month tobacco use than those who do not follow tobacco influencers [20]. Youths are especially vulnerable to social media–based tobacco marketing in part due to their high level of internet and social media use [21], with youths aged 13-18 years spending over 8 hours on the internet every day [22]. Exposure and engagement with social media–based tobacco marketing, including social media promotions, are associated with tobacco product use among US youths [23,24].

Health warning statements serve as a source of health information, increase perceptions of risk, promote smoking cessation, and have the potential to lower smoking initiation among youths [25,26]. Health warnings are required to be displayed on the packaging and advertising for all tobacco products [27]. Even though the law also mandates the inclusion of health warning labels on cigars and pipe tobacco [28], the US District Court for the District of Columbia has issued an order vacating these requirements for cigar and pipe products. As a result, cigar and pipe tobacco firms may opt to voluntarily comply with the health warning provisions set by the US Food and Drug Administration (FDA). However, the 7 largest cigar companies in the United States must still display health warnings in both their advertising and packaging due to an existing consent agreement with the Federal Trade Commission [29]. These companies include Swisher International, Inc (producer of Swisher Sweets cigars) and Altadis U.S.A. (producer of Backwoods and Dutch Master cigars) [29]. The FDA mandates that health warning statements on advertising for covered, roll-your-own, and cigarette tobacco products must (1) appear on the upper portion of the advertisement within the trim area and (2) occupy at least 20% of the area of the advertisement [30]. However, social media advertising, including influencer promotional posts, has not been specified in the health warning requirements for any tobacco product.

Instagram is one of the most popular social media platforms among youth, with 72% of youths reporting Instagram use in 2018 [21]. The photo-oriented nature of Instagram makes it an ideal platform for influencer promotions, since images convey more emotions and intimate feelings than text-oriented platforms such as Twitter [31]. Health warning labels in celebrity-endorsed e-cigarette Instagram advertisements decreased viewers' intention to use e-cigarettes [32]. Tweets that include health warnings for e-cigarettes were found to elicit more negative health perceptions of the e-cigarette brand than those without warnings [33]. The inclusion of FDA-mandated nicotine warning statements in Instagram e-cigarette promotions decreased the appeal of the posts [34]. However, little is known about the use of health warning labels in LCC influencer posts on Instagram and whether the presence of health warnings is associated with user engagement with LCC influencer promotions. This study identified LCC influencers using innovative methods and performed computer vision analysis to investigate the current state of LCC influencer promotions on Instagram and the use of health warning labels in LCC influencer promotional posts. We report here the development and use of an innovative computer vision method. We then use our computer vision algorithm to evaluate the effect of warning labels of leading LCC Instagram influencers on post engagement.

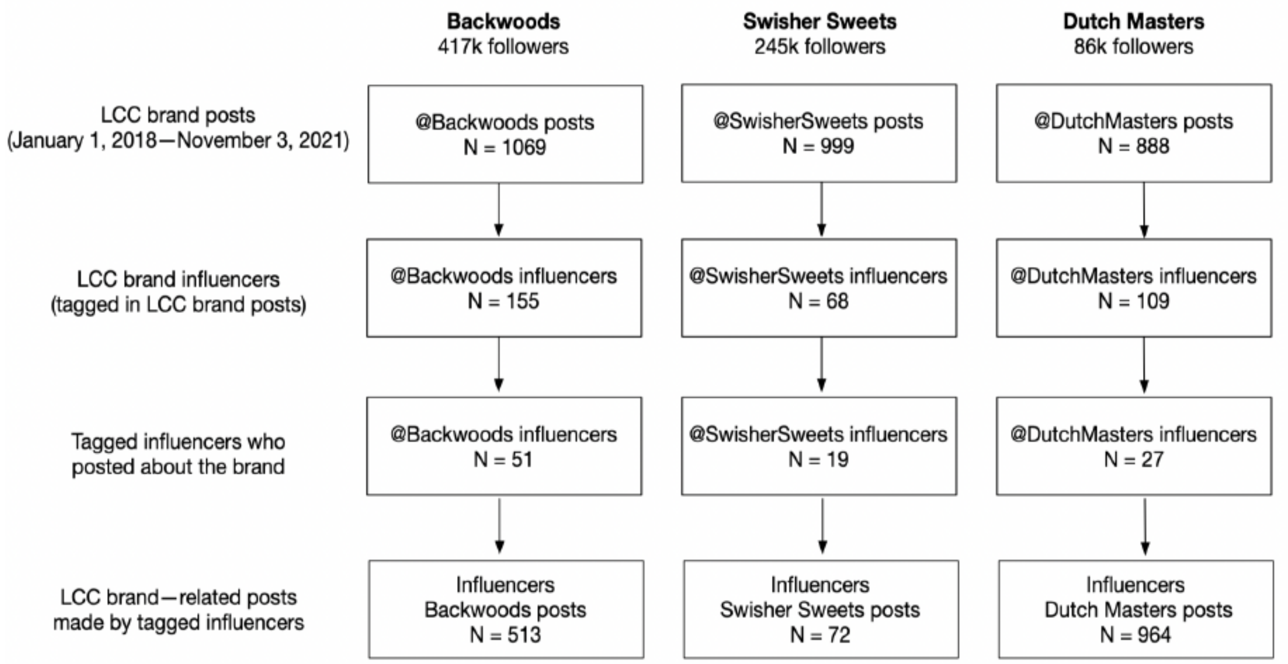
Methods

Data Collection

We focused on 3 LCC brands with the leading market shares in the United States, including Backwoods, Swisher Sweets, and Dutch Masters, which frequently feature influencers on their Instagram brand pages [35,36]. At the time of data collection, these 3 brands were also the most followed LCC brands on Instagram.

In violation of Federal Trade Commission guidance [37], many influencers do not use methods to disclose that they have a "material connection" with the brand, such as including hashtags such as #ad or #sponsored. As a result, paid influencer posts can be difficult to identify and study. Thus, influencers were identified as individuals tagged by one of the 3 LCC brands. First, we scraped all Instagram posts from the 3 LCC brands, which were posted between January 1, 2018, and November 3, 2021. Then, we used the string-matching function in the R software to identify all handles tagged in the captions of the collected LCC brand posts. During this time, Backwoods tagged 155 unique Instagram users, Swisher Sweets tagged 68 Instagram users, and Dutch Masters tagged 109 Instagram users. We collected posts referencing LCCs published by each identified influencer from the users' Instagram pages in November 2021. In total, we identified 51 Backwoods influencers who posted 513 Backwoods-related Instagram posts, 19 Swisher Sweet influencers who posted 72 Swisher Sweets-related Instagram posts, and 27 Dutch Masters influencers who posted 964 Dutch Masters–related Instagram posts (Figure 1). Pictures or videos of brand-related influencer posts were manually downloaded. Engagement metrics of each post, including the number of likes and comments, were also recorded. Collected data were stored in a password-protected computer and were only accessible to the authors.

Figure 1. Data sampling procedure. LCC: little cigar and cigarillo.



Manual Coding

Because not all influencer posts that mentioned brand names contained content related to cigar smoking or LCC brand–sponsored events, we manually coded whether the identified influencer LCC posts were relevant to cigar smoking and LCC brands. Two coders were trained to determine (1) if an influencer post was relevant (ie, it pertained to cigar smoking or LCC-sponsored events such as the Swisher Sweets Artist Project) and (2) among relevant posts, if the post contained a health warning. For video posts, coders watched the entire video and took screenshots if health warnings appeared.

To determine inter-coder reliability, 2 coders were trained on 20 images, and questions regarding coding criteria were discussed. Next, the 2 coders independently evaluated 150 posts (50 posts for each brand), after which Cohen kappa values were calculated to determine the intercoder reliability on the 2 coding questions. The average Cohen kappa values were 0.954 for question 1, that is, whether an influencer post was relevant (pertained to cigar smoking or LCC-sponsored events), and 0.966 for question 2, that is, whether the post contained a health warning (among LCC brand–relevant posts). The high Cohen kappa values indicate a high level of intercoder agreement between the 2 coders [38]. Two coders independently coded the remaining images in the sample.

The Warning Label Multi-Layer Image Identification Algorithm

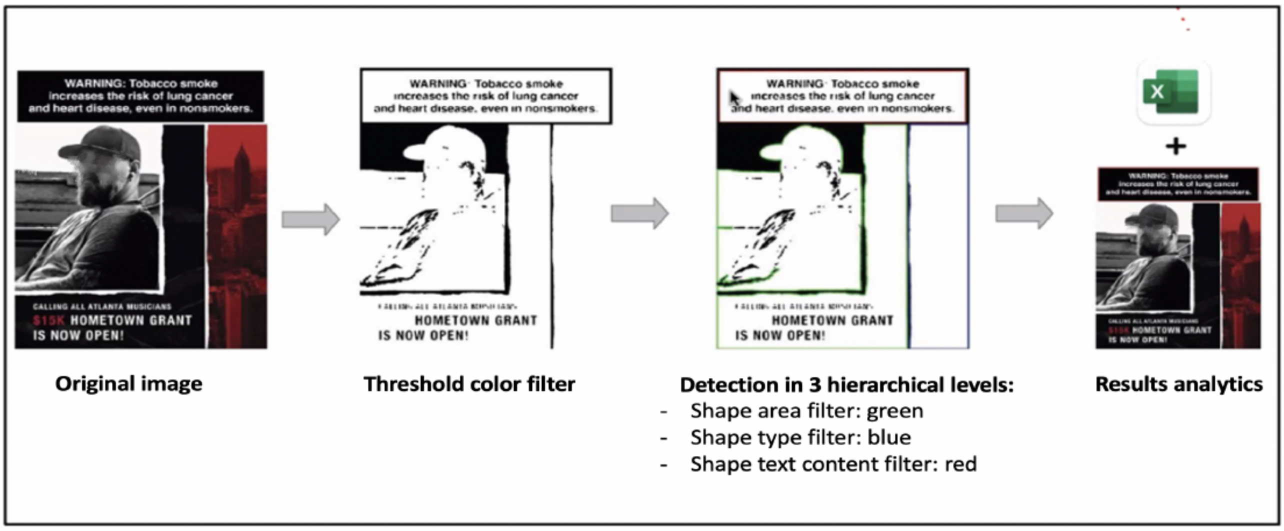
To determine the presence and properties of a warning label inside an image, we developed the Warning Label Multi-Layer Image Identification (WaLi) computer vision algorithm by integrating the computer vision library OpenCV [39] and the open-source OCR (Optical Character Recognition) engine Tesseract [40]. WaLi was developed to specifically identify compliance with 2 FDA guidelines for advertising of covered, roll-your-own, and cigarette tobacco products, which state that health warnings must (1) appear on the upper portion of the advertisement within the trim area and (2) occupy at least 20% of the area of the advertisement. In addition to detecting the word "warning" in the image, we also analyzed the image over 4 different levels: pixel color, pixel contours area, pixel contours shape, and text (using OCR). As a preliminary step, to preprocess the image and allow the algorithm to better detect the warning statement (if present), we applied a black and white color transformation and multiple blurring and morphological functions (erosion and dilation) to remove small noise components.

Since warning labels are required to have black borders, we applied a binary filter to select only the dark-colored pixels of interest and filter out all the others. The selected pixels were grouped into contours to analyze the shape and area. If the area of the contour was between 2 specified thresholds and the shape of the contour was described as a quadrilateral, the selected contour was passed to the last step of the analysis. The 2 area thresholds were selected using empirical experiments on the data available and on the basis of assumptions about how the warning label should look according to the FDA directives. In particular, in order to be able to track both valid and invalid warning labels (with respect to FDA regulations), we chose a minimum threshold of 600 pixels, which corresponds to the minimum size required for reading the words by the text OCR (more details about this method are provided in the next section) and a maximum threshold of half the image size.

Finally, we evaluated whether the image area identified by the contour contained the word "warning," as mandated by the FDA for the health warning labels in advertising of roll-your-own and cigarette tobacco products [30]. We used the tesseract text OCR to extract the text content in the selected image area. If the keyword "warning" was found, then the algorithm returned the position of the warning label and area of the health warning as results. To increase the flexibility of this process, different variations of the described color filters and image preprocessing functions (including blurring and morphology) were used in the initial step of the analysis, allowing for the identification of nonstandard warning labels.

For all post images, the WaLi computer vision algorithm was used to determine (1) if the image contained a health warning, (2) the area (in pixel) of the warning label in relation to the image area, and (3) the placement of the health warning (upper portion of the image; Figure 2). In Figure 2, we highlight the 4 main steps of the warning labels' detection process. The second image in the process describes the pixel color filter using binary thresholding with an example of the output given the original image as the input. The third image shows the output after combining the area, shape, and OCR filter (highlighted in the image with 3 different colors). Finally, the final output shows the original image with only the area that corresponds to the filter criteria highlighted in red. Our code is publicly available on GitHub [41].

Figure 2. Computer vision process.



Statistical Analysis

We used negative binomial regression models to assess the association of health warnings with the engagement (numbers of likes and comments) of an LCC influencer post. Negative binomial models account for overdispersion in count data [42] and have been used in prior research evaluating social media engagement (eg, post "likes") [43]. We also included influencers as random effects in the models as observations were nonindependent. Specifically, over half of the influencers published >1 LCC Instagram post. We also adjusted for follower counts and LCC brand, which can potentially affect the engagement of influencer posts on Instagram. Negative binomial models were fitted using the glmmTMB package in R (version 4.1.0; The R Foundation). The exponentiated regression coefficients in the negative binomial model are reported as incidence rate ratios (IRRs). Corresponding 95% CIs and 2-sided *P* values are also reported. Variance inflation factor scores for all independent variables were within the range of 0 to 2, indicating no substantive multicollinearity.

Ethical Considerations

As only publicly available data were used, the author's institutional review board determined that this study did not meet the definition of human participants research.

Results

Health Warnings in LCC Influencer Posts

In total, we identified 1549 LCC posts from influencers who were tagged by the 3 leading LCC brand Instagram pages. Manual coding revealed that 889 posts (470 Backwoods influencer posts, 60 Swisher Sweets influencer posts, and 359 Dutch Masters influencer posts) featured either cigar smoking or LCC brand–related events. Among the retained 889 LCC influencer posts, manual coding identified 79 posts that contained health warnings. When using the WaLi computer vision algorithm for the 889 LCC branded influencer posts, it successfully captured health warnings in 73 out of the 79 coder-identified posts. The computer vision method failed to detect health warnings in only 6 posts because the health warnings were either too blurry or too small (Figure 3). Thus, the overall accuracy of using computer vision to detect the presence or absence of health warnings in our data set was 99.3%. Subsequent analyses were based on the results reported by the computer vision analysis.

Only 8.2% (n=73) of the LCC influencer posts contained a health warning. Among these 73 posts, the health warning label occupied an average of 8.2% of the area of the post. Specifically, only 4.1% (n=3) of LCC influencer posts included a health warning that occupied at least 20% of the total post, as required by the FDA for advertisements of roll-your-own tobacco and cigarettes [30]. Only 23.3% (n=17) of the posts placed a health warning in the upper area of the post per FDA requirement. Overall, among the 889 identified influencer posts of LCCs on Instagram, only 1 (0.1%) fully met the FDA requirements for health warning labels in tobacco advertising for roll-your-own tobacco and cigarettes, containing a health warning that constitutes at least 20% of the area of the post, and placing the label in the upper portion of the post.

Figure 3. Examples of posts in which Computer Vision failed to detect the presence of health warnings.

Association of Health Warnings With LCC Influencer Post Engagement

In negative binomial analyses, influencer posts that contained health warnings received fewer likes (IRR 0.59, *P*<.001, 95% CI 0.48-0.71) and fewer comments (IRR 0.46, *P*<.001, 95% CI 0.31-0.67) than those that did not contain health warnings, after adjusting for the number of account followers and LCC brand. Holding the number of followers and LCC brand constant, the presence of a health warning in an LCC influencer post was associated with a 41% decrease in the rate of likes and a 54% decrease in the rate of comments.

Discussion

Developing and using an innovative computer vision method—WaLi—our study evaluated the use of health warnings in LCC influencer promotions on Instagram and the association of health warnings with post engagement. We found that few LCC influencers' promotional posts contained a health warning. Additionally, we evaluated the location and size of warnings in LCC influencer posts in accordance with the FDA's warning requirements for roll-your-own tobacco and cigarette advertising. Even though LCC influencer promotions currently fall outside the scope of the FDA health warning requirements, our findings reveal that a very small number of LCC influencer posts met the requirement for using health warnings that constitute 20% of the advertisement's area and are located in the upper portion of the advertisement. Notably, the presence of health warnings in LCC influencer promotional posts was associated with less post engagement, including fewer likes and comment counts.

The current federal laws require warnings on advertisements for tobacco products but do not specify whether the rules apply to social media advertising [29,30]. Our findings support the use of health warnings in tobacco branded content and influencer promotions on social media. We found that the presence of health warnings in LCC influencer posts was associated with lower post engagement (ie, likes and comments). Our research builds upon previous research, which demonstrated that exposure to and engagement with social media–based tobacco marketing, including social media marketing, is associated with tobacco product use among US youths [23,24].

Our study is among the first to demonstrate the relation of health warning statements with engagement of LCC-related content posted by influencers tagged by LCC brands. The LCC brands deliberately tagged an influencer account that has a wide reach on social media, which allows its users to engage with the tagged influencer account. It is possible for influencers to transition to brand ambassadors and establish a lasting partnership with a brand [44]. One example of this is Swisher Sweets and their use of the Artist Project to cultivate partnerships with various influencers within the hip-hop music industry [45]. Those influencers then generate complimentary branded content on their own Instagram pages and engage their followers with branded promotional content. Yet, very few LCC branded influencer posts contained a health warning. When health warnings were present, the posts were associated with less engagement. Given that compared to nonfollowers, followers of tobacco influencers are more likely to be younger and have a lower level of education, making the followers more susceptible to tobacco influencer promotions [20]. Cigar brands frequently employ influencers to market their products on Instagram, with the majority of the influencers being people of color from the music industry, who are particularly appealing to younger, African Americans [35]. Future research is needed to examine whether Black individuals are more likely to encounter LCC influencer promotions on social media, as well as the effects of such exposure on their beliefs and cigar use.

Our study suggests that the incorporation of health warning requirements in social media posts promoting tobacco products could help reduce engagement with promotions of tobacco products on social media. Future studies are needed to identify the most effective implementation strategy of health warnings in social media promotional posts of tobacco to decrease the use and uptake of tobacco products among youths [33]. For example, irrespective of whether the effect of health warning labels on tobacco perceptions and initiation depends upon the size and design [25], it is unclear if previous guidelines on size and design, which focused on health warnings for product packaging [46], are effective in a social media environment.

Our study has several limitations. We only analyzed 3 LCC brands on one social media platform (Instagram). Our findings may not be generalizable to other tobacco products and social media platforms. We also do not know the tobacco use status of the individuals who engaged with the LCC influencer posts; thus, we cannot demonstrate the causal effects of health warnings on the audience’s attitudes and behaviors toward LCCs. Future research is needed to investigate the effects of health warnings in social media promotions of LCCs on youths’ attitudes toward LCCs, their onset, and use. Lastly, despite the fact that the identified influencers were tagged by LCC brand posts, we cannot verify if those individuals are paid by the LCC brands. Influencers can be compensated not only monetarily but also through performance and collaboration opportunities. It is important to conduct further research on how the tobacco industry uses these strategies to attract influencers and promote tobacco products in order to inform future advertising policies in the digital age.

Our study used an innovative strategy to identify influencers who were tagged by the leading LCC brands. This method of identifying influencers can be used to study influencer promotions of other tobacco products such as e-cigarettes, which are frequently advertised on social media [47]. We also implemented a state-of-the-art computer vision method to detect and evaluate health warning statement properties in Instagram influencer promotions. Compared to traditional manual coding of images, WaLi is much more efficient at providing an accurate measurement of the size and location of health warnings in visual tobacco advertisements. Such efficiency and accuracy can be particularly useful for monitoring compliance with health warning regulations in tobacco advertising and for evaluating the effectiveness of such policies in reducing the appeal of tobacco products. Thus, WaLi can be readily used for policy evaluation and surveillance of health warning compliance in various social media–based tobacco advertisements such as email and website advertisements.

In conclusion, our study lends support for the requirement of health warning statements in brand-related influencer promotions of LCCs on social media. Extending health warning requirements to social media advertisements of tobacco products warrants further research.

Acknowledgments

Research reported in this publication was supported, in part, by the National Heart, Lung, And Blood Institute of the National Institutes of Health (award U54HL120163). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. JW and LRR report receiving grants from American Heart Association. JLF reports receiving grants from the American Heart Association and the National Heart, Lung, and Blood Institute (K01 HL143142).

Authors' Contributions

JLF and TH were responsible for conceptualization and investigation. JW was responsible for data collection. JLF, TH, and JW were responsible for developing the study methodology. JMO and DW were responsible for computer vision analysis. JLF, TH, and JW were responsible for drafting the original manuscript. JLF and TH were responsible for supervising the study. LRR, AB, RMR, ZX, DW, and JMO were responsible for review and editing the manuscript draft. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Gentzke AS, Wang TW, Cornelius M, Park-Lee E, Ren C, Sawdey MD, et al. Tobacco product use and associated factors among middle and high school students - National Youth Tobacco Survey, United States, 2021. MMWR Surveill Summ 2022 Mar 11;71(5):1-29 [FREE Full text] [doi: 10.15585/mmwr.ss7105a1] [Medline: 35271557]
2. Kowitt SD, Goldstein AO, Sutfin EL, Osman A, Meernik C, Heck C, et al. Adolescents' first tobacco products: associations with current multiple tobacco product use. PLoS One 2019;14(5):e0217244 [FREE Full text] [doi: 10.1371/journal.pone.0217244] [Medline: 31120972]
3. Cornacchione Ross J, Sutfin E, Suerken C, Walker S, Wolfson M, Reboussin B. Longitudinal associations between marijuana and cigar use in young adults. Drug Alcohol Depend 2020 Mar 23;211:107964 [FREE Full text] [doi: 10.1016/j.drugalcdep.2020.107964] [Medline: 32251920]
4. National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. Preventing Tobacco Use Among Youth and Young Adults. A Report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention; 2012.
5. Lovato C, Linn G, Stead LF, Best A. Impact of tobacco advertising and promotion on increasing adolescent smoking behaviours. Cochrane Database Syst Rev 2003(4):CD003439. [doi: 10.1002/14651858.CD003439] [Medline: 14583977]
6. Portnoy DB, Wu CC, Tworek C, Chen J, Borek N. Youth curiosity about cigarettes, smokeless tobacco, and cigars: prevalence and associations with advertising. Am J Prev Med 2014 Aug;47(2 Suppl 1):S76-S86 [FREE Full text] [doi: 10.1016/j.amepre.2014.04.012] [Medline: 25044199]
7. Li X, Borodovsky JT, Kasson E, Kaiser N, Riordan R, Fentem A, et al. Exploring how tobacco advertisements are associated with tobacco use susceptibility in tobacco naive adolescents from the PATH study. Prev Med 2021 Dec;153:106758 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106758] [Medline: 34358594]
8. Bach L. Tobacco product marketing on the internet. Campaign for Tobacco-Free Kids. 2021. URL: https://www.tobaccofreekids.org/assets/factsheets/0081.pdf [accessed 2023-02-21]
9. Pierce JP, Gilpin EA. How did the Master Settlement Agreement change tobacco industry expenditures for cigarette advertising and promotions? Health Promot Pract 2004 Jul;5(3 Suppl):84S-90S. [doi: 10.1177/1524839904264600] [Medline: 15231101]
10. The Master Settlement Agreement. National Association of Attorneys General. URL: https://www.naag.org/our-work/naag-center-for-tobacco-and-public-health/the-master-settlement-agreement/ [accessed 2021-01-18]
11. Family Smoking Prevention and Tobacco Control Act - An Overview. US Food and Drug Administration. 2020. URL: https://www.fda.gov/tobacco-products/rules-regulations-and-guidance/family-smoking-prevention-and-tobacco-control-act-overview [accessed 2022-01-18]
12. Lewis MJ, Yulis SG, Delnevo C, Hrywna M. Tobacco industry direct marketing after the Master Settlement Agreement. Health Promot Pract 2004 Jul 25;5(3 Suppl):75S-83S. [doi: 10.1177/1524839904264596] [Medline: 15231100]
13. Federal Trade Commission Cigarette Report for 2020 and Smokeless Tobacco Report for 2020. Federal Trade Commission. 2021. URL: https://www.ftc.gov/reports/federal-trade-commission-cigarette-report-2020-smokeless-tobacco-report-2020 [accessed 2021-01-18]
14. Richardson A, Ganz O, Vallone D. Tobacco on the web: surveillance and characterisation of online tobacco and e-cigarette advertising. Tob Control 2015 Jul;24(4):341-347. [doi: 10.1136/tobaccocontrol-2013-051246] [Medline: 24532710]
15. Jackler RK, Li VY, Cardiff RAL, Ramamurthi D. Promotion of tobacco products on Facebook: policy versus practice. Tob Control 2019 Jan;28(1):67-73. [doi: 10.1136/tobaccocontrol-2017-054175] [Medline: 29622602]
16. O'Brien EK, Hoffman L, Navarro MA, Ganz O. Social media use by leading US e-cigarette, cigarette, smokeless tobacco, cigar and hookah brands. Tob Control 2020 Dec 26;29(e1):e87-e97. [doi: 10.1136/tobaccocontrol-2019-055406] [Medline: 32217772]
17. New Investigation Exposes How Tobacco Companies Market Cigarettes on Social Media in the U.S. and Around the World. Campaign for Tobacco-Free Kids. 2018. URL: https://www.tobaccofreekids.org/press-releases/2018_08_27_ftc [accessed 2023-02-21]
18. Calder BJ, Malthouse EC, Schaedel U. An experimental study of the relationship between online engagement and advertising effectiveness. J Interact Mark 2009 Nov 01;23(4):321-331. [doi: 10.1016/j.intmar.2009.07.002]

19. Daniel ES, Crawford Jackson EC, Westerman DK. The influence of social media influencers: understanding online vaping communities and parasocial interaction through the lens of Taylor's six-segment strategy wheel. J Interact Advert 2018 Oct 18;18(2):96-109. [doi: 10.1080/15252019.2018.1488637]
20. Chu K, Majmundar A, Allem J, Soto DW, Cruz TB, Unger JB. Tobacco use behaviors, attitudes, and demographic characteristics of tobacco opinion leaders and their followers: Twitter analysis. J Med Internet Res 2019 Jun 04;21(6):e12676 [FREE Full text] [doi: 10.2196/12676] [Medline: 31165716]
21. Anderson M, Jiang J. Teens, Social Media and Technology 2018. Pew Research Center. 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2021-07-03]
22. The Common Sense Census: Media Use by Tweens and Teens, 2021. Common Sense Media. 2022. URL: https://www.commonsensemedia.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2021 [accessed 2023-02-21]
23. Soneji S, Pierce JP, Choi K, Portnoy DB, Margolis KA, Stanton CA, et al. Engagement with online tobacco marketing and associations with tobacco product use among U.S. youth. J Adolesc Health 2017 Jul;61(1):61-69 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.01.023] [Medline: 28363720]
24. Soneji S, Yang J, Knutzen KE, Moran MB, Tan AS, Sargent J, et al. Online tobacco marketing and subsequent tobacco use. Pediatrics 2018 Feb 02;141(2):e20172927 [FREE Full text] [doi: 10.1542/peds.2017-2927] [Medline: 29295893]
25. Hammond D. Health warning messages on tobacco products: a review. Tob Control 2011 Sep;20(5):327-337. [doi: 10.1136/tc.2010.037630] [Medline: 21606180]
26. Moodie C, MacKintosh AM, Gallopel-Morvan K, Hastings G, Ford A. Adolescents' perceptions of an on-cigarette health warning. Nicotine Tob Res 2017 Oct 01;19(10):1232-1237 [FREE Full text] [doi: 10.1093/ntr/ntw165] [Medline: 27613916]
27. Labeling and Warning Statements for Tobacco Products. US Food and Drug Administration. 2021. URL: https://www.fda.gov/tobacco-products/products-guidance-regulations/labeling-and-warning-statements-tobacco-products [accessed 2022-01-30]
28. Cigar Labeling and Warning Statement Requirements. US Food and Drug Administration. 2020. URL: https://www.fda.gov/tobacco-products/labeling-and-warning-statements-tobacco-products/cigar-labeling-and-warning-statement-requirements [accessed 2022-01-30]
29. Nationwide Labeling Rules for Cigar Packaging and Ads Take Effect Today. Federal Trade Commission. 2001. URL: https://www.ftc.gov/news-events/press-releases/2001/02/nationwide-labeling-rules-cigar-packaging-and-ads-take-effect [accessed 2021-01-30]
30. Advertising and Promotion. US Food and Drug Administration. 2020. URL: https://www.fda.gov/tobacco-products/products-guidance-regulations/advertising-and-promotion [accessed 2021-10-31]
31. Hjorth L, Burgess J. Intimate Banalities The Emotional Currency of Shared Camera Phone Images during the Queensland Flood Disaster. In: The Routledge Companion to Mobile Media. Oxon: Routledge; 2014.
32. Phua J, Lim D. Can warning labels mitigate effects of advertising message claims in celebrity-endorsed Instagram-based electronic cigarette advertisements? Influence on social media users' E-cigarette attitudes and behavioral intentions. J Mark Commun 2022 Feb 09:1-21. [doi: 10.1080/13527266.2022.2037008]
33. Guillory J, Kim AE, Fiacco L, Cress M, Pepper J, Nonnemaker J. An experimental study of nicotine warning statements in e-cigarette tweets. Nicotine Tob Res 2020 Apr 21;22(5):814-821 [FREE Full text] [doi: 10.1093/ntr/ntz029] [Medline: 30820571]
34. Laestadius LI, Penndorf KE, Seidl M, Cho YI. Assessing the appeal of Instagram electronic cigarette refill liquid promotions and warnings among young adults: mixed methods focus group study. J Med Internet Res 2019 Nov 25;21(11):e15441 [FREE Full text] [doi: 10.2196/15441] [Medline: 31763987]
35. Navarro MA, O'Brien EK, Ganz O, Hoffman L. Influencer prevalence and role on cigar brand Instagram pages. Tob Control 2021 Nov;30(e1):e33-e36. [doi: 10.1136/tobaccocontrol-2020-055994] [Medline: 33046583]
36. Delnevo CD, Miller Lo E, Giovenco DP, Cornacchione Ross J, Hrywna M, Strasser AA. Cigar sales in convenience stores in the US, 2009-2020. JAMA 2021 Dec 21;326(23):2429-2432 [FREE Full text] [doi: 10.1001/jama.2021.19692] [Medline: 34932088]
37. Fair L. Disclosures 101: New FTC resources for social media influencers. Federal Trade Commission. 2019. URL: https://www.ftc.gov/news-events/blogs/business-blog/2019/11/disclosures-101-new-ftc-resources-social-media-influencers [accessed 2021-07-03]
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]
39. OpenCV. URL: https://opencv.org/ [accessed 2023-02-21]
40. Weil S. Tesseract open source OCR engine (main repository). GitHub. 2022. URL: https://github.com/tesseract-ocr/tesseract/releases/ [accessed 2023-02-21]
41. Warning-Label-Multi-Layer-Image-Identification (WaLi). GitHub. URL: https://github.com/JMOriggi/Warning-Label-Multi-Layer-Image-Identification [accessed 2023-02-22]
42. Chin HC, Quddus MA. Modeling count data with excess zeroes. Sociol Methods Res 2016 Jun 29;32(1):90-116. [doi: 10.1177/0049124103253459]

43. Bhattacharya S, Srinivasan P, Polgreen P. Social media engagement analysis of U.S. Federal health agencies on Facebook. BMC Med Inform Decis Mak 2017 Apr 21;17(1):49 [FREE Full text] [doi: 10.1186/s12911-017-0447-z] [Medline: 28431582]
44. Richardson A, Ganz O, Vallone D. The cigar ambassador: how Snoop Dogg uses Instagram to promote tobacco use. Tob Control 2014 Jan;23(1):79-80. [doi: 10.1136/tobaccocontrol-2013-051037] [Medline: 23748187]
45. Ganz O, Rose SW, Cantrell J. Swisher Sweets 'Artist Project': using musical events to promote cigars. Tob Control 2018 Jul 08;27(e1):e93-e95. [doi: 10.1136/tobaccocontrol-2017-054047] [Medline: 29439208]
46. White V, Webster B, Wakefield M. Do graphic health warning labels have an impact on adolescents' smoking-related beliefs and behaviours? Addiction 2008 Sep;103(9):1562-1571. [doi: 10.1111/j.1360-0443.2008.02294.x] [Medline: 18783508]
47. Huang J, Duan Z, Kwok J, Binns S, Vera LE, Kim Y, et al. Vaping versus JUULing: how the extraordinary growth and marketing of JUUL transformed the US retail e-cigarette market. Tob Control 2019 Mar;28(2):146-151 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054382] [Medline: 29853561]

Abbreviations

FDA: US Food and Drug Administration
IRR: incidence rate ratio
LCC: little cigar and cigarillo
OCR: Optical Character Recognition
WaLi: Warning Label Multi-Layer Image Identification

Edited by T Mackey; submitted 16.08.22; peer-reviewed by O Ganz, J Kapp; comments to author 18.01.23; revised version received 30.01.23; accepted 06.02.23; published 14.03.23.

Please cite as:
Wu J, Origgi JM, Ranker LR, Bhatnagar A, Robertson RM, Xuan Z, Wijaya D, Hong T, Fetterman JL
Compliance With the US Food and Drug Administration's Guidelines for Health Warning Labels and Engagement in Little Cigar and Cigarillo Content: Computer Vision Analysis of Instagram Posts
JMIR Infodemiology 2023;3:e41969
URL: https://infodemiology.jmir.org/2023/1/e41969
doi:10.2196/41969
PMID:37113379

Original Paper

Characterizing the Discourse of Popular Diets to Describe Information Dispersal and Identify Leading Voices, Interaction, and Themes of Mental Health: Social Network Analysis

Melissa C Eaton[1], BSc, MNutr&Diet; Yasmine C Probst[1], MHlthInfo, MSc, PhD; Marc A Smith[2], BS, MPhil, PhD

[1]School of Medical, Indigenous and Health Sciences, University of Wollongong, Wollongong, Australia

[2]Social Media Research Foundation, Redwood City, CA, United States

Corresponding Author:
Melissa C Eaton, BSc, MNutr&Diet
School of Medical, Indigenous and Health Sciences
University of Wollongong
Northfields Ave
Wollongong, 2522
Australia
Phone: 61 242215302
Email: mce527@uowmail.edu.au

Abstract

Background: Social media has transformed the way health messages are communicated. This has created new challenges and ethical considerations while providing a platform to share nutrition information for communities to connect and for information to spread. However, research exploring the web-based diet communities of *popular diets* is limited.

Objective: This study aims to characterize the web-based discourse of popular diets, describe information dissemination, identify influential voices, and explore interactions between community networks and themes of mental health.

Methods: This exploratory study used Twitter social media posts for an online social network analysis. Popular diet keywords were systematically developed, and data were collected and analyzed using the NodeXL metrics tool (Social Media Research Foundation) to determine the key network metrics (vertices, edges, cluster algorithms, graph visualization, centrality measures, text analysis, and time-series analytics).

Results: The vegan and ketogenic diets had the largest networks, whereas the zone diet had the smallest network. In total, 31.2% (54/173) of the top users endorsed the corresponding diet, and 11% (19/173) claimed a health or science education, which included 1.2% (2/173) of dietitians. Complete fragmentation and hub and spoke messaging were the dominant network structures. In total, 69% (11/16) of the networks interacted, where the ketogenic diet was mentioned most, with depression and anxiety and eating disorder words most prominent in the "zone diet" network and the least prominent in the "soy-free," "vegan," "dairy-free," and "gluten-free" diet networks.

Conclusions: Social media activity reflects diet trends and provides a platform for nutrition information to spread through resharing. A longitudinal exploration of popular diet networks is needed to further understand the impact social media can have on dietary choices. Social media training is vital, and nutrition professionals must work together as a community to actively reshare evidence-based posts on the web.

(JMIR Infodemiology 2023;3:e38245) doi:10.2196/38245

KEYWORDS

social media; popular diets; nutrition; public health; social network analysis

Introduction

Background

Social media, consisting of web-based networking sites such as Twitter, Facebook, and Instagram, has changed the way health messages are communicated [1], creating new challenges and ethical considerations [2]. In 2016, a total of 3.5 billion people worldwide were regular internet users, with more than two-thirds using social media [3]. In 2019, Facebook was the most popular global platform with 2.38 billion users, followed by YouTube, Instagram, TikTok, and Twitter [3]. The emergence of web-based social networking apps has allowed for novel research exploring social interactions, with data from the social networking site to be used [4]. Twitter, for example, enables the access to and analysis of big data over time, creating a means to study human behavior, interactions, and patterns through social network analysis (SNA) [4-6].

The SNA research methodology is primarily based on social network theory [5], suggesting that we understand social connectedness by analyzing the components of related experiences [7]. Over the last decade, SNA has been applied to health research related to physical activity, obesity, and policy change by way of understanding behavior and the transmission of information [8,9]. Social media has opened up the opportunity for SNA to be applied to web-based networks [8].

The web-based SNA methodology has been used in the disciplines of computer science, politics, climate change, education, and health [10-12]. In health, topic areas have included health-related conspiracy theories, public health messaging, and exploring the web-based discourse surrounding topical health and diseases such as COVID-19 [8,12,13]. However, to the best of our knowledge, there is no existing SNA in nutrition aimed at describing web-based discourse related to popular diets. Current research on social media and nutrition consists predominately of intervention, descriptive, and content analysis methods within target populations [14]. Although important, the growing influence and use of social media [3] and public nutrition messages [15] suggest that it is beneficial to consider the SNA methodology to allow for the exploration of patterns to monitor nutrition messages and to gain insights into how nutrition information is spread.

Credible public nutrition messages exist on the web [16]; however, misinformation is regularly shared, including potentially dangerous health messages [15,17,18] and pseudoscientific recommendations [19]. Furthermore, health and diet are 2 of the most common categories of misinformation on the web [18]. In particular, restrictive and popular dietary patterns, including fad diets [20,21], are forms of dietary misinformation that are regularly dispersed on the web [15]. Although there is evidence to support some popular diets in specific population groups [22,23], concern arises when restrictive diets are promoted to the general population using a one-size-fits-all approach.

The promotion of unbalanced nutrition information may be exaggerated on social media, with algorithms designed to show users similar content to what they interact with [24]. This may contribute to the development of an "echo chamber" [25], which can result in unbalanced views on a topic [24]. Kulshrestha et al [26] found this to be true for user intake of diet-related information on Twitter, with the diet content of interest heavily focused on only 1 or 2 topics [26]. As diet-related information consumed on the web can influence dietary choices [27,28], algorithms and "echo chambers" may enhance the promotion of popular diets [29]. This highlights the important role of dietitians and other qualified nutrition professionals on social media. It is vital for such professionals to be the trusted voice of nutrition in the web-based landscape, by sharing evidence-based health information, and to help identify and rectify dietary misinformation [30-32].

Furthermore, restrictive diets have been linked to negative physiological and psychological health outcomes, including eating disorders and depression and anxiety [21,33-35]. There is evidence to support restrictive dieting and body dissatisfaction as risk factors for depression [33,36,37], and anxiety has been associated with extreme dieting behaviors and binge eating [38-40]. These eating behaviors are also well-known risk factors for the development of both disordered eating and clinical eating disorders [41-43]. Therefore, it is important to explore how themes of mental health may exist within web-based diet networks.

Previous research has demonstrated the potential of social media to detect, identify, monitor, and classify mental health conditions [44-47]. De Choudhury et al [48] showed that social networking sites may be used to detect and identify populations with depression [48]. Karami et al [49] identified mental health as a common subtopic of diet-related conversations. In addition, Wilksche [50] found a negative association between social media and disordered eating behaviors in young adolescents following a content analysis of Twitter, whereas Zhou et al [44] found web-based eating disorder behavior described on Twitter to reflect that of offline eating disorder psychopathology. Therefore, as the web-based SNA methodology allows for the exploration of themes, text, and keywords within a network [51,52], this study aimed to explore the mental health themes within existing web-based popular diet networks.

Objectives

This study provides novel insights into how nutrition information is dispersed, the key influential voices of popular diet networks on Twitter, describes how users interact, and explores any related themes of mental health. Owing to the exploratory nature, hypotheses will be generated from the outcomes of the analyses, adding insight for future research. The study objectives were to (1) explore network dissemination and how messages may spread, (2) identify key influential voices of each network, (3) explore the interaction between popular diet networks, and (4) explore the interaction between popular diet networks and mental health. Table 1 summarizes the objectives and their associated outcome measures.

Table 1. Summary of the research objectives and their associated outcome measures.

Research objectives	Associated outcome measures
Explore network dissemination and how messages may spread	• Network size and duration of data • Cluster algorithm and graph visualization
Identify key influential voices of each network	• Betweenness centrality and out degree
Explore the interaction between popular diet networks	• Text analysis of high-frequency hashtags
Explore the interaction between popular diet networks and mental health	• Text analysis of mental health word list

Methods

Overview

This exploratory study used social media posts for an online SNA. Data were accessed from the social networking platform—Twitter, using the NodeXL SNA metrics tool [51]. NodeXL is a software plug-in for Microsoft Excel that allows the extraction, storage, analysis, and visualization of social network data [5,52,53]. NodeXL creates maps and visualizations of public conversations and connections between Twitter users [52]. Although ethically low risk as public data were extracted from Twitter, user names were deidentified by removing Twitter handles, and accounts were categorized using public profile data extracted via NodeXL. The categories included Twitter user (account identity), web-based business (business type), video-sharing platform, government initiative, community initiative, nonprofit initiative, television personality or host (related show or identity), actor, brand (brand type), web-based marketing company, personal brand (brand identity), politician, science- or health-related occupation (eg, Dr, Medical Doctor, PhD, and dietitian), band (band name), and unknown.

Keyword Development

A systematic search strategy was used to source tweets (posts) from Twitter. The strategy was developed from a review by Ge et al [54] and Obert et al [55] to devise a list of keywords representing popular diets. To allow for potential colloquialisms associated with diets when discussed on the web, keywords derived from the coding used by Ramachandran et al [15] were also included.

To identify popular diets that were more restrictive in nature, the diet list was assessed by 2 independent reviewers (ME and YP) against a published definition. Diets were included if they promoted promises of weight loss [21] and either (1) suggested intake of macronutrients in particular proportions or (2) avoidance of particular foods or macronutrients [20].

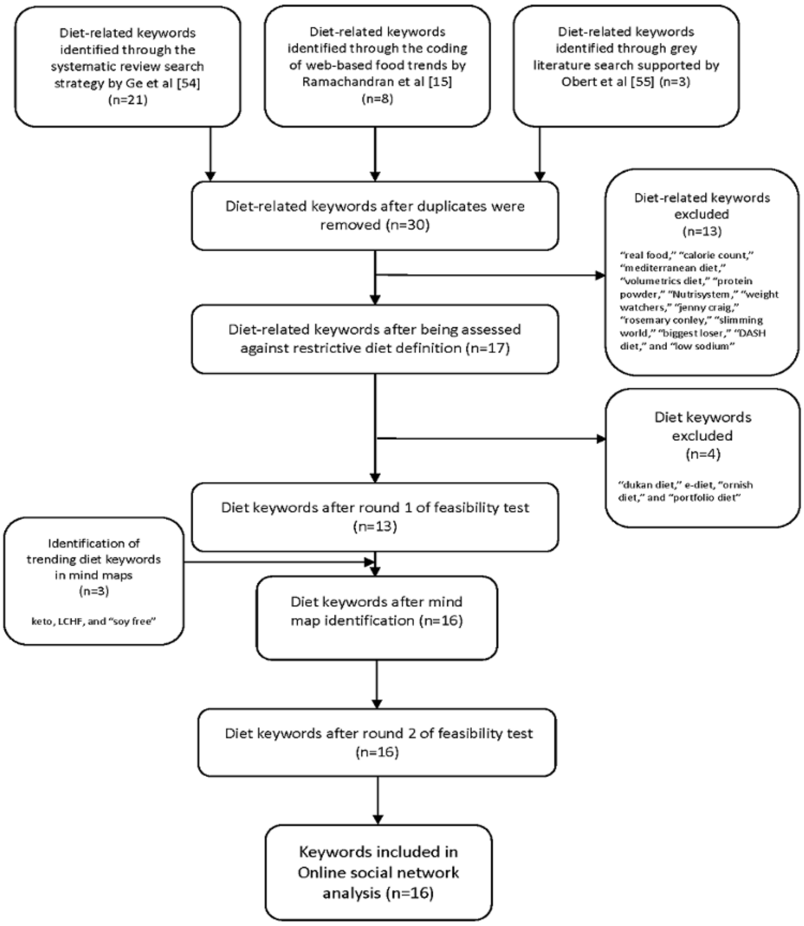
The resulting popular diets formed the initial keywords that underwent 2 stages of feasibility testing using NodeXL. First, each keyword was individually searched at a rate limit of 1000 tweets. The keywords were excluded if the search (1) identified "no users in the network" (which would result in no edges) or (2) produced limited results consisting of <5 unique edges, which signaled a unique connection or interaction between 2 vertices (users).

To ensure that current and trending web-based popular diet discussions were captured and to identify gaps that were not previously captured in the literature, the top 10 hashtags of all the included keywords were identified using a network map (Multimedia Appendix 1 [52,56]; Figure 1). Hashtags were selected over "top words" because of their ability to categorize and pinpoint key messages on a topic [57-59]. Hashtags that were unrelated (eg, #catsoftwitter and #dinosaur) or that were more generally related to diet and health (eg, #diet, #weightloss, #health, #recipe, #healthy, and #food) were excluded. New keywords were included if they (1) fulfilled the inclusion criteria listed above in *Keyword Development* and (2) had at least 2 connections to the previously established keywords.

The top hashtags were also used to determine the validity of words that could represent popular diets chosen for ethical or medical reasons, such as vegan, "gluten free," and "dairy free." These keywords were considered appropriate for inclusion if they had at least 1 connection in the network map with the previously established keywords (Multimedia Appendix 1; Figure 1).

Second, individual keywords were searched in NodeXL at a rate limit of 5000 tweets to assess the web-based network for specificity and relevance. The keyword strategy outlined above in *Keyword Development* was further refined if the extracted data contained polysemy words, resulting in irrelevant topics (ie, words that had multiple meanings and were therefore not relevant). Two reviewers (ME and MS) independently reviewed the updated keyword strategy. A flow diagram outlining the keyword development process is shown in Figure 1.

Figure 1. Flow diagram of keyword development.



Data Extraction

After verifying the keywords, data from Twitter were extracted using NodeXL at a predetermined rate limit of 20,000 tweets [51] to ensure that the networks were comparable. The keywords selected for data collection included the following: paleo (food OR diet OR meal OR dine OR eat OR keto OR “low carb” OR “gluten free” OR “weight loss” OR “healthy” OR “recipes”); “raw food”; vegan; “sugar free”; “dairy free”; “gluten free”; “low carb”; “low fat”; “zone diet”; “atkins diet”; “south beach diet”; “keto”; “intermittent fasting”; “detox diet”; “LCHF”; and “soy free.” Each keyword was searched individually using NodeXL, and each network included all data that were publicly available at the time of extraction. This included Twitter handles, usernames, user bios, user tweets, user mentions and reshares, URLs, hashtags within tweets, user profile images, tweet date and time, and user country.

Data were collected daily in July 2020 from tweets spanning May 12 to July 20, 2020. The timeframe was dictated by the availability of tweets in each network (ie, with some spanning longer or shorter periods depending on the network size). Of note, the smaller networks spanned over a longer period, with more time needed to reach the predetermined threshold of 20,000. In some cases, data collection did not reach the threshold. The threshold was selected because it allowed for processing on available desktop and laptop computing resources and created multiweek and overlapping data sets [60].

The public Twitter application programming interface (API) was used because access to other APIs was limited at the time of extraction. It should be noted that although this API has limitations, it does not contain false positives, and limitations are present only in the form of false negatives [61,62]. To address this issue, repeated daily data collection was performed.

Data Analysis

Data were analyzed using NodeXL through the application of network metrics [53]. All network metrics serve the function of analyzing connections and patterns that exist within the data set [52]. The specific metrics include vertices, edges, cluster algorithms, graph visualization, text analysis, centrality measures (in-degree, out degree, and betweenness), and time-series analytics. For definitions of the network metrics, refer to Multimedia Appendix 1.

Network Message Dissemination

Vertices and time-series analytics were used to determine the size of the individual networks. Vertices or vertex, otherwise known as social media handles, correspond to the number of users within a network, whereas time-series metrics identify the time frame of data collection. A scatter plot was created to determine the size of the network across both dimensions.

Cluster algorithms and graph visualization were used to analyze the social connections between users [53]. An edge, also known as a relationship, tie, or link, represents the connection or interaction between 2 vertices (users)—visualized as a line between users [53]. Graph metrics visualize interactions, and from this, structures emerge by forming patterns and relationships that can be analyzed [53]. Social media form a range of network structures that reflect the social process that generates them. The divided, unified, fragmented, clustered, and in-hub and out-hub patterns each capture a common structure found on social media platforms such as Twitter [63]. Refer to Multimedia Appendix 1 for figure structure visualizations. Cluster algorithms identify, group, and analyze network vertices (users) that have shared characteristics [51]. The Clauset-Newman-Moore cluster algorithm, designed to extract community structures from networks [64], was used to group the network vertices (users). The vertices were then arranged using the Harel-Koren Fast Multiscale layout algorithm [65].

Influential Voices of Each Network

Betweenness and outdegree centrality measures were used to determine the strength of relationships within the network and identify the top 10 users who were the most influential according to the social network theory [53]. Betweenness centrality identifies the center of each network and the users who have the greatest importance and "influence," measured by the behavior and connectivity between the user and others within the network [30]. Degree centrality measures the number of unique connections linked to a vertex (user) and identifies the users that are "most popular" in the network [52,66]. In a directed graph, this is measured as in-degree (the number of connections directed toward the user, ie, another user posting about the user of interest) or outdegree (the number of connections directed away from the user, ie, user of interest actively posting) [52].

To identify the most influential "active influencers" (those who were actively posting and sharing information), the users with both the highest betweenness score and an outdegree score of at least 1.0, representing users who actively tweeted and were not only being "tweeted about," were identified. These measures were used as users at the center of the network with the greatest number of connections served as a bridge connecting other users within each network. If these users were removed, the sharing and spreading of information would be affected [52].

Finally, user bio information and tweet content captured by NodeXL were used to categorize the accounts of the top 10 key users of each network (discussed earlier) and identify whether they actively supported the related popular diet. This was achieved by categorizing their account identity and identifying whether their tweets were suggestive of or actively promoting the diet to the general population.

Interaction Between Popular Diet Networks

Text analysis metrics generated by NodeXL were used to identify the most common hashtags appearing in all tweets within each network [52]. This metric explored possible interactions between 16 popular diet networks by identifying common, recurring, and overlapping hashtags. Hashtags, as opposed to top words (also extracted by NodeXL), were used as they categorized and pinpointed key messages on precise topics [57-59].

Interaction Between Networks and Mental Health

Text analysis was conducted using the NodeXL software to explore the potential associations between web-based popular diet networks and mental health. The NodeXL text analysis feature analyzes the words in each tweet and identifies whether the word is present in 1 of the 2 different word lists [52]. In the interest of our study, text analysis was used to explore the themes of mental health.

Two text classification word lists were created for (1) depression and anxiety and (2) eating disorders, based on the preexisting literature that used a combination of manual and lexical approaches [67]. Therefore, this process involved collating words from preexisting literature, which has demonstrated potential in measuring, classifying, identifying, and detecting various mental health conditions on social media [44-48,68-73]. The full list of text analysis words used can be found in Multimedia Appendix 2 [44-48,68-71,73].

To differentiate between each list, ensure accuracy, and comply with the NodeXL software [52], the following considerations were made when developing the lists: (1) removing all duplicate words from within each list, (2) ensuring that there were no words that overlapped between lists, and (3) removing words that appeared in the "stop" word list in NodeXL.

A sensitivity analysis was performed to account for words that related to popular diet keywords. For words considered in this analysis, refer to Multimedia Appendix 2. To account for words that overlapped between lists, the number of times that the word had appeared in the literature was considered (refer to Multimedia Appendix 3 for further details). If this did not promote a clear decision, a pilot test in NodeXL was used to identify the number of times that the word appeared in the corresponding network (Multimedia Appendix 3). In addition, as suggested by another study [68], superficially innocuous words such as "eating" and "exercise" were also excluded from the eating disorder list because of their potential ambiguity and inflation of results.

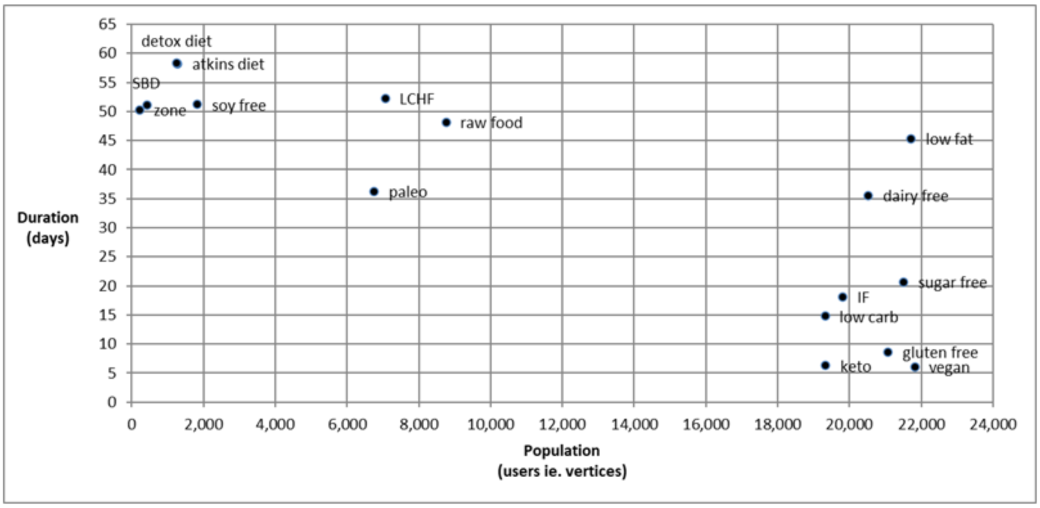
After both lists were created, a feasibility test was conducted to identify any additional words and symbols that needed to be added to the NodeXL "stop word" list. A number of words were identified and added if they had a word count of >100, under the premise that they would influence the salience (Multimedia Appendix 2). In addition, as "soy free" holds a different meaning in Spanish (soy free meaning "I am free"), Spanish was hidden from the "soy free" network analytics.

Results

Network Dissemination and How Messages Spread

The size and duration of each of the 16 networks are shown in Figure 2. Each network comprised all users and user interactions that occurred at the time of data extraction and has been labeled based on the popular diet being addressed. The vegan network had the largest network with data collected over the shortest duration (6 days 1 hour; n=21,819 users), followed by the ketogenic network (6 days 8 hours; n=19,336 users) and gluten-free network (8 days 14 hours; n=21,069 users).

Figure 2. Scatter plot of the total population of each network over the time period of data collection. IF: intermittent fasting; LCHF: low-carbohydrate, high-fat; SBD: South Beach Diet; zone: zone diet.



Paleo (duration=36 days 7 hours; n=6750 users); raw food (duration=48 days 5 hours; n=8760 users); and low-carb, high-fat (LCHF); duration=52 days 6 hours; n=7060 users) had medium-sized networks over longer durations, whereas, Atkins diet (duration=58 days 4 hours; n=1283 users) and detox diet (duration=58 days 6 hours; n=1259 users) had smaller networks over long durations, with zone diet (duration=50 days 6 hours; n=211 users) and south beach diet (duration=51 days 2 hours; n=417 users) being the 2 smallest networks.

Visualization was created for all networks (Multimedia Appendix 4 [56]), which is a visual representation of all user interactions including tweets, mentions, and retweets. Overall, there were 2 dominant structures in all networks: complete fragmentation and hub and spoke. In most networks, the complete fragmentation structure was dominant, followed by the hub and spoke clusters. However, there was evidence of community interaction in the selected networks (keto, dairy free, intermittent fasting, vegan, and LCHF). LCHF was the only network where a hub and spoke structure was prominent.

Influential Voices of Each Network

Of the 16 networks, 160 top 10 user accounts were identified through the betweenness centrality metric alone, with an additional 23 users after adjusting for the outdegree metric (N=183 users) and removing 10 duplicates (total n=173 users). User follower counts ranged from 4 to 72 million. A total of 8 users overlapped between the networks (Multimedia Appendix 5).

Using the information supplied in user bios, 11% (19/173) of the users were identified as claiming a health or science education, background, or profession, including medical doctors, MD (n=9; networks=paleo, low carb, keto, intermittent fasting, and LCHF), PhD, public health (n=1; network=paleo), dietitians (n=2; networks=Atkins diet and LCHF), nutritionists *(*n=3; networks=dairy free and zone diet), and other science education (n=4; networks=low carb, low fat, intermittent fasting, and LCHF). When considering the information collected from both user bios and tweets, 1 user appeared to actively oppose the diet (dietitian and Atkins diet), and 14 users actively endorsed the diet (paleo, vegan, low carb, zone diet, keto, intermittent fasting, and LCHF). LCHF (n=5 users) and intermittent fasting (n=3 users) contained the greatest number of health or science professional users endorsing the diet (Multimedia Appendix 5).

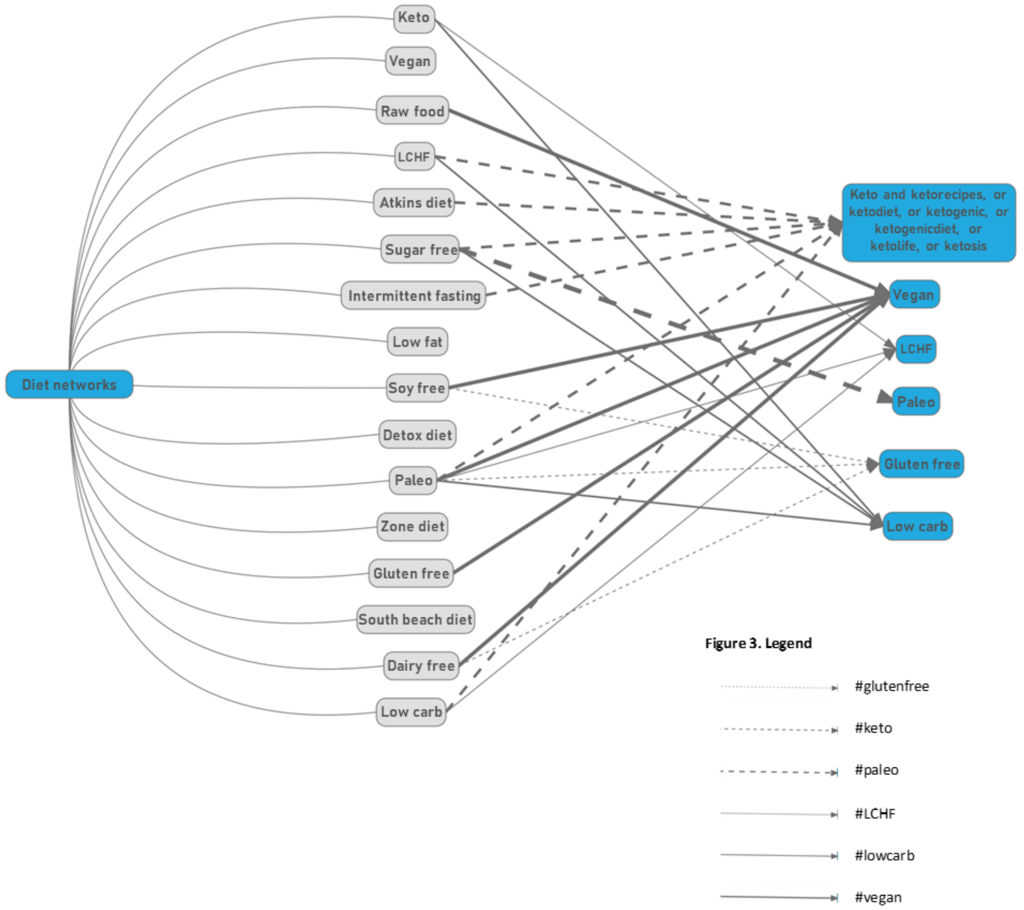
Of the top 10 user accounts, 31.2% (54/173) of the users across all networks openly supported the corresponding diet of the network they belonged to. Networks with the greatest number of active endorsers included LCHF (9/10, 90%), intermittent fasting (8/10, 80%), low carb (8/12, 67%), and keto (7/14, 50%). Networks with no endorsers were raw food, low fat, zone diet, and Atkins diet networks.

Interaction Between Popular Diet Networks

The top 10 hashtags collected from all tweets within each network are presented in Multimedia Appendix 6. In total, 69% (11/16) of the popular diet networks (paleo, raw food, sugar free, dairy free, gluten free, low carb, Atkins diet, keto, intermittent fasting, LCHF, and soy free) displayed some form of interaction demonstrated by overlapping hashtags (Figure 3), although only 6 of the popular diets were referred to. The paleo network referenced the greatest number of diets, with 6 of its top 10 hashtags (keto, ketodiet, lowcarb, gluten free, vegan, and LCHF) representing 5 other networks. Keto had the greatest number of mentions across all networks (n=6; paleo, sugar free, low carb, Atkins diet, intermittent fasting, and LCHF). Of all the networks, LCHF referred to keto the most, with 7 of its 10 top hashtags. The vegan network made reference to 0 other diets.

Across all networks, #weightloss was the most prevalent hashtag used with 9 mentions (paleo, low carb, low fat, zone diet, Atkins diet, keto, intermittent fasting, detox diet, and LCHF), followed by #keto with 6 mentions. All diet networks made reference to at least 1 of the top-ranking hashtags except for the South Beach Diet that referenced 0 hashtags.

Figure 3. Popular diet (colored gray) network interactions as represented by top 10 hashtags (colored blue). LCHF: low-carbohydrate, high-fat.



Interaction Between Networks and Mental Health Frequency

Assessed using the content of all tweets within each network, depression and anxiety (words=113/3929; frequency: 0.029, 2.9%) and eating disorder (words=133/3929; frequency: 0.034, 3.4%) word frequency was the greatest in the "zone diet" network. In the depression and anxiety analysis, "zone diet" was followed by "Atkins diet" (words=493/24,760; frequency: 0.02, 2%), and word frequency was lowest in "soy free" network (words=240/43,874; frequency: 0.005, 0.5%). In the eating disorder analysis, "zone diet" was followed by paleo (words=6574/212,492; frequency: 0.031, 3.1%), and word frequency was equally the lowest in vegan (words=1521/446,955; frequency: 0.003, 0.3%), "dairy free" (words=1541/440,804; frequency: 0.003, 0.3%), and "gluten-free" networks (words=1578/451,842; frequency: 0.003, 0.3%). For all text analyses of mental health word lists, refer to Multimedia Appendices 5 and 7.

Discussion

Principal Findings

To our knowledge, this exploratory study is the first online SNA to explore and characterize the web-based discourse of popular diet networks. Our data provide novel insights into the dissemination of popular diet nutrition information on the web, key influential voices, interactions between web-based diet networks, and associations with mental health themes. The key findings of this study demonstrate that (1) i. social media activity reflects popular diet trends, and ii. nutrition information related to these diets is primarily dispersed through resharing information; (2) i. users claiming a background in science and health are among the most influential voices sharing nutrition information related to the popular diets explored, and ii. follower count does not necessarily affect influence on Twitter; (3) popular diet networks interact and connect through common dietary themes; and (4) there is evidence to suggest that a relationship may exist between popular diets and mental health in the context of web-based social networks.

Network Dissemination and How Messages Spread

Popular diet trends change and evolve over time [21], and related dietary advice and information are regularly dispersed on the web [15]. Our results suggest that social media activity may reflect these trends, with the population and duration varying broadly for each of the diet networks. These findings indicate that each network is unique and imply that not every topic receives the same level of attention at the same point in time. In addition, it is plausible to assume that these dimensions will shift over time and that the size of the network may reflect diet popularity.

To our knowledge, there is limited longitudinal social media research exploring "popular" international diet trends over time; however, other research and Google worldwide trends demonstrate that the zone and Atkins diet were at their peak of popularity in the 1990s [21], the South Beach Diet in 2004, and the detox diet in 2007 [74], all of which have smaller networks.

Larger networks, including vegan and keto, hit their peak in 2019, and gluten-free network has remained popular since 2013 [74]. In addition, the International Food Informational Council Food and Health Survey, released June 2020, found that keto and intermittent fasting were among the top 3 most common diets followed at this time, with low-carbohydrate and gluten-free diets following closely behind [75]. Conversely, despite being our largest network, the vegan diet was reported as 1 of the least common [75], which may reflect ethical motivations for veganism as a "lifestyle" rather than a diet trend [76,77]. Conversely, larger diet networks that seemingly never had a "peak," such as sugar free [74], may predict the emergence of new popular diets. They may also reflect common keywords spanning a number of networks, or they may simply be representative of dietary descriptors and ethos (eg, sugar-free recipes). In addition, smaller and medium-sized networks may reflect trends that have passed, or keywords may not be representative of the way the diets are discussed on the web. The results of this study highlight the role of Twitter in dispersing nutrition information on the web, with larger networks allowing for greater amplification and dissemination of dietary messages. It also demonstrates that messages on social media can continue to spread even decades after they were most relevant.

The structures that emerged through visualization provided additional insights into the dissemination of information [51]. The dominant structure suggests that most users were disconnected and talking about the diet rather than talking *to* each other. Practically, this indicates that influencers were getting retweeted by users without connecting through mutual conversation. To our knowledge, there are currently no network analyses of popular diet networks on Twitter; however, a SNA of the social media platform—Reddit found that users formed close relationships with dense interactions in 3 weight loss–focused networks [78]. Comparisons should be made with caution; however, owing to the presumed connectedness of social media [79], a similar observation was expected from our results. With this in mind, LCHF supported this assumption, dominated by a hub and spoke structure, suggesting a more complex community interaction. Interestingly, LCHF was selected through feasibility testing rather than academic literature. This finding suggests that web-based communities develop their own language to connect on the web. This concept may be similar to that of eating disorder communities, where 1 study found that users connected through specific hashtags [68].

These findings provide valuable insights into the dispersal of "popular" diet-related information on the web and suggest that nutrition professionals could create an impact by resharing key information from within the health community. Further research on the application of resharing nutrition content could inform strategies as part of community and government initiatives in the future. In addition, to gain a deeper understanding, we recommend that future social network analyses and longitudinal research explore how web-based diet networks change over time.

Influential Voices of Each Network

Information dissemination was affected by the most influential voices of each popular diet. In accordance with the social network theory, they form the center of the network and act as a mediator of information between users [7,51]. Our findings identified regular Twitter users, businesses, brands, and users claiming a background in science and health among the most influential endorsers of popular diets, all having a considerably different follower count. This result was somewhat unexpected, with the follower count of other social network platforms a determination of influence [80]. However, this finding is similar to that of a study that explored user influence on Twitter, which found that retweets and mentions exhibited more influence than followers [81].

Furthermore, to the best of our knowledge, no research has identified the key influential voices promoting popular diets on the web. Studies that explore the web-based communication of health messages may provide some insights into their position within the network. With this in mind, studies have found that heroic language increases perceived authenticity [82], and "experts" are perceived as more trustworthy [83]. Although direct comparisons cannot be made, this may be because of the language used by users and perceived expertise as contributing factors for posts being reshared. Public trust in "experts" [84] may have contributed to our findings, where 11% (19/173) of the top influencers identified as claiming a background in science or health. Notably, most of these top "health and science" influencers endorsed the corresponding diet, this included only 2 dietitians, one of whom was actively opposing the popular diet (speaking out about misinformation) and another who was an academic researching the popular diet (Multimedia Appendix 5). Although we did not assess the validity of the diet claims made in this study, the endorsement of these diets and the contribution of nutrition-related information from unqualified users must be acknowledged. This may also reflect a study that conducted an analysis of health-related tweets and found that more than half were not evidence based [85]. It would be beneficial for future research to assess and validate the reliability of tweets by both health professionals and unqualified users.

Finally, it must also be acknowledged that the data collected from the popular diet communities may not reflect those of professional networks. However, our findings identify the need for more qualified nutrition professionals, such as dietitians, to gain a presence on the web. To encourage social media use and overcome potential barriers [30], educational institutions and professional organizations, such as Dietitians Australia, the Association of UK Dietitians, and the Academy of Nutrition and Dietetics, could expand their media training to include more comprehensive social media training. This may also be extended to other health professionals to encourage its use in a meaningful and safe way. In addition, monitoring how key influencers may change over time and a content analysis of tweets may provide insights into user motivation.

Interaction Between Popular Diet Networks

Associations between various popular diets may be observed through the exploration of top hashtags, which are commonly used to ascertain related themes [58]. These themes were evident in our study, with networks connected through shared dietary ethos and topics such as weight loss and health. To our knowledge, there is currently no literature that explores web-based interactions between various diet networks; however, popular diets, such as fad diets, are often categorized by their distribution of macronutrients [21]. Although direct comparisons cannot be made, this concept may provide insight into our results; interestingly, although 69% (11/16) of the networks interacted, only 6 were referred to, most of which were characterized by a high-fat, low-carbohydrate eating pattern [86,87]. Notably, however, with the exception of the vegan and gluten free networks, which may reflect veganism as both a fad and ethical approach to eating [76,88] or gluten free as another way to describe a low-carbohydrate eating pattern, or as an allergy and dietary descriptor (eg, gluten-free bread) [87].

In addition, and as explored earlier, dietary interaction may also reflect trending diets, with keto and vegan being the most referenced [74]. Our findings suggest that popular diets interact on the web and connect through similar themes and ethos. Identifying these connective themes may assist dietitians and nutrition professionals in directing their messaging and addressing potential misinformation being shared on the web.

Interaction Between Networks and Mental Health

As web-based diet content can influence food choice [27] and social media can provide an "echo chamber" of similar information, [24] exposure to web-based diet content may lead to unnecessary restriction by those consuming the information. Restrictive diets have also been linked to potentially negative physiological and psychological health outcomes including eating disorders and depression and anxiety [20,21]. Therefore, this study explored the themes of mental health and their interaction with web-based diet networks. Despite these known associations, to the best of our knowledge, there is currently no literature exploring web-based mental health word frequency within popular diet networks. Although word frequency must be interpreted with caution owing to the complexity of social media [89], related research has identified the potential for detecting mental health symptoms on the web [46,68]. With this in mind, the results of our study identified several networks where words suggestive of mental health concerns were present.

The novel nature of this exploration prevents a deeper interpretation of these findings; however, from a social media perspective, identifying communities where mental health narrative is strongest may allow social networks and organizations to target specific mental health–related support or relevant advertisements (such as mental health–specific helplines) to reach the people that need it most. This concept was explored in a mixed methods survey, where 60% (96/160) of participants endorsed the idea of using social media technology to improve targeting of mental health services [90]. Although it must be interpreted with caution, our results suggest targeting "web-based diet communities" that promote eating patterns that contain rules about carbohydrate intake, with support specific to depression, anxiety, and eating disorders. From a nutritional perspective, although there may be some association between mental health word frequency and specific diet patterns, direct comparisons cannot be made. However, our results reflect what is known in the current dietary literature, where mental health word frequency was found to be highest in a diet with rules about carbohydrate intake [20,21], whereas diets rich in carbohydrate-containing foods, such as whole grains and vegetables, are associated with beneficial effects on mental well-being [91]. Similarly, eating disorder word frequency was found to be the highest in 2 dietary patterns governed by a set of rules and restrictions, which may also correspond with eating disorder symptomology [92,93].

Although these results must be interpreted with caution, these novel insights are suggestive of an existing interaction between popular diet networks and mental health. More in-depth exploration is needed in future research to improve our understanding of this finding, including the directionality of the relationships and the importance of the word frequency. Content analysis of tweets and survey research should also explore consumer experiences and the perceived influence of social media on food choice. Network analysis research may also be used to identify topics that mental health consumers engage with on the web, and in the future, results may help to inform targeted web-based diet and mental health–related initiatives.

Limitations

Some additional limitations of this research must be considered, including (1) using Twitter posts only, (2) the removal of duplicate keywords from text word lists (Multimedia Appendices 2 and 3) that may have resulted in lower word frequency than presented, (3) errors of omission because of analyzing data from only 1 period, and (4) unknown gaps in data owing to access challenges such as privacy settings.

Conclusions

In conclusion, our study provides novel insights into the web-based discourse of popular diet networks and has paved a way for similar research in the future. Our findings are important for health professionals and related organizations to enhance their understanding of the web-based nutrition space and to help inform the effective dissemination of nutrition messages by qualified professionals on the web. To build on our findings, further network analysis and longitudinal and survey research are needed to explore popular diet trends on the web over time and to understand the impact that social media can have on dietary choices. Finally, to encourage qualified nutrition professionals, such as dietitians, to be the leading voices of nutrition information on the web, we recommend social media training for health professionals and that dietitians and nutrition professionals work together as a community by actively resharing posts.

Acknowledgments

The authors would like to thank Dr Alina Arseniev for kindly sharing her codebook of keywords (a previously defined list of words representative of the #proana community on Twitter), which we used in our text analysis of mental health word lists.

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. *JMIR Publications* provided APF support for the publication of this article.

Authors' Contributions

MCE and YCP conceived and designed the study. MCE drafted the original version with input from all authors for the final manuscript. MAS supported the implementation of the methodology and interpretation of the results. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Top 10 hashtags, network metrics, and figure visualizations.
[DOCX File , 2046 KB - infodemiology_v3i1e38245_app1.docx]

Multimedia Appendix 2
Text analysis of mental health words list.
[DOCX File , 24 KB - infodemiology_v3i1e38245_app2.docx]

Multimedia Appendix 3
Mental health word list frequency and overlapping words.
[DOCX File , 14 KB - infodemiology_v3i1e38245_app3.docx]

Multimedia Appendix 4
Graph visualization.
[DOCX File , 14643 KB - infodemiology_v3i1e38245_app4.docx]

Multimedia Appendix 5
Results summary table.
[DOCX File , 58 KB - infodemiology_v3i1e38245_app5.docx]

Multimedia Appendix 6
Top 10 hashtags.
[DOCX File , 15 KB - infodemiology_v3i1e38245_app6.docx]

Multimedia Appendix 7
Text analysis for mental health word lists.
[DOCX File , 20 KB - infodemiology_v3i1e38245_app7.docx]

References

1. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
2. Terrasse M, Gorin M, Sisti D. Social media, e-health, and medical ethics. Hastings Cent Rep 2019 Jan;49(1):24-33. [doi: 10.1002/hast.975] [Medline: 30790306]
3. Ortiz-Ospina E. The rise of social media. Our World in Data. 2019 Sep 18. URL: https://ourworldindata.org/rise-of-social-media [accessed 2022-01-20]
4. McCormick TH, Lee H, Cesare N, Shojaie A, Spiro ES. Using Twitter for demographic and social science research: tools for data collection and processing. Sociol Methods Res 2017 Aug;46(3):390-421 [FREE Full text] [doi: 10.1177/0049124115605339] [Medline: 29033471]

5. Rowe AD, Pitfield DE. The challenge facing existing airport campaign groups when incorporating social media into their campaign: a social network analysis of airport watch's social media utilisation. Geoforum 2018 Nov;96:236-247 [FREE Full text] [doi: 10.1016/j.geoforum.2018.08.022]
6. Cesare N, Grant C, Nsoesie EO. Detection of user demographics on social media: a review of methods and recommendations for best practices. arXiv 2017 Feb 6:1-25.
7. Quatman C, Chelladurai P. Social network theory and analysis: a complementary lens for inquiry. J Sport Manag 2008 May 01;22(3):338-360 [FREE Full text] [doi: 10.1123/jsm.22.3.338]
8. Valente TW, Pitts SR. An appraisal of social network theory and analysis as applied to public health: challenges and opportunities. Annu Rev Public Health 2017 Mar 20;38(1):103-118 [FREE Full text] [doi: 10.1146/annurev-publhealth-031816-044528] [Medline: 27992729]
9. Serrat O. Social network analysis. In: Serrat O, editor. Knowledge Solutions. Singapore, Singapore: Springer; 2017:39-43.
10. Can U, Alatas B. A new direction in social network analysis: online social network analysis problems and applications. Physica A 2019 Dec;535:122372 [FREE Full text] [doi: 10.1016/j.physa.2019.122372]
11. Kim J, Hastak M. Social network analysis: characteristics of online social networks after a disaster. Int J Inf Manage 2018 Feb;38(1):86-96 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2017.08.003]
12. Wood MJ. Propagating and debunking conspiracy theories on Twitter during the 2015-2016 Zika virus outbreak. Cyberpsychol Behav Soc Netw 2018 Aug;21(8):485-490 [FREE Full text] [doi: 10.1089/cyber.2017.0669] [Medline: 30020821]
13. Pascual-Ferrá P, Alperstein N, Barnett DJ. Social network analysis of COVID-19 public discourse on Twitter: implications for risk communication. Disaster Med Public Health Prep 2022 Apr;16(2):561-569 [FREE Full text] [doi: 10.1017/dmp.2020.347] [Medline: 32907685]
14. Dumas AA, Lapointe A, Desroches S. Users, uses, and effects of social media in dietetic practice: scoping review of the quantitative and qualitative evidence. J Med Internet Res 2018 Feb 20;20(2):e55 [FREE Full text] [doi: 10.2196/jmir.9230] [Medline: 29463487]
15. Ramachandran D, Kite J, Vassallo AJ, Chau JY, Partridge S, Freeman B, et al. Food trends and popular nutrition advice online - implications for public health. Online J Public Health Inform 2018 Sep 21;10(2):e213 [FREE Full text] [doi: 10.5210/ojphi.v10i2.9306] [Medline: 30349631]
16. Australian dietary guidelines 1 - 5. National Health and Medical Research Council. 2020. URL: https://www.eatforhealth.gov.au/guidelines/australian-dietary-guidelines-1-5 [accessed 2023-02-18]
17. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
18. Bessi A, Zollo F, Del Vicario M, Scala A, Caldarelli G, Quattrociocchi W. Trend of narratives in the age of misinformation. PLoS One 2015 Aug 14;10(8):e0134641 [FREE Full text] [doi: 10.1371/journal.pone.0134641] [Medline: 26275043]
19. Adamski M, Truby H, M Klassen K, Cowan S, Gibson S. Using the internet: nutrition information-seeking behaviours of lay people enrolled in a massive online nutrition course. Nutrients 2020 Mar 12;12(3):750 [FREE Full text] [doi: 10.3390/nu12030750] [Medline: 32178291]
20. Saltzman E, Thomason P, Roberts SB. Fad diets: a review for the primary care provider. Nutr Clinical Care 2001 Sep;4(5):235-242 [FREE Full text] [doi: 10.1046/j.1523-5408.2001.00003.x]
21. Khawandanah J, Tewfik I. Fad diets: lifestyle promises and health challenges. J Food Res 2016 Nov 15;5(6):80 [FREE Full text] [doi: 10.5539/jfr.v5n6p80]
22. Falsaperla R, D'Angelo G, Praticò AD, Mauceri L, Barbagallo M, Pavone P, et al. Ketogenic diet for infants with epilepsy: a literature review. Epilepsy Behav 2020 Nov;112:107361. [doi: 10.1016/j.yebeh.2020.107361] [Medline: 33181904]
23. Muir JG, Varney JE, Ajamian M, Gibson PR. Gluten-free and low-FODMAP sourdoughs for patients with coeliac disease and irritable bowel syndrome: a clinical perspective. Int J Food Microbiol 2019 Feb 02;290:237-246. [doi: 10.1016/j.ijfoodmicro.2018.10.016] [Medline: 30388591]
24. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M. The echo chamber effect on social media. Proc Natl Acad Sci U S A 2021 Mar 02;118(9):e2023301118 [FREE Full text] [doi: 10.1073/pnas.2023301118] [Medline: 33622786]
25. Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, et al. The spreading of misinformation online. Proc Natl Acad Sci U S A 2016 Jan 19;113(3):554-559 [FREE Full text] [doi: 10.1073/pnas.1517441113] [Medline: 26729863]
26. Kulshrestha J, Zafar M, Noboa L, Gummadi K, Ghosh S. Characterizing information diets of social media users. In: Proceedings of the 9th International AAAI Conference on Web and Social Media. 2015 Presented at: ICWSM '15; May 26-29, 2015; Oxford, UK p. 218-227 URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14634 [doi: 10.1609/icwsm.v9i1.14634]
27. Vaterlaus JM, Patten EV, Roche C, Young JA. #Gettinghealthy: the perceived influence of social media on young adult health behaviors. Comput Human Behav 2015 Apr;45:151-157 [FREE Full text] [doi: 10.1016/j.chb.2014.12.013]
28. Baldwin HJ, Freeman B, Kelly B. Like and share: associations between social media engagement and dietary choices in children. Public Health Nutr 2018 Dec;21(17):3210-3215. [doi: 10.1017/S1368980018001866] [Medline: 30086811]

29. Nguyen CT. Echo chambers and epistemic bubbles. Episteme 2020 Jun;17(2):141-161 [FREE Full text] [doi: 10.1017/epi.2018.32]
30. Probst YC, Peng Q. Social media in dietetics: insights into use and user networks. Nutr Diet 2019 Sep;76(4):414-420. [doi: 10.1111/1747-0080.12488] [Medline: 30370651]
31. Shine D, Minehan M, Knight-Agarwal C. #Healthpromotion: a qualitative exploration of how dietitians can use social media to positively influence women aged 18-35 years. Nutr Diet 2022 Sep;79(4):489-496 [FREE Full text] [doi: 10.1111/1747-0080.12765] [Medline: 36082525]
32. Wu JT, McCormick JB. Why health professionals should speak out against false beliefs on the internet. AMA J Ethics 2018 Nov 01;20(11):E1052-E1058 [FREE Full text] [doi: 10.1001/amajethics.2018.1052] [Medline: 30499434]
33. Goldschmidt AB, Wall M, Choo TH, Becker C, Neumark-Sztainer D. Shared risk factors for mood-, eating-, and weight-related health outcomes. Health Psychol 2016 Mar;35(3):245-252 [FREE Full text] [doi: 10.1037/hea0000283] [Medline: 26690639]
34. Liechty JM, Lee MJ. Longitudinal predictors of dieting and disordered eating among young adults in the U.S. Int J Eat Disord 2013 Dec;46(8):790-800. [doi: 10.1002/eat.22174] [Medline: 23983018]
35. Johnson JG, Cohen P, Kotler L, Kasen S, Brook JS. Psychiatric disorders associated with risk for the development of eating disorders during adolescence and early adulthood. J Consult Clin Psychol 2002 Oct;70(5):1119-1128. [doi: 10.1037//0022-006x.70.5.1119] [Medline: 12362962]
36. Johnson F, Wardle J. Dietary restraint, body dissatisfaction, and psychological distress: a prospective analysis. J Abnorm Psychol 2005 Feb;114(1):119-125. [doi: 10.1037/0021-843X.114.1.119] [Medline: 15709818]
37. Stice E, Hayward C, Cameron RP, Killen JD, Taylor CB. Body-image and eating disturbances predict onset of depression among female adolescents: a longitudinal study. J Abnorm Psychol 2000 Aug;109(3):438-444. [Medline: 11016113]
38. Schaefer JT, Magnuson AB. A review of interventions that promote eating by internal cues. J Acad Nutr Diet 2014 May;114(5):734-760. [doi: 10.1016/j.jand.2013.12.024] [Medline: 24631111]
39. Stewart TM, Williamson DA, White MA. Rigid vs. flexible dieting: association with eating disorder symptoms in nonobese women. Appetite 2002 Feb;38(1):39-44. [doi: 10.1006/appe.2001.0445] [Medline: 11883916]
40. Smith T, Hawks SR. Intuitive eating, diet composition, and the meaning of food in healthy weight promotion. Am J Health Educ 2006;37(3):130-136 [FREE Full text] [doi: 10.1080/19325037.2006.10598892]
41. Patton GC, Selzer R, Coffey C, Carlin JB, Wolfe R. Onset of adolescent eating disorders: population based cohort study over 3 years. BMJ 1999 Mar 20;318(7186):765-768 [FREE Full text] [doi: 10.1136/bmj.318.7186.765] [Medline: 10082698]
42. Hilbert A, Pike KM, Goldschmidt AB, Wilfley DE, Fairburn CG, Dohm FA, et al. Risk factors across the eating disorders. Psychiatry Res 2014 Dec 15;220(1-2):500-506 [FREE Full text] [doi: 10.1016/j.psychres.2014.05.054] [Medline: 25103674]
43. Stice E, Gau JM, Rohde P, Shaw H. Risk factors that predict future onset of each DSM-5 eating disorder: predictive specificity in high-risk adolescent females. J Abnorm Psychol 2017 Jan;126(1):38-51 [FREE Full text] [doi: 10.1037/abn0000219] [Medline: 27709979]
44. Zhou S, Bian J, Zhao Y, Haynos AF, Rizvi R, Zhang R. Analysis of Twitter to identify topics related to eating disorder symptoms. IEEE Int Conf Healthc Inform 2019 Jun;2019:1-4 [FREE Full text] [doi: 10.1109/ichi.2019.8904863] [Medline: 32030368]
45. Wang T, Brede M, Ianni A, Mentzakis E. Detecting and characterizing eating-disorder communities on social media. In: Proceedings of the 10th ACM International Conference on Web Search and Data Mining. 2017 Presented at: WSDM '17; February 6-10, 2017; Cambridge, UK p. 91-100 URL: https://dl.acm.org/doi/10.1145/3018661.3018706 [doi: 10.1145/3018661.3018706]
46. Mowery DL, Park A, Bryan C, Conway M. Towards automatically classifying depressive symptoms from Twitter data for population health. In: Proceedings of the Workshop on Computational Modeling of People’s Opinions, Personality, and Emotions in Social Media. 2016 Presented at: PEOPLES '16; December 12, 2016; Osaka, Japan p. 182-191 URL: https://aclanthology.org/W16-4320.pdf
47. Kumar A, Sharma A, Arora A. Anxious depression prediction in real-time social data. In: Proceedings of the 2019 International Conference on Advances in Engineering Science Management & Technology. 2019 Presented at: ICAESMT '19; January 27-29, 2019; Gwalior, India p. 1-7 URL: https://tinyurl.com/mr2uejw5 [doi: 10.2139/ssrn.3383359]
48. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. In: Proceedings of the 7th International AAAI Conference on Weblogs and Social Media. 2013 Presented at: ICWSM '13; July 8-11, 2013; Cambridge, MA, USA p. 128-137 URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14432 [doi: 10.1609/icwsm.v7i1.14432]
49. Karami A, Dahl AA, Turner-McGrievy G, Kharrazi H, Shaw Jr G. Characterizing diabetes, diet, exercise, and obesity comments on Twitter. Int J Inf Manage 2018 Feb;38(1):1-6 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2017.08.002]
50. Wilksch SM, O'Shea A, Ho P, Byrne S, Wade TD. The relationship between social media use and disordered eating in young adolescents. Int J Eat Disord 2020 Jan;53(1):96-106. [doi: 10.1002/eat.23198] [Medline: 31797420]
51. Smith MA, Shneiderman B, Milić-Frayling N, Rodrigues EM, Barash V, Dunne C, et al. Analyzing (social media) networks with NodeXL. In: Proceedings of the 4th International Conference on Communities and Technologies. 2009 Presented at: C&T '09; June 25-27, 2009; Pennsylvania, PA, USA p. 255-264 URL: https://dl.acm.org/doi/10.1145/1556460.1556497 [doi: 10.1145/1556460.1556497]

52. Hansen DL, Shneiderman B, Smith MA. Analyzing Social Media Networks with NodeXL: Insights from a Connected World. Burlington, MA, USA: Morgan Kaufmann; 2010.
53. Smith MA. NodeXL: simple network analysis for social media. In: Proceedings of the 2013 International Conference on Collaboration Technologies and Systems. 2013 Presented at: CTS '13; May 20-24, 2013; San Diego, CA, USA p. 89-93 URL: https://ieeexplore.ieee.org/document/6567211 [doi: 10.1109/cts.2013.6567211]
54. Ge L, Sadeghirad B, Ball GD, da Costa BR, Hitchcock CL, Svendrovski A, et al. Comparison of dietary macronutrient patterns of 14 popular named dietary programmes for weight and cardiovascular risk factor reduction in adults: systematic review and network meta-analysis of randomised trials. BMJ 2020 Apr 01;369:m696 [FREE Full text] [doi: 10.1136/bmj.m696] [Medline: 32238384]
55. Obert J, Pearlman M, Obert L, Chapin S. Popular weight loss strategies: a review of four weight loss techniques. Curr Gastroenterol Rep 2017 Nov 09;19(12):61. [doi: 10.1007/s11894-017-0603-8] [Medline: 29124370]
56. Smith MA. Identifying and shifting social media network patterns with NodeXL. In: Proceedings of the 2014 International Conference on Collaboration Technologies and Systems. 2014 Presented at: CTS '14; May 19-23, 2014; Minneapolis, MN, USA p. 3-8. [doi: 10.1109/cts.2014.6867534]
57. Skaza J, Blais B. Modeling the infectiousness of Twitter hashtags. Physica A 2017 Jan;465:289-296 [FREE Full text] [doi: 10.1016/j.physa.2016.08.038]
58. Reuter J, Pereira-Martins J, Kalita J. Segmenting Twitter hashtags. Int J Nat Lang Comput 2016 Aug 30;5(4):23-36 [FREE Full text] [doi: 10.5121/ijnlc.2016.5402]
59. Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S. #WhyWeTweetMH: understanding why people use Twitter to discuss mental health problems. J Med Internet Res 2017 Apr 05;19(4):e107 [FREE Full text] [doi: 10.2196/jmir.6173] [Medline: 28381392]
60. Data handling: how large a network can NodeXL handle? Social Media Research Foundation. 2016. URL: https://www.smrfoundation.org/nodexl/faq/how-large-a-network-can-nodexl-handle/ [accessed 2022-01-18]
61. Twitter API. Twitter. 2022. URL: https://developer.twitter.com/en/docs/twitter-api [accessed 2022-01-18]
62. Search API: enterprise. Twitter Developer Platform. 2022. URL: https://developer.twitter.com/en/docs/twitter-api/enterprise/search-api/guides/changelog [accessed 2022-01-18]
63. Smith MA, Rainie L, Shneiderman B, Himelboim I. Mapping Twitter topic networks: from polarized crowds to community clusters. Pew Research Center. 2014 Feb 20. URL: https://www.pewresearch.org/internet/2014/02/20/mapping-twitter-topic-networks-from-polarized-crowds-to-community-clusters/ [accessed 2022-12-12]
64. Newman ME. Fast algorithm for detecting community structure in networks. Phys Rev E Stat Nonlin Soft Matter Phys 2004 Jun;69(6 Pt 2):066133. [doi: 10.1103/PhysRevE.69.066133] [Medline: 15244693]
65. Harel D, Koren Y. A fast multi-scale method for drawing large graphs. In: Proceedings of the 8th International Symposium on Graph Drawing. 2000 Presented at: ISGD' 00; September 20-23, 2000; Colonial Williamsburg, VA, USA p. 183-196. [doi: 10.1142/9789812796608_0010]
66. Golbeck J. Introduction to Social Media Investigation: A Hands-on Approach. Oxford, UK: Syngress; 2015.
67. Sigala M, Gretzel U. Advances in Social Media for Travel, Tourism and Hospitality: New Perspectives, Practice and Cases. Oxfordshire, UK: Routledge; 2017.
68. Arseniev-Koehler A, Lee H, McCormick T, Moreno MA. #Proana: pro-eating disorder socialization on Twitter. J Adolesc Health 2016 Jun;58(6):659-664. [doi: 10.1016/j.jadohealth.2016.02.012] [Medline: 27080731]
69. Cavazos-Rehg PA, Krauss MJ, Costello SJ, Kaiser N, Cahn ES, Fitzsimmons-Craft EE, et al. "I just want to be skinny.": a content analysis of tweets expressing eating disorder symptoms. PLoS One 2019 Jan 16;14(1):e0207506 [FREE Full text] [doi: 10.1371/journal.pone.0207506] [Medline: 30650072]
70. Shen JH, Rudzicz F. Detecting anxiety through reddit. In: Proceedings of the 4th Workshop on Computational Linguistics and Clinical Psychology. 2017 Presented at: CLPsych '17; August 3, 2017; Vancouver, Canada p. 58-65 URL: https://aclanthology.org/W17-3107.pdf [doi: 10.18653/v1/w17-3107]
71. Mowery D, Smith H, Cheney T, Stoddard G, Coppersmith G, Bryan C, et al. Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. J Med Internet Res 2017 Feb 28;19(2):e48 [FREE Full text] [doi: 10.2196/jmir.6895] [Medline: 28246066]
72. de Choudhury M, Sharma SS, Kıcıman E. Characterizing dietary choices, nutrition, and language in food deserts via social media. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing. 2016 Presented at: CSCW '16; February 27- March 2,2016; San Francisco, CA, USA p. 1157-1170 URL: https://dl.acm.org/doi/10.1145/2818048.2819956 [doi: 10.1145/2818048.2819956]
73. de Choudhury M, Counts S, Horvitz EJ. Social media as a measurement tool of depression in populations. In: Proceedings of the 5th Annual ACM Web Science Conference. 2013 Presented at: WebSci '13; May 2-4, 2013; Paris, France p. 47-56 URL: https://dl.acm.org/doi/10.1145/2464464.2464480 [doi: 10.1145/2464464.2464480]
74. Google Trends. Google. 2020. URL: https://trends.google.com/trends [accessed 2020-10-05]
75. Food Insight. 2020 food and health survey. International Food Information Council. 2020 Jun 9. URL: https://foodinsight.org/2020-food-and-health-survey/ [accessed 2022-12-12]

76. Janssen M, Busch C, Rödiger M, Hamm U. Motives of consumers following a vegan diet and their attitudes towards animal agriculture. Appetite 2016 Oct 01;105:643-651. [doi: 10.1016/j.appet.2016.06.039] [Medline: 27378750]
77. Kerschke-Risch P. Vegan diet: motives, approach and duration. Ernahr Umsch 2015;62(6):98-103 [FREE Full text] [doi: 10.4455/eu.2015.016]
78. Enes KB, Brum PP, Cunha TO, Murai F, da Silva AP, Pappa GL. Reddit weight loss communities: do they have what it takes for effective health interventions? In: Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence. 2018 Presented at: WI '18; December 3-6, 2018; Santiago, Chile p. 508-513. [doi: 10.1109/wi.2018.00-45]
79. Kietzmann JH, Hermkens K, McCarthy IP, Silvestre BS. Social media? Get serious! Understanding the functional building blocks of social media. Bus Horiz 2011 May;54(3):241-251 [FREE Full text] [doi: 10.1016/j.bushor.2011.01.005]
80. Segev N, Avigdor N, Avigdor E. Measuring influence on Instagram: a network-oblivious approach. In: Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 2018 Presented at: SIGIR '18; July 8-12, 2018; Ann Arbor, MI, USA p. 1009-1012. [doi: 10.1145/3209978.3210134]
81. Cha M, Haddadi H, Benevenuto F, Gummadi K. Measuring user influence in Twitter: the million follower fallacy. In: Proceedings of the 4th International AAAI Conference on Weblogs and Social Media. 2010 May 16 Presented at: ICWSM '10; May 23-26, 2010; Washington, DC, USA p. 10-17 URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14033 [doi: 10.1609/icwsm.v4i1.14033]
82. Jenkins EL, Ilicic J, Molenaar A, Chin S, McCaffrey TA. Strategies to improve health communication: can health professionals be heroes? Nutrients 2020 Jun 22;12(6):1861 [FREE Full text] [doi: 10.3390/nu12061861] [Medline: 32580495]
83. Jenkins EL, Ilicic J, Barklamb AM, McCaffrey TA. Assessing the credibility and authenticity of social media content for applications in health communication: scoping review. J Med Internet Res 2020 Jul 23;22(7):e17296 [FREE Full text] [doi: 10.2196/17296] [Medline: 32706675]
84. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 02;41:433-451 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
85. Alnemer KA, Alhuzaim WM, Alnemer AA, Alharbi BB, Bawazir AS, Barayyan OR, et al. Are health-related tweets evidence based? Review and analysis of health-related tweets on Twitter. J Med Internet Res 2015 Oct 29;17(10):e246 [FREE Full text] [doi: 10.2196/jmir.4898] [Medline: 26515535]
86. Modi N, Priefer R. Effectiveness of mainstream diets. Obes Med 2020 Jun;18:100239 [FREE Full text] [doi: 10.1016/j.obmed.2020.100239]
87. Norwood R, Cruwys T, Chachay VS, Sheffield J. The psychological characteristics of people consuming vegetarian, vegan, paleo, gluten free and weight loss dietary patterns. Obes Sci Pract 2019 Apr;5(2):148-158 [FREE Full text] [doi: 10.1002/osp4.325] [Medline: 31019732]
88. Lundahl O. Dynamics of positive deviance in destigmatisation: celebrities and the media in the rise of veganism. Consum Mark Cult 2018 Aug 29;23(3):241-271 [FREE Full text] [doi: 10.1080/10253866.2018.1512492]
89. Sánchez-Rada JF, Iglesias CA. Social context in sentiment analysis: formal definition, overview of current trends and framework for comparison. Inf Fusion 2019 Dec;52:344-356 [FREE Full text] [doi: 10.1016/j.inffus.2019.05.003]
90. Ford E, Curlewis K, Wongkoblap A, Curcin V. Public opinions on using social media content to identify users with depression and target mental health care advertising: mixed methods survey. JMIR Ment Health 2019 Nov 13;6(11):e12942 [FREE Full text] [doi: 10.2196/12942] [Medline: 31719022]
91. Gopinath B, Flood VM, Burlutksy G, Louie JC, Mitchell P. Association between carbohydrate nutrition and prevalence of depressive symptoms in older adults. Br J Nutr 2016 Dec;116(12):2109-2114. [doi: 10.1017/S0007114516004311] [Medline: 28065177]
92. Shahan M, Jordan K, Siegfried N, Douglas C, Lawrence J. Eating disorder risk and severity as a function of following self-administered popular diets. Curr Dev Nutr 2020 Jun;4(Suppl 2):565 [FREE Full text] [doi: 10.1093/cdn/nzaa046_065]
93. Nicholas JK, Cusack CE, Levinson CA. Eating disorder symptom and fear change trajectories during imaginal exposure therapy: a slope network analysis. Behav Ther 2023 Mar;54(2):346-360. [doi: 10.1016/j.beth.2022.09.006] [Medline: 36858764]

Abbreviations

API: application programming interface
LCHF: low-carb, high-fat
SNA: social network analysis

Edited by T Mackey; submitted 25.03.22; peer-reviewed by K Reuter, Y Tao, M Abrahams; comments to author 19.07.22; revised version received 13.12.22; accepted 10.01.23; published 05.05.23.

Please cite as:
Eaton MC, Probst YC, Smith MA
Characterizing the Discourse of Popular Diets to Describe Information Dispersal and Identify Leading Voices, Interaction, and Themes of Mental Health: Social Network Analysis
JMIR Infodemiology 2023;3:e38245
URL: https://infodemiology.jmir.org/2023/1/e38245
doi:10.2196/38245
PMID:37159259

Original Paper

News Coverage of Face Masks in Australia During the Early COVID-19 Pandemic: Topic Modeling Study

Pritam Dasgupta[1*], MBBS, MPH; Janaki Amin[1*], BSc, MPH, PhD; Cecile Paris[2*], BA, PhD; C Raina MacIntyre[3*], MBBS, MAppEpid, PhD

[1]Department of Health Sciences, Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, New South Wales, Australia
[2]Commonwealth Scientific and Industrial Research Organisation Data61, Sydney, New South Wales, Australia
[3]Biosecurity Program, Kirby Institute, University of New South Wales, Sydney, New South Wales, Australia
[*]all authors contributed equally

Corresponding Author:
Pritam Dasgupta, MBBS, MPH
Department of Health Sciences
Faculty of Medicine, Health and Human Sciences
Macquarie University
Level 3,
75 Talavera Rd
Sydney, New South Wales, 2109
Australia
Phone: 61 480148227
Email: pritam.dasgupta@students.mq.edu.au

Abstract

Background: During the COVID-19 pandemic, web-based media coverage of preventative strategies proliferated substantially. News media was constantly informing people about changes in public health policy and practices such as mask-wearing. Hence, exploring news media content on face mask use is useful to analyze dominant topics and their trends.

Objective: The aim of the study was to examine news related to face masks as well as to identify related topics and temporal trends in Australian web-based news media during the early COVID-19 pandemic period.

Methods: Following data collection from the Google News platform, a trend analysis on the mask-related news titles from Australian news publishers was conducted. Then, a latent Dirichlet allocation topic modeling algorithm was applied along with evaluation matrices (quantitative and qualitative measures). Afterward, topic trends were developed and analyzed in the context of mask use during the pandemic.

Results: A total of 2345 face mask–related eligible news titles were collected from January 25, 2020, to January 25, 2021. Mask-related news showed an increasing trend corresponding to increasing COVID-19 cases in Australia. The best-fitted latent Dirichlet allocation model discovered 8 different topics with a coherence score of 0.66 and a perplexity measure of –11.29. The major topics were T1 (mask-related international affairs), T2 (introducing mask mandate in places such as Melbourne and Sydney), and T4 (antimask sentiment). Topic trends revealed that T2 was the most frequent topic in January 2021 (77 news titles), corresponding to the mandatory mask-wearing policy in Sydney.

Conclusions: This study demonstrated that Australian news media reflected a wide range of community concerns about face masks, peaking as COVID-19 incidence increased. Harnessing the news media platforms for understanding the media agenda and community concerns may assist in effective health communication during a pandemic response.

(JMIR Infodemiology 2023;3:e43011) doi:10.2196/43011

KEYWORDS

face masks; mask; COVID-19; web-based news; community sentiment; topic modeling; latent Dirichlet allocation

Introduction

In response to the initial phase of the COVID-19 pandemic, Australia implemented widespread testing, strict lockdown, and hotel quarantine measures [1,2]. In June 2020, most of the measures were relaxed and accompanied by less disruptive measures, such as increased testing capacity and contact tracing [2]. Along with the timely policy recommendations, the Australian community participated in health-protective behaviors, at least in the earlier stage of the pandemic [3]. A survey over the internet among Australian residents identified the determinants of their level of engagement with health-protective behaviors. Higher risk perception and following media coverage about COVID-19 were the predictors of greater engagement with protective behaviors [3]. Despite changing mask policy of international organizations, most of the respondents (79.7%) identified the policy recommendation at that time (face masks were only for sick people) [3]. At the same time, the respondents also reported that the mainstream news media was the most popular source of information in Australia [3,4] and they consumed more of the news on the internet than the offline version [4]. Overall, timely policy adoption, health communication, and widespread community adherence contributed to the early success of controlling the pandemic in Australia.

The SARS-CoV-2 is the viral pathogen that caused the COVID-19 pandemic in 2020 [5]. Face masks appear to be an effective tool in mitigating community transmission of SARS-CoV-2 [6]. While providing source control by trapping the infectious droplets and aerosols from an infected wearer, face masks also provide outward protection that reduces the viral inoculum from the environment to well-wearers [5]. This is particularly important for SARS-CoV-2 transmission, which may occur from symptomatic, minimally symptomatic, presymptomatic, and even asymptomatic people [5]. Evidence suggests that approximately 40%-45% of cases of COVID-19 are asymptomatic, which can silently spread the virus throughout the community [7]. In a systematic review and meta-analysis, funded by World Health Organization (WHO), researchers synthesized 172 observational studies related to physical distancing, face masks, and eye protection to prevent the transmission of COVID-19 in both health care and community settings [8]. The findings suggested that mask-wearing can reduce COVID-19 transmission by up to 85% and is often better with higher-end masks such as N95 and similar respirators [8].

In line with existing supportive evidence on community use of face masks, many countries that historically do not have a mask-wearing culture have initially adopted a mask-wearing policy, while other countries with an established mask-wearing culture have continued to use masks [9]. Australia did not mandate mask-wearing until community outbreaks of COVID-19 [10]. In response to a local outbreak of COVID-19, the State Government of Victoria introduced social distancing, a stay-at-home order, and a mask mandate in metropolitan Melbourne and the Mitchell Shire from July 22, 2020 [11]. A recent study showed that the mandatory mask-wearing policy increased community adherence to face masks and significantly reduced the number of COVID-19 cases during that outbreak [12].

During national crises and health emergencies, news media plays a central role in a contested landscape of uncertainty [13-15]. News media can communicate risks and preventive measures, and can shape public perceptions but may also contribute to disinformation. News stories during a health crisis often include public debate and policy responses to conflicting priorities [13,15]. In a complex situation such as a pandemic, news media allows health professionals, policymakers, and the general public to interact and exchange information [13,16]. During the early COVID-19 pandemic, Australian media covered a wide range of topics related to COVID-19 and consistently prioritized "mask-wearing" and "mental health" topics in their news stories. Health professionals and academic experts, along with political leaders, received coverage and therefore influenced the media agenda and were able to use media for advocating pandemic prevention strategies [17]. However, previous pandemics have shown that inadequate quantity and scientific value of the news content can also limit the effectiveness of public health policy responses [16]. In recent research, more than two-thirds of the respondents expressed concern about polarized media agenda, based on the sponsors and political ideology [18]. A higher media literacy can aid to distinguish factual news from fake news, advertisement, and poor journalism [18]. So, it is imperative to constantly scrutinize the news contents, particularly during a health crisis.

Agenda-setting theory has been considered one of the most popular conceptual frameworks in communication research [19], referring to the strong correlation between media coverage on certain issues and people's perception of the importance of these issues [20]. Although the application of agenda-setting theory has been apparent in political contexts [21], the theory has long been applied in health issues [22] and crisis communication research [23]. In the context of COVID-19, agenda-setting became more diverse and interrelated where the media agenda was mostly shaped by the government and the public [24]. During the early stage of the pandemic in Australia, the national media coverage synchronized with the government-led agenda, including daily press conferences from State and Federal health leaders, and continuously relayed information from the health officials to the grassroots level [17].

While news media exercises the power of considerable discretion in choosing content for its storylines, media reporting on face masks can be designed within the skeleton of agenda-setting theory [20]. An ideal environment to shape public opinion and guide them toward informed decision-making mostly depends on the news reporting of mask-related agendas and ensuring their availability and repeated appearance on web-based news platforms. Based on the theoretical framework of agenda setting, the media is and has always been, a powerful tool of persuasion to facilitate desired health-protective behavior in the community [25]. Considering the importance of mask-wearing during the COVID-19 pandemic, it is useful to identify the media agenda and examine the news media contents, relevant topics, and their trends on face masks. It can provide useful insights into media narratives related to mask messaging.

At the beginning of the COVID-19 pandemic, Thomas et al [26] identified that a subtopic in Australian print media platforms was mask related. Basch et al [27] also reported similar findings in international news videos related to COVID-19 and found that only 6.2% of the video clips highlighted mask-wearing while caring for ill persons. Following the application and evaluation of a topic modeling algorithm in Chinese news data from the early COVID-19 pandemic, Liu et al [28] identified topics that emphasized mask use for both medical professionals and the general population. Yang et al [29] comparatively analyzed textual news media data and found that the China Daily newspaper paid more attention to mask-wearing throughout the study period than its counterparts from the United Kingdom and the United States. Moreover, multiple studies in this category analyzed news media contents from Iran [30], Brazil [31], and Italy [32]. Face masks appeared in these studies only as a subtopic under the broad topic of prevention and control measures, indicating a lack of media focus.

Lee et al [33] and Suh et al [34] attempted to investigate topics in mask-related news reports. In an earlier study, Suh et al [34] examined Korean news articles related to mask-wearing during 3 waves of COVID-19 in Korea from January 2020 to November 2020. By using a latent Dirichlet allocation (LDA)–based topic model, Suh et al [34] identified the major topics during the first and second waves of the COVID-19 pandemic in Korea. Additionally, Lee et al [33] applied a structural topic model algorithm on mask-related international news media data and identified underlying topics, and intertopic correlations.

Manual thematic analysis has limitations, especially in terms of scale. Natural language processing–based computational topic modeling techniques, in contrast, can handle complex and large amounts of data [35]. As part of the mixed method approach, topic modeling advances further qualitative interpretations and a deeper understanding of the topics [36].

The aim of the study was to examine news related to face masks as well as to identify related topics and temporal trends in Australian web-based news media during the early COVID-19 pandemic period.

Methods

Data Sources and Data Collection

Relevant data for the analysis were collected, both retrospectively and prospectively, from keyword-specific Google News search and Google News alert. Retrospective data were collected from January 25, 2020, to September 11, 2020, via Google News search. Subsequently, a customized Google News alert was set to collect prospective data from September 12, 2020, to January 25, 2021. For the retrospective and prospective data extraction, the keyword "mask" was used in conjunction with the ~ (tilde) symbol. That means the generated results, for both the Google News search and Google News alerts, contained the keyword or the synonyms (or both).

The results were then joined to form a combined data repository and named as "News topic data set." In that repository, news titles containing the search keyword and its synonyms were regarded as most relevant, and therefore only those news titles were selected for the analysis. Additional eligibility criteria were established to include consistent, reliable, and uniform data.

Selection Criteria

Inclusion Criteria

The inclusion criteria were (1) news titles in the English language and (2) news published by Australian news publishers only.

Exclusion Criteria

The study excluded irrelevant "mask"-related news articles such as beauty mask and sheet mask.

Face Mask–Related News Trend

As part of the descriptive data analysis, the trend in the number of mask-related items published in news on the internet was explored by plotting the frequency of occurrence in relation to key events during the COVID-19 pandemic, face mask–related incidents, and policy announcements in Australia (as described in the Result section). The key events were obtained from the Australian Government's website and national news reports.

Exploratory Data Analysis by Topic Modeling

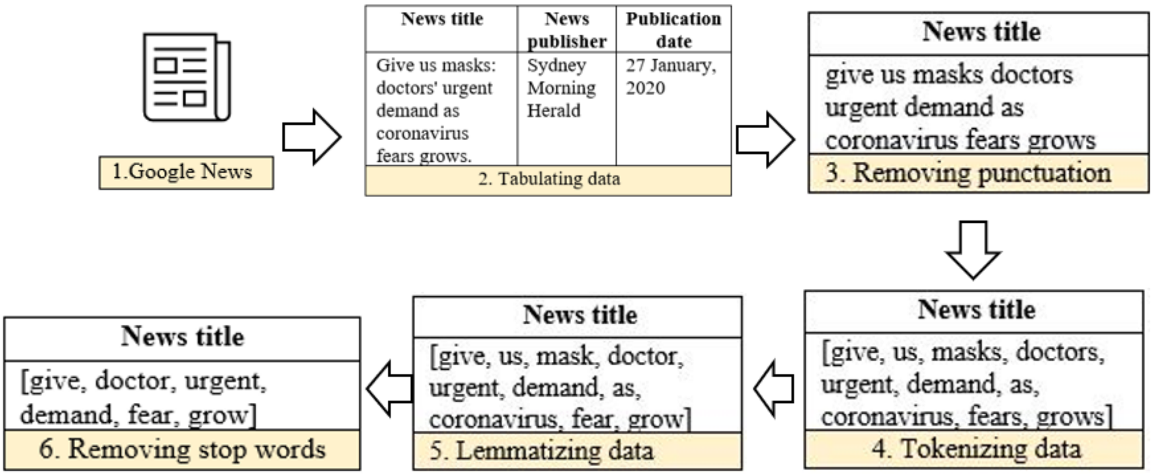
Overview

Exploratory data analysis of the news titles in the "News topic data set" was done by topic modeling technique. Topic modeling is a probabilistic computational method that identifies latent semantic patterns of word co-occurrence, also known as topic, in a collection of digital documents [37]. There are different topic modeling algorithms, particularly for short-text data such as news titles. They can be unsupervised, supervised, or semisupervised [38]. There are 5 frequently used topic modeling methods: LDA, latent semantic analysis, nonnegative matrix factorization, principal component analysis, and random projection [38]. LDA has the advantage of being an unsupervised model and some also consider the topics from LDA as easily interpretable. In a study that compared the various topic modeling algorithms, both LDA and nonnegative matrix factorization outperformed other methods and detected meaningful topics in 2 different textual data sets [38]. LDA is the mostly applied topic-modeling algorithm for textual data [39]. We used the LDA model as described by Blei et al [39] in 2003.

Data Preprocessing

Preprocessing of the news titles was performed in Python (version 3; Python software foundation) [40]. The data processing is illustrated with examples in Figure 1.

Figure 1. Steps of data preprocessing for latent Dirichlet allocation topic modeling.



Vectorization

Meaningful numerical representation of the features in the corpus is called vectorization. It represents the textual data for better understanding by the machine and is often considered to be imperative in natural language processing applications [41]. Term frequency–inverse document frequency is the commonly used method for vectorization of textual corpus [41] and was hence used in this study.

LDA Topic Model Algorithm

We again used Python to develop the LDA model [40]. LDA requires the number of topics (k) to be predetermined by the researchers.

Evaluation Matrices of LDA Models

Overview

As there is no statistical test to determine the number of topics (k), evaluation matrices can be used as a guide to identifying the best fit for the data set. We developed our model with different numbers of topics (k=1 to 20) and evaluated them according to both quantitative and qualitative matrices. Initially, we selected the top 3 models according to their performances on the quantitative evaluation scale. We then qualitatively analyzed those top 3 models and selected the best model for our data set.

Quantitative Matrices

Two quantitative matrices were used to evaluate the models.

Perplexity Measure

Perplexity is a statistical measure to estimate the performance of a probability model in predicting a sample. A lower perplexity score indicates a better model for predicting the existing topics in the data set [39]. However, sometimes perplexity does not correlate with human interpretability, as a result, additional measures are required to determine the best-fitted model.

Coherence Score

The coherence score refers to the degree of semantic similarity among high-scoring words in a topic [42,43]. It is proportionately related to the quality of the topics and their human interpretability [44]. A higher coherence score indicates a better-performing model [42]. The coherence score is considered the best quantitative evaluation matrix for topic modeling. So, we used it to evaluate our LDA topic models.

Both the perplexity measures and the coherence scores for our LDA topic models (with different numbers of topics) are shown in the *Results* section.

Qualitative Matrices

Interactive Topic Visualization

To visualize the topics, we used LDAvis [45], a bubble chart that presents the topics in a 2D space created by PC1 and PC2 (principal components 1 and 2). Each bubble represents 1 topic with the volume proportional to the frequency of the topic in the corpus. A well-designed topic model is indicated by nonoverlapping bubbles spread throughout the 2D plane. To identify the best-fitted LDA model, we compared the Intertopic distance maps of the top 3 models with their unique value of k.

Word Cloud

A word cloud for each topic shows dominating terms for the topic. The word cloud is constructed using the topic terms and their probabilities in the topic construction. Dominating terms can be easily identified and compared contextually, both within and outside the topic. The word cloud of our final LDA model is visualized in the Results section.

Topic in Context of News Titles

For a contextual analysis of the topics, we tabulated the most-representative news title for each topic and the proportion of the topic in that particular news title.

Temporal Topic Trends

Upon selection of the best-fitted model, the model can be examined at both macro and micro levels. At the macro level, the model provides an overview of the individual topic throughout the observation period. In contrast, microlevel observation can use the temporal variation of the data and illustrate the trends of the topics. In this study, the observation period was separated per month, resulting in 13 time slices, spanning January 25, 2020, to January 25, 2021. Each news title was tagged with its dominant topic (calculated from the corresponding probability distribution). Then, the distribution of the dominant topics for news titles was plotted for the 13 time slices. The evolution of the topics is illustrated in a heat map.

Results

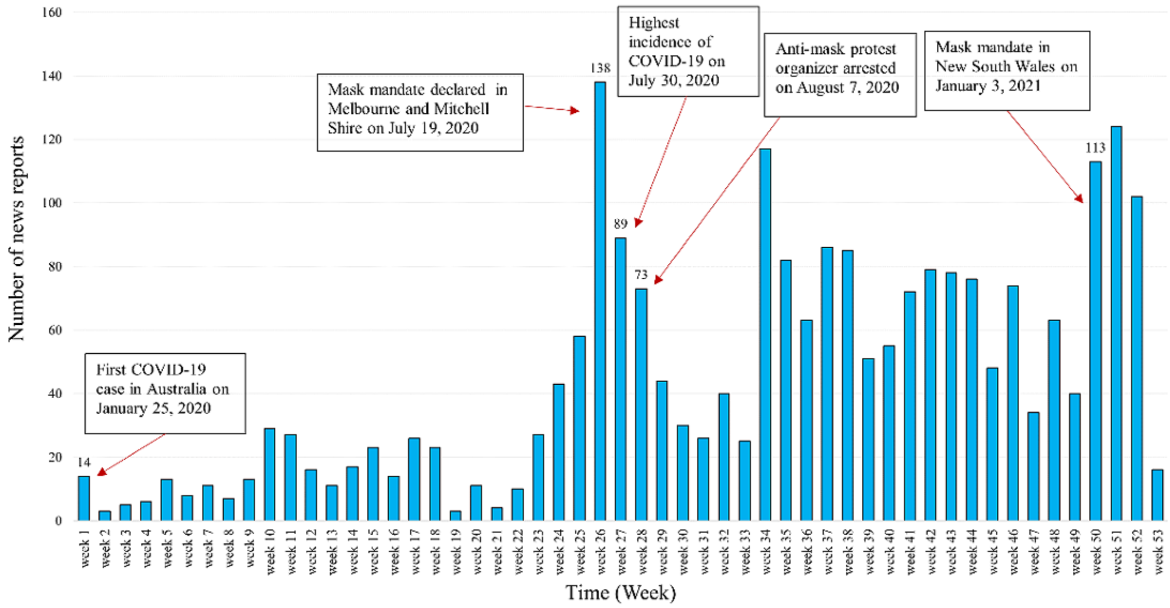
Overview

A total of 2404 mask-related Google News titles were identified in Australian web-based news websites from January 25, 2020, through to January 25, 2021. After excluding 59 irrelevant items (2.45%), a data set of 2345 mask-related news titles was finalized. The leading publishers were the national mainstream news media, such as ABC News, 7 News, and Sydney Morning Herald.

News Trend Analysis

Figure 2 presents the time trend of weekly published mask-related web-based news articles on Australian news websites for the duration of 2020-2021. The highest number (138 news articles) of news items (per week) was recorded on the 26th week (July 18 to 24, 2020). It corresponds to a pandemic wave in Melbourne, Victoria, and a mask mandate introduced in Melbourne and Mitchell shires on July 19, 2020.

Figure 2. Weekly number of mask-related web-based news articles in Australia, 2020-2021 (Week 1 commences January 25, 2020).

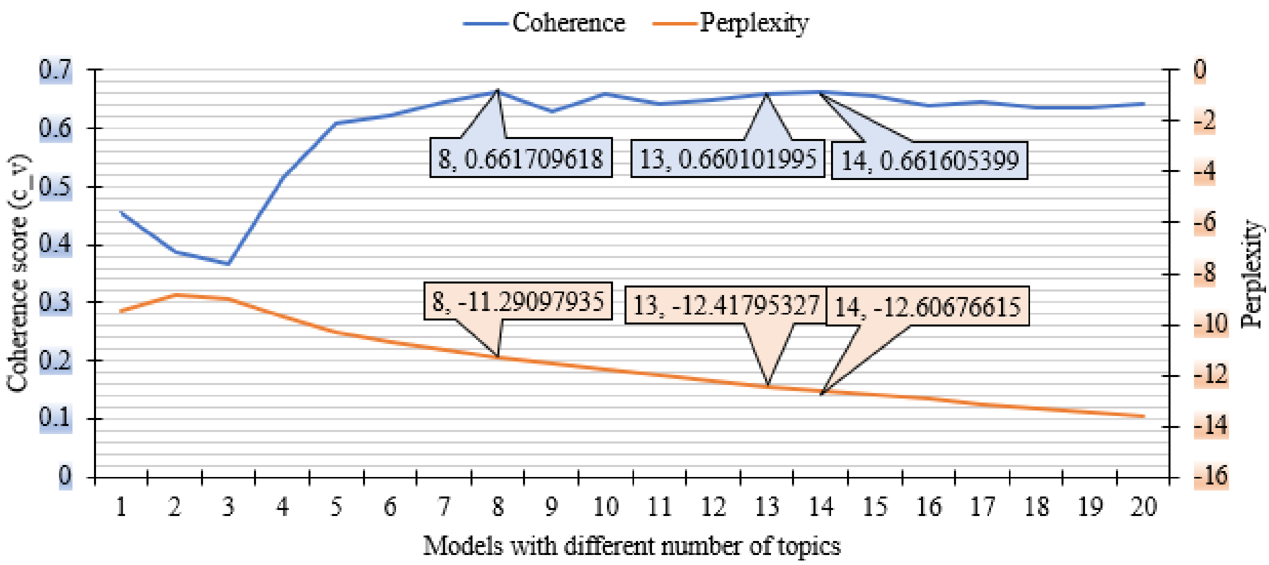


Evaluation of the LDA Topic Model

Quantitative Evaluation

Figure 3 contains line graphs for the coherence scores and perplexity measures of LDA models with 20 different numbers of topics. Based on the quantitative evaluation matrices, the selected models were the model with 8 topics, 13 topics, and 14 topics.

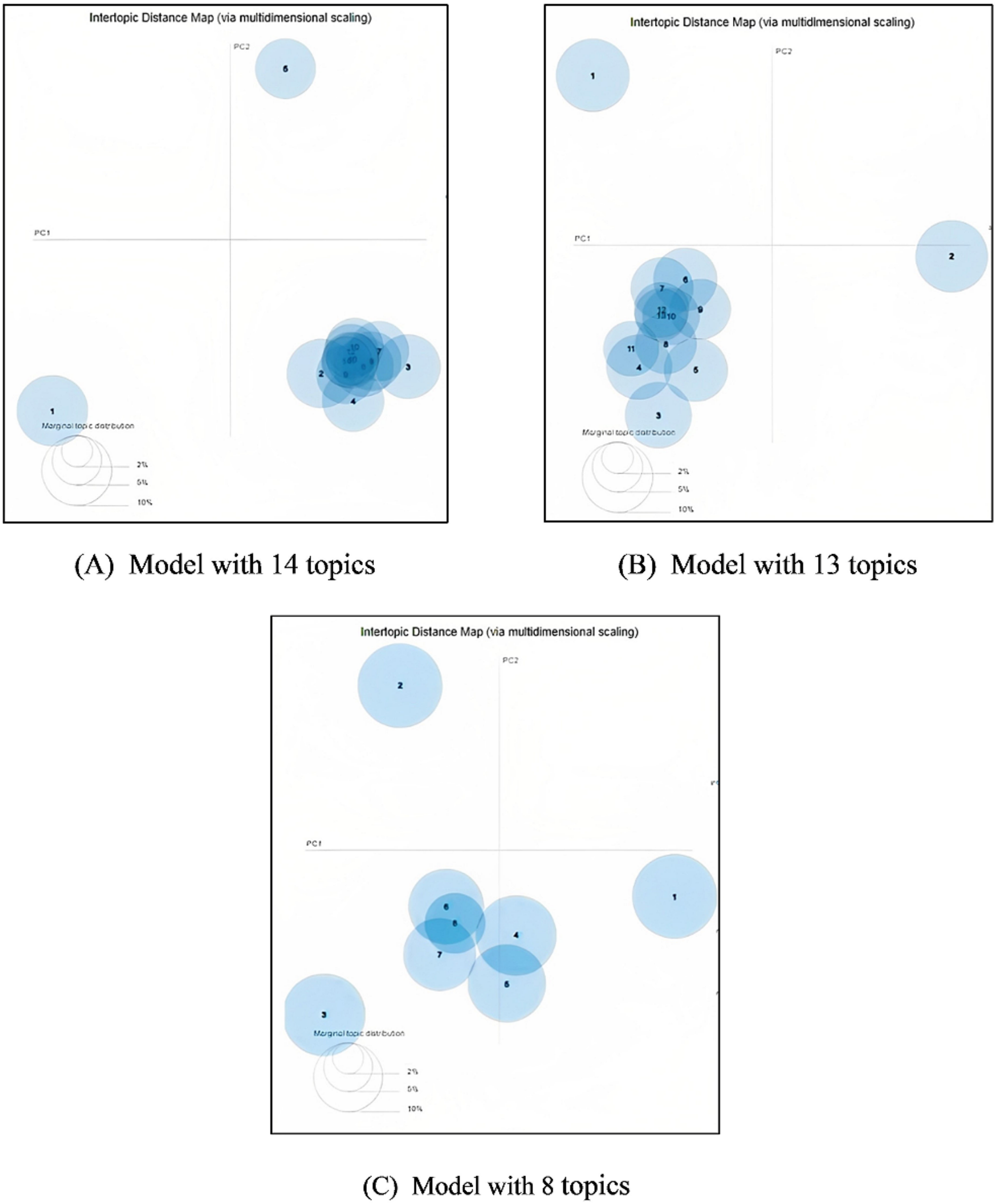
Figure 3. Perplexity measures and coherence scores of 20 topic models.



Qualitative Evaluation

As part of the qualitative evaluation, we used the Intertopic distance maps of the models. Figure 4 visualizes the Intertopic distance maps for topic models with 14, 13, and 8 topics. Overall, the model with 8 topics shows less overlap and more evenly distributed topics. This is the model we selected as the best-fitted model for our “News topic data set.”

Figure 4. Intertopic distance maps of topic models. For a higher-resolution version of this figure, see Multimedia Appendix 1.



Interpreting the Resulting Topics

Multimedia Appendix 2 shows a visualization of the topics obtained, using word clouds. As already mentioned, the size of the topic terms depends on their probabilities in the topic. The figure also includes the proportion of overall tokens for each topic, whereas the term “token” represents an individual word in the corpus. The bigger the proportion of the tokens for each topic, the better it can be interpreted.

We named the topics according to their dominant terms: mask-related international affairs (T1), introducing mask mandate in places such as Victoria and Sydney (T2), implementing mask mandate in Sydney (T3), antimask sentiment (T4), miscellaneous topic (T5), mask use in public places (T6), promoting mask-wearing (T7), and mask-wearing after the lifting of restriction (T8).

The miscellaneous topic (T5) included news articles related to the challenges of mask-wearing in western culture. In addition to the example in Table 1, another news title under this topic is, “Video shows GOP Sen. Dan Sullivan and Democratic Sen. Sherrod Brown arguing after the Republican refused to wear a mask while speaking in the Senate.” T5 also identified face masks as a canvas for individual and community expression on local and global issues. A few examples of news titles are, “Twins Max Kepler sorry for Blue Lives Matter mask amid Minneapolis protests” and “Ai Weiwei masks flip the bird, a kebab knit and gorgons: corona art.” Table 1 contains a table of the topics, their keywords, and most representative news titles with the probability of that given topic in that particular news title.

Table 1. The most representative news of 8 topics.

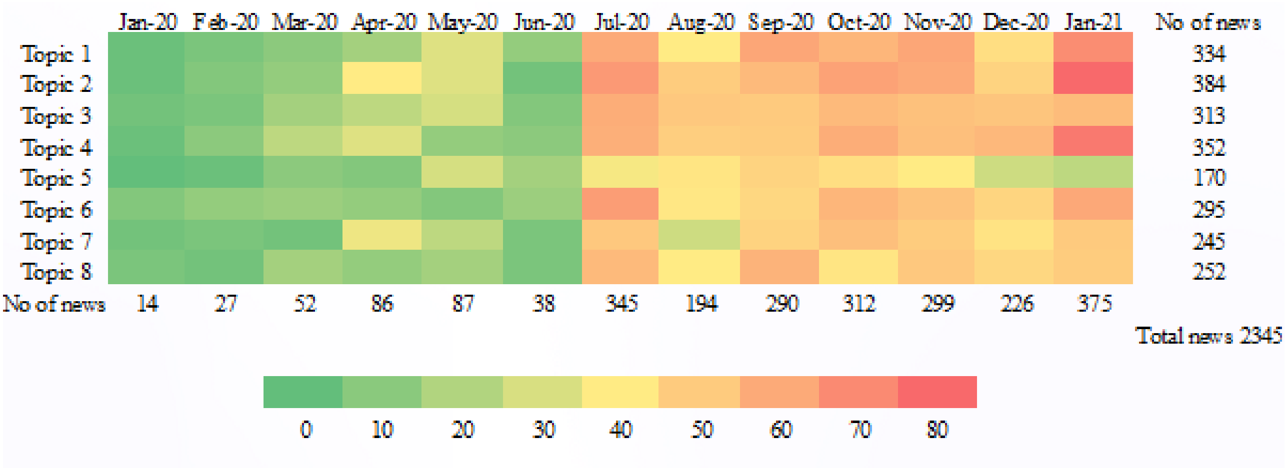
Topic number	Topic keywords	Most representative news title	Topic contribution in the title, %
1	restriction, international, eased, biden, melbourne, lifted, shopper, dont, open, one	Melania Trump votes without a mask on US election day as Donald Trump, Joe Biden, and Kamala Harris make last-minute pitches.	65.58
2	new, mandatory, fine, victoria, greater, premier, sydney, 2021, victorian, officer	Bow Wow performed in a packed Houston nightclub for scores of mask-less attendees—and people have a lot to say about it.	61.42
3	sydney, mandated, ease, must, resident, fined, strain, public, mandatory, could	NSW Health Minister Brad Hazzard apologizes for calling Labor leader Jodi McKay “quite stupid” in mask stoush.	63.99
4	mandate, charged, flight, police, supermarket, antimask, right, woman, barty, refusing	Delta has now banned some 350 passengers for refusing to wear masks during flights — and it is adding 100 people a month to its no-fly list.	65.51
5	man, 10, warning, set, best, required, latest, biting, alert, plan	HHS secretary Alex Azar says wearing masks is “a difficult message for all Western democracies.”	58.82
6	airport, brisbane, test, case, office, wa, opera, stock, get, around	Armidale’s Isabelle and Lillie Kelly go from homemade scrunchies and hair ties to face masks.	60.44
7	rule, smart, pushup, likely, confirms, made, without, shopping, half, day	Queen Elizabeth wears a face mask in public for ceremony to mark the centenary of burial of the Unknown Warrior.	60.24
8	nsw, masked, buy, across, issue, men, bali, tourist, mandate, order	Coronavirus in Ohio: Parents sue health department director Lance Himes over K-12 school mask mandate.	62.22

Temporal Topic Trends

Figure 5 is a heat map of the temporal trends of the topics identified in our model.

The heat map distinctly represents both similar and opposite trends of topic evolution. Few mask-related news items were published in the initial 6 months (January 2020 to June 2020), followed by a sharp increase afterward for all the topics. In the first half of our analysis, January 2020 to June 2020, news relating to T2, T4, and T7 appeared more frequently. Then T2 and T4 were predominant with a peak in January 2021 (77 items for T2 and 70 items for T4).

Figure 5. Heat map of the temporal distribution of mask-related topics.



Discussion

Principal Findings

During the COVID-19 pandemic, national mainstream media organizations were at the forefront of reporting mask-related stories in Australia. We conducted a series of LDA analyses of our corpus on mask-related news items, using quantitative and qualitative measures to identify the best-fitted model (the best number of topics). Even though both coherence score and perplexity measure were used to quantitatively assess our LDA model performances, the utility of perplexity measure was deemed negligible in our LDA model selection. Based on the evaluation matrices, we identified 8 topics related to face masks, with 3 key topics: mask-related international affairs (T1), introducing mask mandate in places such as Victoria and Sydney (T2), and antimask sentiment (T4). Both T2 and T4 were dominant and peaked in January 2021. This study demonstrated the utility of a data-driven approach to analyzing news media and identified how it communicated health-related information during a public health crisis.

During the study period, there were 2 distinct waves of mask-related news publications in Australia. The second wave was much stronger than the first and was mostly caused by local spikes in COVID-19 cases and the introduction of mandatory mask-wearing in different states. In a comparative cluster analysis, Yang et al [29] found a similar trend in the United Kingdom and the United States: both countries initially experienced relatively few news items related to mask, followed by an increase in response to local COVID-19 outbreak. In contrast, Chinese news media consistently published mask-related news since the pandemic started [29].

Similar to our study methodology, Suh et al [34] used an LDA topic model for their Korean news data set on face masks. However, they only used a quantitative method to evaluate their model, and their semantic similarity score of topic keywords is smaller (0.52) than the one in our study (0.66). It means that the topics in our study are more clearly defined and more easily interpretable.

Topic wise, T1 (mask-related stories in international news) was one of the major topics identified in our corpus, with 334 news titles. The most representative news title under this topic showed that Australian web-based media coverage of mask-related international events was mostly from the US political arena. Considering the escalating conflicts between politicians and the later politicization of mask use in the US election, this topic reflects the centralized role of the United States in overall "mask madness" [46]. Lee et al [33] identified a similar topic in their study, which they labeled "President election." On topic-trend analysis in our study, T1 constantly appeared from July 2020 onward. Following the recommendation of the US Centers for Disease Control and Prevention (CDC) on public masking, this topic covered an increasing number of news reports on public gatherings (protests, political campaigns) that might worsen the ongoing pandemic.

T2 of our LDA model was about introducing mask mandates, mainly in response to the new cases of COVID-19 in Victoria and Greater Sydney. This topic gained the highest media coverage, with a total of 384 news titles. On the temporal trend analysis, this topic reflected the government-led agenda on mask-wearing and evolved along with the mask-relevant policy changes across the states in Australia. The highest amount of T2 specific news coverage (77 items) was recorded during the mask mandate in the Greater Sydney region in January 2021, followed by the Melbourne mask mandate (56 items) in July 2020. The predominance of the second topic around the policy change implied that the reporting by Australian media was timely and in line with a policy maker's guideline to promote mask-wearing [6]. Suh et al [34] identified a similar topic during the second outbreak of COVID-19 in Korea. Lee et al [33] also identified this topic in a large data set of international news about mask-wearing.

Our third topic (T3) was related to mask mandates in the Greater Sydney area, reflecting mostly actions taken by law-enforcement agencies. This topic also depicted different attitudes of the political leaders and public figures toward face masks and associated policy development. Over time, T3 was consistent from the time of the Melbourne outbreak in July 2020. Following the outbreak in the Greater Sydney region in January 2021, this topic exhibited an increased number of related news publications.

Besides T2, the next major topic was related to antimask sentiments (T4). In our study, a total of 352 news titles were recorded under this topic. Lee et al [33] had a similar finding in their analysis, concluding that the "Protesting movement" topic in mask-related news predominantly appeared after WHO's pandemic announcement. This topic illustrates the media messaging on the conflict between the government agenda and public agenda in terms of mass masking strategy. The highest number of news items for this topic was related to the Greater Sydney mask mandate in January 2021 (70 news titles), followed by the Melbourne mask mandate in July 2020 (46 news titles).

With 295 relevant news articles, T6 was about news encouraging mask-wearing in public places and at events. Similarly, T8 described mask-related policy implementation while reopening offices and educational institutions. With 252 mask-related articles, this topic drew more attention during September 2020. Overall, T6 and T8 focused on the universal adoption of face masks and integrating the concept of mask-wearing in regular and recreational events. These topics suggest the importance of mask-wearing to minimize COVID-19 transmission while getting "back to normal." In the United States, mask-wearing policies related to school-reopening strategies effectively reduced the community transmission of COVID-19 [47]. In international news media data, Lee et al [33] also identified multiple topics about mask-wearing in educational institutions, local businesses, sports, and family events.

The next topic in our study, T7, was about the innovative promotion of mask-wearing. With 245 news titles, this topic developed since the recommendation of the US CDC on wearing nonmedical face masks in public places. During the second phase of our study, from July 2020, T7 evolved constantly and might contribute to encouraging mask-wearing by framing positive social messages in the news content. Furthermore, these positive social messages can contribute to developing positive social cognitions in the community [48].

Last, T5 was regarded as a miscellaneous topic with the least news coverage. Topic-specific keywords did not cluster under any specific topic. On the temporal analysis, the number of such miscellaneous news gradually decreased, making a place for more structured and topic-specific media messaging in Australian web-based news platforms. In contrast, Suh et al [34] identified an increasing number of news items under miscellaneous topics during the third wave of COVID-19 in Korea.

Our study has several strengths that can be attributed to both the methodology and the outcome of the study. Methodological strengths include the inclusion of both national and local news media reports on face masks in Australia. Although international news articles rarely reflect the Australian national agenda, selective news websites may not represent all the national issues. Additionally, the data collection tool Google News is distinctly advantageous over other news databases such as LexisNexis. Google News provides more reliable data and can capture more local news and nonprint articles that are often overlooked in other database searches [49]. Moreover, Google News is a free data collection tool that permits a wide range of researchers to access and analyze news data, and therefore analyses are amenable to assess for replicability.

Limitations

There are several limitations that can be attributed to this study. One of the major limitations is analyzing only the news titles. They might not provide adequate context to capture appropriate topics for corresponding news articles. However, considering the structural integrity of news titles in summarizing news themes, it is likely still meaningful. In future research, we could aim at capturing both the title and the snippet returned by the search engine. By using images from the news and their associated titles, visual topic modeling could also be a novel approach to address the limitations of only news title analysis [50]. Furthermore, there are certain limitations to Google News. It sometimes restricts access to news that requires a subscription. Methodologically, LDA may not necessarily perform well when analyzing short news titles. For example, the most representative news title for T2 does not represent the topic in our LDA model (see Table 1). Additionally, we did not perform a sentiment analysis on the news title. Such an analysis could be included in future work [51,52]. Finally, our study of news items could be complemented by an analysis of social media discourse.

Conclusions

Appropriate preventive strategies, health communication, public education, and active community participation have always been at the center of effective pandemic responses. News media platforms, as primary sources of updated health information, could serve the interests of the government and the community by disseminating information about health policies and healthy behaviors. However, the media can also amplify dissent and uncertainty. This study highlighted the public health recommendations about mask wearing and diverse aspects of mask use (or not) in the media. Timely, targeted, and transformative media messages can rapidly spread around the world, and therefore those messages can facilitate local and global community adoption of mask-wearing which will benefit all.

Acknowledgments

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided partial article processing fee (APF) support for the publication of this article.

Conflicts of Interest

CRM has received funding for antimicrobial testing of masks from Detmold, and has been on an advisory board for mask manufacturer Ascend.

Multimedia Appendix 1
Intertopic distance maps of topic models.
[PNG File , 1158 KB - infodemiology_v3i1e43011_app1.png]

Multimedia Appendix 2
Word clouds of 8 topics.
[PNG File , 421 KB - infodemiology_v3i1e43011_app2.png]

References

1. O'Sullivan D, Rahamathulla M, Pawar M. The impact and implications of COVID-19: an Australian perspective. Int J Community Soc Dev 2020;2(2):134-151 [FREE Full text] [doi: 10.1177/2516602620937922]
2. Saul A, Scott N, Crabb BS, Majumdar SS, Coghlan B, Hellard ME. Impact of Victoria's stage 3 lockdown on COVID-19 case numbers. Med J Aust 2020;213(11):494-496.e1 [FREE Full text] [doi: 10.5694/mja2.50872] [Medline: 33230817]

3. Faasse K, Newby J. Public perceptions of COVID-19 in Australia: perceived risk, knowledge, health-protective behaviors, and vaccine intentions. Front Psychol 2020;11:551004 [FREE Full text] [doi: 10.3389/fpsyg.2020.551004] [Medline: 33117223]
4. Park S, Fisher C, Lee JY, McGuinness K. COVID-19: Australian News and Misinformation. Canberra, Australia: News and Media Research Centre, University of Canberra; 2020.
5. Mittal R, Ni R, Seo JH. The flow physics of COVID-19. J. Fluid Mech 2020;894:F2 [FREE Full text] [doi: 10.1017/jfm.2020.330]
6. Promoting mask-wearing during the COVID-19 pandemic: a policymaker's guide. Vital Strategies. 2021. URL: https://preventepidemics.org/wp-content/uploads/2020/08/Promoting-Mask-Wearing-During-COVID-19.pdf [accessed 2023-06-30]
7. Oran DP, Topol EJ. Prevalence of asymptomatic SARS-CoV-2 infection : a narrative review. Ann Intern Med 2020;173(5):362-367 [FREE Full text] [doi: 10.7326/M20-3012] [Medline: 32491919]
8. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
9. Laestadius L, Wang Y, Ben Taleb Z, Kalan ME, Cho Y, Manganello J. Online National Health Agency mask guidance for the public in light of COVID-19: content analysis. JMIR Public Health Surveill 2020;6(2):e19501 [FREE Full text] [doi: 10.2196/19501] [Medline: 32427103]
10. Campbell K, Vines E, Phillips J, Corrigan N, Biggs A. COVID-19: a chronology of state and territory government announcements (up until 30 June 2020). Parliament of Australia, Parliament Library Research Paper Series. 2021. URL: https://parlinfo.aph.gov.au/parlInfo/download/library/prspub/8027137/upload_binary/8027137.pdf [accessed 2023-06-30]
11. Face coverings mandatory for Melbourne and Mitchell Shire. Victoria State Government, Health and Human Services. 2020. URL: https://tinyurl.com/yc2pm7ds [accessed 2023-06-30]
12. Scott N, Saul A, Spelman T, Stoove M, Pedrana A, Saeri A, et al. The introduction of a mandatory mask policy was associated with significantly reduced COVID-19 cases in a major metropolitan city. PLoS One 2021;16(7):e0253510 [FREE Full text] [doi: 10.1371/journal.pone.0253510] [Medline: 34288910]
13. Laing A. The H1N1 crisis: roles played by government communicators, the public and the media. J Prof Commun 2021;1(1):123-149 [FREE Full text] [doi: 10.15173/jpc.v1i1.88]
14. Klemm C, Das E, Hartmann T. Swine flu and hype: a systematic review of media dramatization of the H1N1 influenza pandemic. J Risk Res 2014;19(1):1-20. [doi: 10.1080/13669877.2014.923029]
15. Pieri E. Media framing and the threat of global pandemics: the Ebola crisis in UK media and policy response. Sociol Res Online 2019;24(1):73-92 [FREE Full text] [doi: 10.1177/1360780418811966] [Medline: 34285639]
16. Hoffman SJ, Justicz V. Automatically quantifying the scientific quality and sensationalism of news records mentioning pandemics: validating a maximum entropy machine-learning model. J Clin Epidemiol 2016;75:47-55 [FREE Full text] [doi: 10.1016/j.jclinepi.2015.12.010] [Medline: 26854419]
17. Nolan D, Hanna C, McGuinness K, McCallum K. Covering COVID-19: How Australian Media Reported the Coronavirus Pandemic in 2020. Canberra, Australia: News and Media Research Centre, University of Canberra; 2021.
18. News in Australia-impartiality and commercial influence: review of literature and research. Centre for Media Transition. 2020. URL: https://tinyurl.com/2xktvppn [accessed 2023-06-30]
19. Luo Y, Burley H, Moe A, Sui M. A meta-analysis of news media's public agenda-setting effects, 1972-2015. Journal Mass Commun Q 2018;96(1):150-172 [FREE Full text] [doi: 10.1177/1077699018804500]
20. McCombs ME, Shaw DL. The agenda-setting function of mass media. Public Opin Q 1972;36(2):176-187. [doi: 10.1086/267990]
21. Conway-Silva BA, Filer CR, Kenski K, Tsetsi E. Reassessing Twitter's agenda-building power: an analysis of intermedia agenda-setting effects during the 2016 presidential primary season. Soc Sci Comput Rev 2017;36(4):469-483. [doi: 10.1177/0894439317715430]
22. Ogata Jones K, Denham BE, Springston JK. Effects of mass and interpersonal communication on breast cancer screening: advancing agenda-setting theory in health contexts. J Appl Commun Res 2006;34(1):94-113. [doi: 10.1080/00909880500420242]
23. Su Y, Borah P. Who is the agenda setter? Examining the intermedia agenda-setting effect between Twitter and newspapers. J Inf Technol Politics 2019;16(3):236-249. [doi: 10.1080/19331681.2019.1641451]
24. Zhou S, Zheng X. Agenda dynamics on social media during COVID-19 pandemic: interactions between public, media, and government agendas. Commun Stud 2022;73(3):211-228. [doi: 10.1080/10510974.2022.2082504]
25. Medina LM, Rodriguez JR, Sarmiento PJD. Shaping public opinion through the lens of agenda setting in rolling out COVID-19 vaccination program. J Public Health (Oxf) 2021;43(2):e389-e390 [FREE Full text] [doi: 10.1093/pubmed/fdab100] [Medline: 33783540]
26. Thomas T, Wilson A, Tonkin E, Miller ER, Ward PR. How the media places responsibility for the COVID-19 pandemic-an Australian media analysis. Front Public Health 2020;8:483 [FREE Full text] [doi: 10.3389/fpubh.2020.00483] [Medline: 32974266]

27. Basch CH, Hillyer GC, Meleo-Erwin Z, Mohlman J, Cosgrove A, Quinones N. News coverage of the COVID-19 pandemic: missed opportunities to promote health sustaining behaviors. Infect Dis Health 2020;25(3):205-209 [FREE Full text] [doi: 10.1016/j.idh.2020.05.001] [Medline: 32426559]
28. 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;22(4):e19118 [FREE Full text] [doi: 10.2196/19118] [Medline: 32302966]
29. Yang S, Zhang W, Yuan Z. Media reports of the covid-19 pandemic: a computational text analysis of english reports in China, the UK, the US. Adv Journal Commun 2021;9(2):47-62 [FREE Full text] [doi: 10.4236/ajc.2021.92004]
30. Kaveh-Yazdy F, Zarifzadeh S. Track Iran's national COVID-19 response committee's major concerns using two-stage unsupervised topic modeling. Int J Med Inform 2021;145:104309 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104309] [Medline: 33181447]
31. de Melo T, Figueiredo CMS. Comparing news articles and tweets about COVID-19 in Brazil: sentiment analysis and topic modeling approach. JMIR Public Health Surveill 2021;7(2):e24585 [FREE Full text] [doi: 10.2196/24585] [Medline: 33480853]
32. Ophir Y, Walter D, Arnon D, Lokmanoglu A, Tizzoni M, Carota J, et al. The framing of COVID-19 in Italian media and its relationship with community mobility: a mixed-method approach. J Health Commun 2021;26(3):161-173. [doi: 10.1080/10810730.2021.1899344] [Medline: 33787462]
33. Lee KR, Kim B, Nan D, Kim JH. Structural topic model analysis of mask-wearing issue using international news big data. Int J Environ Res Public Health 2021;18(12):6432 [FREE Full text] [doi: 10.3390/ijerph18126432] [Medline: 34198600]
34. Suh YR, Koh KP, Lee J. An analysis of the change in media's reports and attitudes about face masks during the COVID-19 pandemic in South Korea: a study using Big Data latent dirichllocation (LDA) topic modelling. J Korea Inst Inf Commun Eng 2021;25(5):731-740. [doi: 10.6109/jkiice.2021.25.5.731]
35. Andreotta M, Nugroho R, Hurlstone MJ, Boschetti F, Farrell S, Walker I, et al. Analyzing social media data: a mixed-methods framework combining computational and qualitative text analysis. Behav Res Methods 2019;51(4):1766-1781 [FREE Full text] [doi: 10.3758/s13428-019-01202-8] [Medline: 30941697]
36. Rodriguez MY, Storer H. A computational social science perspective on qualitative data exploration: using topic models for the descriptive analysis of social media data. J Technol Hum Serv 2019;38(1):54-86. [doi: 10.1080/15228835.2019.1616350]
37. Jacobi C, van Atteveldt W, Welbers K. Quantitative analysis of large amounts of journalistic texts using topic modelling. Digit Journal 2015;4(1):89-106. [doi: 10.1080/21670811.2015.1093271]
38. Albalawi R, Yeap TH, Benyoucef M. Using topic modeling methods for short-text data: a comparative analysis. Front Artif Intell 2020;3:42 [FREE Full text] [doi: 10.3389/frai.2020.00042] [Medline: 33733159]
39. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003;3:993-1022 [FREE Full text] [doi: 10.5555/944919.944937]
40. Van Rossum G, Drake FL. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace; 2009.
41. Singh AK, Shashi M. Vectorization of text documents for identifying unifiable news articles. Int J Adv Comput Sci Appl 2019;10(7):305-310 [FREE Full text] [doi: 10.14569/ijacsa.2019.0100742]
42. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. Proceedings of the Conference on Empirical Methods in Natural Language Processing; Edinburgh, United Kingdom: Association for Computational Linguistics; 2011 Presented at: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing; July 27–31, 2011; Edinburgh, Scotland, UK p. 262-272.
43. DiMaggio P, Nag M, Blei D. Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. government arts funding. Poetics 2013;41(6):570-606. [doi: 10.1016/j.poetic.2013.08.004]
44. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. NY, US: Association for Computing Machinery; 2015 Presented at: WSDM 2015: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining; February 2 - 6, 2015; Shanghai, China p. 399-408. [doi: 10.1145/2684822.2685324]
45. Sievert C, Shirley K. LDAvis: a method for visualizing and interpreting topics. : Association for Computational Linguistics; 2014 Presented at: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces; June 27, 2014; Baltimore, Maryland, USA p. 63-70. [doi: 10.3115/v1/w14-3110]
46. Kahane LH. Politicizing the mask: political, economic and demographic factors affecting mask wearing behavior in the USA. East Econ J 2021;47(2):163-183 [FREE Full text] [doi: 10.1057/s41302-020-00186-0] [Medline: 33424048]
47. Chernozhukov V, Kasahara H, Schrimpf P. The association of opening K-12 schools with the spread of COVID-19 in the United States: county-level panel data analysis. Proc Natl Acad Sci U S A 2021;118(42):e2103420118 [FREE Full text] [doi: 10.1073/pnas.2103420118] [Medline: 34642247]
48. Yao Z, Yu R. The spreading of social energy: how exposure to positive and negative social news affects behavior. PLoS One 2016;11(6):e0156062 [FREE Full text] [doi: 10.1371/journal.pone.0156062] [Medline: 27253877]
49. Weaver DA, Bimber B. Finding news stories: a comparison of searches using Lexisnexis and Google News. Journal Mass Commun Q 2008;85(3):515-530. [doi: 10.1177/107769900808500303]

50. Pivovarova L, Zosa E. Visual topic modelling for NewsImage Task at MediaEval 2021. MediaEval Benchmarking Initiative for Multimedia Evaluation. 2021. URL: https://ceur-ws.org/Vol-3181/paper37.pdf [accessed 2023-06-30]
51. Kim JH, Nan D, Kim Y, Park MH. Computing the user experience via big data analysis: a case of Uber services. Comput Mater Contin 2021;67(3):2819-2829 [FREE Full text] [doi: 10.32604/cmc.2021.014922]
52. Kim JH, Jung HS, Park MH, Lee SH, Lee H, Kim Y. Exploring cultural differences of public perception of artificial intelligence via Big Data Approach. : Springer, Cham; 2022 Presented at: HCI International 2022 Posters: 24th International Conference on Human-Computer Interaction, HCII 2022; June 26 – July 1, 2022; Virtual Event p. 427-432. [doi: 10.1007/978-3-031-06417-3_57]

Abbreviations

CDC: Centers for Disease Control and Prevention
LDA: latent Dirichlet allocation
WHO: World Health Organization

Edited by T Mackey; submitted 27.09.22; peer-reviewed by D Nan, T Purnat; comments to author 25.02.23; revised version received 20.04.23; accepted 02.05.23; published 16.08.23.

Please cite as:
Dasgupta P, Amin J, Paris C, MacIntyre CR
News Coverage of Face Masks in Australia During the Early COVID-19 Pandemic: Topic Modeling Study
JMIR Infodemiology 2023;3:e43011
URL: https://infodemiology.jmir.org/2023/1/e43011
doi:10.2196/43011
PMID:37379362

Original Paper

Content and User Engagement of Health-Related Behavior Tweets Posted by Mass Media Outlets From Spain and the United States Early in the COVID-19 Pandemic: Observational Infodemiology Study

Miguel Angel Alvarez-Mon[1,2,3], MD, PhD; Victor Pereira-Sanchez[4], MD, PhD; Elizabeth R Hooker[5,6], MPH, MS; Facundo Sanchez[7,8], MD; Melchor Alvarez-Mon[1,3], MD, PhD; Alan R Teo[5,9], MS, MD

[1]Department of Medicine and Medical Specialties, Faculty of Medicine and Health Sciences, University of Alcala, Alcala de Henares, Spain

[2]Department of Psychiatry and Mental Health, University Hospital Infanta Leonor, Madrid, Spain

[3]Ramón y Cajal Institute of Sanitary Research (IRYCIS), Madrid, Spain

[4]Department of Child and Adolescent Psychiatry, NYU Grossman School of Medicine, New York, NY, United States

[5]VA Portland Health Care System, Health Services Research & Development Center to Improve Veteran Involvement in Care, Portland, OR, United States

[6]OHSU-PSU School of Public Health, Oregon Health and Science University, Portland, OR, United States

[7]Lincoln Medical and Mental Health Center, New York, NY, United States

[8]Devers Eye Institute, Portland, OR, United States

[9]Department of Psychiatry, Oregon Health & Science University, Portland, OR, United States

Corresponding Author:
Miguel Angel Alvarez-Mon, MD, PhD
Department of Medicine and Medical Specialties
Faculty of Medicine and Health Sciences
University of Alcala
Campus Universitario – C/ 19, Av de Madrid, Km 33,600.
Alcala de Henares, 28871
Spain
Phone: 34 918854505
Email: maalvarezdemon@icloud.com

Abstract

Background: During the early pandemic, there was substantial variation in public and government responses to COVID-19 in Europe and the United States. Mass media are a vital source of health information and news, frequently disseminating this information through social media, and may influence public and policy responses to the pandemic.

Objective: This study aims to describe the extent to which major media outlets in the United States and Spain tweeted about health-related behaviors (HRBs) relevant to COVID-19, compare the tweeting patterns between media outlets of both countries, and determine user engagement in response to these tweets.

Methods: We investigated tweets posted by 30 major media outlets (n=17, 57% from Spain and n=13, 43% from the United States) between December 1, 2019 and May 31, 2020, which included keywords related to HRBs relevant to COVID-19. We classified tweets into 6 categories: mask-wearing, physical distancing, handwashing, quarantine or confinement, disinfecting objects, or multiple HRBs (any combination of the prior HRB categories). Additionally, we assessed the likes and retweets generated by each tweet. Poisson regression analyses compared the average predicted number of likes and retweets between the different HRB categories and between countries.

Results: Of 50,415 tweets initially collected, 8552 contained content associated with an HRB relevant to COVID-19. Of these, 600 were randomly chosen for training, and 2351 tweets were randomly selected for manual content analysis. Of the 2351 COVID-19–related tweets included in the content analysis, 62.91% (1479/2351) mentioned at least one HRB. The proportion of COVID-19 tweets mentioning at least one HRB differed significantly between countries (*P*=.006). Quarantine or confinement was mentioned in nearly half of all the HRB tweets in both countries. In contrast, the least frequently mentioned HRBs were disinfecting objects in Spain 6.9% (56/809) and handwashing in the United States 9.1% (61/670). For tweets from the United

States mentioning at least one HRB, disinfecting objects had the highest median likes and retweets, whereas mask-wearing– and handwashing-related tweets achieved the highest median number of likes in Spain. Tweets from Spain that mentioned social distancing or disinfecting objects had a significantly lower predicted count of likes compared with tweets mentioning a different HRB (*P*=.02 and *P*=.01, respectively). Tweets from the United States that mentioned quarantine or confinement or disinfecting objects had a significantly lower predicted number of likes compared with tweets mentioning a different HRB (*P*<.001), whereas mask- and handwashing-related tweets had a significantly greater predicted number of likes (*P*=.04 and *P*=.02, respectively).

Conclusions: The type of HRB content and engagement with media outlet tweets varied between Spain and the United States early in the pandemic. However, content related to quarantine or confinement and engagement with handwashing was relatively high in both countries.

(JMIR Infodemiology 2023;3:e43685) doi:10.2196/43685

KEYWORDS

COVID-19; health communication; social media; Twitter; health promotion; public health; mass media

Introduction

Harnessing Twitter during the COVID-19 pandemic may have been beneficial [1,2]. In previous disasters (eg, H1N1 outbreak), Twitter was among the most used social platforms [3-5]. Mass media outlets are a key tool during disasters such as the COVID-19 pandemic because they can educate people [6]. Through Twitter alone, millions of users interact each day, and during disasters the number of interactions and tweets increases [7]. Tweets during disasters can inform the public of risk factors, where to ask for help, locations and availability of hospitals, and locations of people who need help (eg, elderly living alone) [8]. If media outlets share accurate and valuable information (eg, encourage people to stay at home) through Twitter, they may contribute to saving lives [9]. Mass media outlet Twitter accounts can become an excellent resource for the public to stay updated on health policies during the COVID-19 pandemic [10].

Some behavioral measures, such as handwashing, mask-wearing, and physical distancing, are among the most effective tools to combat the COVID-19 pandemic [11,12]. These are considered preventive health behaviors. The World Health Organization, the Centers for Disease Control and Prevention, and many other health institutions have addressed their importance in slowing the spread of the coronavirus. Furthermore, shelter in place (stay-at-home orders for the general population), quarantine (separation and restriction of movements of people who have potentially been exposed to COVID-19), and isolation (separation of people who have been diagnosed from people who are not sick) have been mandatory in many countries around the world.

Social media can be used to create indicators of the health environment that are associated with area-level mortality and health behaviors [13,14]. Local area characteristics are increasingly associated with health outcomes. Social processes affect health through the maintenance of social norms, stimulation of new interests, and the dispersal of knowledge. Therefore, comparing Twitter posts from different countries might explain certain differences found among them with regard to responding to the COVID-19 pandemic.

Assessing health-related behaviors (HRBs) on Twitter can help us understand the degree of awareness of the population using these preventive measures. It has been shown that there is an association between the characteristics of the content published in a certain geographic area and the rates of obesity and diabetes mellitus in that area [15]. In fact, a study showed that areas with the most tweets about physical activity or healthy foods (fruits, vegetables, etc) had lower obesity rates [16]. Twitter posts have also been analyzed to track behaviors related to the transmission of infectious diseases such as HIV [17]. Currently, multiple researchers are applying this methodology to better understand the reactions of the population, as well as raise awareness and promote compliance with health measures [18].

We conducted an observational retrospective study analyzing the content of mass media outlets posted on Twitter referring to COVID-19 during the early pandemic. Using Twitter, we sought to analyze the content posted by 30 major media outlets (17 from Spain and 13 from the United States) about COVID-19 during the first 6 months of the pandemic. Our two primary research aims were to (1) describe the extent to which major media outlets in the United States and Spain have tweeted about COVID-19 HRBs and determine if differences exist between major media outlets in the 2 countries and (2) determine the extent of user engagement in response to tweets about HRBs. We aimed to incorporate a cross-cultural perspective for several reasons. First, we sought to assess whether the media in Spain exhibited communication patterns similar to those in the United States. Furthermore, our hypothesis is that the beliefs and practices related to health may differ between the 2 countries because of political or cultural factors (such as differences in the health care system). In addition, the demographic characteristics of each country vary significantly, which can affect how the population experiences the pandemic. In Spain, a larger percentage of the population resides in multi-unit buildings where shared spaces are common, whereas individual houses are more prevalent in the United States.

Methods

Study Design and Overview

In this observational infodemiology study, we used concurrent collection and analysis of qualitative and quantitative data from tweets concerning public health measures relevant to the prevention and mitigation of the spread of the novel coronavirus posted by major media outlets in Spain and the United States.

Data Collection

Tweets were drawn from a total of 30 major media outlets: 17 from Spain (Antena 3, La Sexta, TVE, Telecinco, cadena SER, cadena COPE, Onda Cero, ABC, El Pais, El Mundo, Europa Press, Noticias Cuatro, EFE, El Diario, La Vanguardia, Público, and Info Libre) and 13 from the United States (New York Times, Washington Post, Los Angeles Times, USA Today, The Chicago Tribune, New York Post, Wall Street Journal, Boston Globe, San Francisco Chronicle, MSNBC, CNN, ABC News, and CBS News). The 30 media outlets that we selected were general, had national coverage, had a large audience, and were among those with the highest social influence in their respective countries. We included different modalities, such as radio, newspapers, and television, to be more representative.

Our data collection strategy focused on identifying original tweets related to 4 HRBs: mask-wearing, physical distancing, quarantine or confinement, and hygiene. We included all original tweets posted from the Twitter accounts of the previously mentioned media outlets, thus avoiding the collection of tweets posted by other users, even if they were retweeted or mentioned by these media outlets, that referred to coronavirus (#corona, #coronavirus, #covid-19, #SARS-CoV2, and #2019-nCoV) and contained at least one of the following keywords (or their Spanish equivalents): #handwashing, #selfquarentine, #socialdistancing, #selfisolation, #masking, #mask, #disinfect, #disinfection, #socialgatherings, #sneez, #cough, #physical distancing, #facemask, #facecovering, #clothface, #lavadodemanos, #estornudo, #tos, #cuarentena, #distanciasocial, #distanciafísica, #aislamiento, #desinfectar, #desinfección, and #mascarilla. Original tweets posted between December 1, 2019, and May 31, 2020, were included, and data were collected on May 31, 2020. For each tweet, we extracted text, date, permanent link, and metadata. In addition, we collected user information for each of the media outlets, such as the number of followers and tweets posted with the mentioned hashtags. Finally, we extracted the number of likes and retweets generated by each tweet [19].

Data Processing and Content Analysis

As described in our previous studies involving Twitter content analysis, our search tool Tweet Binder [20] allows access to 100% of all public tweets that match the search criteria (query). In contrast, other search engines based on Twitter's free application programming interface (API) can only access a small sample [21,22]. Tweet Binder has its own data collection system that gathers all publicly available tweets on Twitter and retrieves both tweets and user information from Twitter's API. Initially, the search tool scans the public section of Twitter to collect the IDs of tweets that match the search query. Subsequently, a call is made to the Twitter API to retrieve the tweet and user information. It does not directly access Twitter firehose itself.

Figure 1 summarizes the data collection and analysis steps, along with the number of tweets included and excluded in each step. First, we created a codebook based on our research question. There were 2 primary coders (MAA-M and FS). Coders first used a data set of 600 tweets to explore the content, generate codes, and obtain training. Discrepancies were discussed in regular team meetings. Training for coders was provided by the research team members (VPS and ART) experienced in content analysis and codebook development. After training, coders continued coding using the analytic data set. Interrater reliability was examined periodically to prevent rater drift. In a subset of 300 tweets, the interrater reliability averaged 91.7% (275/300) agreement for Spanish tweets and 83.3% (250/300) agreement for English tweets for different categories. The final codebook comprised six categories of HRB: (1) mask-wearing, (2) physical distancing, (3) handwashing, (4) quarantine or confinement, (5) disinfecting objects, and (6) multiple HRBs (any combination of the prior HRB categories). Examples of tweets are listed in Table 1.

Figure 1. Flowchart.

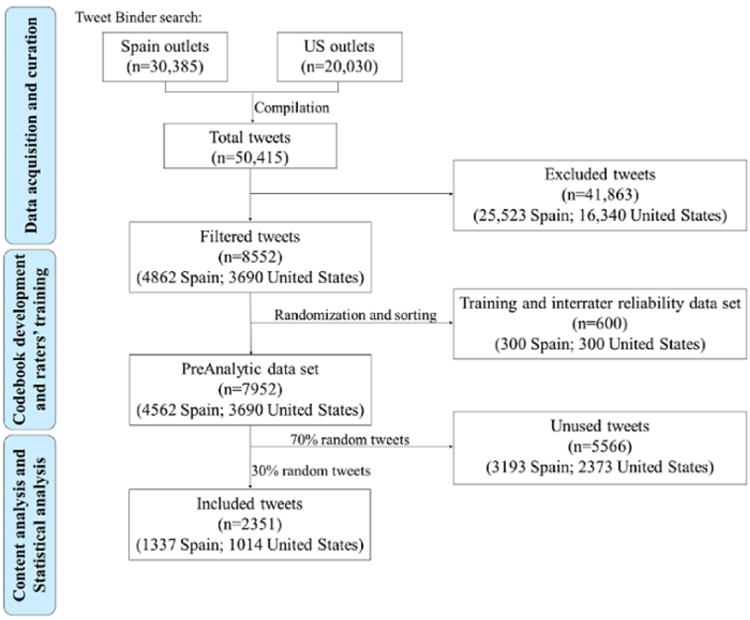


Table 1. Examples of tweets of each health-related behavior (HRB). Tweets that mentioned >1 HRB were included in the Multiple HRBs category. Usernames and personal names were removed. All tweets reported here are in English (tweets in Spanish have been translated).

HRB	Examples of tweets
Quarantine or confinement	• "Twitter is the first major U.S. corporation to strongly encourage its employees to work from home to avoid spreading coronavirus. 'Beginning today, we are strongly encouraging all employees globally to work from home if they're able." • "In R.I., 26 people have quarantined themselves while being monitored for coronavirus." • "Students and staff at a Winnetka school are being told to self-quarantine for 14 days, after a 7th-grader was diagnosed with a "probable" case of COVID-19." • "Olympic rugby player arrested after allegedly breaking Fiji's coronavirus quarantine."
Mask-wearing	• "Globetrotting influencers combatting coronavirus with designer face masks." • "Fashion brands are making face masks, medical gowns for the coronavirus crisis." • "Woman wearing face mask attacked in possible coronavirus hate crime." • "Your hoarding masks could cost me my life—a doctor's view from the coronavirus front lines."
Social distancing	• "DINING AT A DISTANCE: A Swedish couple has opened a 'COVID-safe' restaurant, with one table and one chair located in the middle of a meadow." • "Chicago-area Catholic churches are changing Mass practices to reduce risk of spreading coronavirus. Parishioners are being asked to avoid shaking hands during the Sign of Peace, using holy-water fonts or drinking communion wine." • "Despite state orders to social distance to prevent spread of coronavirus, many funerals in the ultra-Orthodox Jewish communities in Brooklyn continue to be large-scale events, prompting concern from leaders."
Handwashing	• "'Focus on slowing down the spread of COVID-19, the coronavirus. Did I mention: Wash. Your. Hands. Then wash them again,' says epidemiologist Malia Jones." • "The best protection against coronavirus is washing your hands—but you have to do it properly: Wet your hands. Lather, making sure to get soap in all the nooks and crannies. Scrub for 20 seconds. Rinse and dry thoroughly." • "Coronavirus spawns viral TikTok dance about washing your hands."
Disinfecting objects	• "EPA releases list of disinfectant products approved for use against COVID-19 on surfaces-including multiple products from brands such as Clorox and Lysol." • "The Moscow Metro adopts stricter measures for disinfecting trains, as around 6000 metro carriages are being sterilized with UV light and disinfectant. The city is also starting construction on an infectious diseases hospital amid the coronavirus pandemic." • "Public spaces and elderly people's homes sprayed with disinfectant by members of Spanish military as nation fights coronavirus pandemic."
Multiple HRBs	• "Cambridge is sending out 'sound trucks' to remind residents to stay 6 feet apart and wear masks during coronavirus pandemic." • "AHEAD: @[user] joins @[user] live to discuss how she turned her hit song 'Get On Your Feet' into a powerful message about wearing masks and staying home. Plus, why she's bringing attention to minority communities disproportionately impacted by #coronavirus." • "People in many parts of the world are being asked to avoid crowds, limit travel and even work from home to help limit the spread of novel coronavirus, and satellite images suggest they're heeding that advice."

Ethical Considerations

This study was initially reviewed by the Oregon Health & Science University Research Integrity Office and the institutional review board and was determined not to involve human participants. This study was approved to be conducted without continued institutional review board oversight and was compliant with the research ethics principles of the Declaration of Helsinki (7th revision, 2013).

Statistical Analysis

First, we examined the prevalence of tweets with each HRB among tweets with at least one HRB by country, with $P<.05$ indicating a statistically significant difference in the proportion of a specific HRB in Spanish compared with US media sources. All further analyses were stratified by country, with Spain and the United States examined independently. Next, we calculated the measures of central tendency of the number of likes and retweets by HRB type. We assessed the distributions of the outcomes by tweet category and country, including calculating measures of central tendency and generating visualizations, such as histograms. The results showed substantial right skew for both the number of likes and retweets; thus, we presented the median (IQR) rather than mean (SD) and chose a Poisson distribution for our regression models.

Poisson mixed-effects regression models were run both unadjusted and with adjustment for media source number of followers, media source number of tweets, and follow-up time in days between tweet posting and the data collection date. We included media source as a random effect. Results were presented as the estimated counts of likes or retweets for tweets with each HRB compared with either (1) tweets with a different HRB or (2) tweets with no HRB. For ease of interpretation, estimated counts or average marginal effects were used instead of model coefficients. The reported P values indicate the significance of the association between the independent variable and the outcome (number of likes or retweets). Critical values for Bonferroni correction for multiple comparisons were calculated by dividing the α level (.05) by the number of hypotheses [6] and were applied to all results (critical value *P*<.008). All analyses were performed using Stata (version 16; StataCorp).

Results

Distribution of Tweets by HRBs

We collected all the tweets that included a hashtag mentioning COVID-19 (in different ways: #coronavirus, #covid-19, #SARS-CoV2, and #2019-nCoV) and any of the HRBs mentioned previously (including Spanish equivalents). That is, the tweet had to mention COVID and at least one HRB to be collected. With these search criteria, we collected 50,415 tweets but excluded 41,863 because they had nothing to do with any HRB. At that time, it was very common to include these types of hashtags in tweets, even though the content of the tweets had nothing to do with health issues, so we decided to exclude them. Of the remaining 8552 tweets, we randomly selected 600 to design the codebook and train the raters. Finally, 7952 were left, and 30% (2386) were randomly selected for manual analysis.

The number of HRB posts varied considerably among the analyzed mass media accounts (Multimedia Appendix 1). Overall, mass media outlets from Spain showed the largest number of posts. In particular, EFE Noticias, Antena 3, and TVE y La Sexta were the most active media outlets in Spain, whereas CNN, ABC, and New York Post were the most active outlets in the United States. Among the outlets that posted >100 tweets, the New York Post and ABC had the highest proportion of HRB tweets, 78.8% (93/118) and 70.1% (109/154) respectively, whereas TVE and Antena 3 had the lowest proportion, 58.2% (89/153) and 59.4% (95/160) respectively.

As shown in the flowchart in Figure 1, of the 50,415 tweets collected, 2351 were included in the content analysis, and 1479 of them (62.91%) mentioned at least one HRB. As shown in Table 2, the proportion of tweets mentioning HRBs was significantly different between the 2 countries (*P*=.006); 60.51% (809/1337) of all tweets related to COVID-19 posted by media outlets from Spain and 66.07% (670/1014) of tweets posted by US outlets contained at least one HRB (Table 2). In both countries, the distribution of tweets across different categories of HRB was heterogeneous. Quarantine or confinement-related tweets accounted for the highest proportion of tweets in both countries, 48.7% (394/809) in Spain and 49.7% (333/670) in the United States), followed by tweets related to masks and social distancing. In contrast, the least frequent HRB categories were disinfecting objects in Spain, 6.9% (56/809), and handwashing in the United States, 9.1% (61/670). There was no significant difference in the proportion of tweets mentioning each HRB between major Spanish and US media outlets.

Table 2. Descriptive characteristics of the tweets regarding health-related behaviors (HRBs) for the prevention of COVID-19, by category[a] and country.

	Country		*P* value
	Spain	United States	
All COVID-19 tweets, n	1337	1014	N/A[b]
Tweets with 0 HRB, n (%)	528 (39.5)	344 (33.9)	.006
Tweets with ≥1 HRB, n (%)	809 (60.5)	670 (66.1)	N/A
Among COVID-19 tweets with ≥1 HRB, n (%)			
Quarantine or confinement	394 (48.7)	333 (49.7)	.70
Masks	285 (35.2)	227 (33.9)	.59
Social distancing	129 (16)	84 (12.5)	.06
Handwashing	80 (9.9)	61 (9.1)	.61
Disinfecting objects	56 (6.9)	65 (9.7)	.05
Multiple HRBs	98 (12.1)	70 (10.5)	.32

[a]Categories are not mutually exclusive; tweets with multiple HRBs are counted in those categories as well as in the Multiple HRBs category. Subgroup restricted to tweets with ≥1 HRB and >3 days between tweet posts and data collection date.

[b]N/A: not applicable.

Engagement Metrics by HRB and by Country

We investigated engagement with tweets posted by Spanish and US media outlets from social media users by analyzing the number of likes and retweets received. Tweets mentioning at least one HRB received a similar number of likes and retweets as those tweets that did not mention an HRB, as shown in Table 3. In both Spain and the United States, the median number of likes received by each tweet was higher than the median number of retweets. Among tweets with at least one HRB, disinfecting objects had a median (IQR) of 197 (63-486) likes and 90 (33-246) retweets, which constitutes the highest for tweets from US media outlets, and it is twice the number achieved by quarantine or confinement-related tweets (Table 3). Among those posted by Spanish media, mask- and handwashing-related tweets had the highest median number of likes (22 and 21, respectively), whereas social distancing-related tweets had the lowest median number of likes. In contrast, all HRBs had a similar median number of retweets (11 or 12).

Results from mixed Poisson regression analyses in tweets with clustering by media source are presented in Table 4 (Spain) and Table 5 (United States). In adjusted models, tweets posted by mass media outlets from Spain that mentioned social distancing or mentioned disinfecting objects had a significantly lower predicted count of likes compared with tweets mentioning a different HRB (*P*=.02 and *P*=.01, respectively) or tweets related to COVID-19 but not mentioning any HRB at all (*P*=.01 and *P*=.005, respectively). Tweets mentioning multiple HRBs also had a significantly lower count of predicted likes than tweets mentioning just 1 HRB (*P*<.001) or did not mention any HRB (*P*<.001). In respect of retweets, disinfecting objects and tweets mentioning multiple HRBs had a significantly lower predicted number of retweets than tweets mentioning other HRB's or not mentioning any HRB. Other associations that were not significant (*P*<.05 are presented in Table 4.

In regard to tweets posted by US media outlets, in adjusted models those mentioning quarantine or confinement or disinfecting objects had a significantly lower predicted number of likes compared with tweets mentioning a different HRB (*P*<.001), whereas mask- and handwashing-related tweets had a significantly greater predicted number of likes (*P*=.04 and *P*=.02, respectively; Table 5). When compared with tweets related to COVID-19 but not mentioning an HRB, those tweets related to quarantine or confinement or to social distancing had a significantly lower predicted number of likes (*P*=.01 and *P*=.005, respectively). With respect to retweets, quarantine or confinement had a significantly lower predicted number of retweets than tweets mentioning other HRB's (*P*<.001), whereas those mentioning handwashing had a greater probability of being retweeted (*P*=.006). When compared with tweets related to COVID-19 not mentioning an HRB, those tweets related to quarantine or confinement, social distancing, or multiple HRBs had a significantly lower predicted number of retweets (*P*<.001).

Table 3. Distribution of likes and retweets per category, by country.

Category	Spain			United States		
	Value, n	Likes, median (IQR)	Retweets, median (IQR)	Value, n	Likes, median (IQR)	Retweets, median (IQR)
All COVID-19 tweets	1337	19 (8-44)	11 (5-28)	1014	109 (36-316)	54 (18-137)
Tweets with 0 HRB[a]	528	18 (8-41)	12 (5-26)	344	105 (35-294)	57 (19-139)
Tweets with ≥1 HRB	809	20 (8-47)	11 (5-30)	670	110 (37-321)	51 (18-135)
Among COVID-19 tweets with ≥1 HRB						
Quarantine or confinement	394	17 (7-39)	11 (5-28)	333	88 (31-262)	45 (16-102)
Masks	285	22 (10-53)	12 (5-35)	227	127 (41-588)	57 (20-183)
Social distancing	129	16 (9-37)	11 (5-23)	84	171 (44-390)	62 (17-127)
Handwashing	80	21 (8-56)	11 (4-44)	61	118 (58-447)	59 (24-179)
Disinfecting objects	56	18 (10-59)	12 (6-48)	65	197 (63-486)	90 (33-246)
Multiple HRBs	98	13 (7-32)	11 (4-25)	70	191 (58-358)	66 (28-175)

[a]HRB: health-related behavior.

Table 4. Average predicted number of likes by presence or absence of category in Spain. Results presented as predicted mean (95% CI) from adjusted Poisson models with clustering for media source. Adjustment variables include source number of followers, source number of tweets, and days from tweet posting to data collection date.

	Tweet with specific HRB[a] versus tweet with different HRB (n=809)			Tweet with specific HRB versus tweet with no HRB			
	Tweet with a different HRB, mean (95% CI)	Presence of HRB, mean (95% CI)	*P* value[b]	Model, n	Tweet with no HRB, mean (95% CI)	Presence of HRB, mean (95% CI)	*P* value[b]
Likes							
Quarantine or confinement	96.8 (42.5-151.1)	67.3 (51.4-83.1)	.19	922	80.5 (60.1-100.9)	67.3 (51.8-79.6)	.09
Masks	65.4 (51.4-79.3)	115.7 (26.9-204.5)	.16	813	79.7 (59.9-99.5)	119.1 (31.4-206.7)	.18
Social distancing	90.2 (51.2-129.2)	43.5 (29.1-57.8)	.02	657	80.1 (58.5-101.6)	45.3 (32.0-58.6)	.01
Handwashing	85.5 (48.8-122.11)	62.0 (37.1-86.9)	.36	608	82.3 (59.9-104.6)	63.2 (43.2-83.2)	.24
Disinfecting objects	86.5 (51.5-121.4)	44.7 (30.6-58.9)	.01	584	82.3 (60.4-104.3)	49.2 (32.4-65.9)	.005
Multiple HRBs	89.1 (52.8-125.4)	37.7 (28.3-47.1)	<.001	626	80.8 (59.5-102.0)	38.8 (28.2-49.4)	<.001
Retweets							
Quarantine or confinement	60.1 (20.2-100.0)	48.0 (35.8-60.2)	.49	922	53.6 (38.7-68.6)	45.9 (35.8-55.9)	.22
Masks	46.5 (35.7-57.2)	69.4 (3.4-135.4)	.43	813	52.2 (37.9-66.4)	71.5 (5.8-137.3)	.42
Social distancing	57.6 (28.3-86.9)	36.5 (23.4-49.7)	.25	657	52.4 (36.6-68.2)	38.0 (24.9-51.1)	.25
Handwashing	56.1 (29.2-82.9)	41.1 (27.6-54.6)	.36	608	54.3 (38.6-69.9)	42.2 (30.7-53.7)	.20
Disinfecting objects	56.7 (30.8-82.7)	30.3 (18.7-42.0)	.04	584	54.1 (38.7-69.6)	34.5 (20.9-48.2)	.03
Multiple HRBs	57.7 (30.6-84.9)	29.9 (21.2-38.6)	.04	626	53.0 (37.9-68.2)	30.8 (21.3-40.4)	.01

[a]HRB: health-related behavior.

[b]*P* value is for the association (independent variable coefficient) between the presence or absence of a category and the number of likes it received.

Table 5. Average predicted number of likes by presence or absence of category in United States. Results presented as predicted mean (95% CI) from adjusted Poisson models with clustering for media source. Adjustment variables include source number of followers, source number of tweets, and days from tweet posting to data collection date.

	Tweet with specific HRB[a] versus tweet with different HRB (n=669)			Tweet with specific HRB versus tweet with no HRB			
	Tweet with a different HRB, mean (95% CI)	Presence of HRB, mean (95% CI)	*P* value[b]	Model, n	Tweet with no HRB, mean (95% CI)	Presence of HRB, mean (95% CI)	*P* value[b]
Likes							
Quarantine or confinement	709.2 (404.7-1013.7)	331.7 (250.5-414.4)	<.001	677	510.3 (444.9-575.7)	351.3 (259.9-442.7)	.01
Masks	467.1 (288.3-645.9)	620.0 (369.8-870.2)	.04	571	523.4 (439.0-607.8)	590.0 (333.0-847.0)	.66
Social distancing	536.2 (303.7-768.7)	395.3 (258.1-532.4)	.41	428	540.9 (480.5-601.3)	398.2 (314.9-481.5)	.005
Handwashing	448.3 (347.4-551.0)	1154.6 (46.9-2262.4)	.02	405	515.3 (427.5-603.1)	1357.8 (90.2-2625.4)	.06
Disinfecting objects	542.7 (335.5-750.0)	305.8 (250.1-361.5)	<.001	409	539.9 (478.5-601.3)	306.3 (137.3-475.4)	.06
Multiple HRBs	535.4 (319.4-751.3)	364.6 (273.5-455.7)	.20	414	524.0 (458.9-589.2)	393.0 (282.8-503.1)	.08
Retweets							
Quarantine or confinement	263.4 (143.0-383.9)	120.7 (83.5-157.9)	<.001	677	223.1 (197.2-249.1)	129.3 (89.7-169.0)	<.001
Masks	177.8 (102.8-252.8)	212.8 (119.9-305.7)	.26	571	226.6 (192.7-260.4)	201.1 (104.2-298.1)	.68
Social distancing	198.4 (108.4-288.4)	127.8 (94.8-160.8)	.18	428	234.2 (208.2-260.3)	122 (102.7-142.8)	<.001
Handwashing	159.6 (115.8-203.5)	463.0 (30.7-895.2)	.006	405	224.9 (190.6-259.2)	525.4 (60.3-990.5)	.08
Disinfecting objects	192.3 (110.0-274.7)	160.1 (134.5-185.7)	.32	409	233.3 (206.9-259.8)	153.9 (79.9-227.8)	.10
Multiple HRBs	195.8 (111.3-280.2)	134.3 (110.5-158.2)	.12	414	228.3 (200.2-256.4)	141.1 (111.0-171.1)	<.001

[a]HRB: health-related behavior.

[b]*P* value is for the association (independent variable coefficient) between the presence or absence of a category and the number of likes it received.

Discussion

Principal Findings

In this study, we found that major media outlets from Spain and the United States, when posting information on Twitter related to COVID-19, mentioned an HRB in the majority of their tweets, both focusing on quarantine or confinement and masks. Twitter users from both countries showed similar engagement in tweets mentioning an HRB compared with those that did not mention an HRB. Remarkably, the engagement of users following media outlets from Spain was more equally distributed among different HRBs. Furthermore, in Spanish tweets, none of the HRB tweets had a higher probability of being liked or retweeted than others. However, tweets mentioning handwashing or masks had greater probabilities of being liked than tweets mentioning a different HRB or not mentioning any of them when posted from US media outlets. Finally, we observed that media outlets from Spain differed in their Twitter posting patterns quantitatively and qualitatively from US media outlets.

It has previously been shown that entertainment media and social media play a critical role in the behaviors of individuals and have the potential to influence awareness, which is important because adhering to health recommendations is considered a very relevant element for the prevention of COVID-19 infection and overcoming the pandemic [23-25]. Certain health recommendations have been changed since the outbreak of the pandemic [26]. However, majority of health professionals and institutions have promoted some HRBs since the early stages of the pandemic. Social media platforms such as Twitter are increasingly being leveraged by researchers for surveillance and to explore complex social issues, such as perceptions of the public on HRBs, including masks, handwashing, social distancing, and vaccines [27-29]. Furthermore, recognized socially influential agents, such as media outlets or politicians, use Twitter as a dissemination tool for their information, including COVID-19–related news [30]. Social media has become the main source of COVID-19–related information for many people [31]. When media outlets share information, their influence is enormous, particularly in situations such as a pandemic [32,33]. Thus, it is important to analyze the impact of this information on society because exposure to misinformation has been associated with psychological distress, poorer COVID-19 knowledge, and lower adoption of preventive behaviors [34].

Communication Media and HRBs

Our data show that the number of tweets posted by major media outlets from Spain and the United States regarding HRBs is high overall. Interestingly, the number of posts was not homogeneously distributed among the different categories, with quarantine or confinement and masks receiving the highest number of tweets. These results point in the same direction as previous reports, which also found that a great variety of COVID-19–related human behaviors have been discussed on social media, with masks and sheltering in place prevailing over others [35,36]. In fact, Americans initially posted about China, but once COVID-19 became a reality in the United States, their social media posts started to focus on US-centered issues, such as lockdown or stay-at-home recommendations [37]. Nonetheless, those following Twitter accounts of major US media outlets were more interested in handwashing, whereas those following Spanish outlets did not show such a preference.

Several reasons may explain the differences found in the priorities or interests of the media and Twitter users, as well as the differences found between users following media outlets from Spain and the United States. First, it may be in the best interest of media outlets and politicians to focus on issues related to masks (shortage, logistics organization of distribution, legislation, etc) rather than on promoting less controversial matters such as handwashing. Second, the great mediatic and social impact generated by quarantines and confinements in comparison with other measures established to prevent the spread of COVID-19 may explain why this HRB, in particular, is so present in tweets posted by media outlets [38,39]. Third, wide sectors of society, such as political parties in the opposition, may push the media to speak and try to generate buzz around quarantine or confinement and masks rather than promoting healthy habits that are not controversial, such as handwashing. Fourth, many authorities may be especially active in promoting confinement through press news because of its implications. In fact, politicians' announcements related to confinements and quarantines have been disseminated through media outlets on social media accounts. In addition, several studies have suggested that political parties and big companies have a strong influence on the agenda-setting of media outlets and on the information that they distribute. All of these facts may have contributed to our observed prominence of quarantine or confinement and mask tweets, as compared with handwashing, despite the fact that the latter generated more engagement among Twitter users.

In our study, differences were found in the probability of tweets being liked or retweeted according to each HRB. However, it is relevant to note that all HRB tweets achieved a median number of likes and retweets higher than those found in previous articles for tweets posted by US media outlets on diseases with high prevalence and morbidity (such as cancer, Parkinson, depression, and osteoporosis) or on tweets related to other preventive medical measures, such as contraceptives [21,40].

Moreover, it is important to note that these differences in the probabilities of a tweet being liked or retweeted were more pronounced in tweets posted by US media outlets. One possible explanation could be that US media outlets not only have a greater number of followers but also a more international audience. This greater diversity among followers may contribute to greater polarization in their interests. Nevertheless, cultural differences among countries regarding public perceptions and preventive behaviors during the COVID-19 pandemic have been previously described [41]. Furthermore, the relative weight assigned to each HRB by the media outlets, as defined by the percentage of tweets received, was not related to the retweets and likes generated by Twitter users. That is, the HRB that gathered the most attention was handwashing, whereas tweets mentioning multiple HRBs did not generate much attention despite prestigious studies showing strong evidence of each of the HRB mentioned. This finding could be explained by the fact that tweets are short in nature and may be easier to capture user's attention if only focusing on a specific HRB. In addition, Twitter users might have been prone to promoting handwashing, because the benefits of doing so are strongly supported by scientific data and have never been questioned. Furthermore, handwashing, in contrast to masks or quarantine or confinement, has no political connotations, thus allowing Twitter users to share handwashing posts without publicly declaring their political preferences [42]. Moreover, tweets mentioning disinfecting objects despite President Trump's declarations in this regard and all the controversies generated had lower probabilities of being retweeted or liked than tweets mentioning a different HRB.

In addition, we analyzed the proportion of tweets posted by each of the media outlets analyzed, mentioning each HRB. Our data showed that most media outlets, with only 1 exception, focused on quarantine or confinement and masks. Thus, bias in the information related to HRB was not detected. This may indicate that media outlets share common interests. Nonetheless, it is important to highlight that certain differences were observed between them.

The important role that media outlets play in generating popular opinion is well known. Thus, our results suggest that health promotion is not as relevant as generating controversy for media outlets. This is worrying given that measures such as washing hands or maintaining physical distancing are as important as wearing masks or complying with quarantine or confinement, despite the latter being the object of more controversy. The adoption of all HRB is desirable to prevent COVID-19 infection. However, according to our results, controversial measures attract more attention from the media but not from Twitter users.

Limitations

It should be noted that this study has limitations. First, the relevance of Twitter as a social interest marker remains controversial. In addition, the lack of data regarding the geographic location of Twitter users is a limitation in interpreting engagement. Second, the analyzed media outlets do not necessarily reflect the posting pattern of all the press and might have a different set of priorities. Third, our Twitter data were collected according to our selected keywords; thus, we might have missed tweets using different keywords despite discussing the same topic. Fourth, content analysis implies a certain degree of subjectivity. To address this issue, the study comprised a series of steps: initial review, the design of a codebook through a comprehensive process, and the testing of a coder agreement.

Conclusions

To our knowledge, this study is the first to compare media outlet posts related to COVID-19 HRBs from 2 different countries. Media outlets provided more content related to quarantine or confinement and masks, whereas Twitter users, especially those following US media outlets, showed greater engagement with handwashing. Moreover, tweets mentioning multiple HRBs did not result in as much engagement from the Twitter community as those mentioning only 1 HRB. We believe that this finding may have been influenced by the nature of Twitter. Understanding health communication on social media is necessary to design appropriate public health campaigns that might contribute to reducing the rates of contagion and ultimately overcome the COVID-19 pandemic. Future studies could expand the current research by assessing the impact of media publications on the evolution of the pandemic.

Acknowledgments

This work was partially supported by grants from the Fondo de Investigación de la Seguridad Social, the Instituto de Salud Carlos III (PI18/01726; Spain), the Programa de Actividades de I+D de la Comunidad de Madrid en Biomedicina (B2017/BMD-3804; Madrid, Spain), and Helekulani SL.

Disclaimer

The views expressed in this paper are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States Government.

Authors' Contributions

ART, VP-S, and MAA-M were the principal contributors to research design, coordination of data analysis, and manuscript preparation. MAA-M specifically coordinated data acquisition. MAA-M and FS were tweet coders, contributing to codebook development, training, and analysis of the tweets. ERH conducted and reported the statistical analyses. MA-M contributed to the manuscript as a reviewer. ART was the main supervisor in all phases of the project, with special involvement in the study design, interpretation of data, and manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Number of tweets by media outlet, by country. Results presented as n (%) of health-related behavior tweets.
[DOCX File , 28 KB - infodemiology_v3i1e43685_app1.docx]

References

1. Al-Dmour H, Masa'deh R, Salman A, Abuhashesh M, Al-Dmour R. Influence of social media platforms on public health protection against the COVID-19 pandemic via the mediating effects of public health awareness and behavioral changes: integrated model. J Med Internet Res 2020 Aug 19;22(8):e19996. [doi: 10.2196/19996] [Medline: 32750004]
2. 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]
3. Daughton AR, Paul MJ. Identifying protective health behaviors on twitter: observational study of travel advisories and Zika virus. J Med Internet Res 2019 May 13;21(5):e13090 [FREE Full text] [doi: 10.2196/13090] [Medline: 31094347]
4. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
5. Nagar R, Yuan Q, Freifeld CC, Santillana M, Nojima A, Chunara R, et al. A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. J Med Internet Res 2014 Oct 20;16(10):e236 [FREE Full text] [doi: 10.2196/jmir.3416] [Medline: 25331122]
6. Skarpa PE, Garoufallou E. Information seeking behavior and COVID-19 pandemic: a snapshot of young, middle aged and senior individuals in Greece. Int J Med Inform 2021 Jun;150:104465 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104465] [Medline: 33887589]
7. Chen E, Lerman K, Ferrara E. Tracking social media discourse about the COVID-19 pandemic: development of a public coronavirus Twitter data set. JMIR Public Health Surveill 2020 May 29;6(2):e19273 [FREE Full text] [doi: 10.2196/19273] [Medline: 32427106]

8. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194 [FREE Full text] [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
9. Yousuf H, Corbin J, Sweep G, Hofstra M, Scherder E, van Gorp E, et al. Association of a public health campaign about coronavirus disease 2019 promoted by news media and a social influencer with self-reported personal hygiene and physical distancing in the Netherlands. JAMA Netw Open 2020 Jul 01;3(7):e2014323 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.14323] [Medline: 32639569]
10. Fernández-Torres MJ, Almansa-Martínez A, Chamizo-Sánchez R. Infodemic and fake news in Spain during the COVID-19 pandemic. Int J Environ Res Public Health 2021 Feb 12;18(4):1781 [FREE Full text] [doi: 10.3390/ijerph18041781] [Medline: 33673095]
11. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature 2020 Aug;584(7820):262-267. [doi: 10.1038/s41586-020-2404-8] [Medline: 32512578]
12. Lin A, Vittinghoff E, Olgin J, Peyser N, Aung S, Joyce S, et al. Predictors of incident SARS-CoV-2 infections in an international prospective cohort study. BMJ Open 2021 Sep 21;11(9):e052025 [FREE Full text] [doi: 10.1136/bmjopen-2021-052025] [Medline: 34548363]
13. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, et al. Psychological language on Twitter predicts county-level heart disease mortality. Psychol Sci 2015 Feb;26(2):159-169 [FREE Full text] [doi: 10.1177/0956797614557867] [Medline: 25605707]
14. Nguyen QC, Meng H, Li D, Kath S, McCullough M, Paul D, et al. Social media indicators of the food environment and state health outcomes. Public Health 2017 Jul;148:120-128 [FREE Full text] [doi: 10.1016/j.puhe.2017.03.013] [Medline: 28478354]
15. Nguyen QC, Brunisholz KD, Yu W, McCullough M, Hanson HA, Litchman ML, et al. Twitter-derived neighborhood characteristics associated with obesity and diabetes. Sci Rep 2017 Nov 27;7(1):16425 [FREE Full text] [doi: 10.1038/s41598-017-16573-1] [Medline: 29180792]
16. Gore RJ, Diallo S, Padilla J. You are what you tweet: connecting the geographic variation in America's obesity rate to Twitter content. PLoS One 2015;10(9):e0133505 [FREE Full text] [doi: 10.1371/journal.pone.0133505] [Medline: 26332588]
17. Ireland ME, Chen Q, Schwartz HA, Ungar LH, Albarracin D. Action tweets linked to reduced county-level HIV prevalence in the United States: online messages and structural determinants. AIDS Behav 2016 Jun;20(6):1256-1264 [FREE Full text] [doi: 10.1007/s10461-015-1252-2] [Medline: 26650382]
18. Chandrasekaran R, Mehta V, Valkunde T, Moustakas E. Topics, trends, and sentiments of tweets about the COVID-19 pandemic: temporal infoveillance study. J Med Internet Res 2020 Oct 23;22(10):e22624 [FREE Full text] [doi: 10.2196/22624] [Medline: 33006937]
19. Viguria I, Alvarez-Mon MA, Llavero-Valero M, Asunsolo Del Barco A, Ortuño F, Alvarez-Mon M. Eating disorder awareness campaigns: thematic and quantitative analysis using Twitter. J Med Internet Res 2020 Jul 14;22(7):e17626 [FREE Full text] [doi: 10.2196/17626] [Medline: 32673225]
20. Social media management with Tweet Binder. TweetBinder blog. URL: https://www.tweetbinder.com/blog/ [accessed 2023-07-14]
21. Alvarez-Mon M, Asunsolo Del Barco A, Lahera G, Quintero J, Ferre F, Pereira-Sanchez V, et al. Increasing interest of mass communication media and the general public in the distribution of tweets about mental disorders: observational study. J Med Internet Res 2018 May 28;20(5):e205 [FREE Full text] [doi: 10.2196/jmir.9582] [Medline: 29807880]
22. Pereira-Sanchez V, Alvarez-Mon M, Asunsolo Del Barco A, Alvarez-Mon M, Teo A. Exploring the extent of the Hikikomori phenomenon on Twitter: mixed methods study of western language tweets. J Med Internet Res 2019 May 29;21(5):e14167 [FREE Full text] [doi: 10.2196/14167] [Medline: 31144665]
23. Merchant RM, Lurie N. Social media and emergency preparedness in response to novel coronavirus. JAMA 2020 May 26;323(20):2011-2012. [doi: 10.1001/jama.2020.4469] [Medline: 32202611]
24. Yousuf H, Narula J, Zwetsloot PP, Hofstra M, de Levita A, Scherder E, et al. Using entertainment to improve lifestyles and health. Lancet 2019 Jul 13;394(10193):119-120. [doi: 10.1016/S0140-6736(19)30250-8] [Medline: 31305250]
25. Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA 2020 May 19;323(19):1915-1923 [FREE Full text] [doi: 10.1001/jama.2020.6130] [Medline: 32275295]
26. Feng S, Shen C, Xia N, Song W, Fan M, Cowling BJ. Rational use of face masks in the COVID-19 pandemic. Lancet Respir Med 2020 May;8(5):434-436 [FREE Full text] [doi: 10.1016/S2213-2600(20)30134-X] [Medline: 32203710]
27. Kwok SW, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: machine learning analysis. J Med Internet Res 2021 May 19;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: 33886492]
28. 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]

29. Zhou X, Song Y, Jiang H, Wang Q, Qu Z, Zhou X, et al. Comparison of public responses to containment measures during the initial outbreak and resurgence of COVID-19 in China: infodemiology study. J Med Internet Res 2021 Apr 05;23(4):e26518 [FREE Full text] [doi: 10.2196/26518] [Medline: 33750739]
30. Ahmed W, Vidal-Alaball J, Lopez Segui F, Moreno-Sánchez PA. A social network analysis of tweets related to masks during the COVID-19 pandemic. Int J Environ Res Public Health 2020 Nov 07;17(21):8235 [FREE Full text] [doi: 10.3390/ijerph17218235] [Medline: 33171843]
31. Zhao Y, Xu S, Wang L, Huang Y, Xu Y, Xu Y, et al. Concerns about information regarding COVID-19 on the internet: cross-sectional study. J Med Internet Res 2020 Nov 09;22(11):e20487 [FREE Full text] [doi: 10.2196/20487] [Medline: 33095740]
32. Liu PL. COVID-19 information seeking on digital media and preventive behaviors: the mediation role of worry. Cyberpsychol Behav Soc Netw 2020 Oct;23(10):677-682. [doi: 10.1089/cyber.2020.0250] [Medline: 32498549]
33. Shimizu K. 2019-nCoV, fake news, and racism. Lancet 2020 Feb 29;395(10225):685-686 [FREE Full text] [doi: 10.1016/S0140-6736(20)30357-3] [Medline: 32059801]
34. Lee JJ, Kang KA, Wang MP, Zhao SZ, Wong JY, O'Connor S, et al. Associations between COVID-19 misinformation exposure and belief with COVID-19 knowledge and preventive behaviors: cross-sectional online study. J Med Internet Res 2020 Nov 13;22(11):e22205 [FREE Full text] [doi: 10.2196/22205] [Medline: 33048825]
35. Daughton AR, Shelley CD, Barnard M, Gerts D, Watson Ross C, Crooker I, et al. Mining and validating social media data for COVID-19-related human behaviors between January and July 2020: infodemiology study. J Med Internet Res 2021 May 25;23(5):e27059 [FREE Full text] [doi: 10.2196/27059] [Medline: 33882015]
36. 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]
37. Valdez D, Ten Thij M, Bathina K, Rutter LA, Bollen J. Social media insights into US mental health during the COVID-19 pandemic: longitudinal analysis of twitter data. J Med Internet Res 2020 Dec 14;22(12):e21418 [FREE Full text] [doi: 10.2196/21418] [Medline: 33284783]
38. Xue J, Chen J, Chen C, Zheng C, Li S, Zhu T. Public discourse and sentiment during the COVID 19 pandemic: using Latent Dirichlet Allocation for topic modeling on Twitter. PLoS One 2020;15(9):e0239441 [FREE Full text] [doi: 10.1371/journal.pone.0239441] [Medline: 32976519]
39. Xue J, Chen J, Hu R, Chen C, Zheng C, Su Y, et al. Twitter discussions and emotions about the COVID-19 pandemic: machine learning approach. J Med Internet Res 2020 Nov 25;22(11):e20550 [FREE Full text] [doi: 10.2196/20550] [Medline: 33119535]
40. Alvarez-Mon MA, Donat-Vargas C, Llavero-Valero M, Gea A, Alvarez-Mon M, Martinez-Gonzalez MA, et al. Analysis of media outlets on women's health: thematic and quantitative analyses using Twitter. Front Public Health 2021;9:644284 [FREE Full text] [doi: 10.3389/fpubh.2021.644284] [Medline: 34136450]
41. Bowman L, Kwok KO, Redd R, Yi Y, Ward H, Wei WI, et al. Comparing public perceptions and preventive behaviors during the early phase of the COVID-19 pandemic in Hong Kong and the United Kingdom: cross-sectional survey study. J Med Internet Res 2021 Mar 08;23(3):e23231 [FREE Full text] [doi: 10.2196/23231] [Medline: 33539309]
42. Lang J, Erickson WW, Jing-Schmidt Z. #MaskOn! #MaskOff! digital polarization of mask-wearing in the United States during COVID-19. PLoS One 2021;16(4):e0250817 [FREE Full text] [doi: 10.1371/journal.pone.0250817] [Medline: 33909669]

Abbreviations

API: application programming interface
HRB: health-related behavior

Edited by T Mackey; submitted 20.10.22; peer-reviewed by N Sigalo, C Chen; comments to author 10.01.23; revised version received 17.02.23; accepted 30.05.23; published 22.08.23.

Please cite as:
Alvarez-Mon MA, Pereira-Sanchez V, Hooker ER, Sanchez F, Alvarez-Mon M, Teo AR
Content and User Engagement of Health-Related Behavior Tweets Posted by Mass Media Outlets From Spain and the United States Early in the COVID-19 Pandemic: Observational Infodemiology Study
JMIR Infodemiology 2023;3:e43685
URL: https://infodemiology.jmir.org/2023/1/e43685
doi:10.2196/43685
PMID:37347948

Original Paper

Sharing Reliable COVID-19 Information and Countering Misinformation: In-Depth Interviews With Information Advocates

Alexis M Koskan[1*], MA, PhD; Shalini Sivanandam[1*], MD, PhD; Kristy Roschke[2*], PhD; Jonathan Irby[1*], MSc; Deborah L Helitzer[1*], SCD; Bradley Doebbeling[1], MSc, MD

[1]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

[2]Walter Cronkite School of Journalism and Mass Communication, Arizona State University, Phoenix, AZ, United States

[*]these authors contributed equally

Corresponding Author:
Alexis M Koskan, MA, PhD
College of Health Solutions
Arizona State University
425 N. 5th Street
Phoenix, AZ, 85004
United States
Phone: 1 480 884 2533 ext 42533
Fax: 1 602 827 2253
Email: alexis.koskan@asu.edu

Abstract

Background: The rampant spread of misinformation about COVID-19 has been linked to a lower uptake of preventive behaviors such as vaccination. Some individuals, however, have been able to resist believing in COVID-19 misinformation. Further, some have acted as information advocates, spreading accurate information and combating misinformation about the pandemic.

Objective: This work explores highly knowledgeable information advocates' perspectives, behaviors, and information-related practices.

Methods: To identify participants for this study, we used outcomes of survey research of a national sample of 1498 adults to find individuals who scored a perfect or near-perfect score on COVID-19 knowledge questions and who also self-reported actively sharing or responding to news information within the past week. Among this subsample, we selected a diverse sample of 25 individuals to participate in a 1-time, phone-based, semistructured interview. Interviews were recorded and transcribed, and the team conducted an inductive thematic analysis.

Results: Participants reported trusting in science, data-driven sources, public health, medical experts, and organizations. They had mixed levels of trust in various social media sites to find reliable health information, noting distrust in particular sites such as Facebook (Meta Platforms) and more trust in specific accounts on Twitter (X Corp) and Reddit (Advance Publications). They reported relying on multiple sources of information to find facts instead of depending on their intuition and emotions to inform their perspectives about COVID-19. Participants determined the credibility of information by cross-referencing it, identifying information sources and their potential biases, clarifying information they were unclear about with health care providers, and using fact-checking sites to verify information. Most participants reported ignoring misinformation. Others, however, responded to misinformation by flagging, reporting, and responding to it on social media sites. Some described feeling more comfortable responding to misinformation in person than online. Participants' responses to misinformation posted on the internet depended on various factors, including their relationship to the individual posting the misinformation, their level of outrage in response to it, and how dangerous they perceived it could be if others acted on such information.

Conclusions: This research illustrates how well-informed US adults assess the credibility of COVID-19 information, how they share it, and how they respond to misinformation. It illustrates web-based and offline information practices and describes how the role of interpersonal relationships contributes to their preferences for acting on such information. Implications of our findings could help inform future training in health information literacy, interpersonal information advocacy, and organizational information advocacy. It is critical to continue working to share reliable health information and debunk misinformation, particularly since this information informs health behaviors.

(JMIR Infodemiology 2023;3:e47677) doi:10.2196/47677

KEYWORDS

COVID-19; coronavirus; pandemic; infodemic; misinformation; social media; qualitative research; public health; health communication

Introduction

In the United States, the COVID-19 pandemic and the resulting "infodemic" that ensued highlighted the role of the rampant spread of misinformation and its impact on health-related decisions [1]. Survey research conducted in November 2021 by Kaiser Family Foundation found that nearly 80% of respondents believed a common myth about COVID-19 or COVID-19 vaccines (eg, the vaccine causes infertility and the government exaggerated COVID-19 deaths) or were unsure about the accuracy of such myths [2]. Widespread misinformation, in turn, impacted COVID-19 prevention behaviors. For example, individuals exposed to misinformation about COVID-19 vaccines were more likely to reject the evidence-based recommendation of receiving the vaccine [3].

In understanding the spread of misinformation, it is critical to explore which information sources people trust and distrust and which sources they use to find reliable health information. Before the COVID-19 pandemic, people generally trusted medical and scientific sources for their health information. However, Pew Research recently published a report illustrating the waning trust in various authorities, including scientists and medical professionals, with more adults reporting lower levels of trust since the beginning of the pandemic [4]. At a closer look, the levels of trust in science differed based on political party affiliation, with Republicans reporting reduced trust while Democrats reported the same level of trust in scientific authorities [4]. Further, believing COVID-19 misinformation was influenced by which national news media individuals regularly watch or read, with more individuals believing misinformation when they watched or read conservative network news (eg, Fox News and Newsmax) [2].

It is critical to explore how some individuals found correct information about the pandemic at a time when misinformation was widespread. Such strategies can be used for future training in information and media literacy. Additionally, assessing how these individuals with accurate information acted as information advocates, sharing reliable COVID-19 information and combating or debunking misinformation, is crucial. Information advocates are expected to have uncovered effective practices to avoid COVID-19 misinformation and share reliable COVID-19 information. Therefore, the purpose of this study was to explore well-informed (about COVID-19) information advocates' health information sources and behaviors, including their practices of disseminating accurate, trustworthy COVID-19 information and combating misinformation about the pandemic.

Methods

Overview

In order to recruit our sample, we selected individuals who completed our previous internet-based survey, which was hosted on QuestionPro (Survey Analytics LLC) and disseminated via an internet-based research panel in September 2022, about COVID-19 and scored a perfect or near-perfect score on COVID-19 knowledge questions and who actively shared or responded to news information within the past week. The 18 knowledge-based survey items about COVID-19 included 8 questions taken from the Kaiser Family Foundation's survey, which assessed adults' beliefs in common COVID-19 myths [2]. Additionally, 1 team member (an infectious disease epidemiologist) created 20 questions about the safety and effectiveness of COVID-19 prevention and control strategies. A team consisting of an epidemiologist, an emergency medicine physician, a physician specializing in infectious diseases, and a public health professor reviewed these 20 survey items, and to reduce the burden on the participants, recommended including 10 of these questions. Multimedia Appendix 1 lists the final 18 COVID-19 knowledge questions. Further, to identify individuals who actively share information, we included 2 questions that asked participants about how they shared or responded to news information within the past week.

A total of 1498 individuals completed the survey, of whom 603 (40.3%) scored 18/18 correctly, and 261 (17.4%) correctly answered 17/18 COVID-19 knowledge questions. Of these 864 individuals, 765 (88.5%) reported having acted on (eg, sharing, liking, discussing, and debunking) news information within the past week, of whom 486 (63.5%) provided contact information for a follow-up interview. We selected a diverse sample of these individuals to complete interviews for this study and conducted all interviews in October 2022.

We created and pilot-tested a semistructured interview guide with 2 COVID-19 experts from the aforementioned team of scientific and medical experts (see Multimedia Appendix 2 for the interview guide). We messaged participants within 1 month of their survey completion to invite them to participate in a 1-time interview and schedule it. The first author texted all potential participants who had scheduled an interview and reminded them of the upcoming phone interview using Google Voice (Google LLC). On the day of the interview, she called the participants, described the study, obtained their informed consent, and conducted the interview. She emailed all interview participants a US $30 Amazon e-gift card as compensation for their time and shared experiences. All interviews were recorded, and a professional transcription service transcribed the audio files.

Using an inductive approach, 1 study team member read all transcripts and served as the primary codebook creator and editor, expanding and merging codes as needed. To enhance the rigor of the qualitative analysis, she trained another study team member to coconduct the thematic analysis [5]. They met weekly to discuss coding guide definitions, update the coding guide as needed, and reach a consensus on how they coded interview data. They first coded 1 transcript together to test the coding guide and ensure coding similarities. Then, they coded 3 transcripts separately before meeting and resolving coding discrepancies and making further clarifications to the coding guide (expanding code definitions, merging codes, etc). They coded 2 additional transcripts separately before reaching a consensus on coding and splitting the remaining transcripts to hand-code and enter into ATLAS.ti software (ATLAS.ti Scientific Software Development GmbH). They aggregated the ATLAS.ti output, synthesized their findings, and summarized the results by code. Finally, they created a comprehensive summary of the findings, identifying quotes that best illustrated emerging interview themes and subthemes.

Ethics Approval

The Arizona State University Institutional Review Board approved this research study (STUDY00015977).

Results

Overview

The in-depth interview participants represented a diverse sample of 25 US adults of various ages, races, communities (eg, rural, suburban, and urban), and political affiliations. It is important to note that all participants reported vaccinating against COVID-19, although this was not part of the inclusion criteria for the study. See Table 1 for participants' demographic information.

In their interviews, information advocates described the sources of information they trusted for COVID-19 information, the sources they did not trust, their processes for determining information credibility, responses to misinformation, and how they shared reliable COVID-19 information. See Table 2 for themes and subthemes that emerged from the qualitative data and representative quotes that illustrate these findings.

Table 1. Demographic information of interview participants (N=25).

Interview participant demographics	Values
Age (years), mean (SD)	46.1 (13.5)
Sex, n (%)	
Female	10 (40)
Male	15 (60)
Ethnicity, n (%)	
Hispanic	4 (16)
Non-Hispanic	21 (84)
Race or ethnicity, n (%)	
White	17 (68)
Hispanic	3 (12)
Black	2 (8)
Asian	3 (12)
Education, n (%)	
Some high school	1 (4)
Some college	3 (12)
Trade or tech school	2 (8)
College graduate	10 (40)
Some postgraduate work	4 (16)
Postgraduate degree	5 (20)
Employment, n (%)	
Full-time	12 (48)
Part-time	3 (12)
Retired	3 (12)
Not employed	6 (24)
Income (US $), n (%)	
$0-$25,999	4 (16)
$26,000-$51,999	6 (24)
$52,000-$75,000	2 (8)
$75,000-$100,000	6 (24)
$100,000-$200,000	6 (24)
≥$200,000	1(4)
Political affiliation, n (%)	
Democrat	14(56)
Republican	3 (12)
Moderate	1 (4)
Independent	7 (28)
Community, n (%)	
Rural	5 (20)
Town	4 (16)
Suburb	8 (32)
Large city	8 (32)
Insurance status, n (%)	

Interview participant demographics	Values
Insured	24 (96)
PCP[a] status, n (%)	
Has a PCP	24 (96)
COVID-19 vaccines received, n (%)	
Initial doses and booster dose	21 (84)
1+ doses of a COVID-19 vaccine (initial doses only)	4 (16)
Previous COVID-19 diagnosis, n (%)	
Yes	8 (32)
No	17 (68)

[a]PCP: primary care provider.

Table 2. Qualitative responses illustrating the study findings.

	Example quote
Trusted information source	
Neutral information sources	Yeah, not Fox News or MSNBC. Just something neutral. (Participant 23)
News media	Well, even though I don't own a TV, I have a laptop and what I do is go to YouTube and I'll watch certain news programs, like PBS NewsHour. I'll watch from the three major networks CBS, NBC, and ABC, their nightly news broadcasts on YouTube. Also, on my phone, I look at the Apple News app, and I have programmed in there what periodicals I want to look at. So, it might be the New York Times I have on there, the Washington Post, maybe USA Today's on there. (Participant 21)
State and local news (as opposed to national news)	I actually found the New York State Governor's press conferences very informative. I also used the news for my county. I live in upstate New York, [and I look at] their public health Facebook accounts just to keep track of the local trends and statistics. They were very good at updating those regularly. I'm so far away from New York City that we had a totally different wave and a totally different behavior of COVID up here. So, if you were listening just to New York City and the trauma they had down there, you would have a totally different view of what it was like up here.... There were a lot of restrictions put on upstate New York when there didn't need to be. And many things happened in New York City that didn't apply to upstate New York. (Participant 18)
Monitored sources	Whenever it first started, I originally started reading about it on Reddit. And as it started spreading to the U.S. Then I would follow it on a Reddit Megathread that always had updated information. (Participant 7)
Health care provider	Before the shots came out, I talked with doctors who were struggling to find how to fix it. I watched the news sometimes, but mostly, I listened to the doctors who were helping my dad. That was really about it. Just listening to the doctors and the pharmacist. (Participant 14)
Untrusted sources	
Pseudoscience	Sometimes you'll see a really sketchy URL in the search results that seems hokey or not really science-based, more like a holistic medicine or like the woo-woo stuff that you can pretty much tell from the link and the URL that it's pseudoscience. (Participant 5)
Social media	All these platforms that provide information, such as social media, do a very poor job of giving you proper information. And they're so fantastic at throwing you bad information without any censorship or warning. We seem to have a culture that promotes misinformation based on how viral it [the information] could be or how emotionally appealing it can be. When it comes to information, if you're going into emotions, first, you have to be drawn to something that correlates more to your original thought process than the actual truth of the situation. (Participant 20)
News media	Take everything with a grain of salt when you're watching the news. Don't take in everything at face value, I guess. (Participant 14)
Deciding if a source is trustworthy	
Data-driven sources	Yeah. I mean, I feel like certain sources are more data-based. They're less speculative. They run more on actual research and science and data. I know that the Mayo Clinic is like that, and they cite research and specific studies. Things that make them more vetted and reliable as a source of information. The CDC, I know, is data-based and research-based, so I just feel like they're more credible. (Participant 5)
Trust in scientific professionals	These people spend their entire lives - doctors, scientists, and researchers - these people spend their whole lives trying to better humanity, protect us from diseases, and improve our wellness. You need to extend some trust to these people. (Participant 2)
Assess information and sources for bias and a hidden agenda	Lots of times misinformation will be tied to some third-party websites or certain websites funded by certain think tanks or certain groups that are very political-leaning one way. So, I can usually get an indication of where this information comes from and what kind of political agenda it has. (Participant 19)
	Yeah, I would say get the actual source of the information. Most information has quotes compiled from people which are usually compiled from other sources. So, I would say first, find the source of the information that you're receiving. And I would say secondly, find out who is funding the people getting the information and what their intentions may be. So, I think you know in the short term, follow the money. Follow the information and follow the money. Get a good idea as to what initiative has brought you that information. (Participant 20)
Check the information with multiple sources	Check for multiple sources to get a similar result, a similar answer to what you're looking for. Check multiple sources because if it's out there in one source, it doesn't mean it's [the information] true. So, I would just tell them to check. Like if you're unsure, check, and then if you find it, check again because that first one could be an offshoot, and then just maybe three or four [sources]. And think after three or four [sources], you probably have a good idea that it's well-reported and probably legitimate. (Participant 22)
Fact-checking software	I would say if you find something or if you hear something, utilize one of the fact-checking sites. This is what I was doing early on, too. You know, the fact-checking software, like Snopes - just to kind of see if something is true or false. (Participant 8)

	Example quote
Responses to misinformation	
Ignore misinformation	If I see the headline and it's something that I think is absolutely incorrect, I already know the facts, so I don't think I bother with it. (Participant 12)
	I know I've seen videos on YouTube in my feed that are undoubtedly misinformation or lies, and I try to ignore them. I don't have any process for reporting or logging them or anything like that. I try to just filter it out and go about my business. (Participant 5)
Correct misinformation	Well, I hear the information, like okay. I mean, I'm trying to see it from their point of view because I'm not sure I'm just being speculative from my point of view. Like, are they just hesitant? Are they afraid about the outcome of it? And just try to project their opinion and try to spin it in their own sort of way that fits their narrative. But I don't really do anything. I just think, "Okay, that's your opinion." I don't leave comments. I don't "like" it. I don't "dislike" it [referring to the action of responding to information with emojis]. I just move on. (Participant 24)
Action depends on the relationship to the person posting misinformation	My head goes to social media, where I see a lot of misinformation. It depends on the person as to whether I confront them, like if I was close enough to them. But I would try to back up what I am saying if I were to confront somebody about it. But otherwise, I would just ignore the articles and stuff people posted. (Participant 25)
	I don't know whether there are a lot of people like me. I just kind of like back away. I don't want to deal with this. Or, I don't want to potentially ruin a friendship or alienate a family member. (Participant 11)
	So, if it's a friend or a family member, then, of course, I'll tell them [that their information is inaccurate]. But let me rephrase that. If it's a friend or a family member on my dad's side, I'll make sure to correct them about it. If it's a family member on my mom's side, they are absolute loons. My dad and I tried to talk to them about it, but they just absolutely would not hear it. They were these people who thought COVID was made up to influence the election. They also thought 5G networks caused it, even though 5G towers don't even exist where we live. There was just so much nonsense, of the fake information of where I currently lived when the pandemic started, that it was just insane. There was just no talking to them. (Participant 15)
Flagging or reporting misinformation	I believe that there's a lot of misinformation, but what I've done is I've sent in like a flag or something to that effect or tried to come up - not come up, but present a fact that - an actual fact versus whatever the misinformation is. (Participant 3)
Blocking misinformation	If it's online, it depends on how badly it upsets me. Sometimes if I'm on Twitter, I will tweet it and say, "Yeah, no, that's not right. Here's better information." If it's just something that I thought I should follow, then I will immediately unfollow and block [the account that posted misinformation] and move on. I find there's a lot of shouting into the void on Twitter, and I'm not as participatory on Facebook. (Participant 18)
Regret over confronting misinformation	I mean, I had a few experiences where people who were sharing misinformation were really defensive about their bad information. There was one guy who, I forget when this was, maybe halfway through, a little bit before halfway through. He posted on Facebook, and he said, "This is nothing but the flu." And I put a comment saying, "That's reckless. You could be hurting people by putting this kind of stuff out there. You should be more careful about what you're saying." And then, this person started sending me private messages, demanding an apology for insulting him. Yeah, I kind of think in a lot of cases, it's not worth messing with crazy. (Participant 11)
Sharing reliable health information	
Direct forms of interpersonal communication	I basically would tell them, like my mom. And some calls and texts to my friends. If it was a person that I don't know, I would text them. The people I do know, I would tell them in person. And I would also text and email them. But primarily, it was a phone call or a text. (Participant 13)
	I can tell them what I know. It's about all I like to do with that, and I guess I judge by the tone of the conversation if they want to hear it or not. (Participant 6)
	It's usually just talking with neighbors and friends, just talking about it. (Participant 17)
Social media posts	I usually post it on my Facebook page or my Twitter page. And if it's like really good, I'll pin it to the top. So, when the county started detecting COVID in wastewater reports, I thought that was really interesting. They could predict what they would see in a couple of weeks in the hospitals [based on what they detected in wastewater]. So, I thought that was really interesting. (Participant 22)
Resharing information	Online, I would retweet tweets or post on Facebook things that ultimately found reliable. And things like, well, these are the statistics; this is the delta strain; this is what we're trying to deal with right now. Things would have to meet certain criteria for me to retweet it or post it on Facebook. But I feel like some of the people that I talked to aren't as savvy when it comes to scientific literacy. And I don't want to accidentally push them the wrong way. So, I have to have a level of care when I try to promote it [information]. (Participant 17)

Trusted Information Sources

Participants trusted numerous science- and data-driven sources for their COVID-19 information, including scientific professionals such as physicians and pharmacists. Most often, participants cited trusting specific scientific organizations such as the US Centers for Disease Control and Prevention (CDC), the World Health Organization (WHO), and Johns Hopkins for their up-to-date information on the pandemic. Some used these sources to cross-verify information from other sources.

An equal number of participants reported relying on national and local news for their up-to-date information about COVID-19. Among those who reported trusting national news, they received this news information from mainstream television news, newspapers, or articles on news aggregator apps and believed that these news sources also allowed them to understand the global and national scope of the pandemic. Others described trusting state and local governments and health departments to provide the latest updates about COVID-19 and COVID-19 vaccines. Some described information shared on national news as misleading, part of a political agenda, or not representative of what was happening in their local communities. Especially as mandates on social distancing and masks varied by county, some reported relying more heavily on local news.

Participants who trusted social media platforms for their information stated that they followed credible sources or accounts on these sites. For example, many participants followed medical or health professionals' accounts on Twitter. They believed they could differentiate accurate from inaccurate information and also trusted these specific experts' information. Another commonly used social media site was Reddit. Participants described using the Reddit news feed for receiving COVID-19 headline news and trusted specific pages on Reddit (subreddits) whose users engaged in monitoring practices for information posted on the page.

Untrusted Sources

Participants described a lack of trust in information shared on some social media sites, particularly Facebook. They saw this site as one where people more often shared opinions and emotionally-charged information instead of facts or fact-based articles. Further, participants noted how misinformation spreads quickly on social media sites since users, many of whom are not health or science experts, can post or share information. These nonexperts propagated news from unverified sources, particularly misinformation that appealed to people's emotional and fear-based states.

Deciding Whether a Source Is Trustworthy

Participants described their processes for deciding whether certain information is credible and trustworthy. They discussed the importance of examining the sources who posted information about COVID-19. Participants also noted their trust in science and individuals trained in scientific disciplines, particularly since they had dedicated their profession and lives to health and science.

When asked what made a source reliable or trustworthy, participants described using data-driven sources that provided statistical information on COVID-19 incidence, prevalence, infection rate, hospitalizations, and mortality. When they questioned information about the pandemic, participants described examining the source for their potential bias or for a hidden agenda. Some participants described how, upon hearing certain claims, they would examine the source to identify a political agenda or a potential financial gain for posting such information.

Most participants described the need to verify information when they were uncertain of the information's trustworthiness. Often, they cross-referenced the information to either a trustworthy source or information provided by multiple sources. They described critically appraising information that did not align with what other sources were saying. Other times, they described verifying the information using fact-checking software. Participants also described verifying health information with their health care providers, given their expertise in health and medicine.

Some participants assessed whether the information they were exposed to on social media sounded plausible, using common sense to parse fact from fiction. They often provided specific examples of COVID-19 myths, such as chips implanted into individuals who receive COVID-19 vaccines (and tracked via the 5G network) and drinking bleach to help treat COVID-19. These participants recommended critical thinking as a means to reject health misinformation.

Responses to Misinformation

The most commonly reported response to seeing or hearing misinformation was to ignore it. Participants who reported seeing misinformation on social media or web-based stories would scroll past the information, with some elaborating on how misinformation was a distraction and not worth exploring. Others explained how, out of curiosity, they were willing to listen to or read more of the misinformation without correcting the source. These individuals believed that correcting the misinformation was a poor use of their time, particularly if they sensed that the other party was adamant about their beliefs. Others described how the context of the relationship to the person posting or sharing misinformation would influence whether or not they responded. Some preferred ignoring information rather than confronting the individuals propagating misinformation to avoid online controversy and potentially ruin relationships in their social network.

Occasionally participants reported flagging and reporting web-based misinformation. Some participants blocked misinformation spread via social media and felt it was not helpful to address it, particularly on social media sites such as Facebook and Twitter. Others would alert the other person sharing misinformation that their information was inaccurate, posting the correct information as an online response. However, their willingness to correct misinformation depended on the type and closeness of the relationship to the person sharing misinformation. For example, some individuals felt more comfortable correcting misinformation with people with whom they shared a closer relationship or who shared similar beliefs. Others, regardless of the response they received or the context of their relationship, reported having confronted misinformation. They believed that the misinformation, especially if viewed by others, could negatively impact others' health.

Sharing Reliable Health Information

More participants were comfortable sharing reliable health information through in-person discussions or other direct forms of communication such as texting. Some noted elements of the in-person communication (eg, nonverbal communication) that influenced the amount of information they would share, whether they provided any clarification, and the extent to which they would discuss the pandemic.

Others created social media posts to share reliable and interesting COVID-19 information. Few participated in online groups, individuals who self-selected to be in a forum due to similar interests (eg, gaming), health conditions (eg, immunocompromised), or geographic community. Often, participants described vetting information to make sure it was correct and coming from a reliable source before sharing or resharing posts on social media or web-based news articles.

Discussion

Principal Findings

This research moved beyond past research that described the content of misinformation and characteristics of people who believe in misinformation [6,7] to examine the information-related practices among well-informed individuals who also act as information advocates. Findings describe not only the information sources that information advocates trust and distrust to find reliable health information but also the various ways they reacted to misinformation. Results illustrate how well-informed information advocates shared reliable information among members of their social network, both through direct interpersonal communication and on the internet, in more public forums. Results also describe the role of interpersonal relationships and the closeness of social ties when considering whether to respond to misinformation.

Information advocates discussed their trust in science, medicine, physicians, and public health organizations. They neither rejected science nor engaged in conspiratorial thinking, which are the traits of individuals who believe in health misinformation [8]. Instead, similar to past research, study participants preferred receiving COVID-19 information from individuals and organizations who had or represented relevant scientific training or background and whom they believed lacked a hidden agenda for sharing information about the pandemic [9]. It is critical to capitalize on these trusted relationships by offering training for health care providers to, at the interpersonal level, address patients' misinformation. Further, at a community level, such experts can help organizations (including community-based organizations) develop and deploy health messaging [10].

Participants expressed mixed reviews about finding trustworthy information on social media. A few participants described trusting information on specific Twitter accounts and Reddit pages. These pages may have administrators, moderators, or followers who audit and regulate content and act against misinformation or individuals posting it. On the other hand, other information advocates noted their lack of trust in social media sites, particularly Facebook, a site where they believed that opinions and anecdotes, not facts, spread rapidly. Interestingly, a study conducted by Yang et al [11] found surges of low-credibility content posted on both Facebook and Twitter, with Facebook having a greater volume of low-credibility information in January 2020 and on Twitter between April and October 31, 2020, with Twitter posting more misinformation, overall. To combat misinformation, social media platforms heavily invested in content moderation and flagging systems; however, as the tech industry continues to face financial concerns and budget cuts, social media platforms are deprioritizing the fight against misinformation [12]. Passing legislation that requires social media platforms to remove misinformation is 1 systems-level approach that could be a critical step in addressing health-related misinformation on the internet, particularly at a time when their efforts are waning [12]. The US Department of Health and Human Services continues urging researchers to engage in multidisciplinary, multilevel research to identify ways to detect and combat the impact of misinformation, particularly among populations who experience health disparities (NIH RFA-MD-22-008).

Information advocates used a comprehensive range of sources to get the complete picture of the current COVID-19 pandemic state of affairs. They broadened how they searched for information, taking local, national, and global perspectives into consideration to understand the scope of the pandemic. They looked for data-driven sources and cross-referenced information with multiple sources. They questioned the motives of information sources, specifically looking for potential hidden agendas and ulterior motives, a practice recommended for identifying misinformation [13]. Information advocates described gauging the plausibility of the information and using common sense to assess the accuracy of information they heard about COVID-19. They reported their lack of reliance on their own intuitions and emotions to assess whether they believed COVID-19 information seemed correct, practices associated with believing false information and spreading misinformation [14]. Some also described using fact-checking sites to assess information accuracy, which can help individuals differentiate between accurate and inaccurate information. Using such sites and identifying web-based misinformation may lead to their debunking (correction of misinformation) of that information, a behavior signifying their interest in acting as information advocates as opposed to information consumers, alone [15]. The National Association for Media Literacy Education recommends using all of these strategies to encourage critical thinking about information posted on the internet [16]. Instilling critical thinking and media literacy skills may enhance individuals' ability to detect misinformation. However, little agreement exists about the best methods (eg, training courses and gamification of information) to enhance media literacy in a "generationally inclusive manner" [17].

Some information advocates described specific actions (eg, flagging and reporting misinformation) to alert websites to user accounts that were actively spreading inaccurate information. Other times, participants blocked users who spread misinformation. Previously, science communication experts assumed that misinformation is spread because people lack access to factual information or the tools to discern fact from fiction, known as the knowledge deficit model [18]. However, behavioral researchers argue that misinformation is accepted as true and then spread, not as a result of mere ignorance but due to psychological factors such as conspiratorial mentality, individual fears, identity expression, and motivated reasoning (driven by personal or moral values over objective evidence)

[19]. They therefore recommend that information advocates respond to misinformation in the following three ways: (1) directly identify misinformation and counter it with fact-based information, (2) identify and address the fallacies in the misinformation sharer's argument, and (3) question the misinformation and source's credibility [19]. Further, to avoid unintentionally sharing misinformation, researchers recommend "nudging" individuals to assess the accuracy of information before they share it on the internet, a practice that, in 1 experiment, enhanced social media users' intentions to engage in critical thinking prior to sharing information within their social networks [20].

Participants were concerned that challenging misinformation could negatively impact their interpersonal relationships. In considering the spiral of silence communication theory, most people prefer to remain silent when their opinion deviates from the dominating view rather than refuting or challenging misinformation [21]. Ignoring misinformation and remaining nonresponsive to it may help them maintain their online reputation or help them avoid violating the norm of online politeness [22]. Some felt more comfortable correcting misinformation with individuals with whom they perceived their social connection to be secure. This is consistent with previous work that illustrated how individuals were more likely to correct misinformation or share debunking information if the original source of that information was a close tie and someone who shares similar traits (in-group members) [23]. Past research suggests that misinformation correction may be more acceptable (and perceived as a less aggressive form of communication) if the relationship with the person being corrected is strong [24]. Correcting misinformation with a weaker tie may further compromise and weaken their social connection [24]. Future research can identify strategies that allow for correcting misinformation in a way that does not weaken interpersonal relationships. Further, future research can identify ways to increase social norms about correcting misinformation on the internet [25]. Additionally, trusted organizations and experts can engage more actively in correcting misinformation [25], especially because doing so does not negatively influence the organizations' credibility or relationships [26]. The US surgeon general further recommends professional associations to provide information or training to their members to serve as experts in sharing and correcting information, including on web-based forums [10]. Such misinformation correction, particularly when made by expert sources, can lead to observational correction, when other social media users change their own attitudes or understanding about a topic after clarification or correction has been publicly made on the internet [26]. Although past research has warned about a backfire effect (originally used in political science), when individuals strengthen their original belief in misinformation after hearing a counterargument, the prevalence of this effect can be exaggerated and is, in fact, most common among individuals who hold more extreme beliefs about particular topics [27].

Some information advocates based their decision to correct web-based misinformation on how consequential they perceived believing in or acting upon the misinformation could be, acting as "communal guardians" of information. They preferred violating the norm of politeness (by not correcting the misinformation) if the outcome meant protecting other information consumers [22]. For example, 1 participant described deciding to confront misinformation based on her emotional response (outrage) to the misinformation. Although past research describes how web-based misinformation often evokes a strong emotional response, which encourages readers to share this information [14], our study finds that web-based misinformation prompted a similar reaction, leading 1 participant to debunk false information.

Limitations

This study is not without its limitations. Our study only examined the perspectives of individuals with a perfect or near-perfect score on COVID-19 knowledge questions and who acted on news information within the past week. Because of this, we cannot speculate on the media literacy behaviors of those with lower COVID-related knowledge scores. Additional limitations are related to our sample, including the study's small sample size. However, the study reached information saturation, a standard method for determining sample sizes in qualitative research. A strength of our sample was its diversity, with participants representing various ages, educational levels, household incomes, and other demographic factors. However, the majority of our nonprobability sample identified as White (n=17, 68%), highly educated (n=19, 76% had at least a college education), and Democrats (n=14, 56%) and therefore did not necessarily represent the diverse US population. This is a common problem among research panel participants, particularly MTurk [28]. There was little diversity in vaccination status; all participants received at least the original COVID-19 doses. However, this may be expected among people who reported trusting science and having high levels of COVID-19 knowledge. There was also less diversity in political ideology represented. This may be because Democrats demonstrated greater vaccine acceptance [29]. On the other hand, in individuals with conservative political leanings, the perceived threat because of COVID-19 was low and had lower vaccine acceptance [30]. Further, the original survey questions used to recruit our sample focused on sharing news information, not necessarily COVID-19 news. However, research illustrates that people share information and react to misinformation similarly regardless of its content [20]. Another limitation of qualitative interviews can include social desirability bias, when participants answer questions in a way that they believe is socially acceptable [31]. The interviewer used strategies to minimize such bias by hosting telephone interviews (instead of video-based interviews), describing the purpose of the study, discussing their anonymity, and assuring participants that there were no wrong responses.

Conclusions

In this study, information advocates trusted science and scientific sources for their COVID-19 information. They preferred getting their information about the pandemic from renowned public health sources such as the CDC and WHO and also trusted social media accounts of health and scientific professionals. For other populations (who may not be as knowledgeable about COVID-19 or who do not act as information advocates), improving the credibility of government and health care institutions and their ability to share public health messages promptly and transparently is critical. Further, this research illustrated the importance of providing objective, nonpolitically biased health information. Future information and media literacy training can teach effective strategies for assessing the credibility of health information by prompting individuals to not only examine the sources but also to question the motives of those posting information. Other skills involved in information and media literacy training can emphasize similar skills used by our information advocates, such as cross-referencing information, using fact-checking sites to verify information, and talking to health care providers about any questions they have about health information.

Insight from this research illustrated how individuals who are well-informed about COVID-19 and serve as information advocates find, share, avoid, and confront misinformation about COVID-19. More individuals preferred sharing accurate information than countering misinformation, especially on the internet. Our findings describe how participants' perceptions of the type and quality of their interpersonal relationships influenced their willingness to combat misinformation. They addressed web-based misinformation by flagging or reporting it. Offline, they preferred addressing misinformation in a more interpersonal and private manner, such as through in-person conversations or text messaging. Some information advocates feared losing relationships over correcting misinformation, while others, based on their perceptions of how dangerous it could be, perceived addressing such false information as an essential action to help protect members of their social networks.

Populations must have accurate health information about the pandemic and practice media literacy and critical thinking skills, particularly since this may affect their disease-prevention behaviors. Implications of our findings could inform future training in health information literacy, interpersonal information advocacy, and organizational web-based information advocacy. Imparting such skills through media and information literacy training may help others share credible, trustworthy information and avoid misinformation, leading to a more informed and healthier public.

Acknowledgments

Arizona State University College of Health Solutions' Infectious Disease Prevention and Public Response Translational Team members reviewed survey items to assess COVID-19 knowledge, pilot-tested the survey instrument, and reviewed the semistructured interview guide for this research. This research group also provided the internal funds for the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COVID-19 knowledge questions used for selecting study participants.
[DOCX File , 16 KB - infodemiology_v3i1e47677_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[DOCX File , 17 KB - infodemiology_v3i1e47677_app2.docx]

References

1. An overview of infodemic management during COVID-19, January 2020-May 2021. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240035966 [accessed 2023-03-11]
2. Hamel L, Lopes L, Kirzinger A, Sparks G, Stokes M, Brodie M. KFF COVID-19 vaccine monitor: media and misinformation. Kaiser Family Foundation. 2021. URL: https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-media-and-misinformation/ [accessed 2023-03-11]
3. 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 [FREE Full text] [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
4. Kennedy B, Tyson A, Funk C. Americans' trust in scientists, other groups declines. Pew Research Center. 2022. URL: https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/ [accessed 2023-02-27]
5. Saldaña J. The Coding Manual for Qualitative Researchers. Thousand Oaks, CA: Sage Publications; 2013.
6. Khubchandani J, Sharma S, Price JH, Wiblishauser MJ, Sharma M, Webb FJ. COVID-19 vaccination hesitancy in the United States: a rapid national assessment. J Community Health 2021;46(2):270-277 [FREE Full text] [doi: 10.1007/s10900-020-00958-x] [Medline: 33389421]
7. Nguyen LH, Joshi AD, Drew DA, Merino J, Ma W, Lo CH, COPE Consortium. Self-reported COVID-19 vaccine hesitancy and uptake among participants from different racial and ethnic groups in the United States and United Kingdom. Nat Commun 2022;13(1):636 [FREE Full text] [doi: 10.1038/s41467-022-28200-3] [Medline: 35105869]

8. Rutjens BT, Većkalov B. Conspiracy beliefs and science rejection. Curr Opin Psychol 2022;46:101392 [FREE Full text] [doi: 10.1016/j.copsyc.2022.101392] [Medline: 35797743]
9. Abu-Akel A, Spitz A, West R. The effect of spokesperson attribution on public health message sharing during the COVID-19 pandemic. PLoS One 2021;16(2):e0245100 [FREE Full text] [doi: 10.1371/journal.pone.0245100] [Medline: 33534800]
10. Murthy VH. Confronting health misinformation: the U.S. Surgeon General's advisory on building a healthy information environment. U.S. Department of Health and Human Services. 2021. URL: https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf [accessed 2023-03-16]
11. Yang KC, Pierri F, Hui PK, Axelrod D, Torres-Lugo C, Bryden J, et al. The COVID-19 infodemic: Twitter versus Facebook. Big Data Soc 2021;8(1):205395172110138 [FREE Full text] [doi: 10.1177/20539517211013861]
12. Myers SL, Grant N. Combating disinformation wanes at social media giants. New York Times. 2023. URL: https://www.nytimes.com/2023/02/14/technology/disinformation-moderation-social-media.html [accessed 2023-03-16]
13. Lewandowsky S, Ecker UKH, Seifert CM, Schwarz N, Cook J. Misinformation and its correction: continued influence and successful debiasing. Psychol Sci Public Interest 2012;13(3):106-131 [FREE Full text] [doi: 10.1177/1529100612451018] [Medline: 26173286]
14. MacFarlane D, Hurlstone MJ, Ecker UKH. Protecting consumers from fraudulent health claims: a taxonomy of psychological drivers, interventions, barriers, and treatments. Soc Sci Med 2020;259:112790. [doi: 10.1016/j.socscimed.2020.112790] [Medline: 32067757]
15. Roozenbeek J, van der Linden S. How to combat health misinformation: a psychological approach. Am J Health Promot 2022;36(3):569-575 [FREE Full text] [doi: 10.1177/08901171211070958] [Medline: 35164546]
16. Building healthy relationships with media: a parent's guide to media literacy. National Association for Media Literacy Education. 2022. URL: https://namle.net/wp-content/uploads/2022/06/220348-NAMLE-Parent-Guide-202200619-2A3.pdf [accessed 2023-03-16]
17. Dumitru EA, Ivan L, Loos E. A generational approach to fight fake news: in search of effective media literacy training and interventions. In: Human Aspects of IT for the Aged Population. Design, Interaction and Technology Acceptance. 2022 Presented at: 8th International Conference, ITAP 2022, Held as Part of the 24th HCI International Conference, HCII 2022; June 26-July 1, 2022; Virtual Event p. 291-310.
18. Simis MJ, Madden H, Cacciatore MA, Yeo SK. The lure of rationality: why does the deficit model persist in science communication? Public Underst Sci 2016;25(4):400-414. [doi: 10.1177/0963662516629749] [Medline: 27117768]
19. 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 [FREE Full text] [doi: 10.1038/s44159-021-00006-y]
20. 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;31(7):770-780 [FREE Full text] [doi: 10.1177/0956797620939054] [Medline: 32603243]
21. Ejikeme I. The global implication of COVID-19 social media misinformation campaigns. Glob Media J 2022;20(53):316 [FREE Full text] [doi: 10.36648/1550-7521.20.53.316]
22. Malhotra P, Scharp K, Thomas L. The meaning of misinformation and those who correct it: an extension of relational dialectics theory. J Soc Pers Relat 2021;39(5):1256-1276. [doi: 10.1177/02654075211050429]
23. Pasquetto IV, Jahani E, Atreja S, Baum M. Social debunking of misinformation on WhatsApp: the case for strong and in-group ties. Proc ACM Hum Comput Interact 2022;6(CSCW1):1-35. [doi: 10.1145/3512964]
24. Tandoc EC, Lim D, Ling R. Diffusion of disinformation: how social media users respond to fake news and why. Journalism 2019;21(3):381-398. [doi: 10.1177/1464884919868325]
25. Vraga EK, Bode L. Correction as a solution for health misinformation on social media. Am J Public Health 2020;110(S3):S278-S280. [doi: 10.2105/AJPH.2020.305916] [Medline: 33001724]
26. Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. Sci Commun 2017;39(5):621-645. [doi: 10.1177/1075547017731776]
27. Nyhan B. Why the backfire effect does not explain the durability of political misperceptions. Proc Natl Acad Sci U S A 2021;118(15):e1912440117 [FREE Full text] [doi: 10.1073/pnas.1912440117] [Medline: 33837144]
28. Zack ES, Kennedy JM, Long JS. Can nonprobability samples be used for social science research? A cautionary tale. Surv Res Methods 2019;13(2):215-227 [FREE Full text] [doi: 10.18148/srm/2019.v13i2.7262]
29. Milligan MA, Hoyt DL, Gold AK, Hiserodt M, Otto MW. COVID-19 vaccine acceptance: influential roles of political party and religiosity. Psychol Health Med 2022;27(9):1907-1917. [doi: 10.1080/13548506.2021.1969026] [Medline: 34407721]
30. Calvillo DP, Ross BJ, Garcia RJB, Smelter TJ, Rutchick AM. Political ideology predicts perceptions of the threat of COVID-19 (and susceptibility to fake news about it). Soc Psychol Personal Sci 2020;11(8):1119-1128 [FREE Full text] [doi: 10.1177/1948550620940539]
31. Bergen N, Labonté R. "Everything is perfect, and we have no problems": detecting and limiting social desirability bias in qualitative research. Qual Health Res 2020;30(5):783-792. [doi: 10.1177/1049732319889354] [Medline: 31830860]

Abbreviations

CDC: US Centers for Disease Control and Prevention
WHO: World Health Organization

Edited by E Lee; submitted 28.03.23; peer-reviewed by A Hicks, F Medina; comments to author 25.07.23; revised version received 24.08.23; accepted 05.09.23; published 20.10.23.

Please cite as:
Koskan AM, Sivanandam S, Roschke K, Irby J, Helitzer DL, Doebbeling B
Sharing Reliable COVID-19 Information and Countering Misinformation: In-Depth Interviews With Information Advocates
JMIR Infodemiology 2023;3:e47677
URL: https://infodemiology.jmir.org/2023/1/e47677
doi:10.2196/47677
PMID:37862066

Original Paper

Monitoring SARS-CoV-2 Using Infoveillance, National Reporting Data, and Wastewater in Wales, United Kingdom: Mixed Methods Study

Jordan P Cuff[1,2], BSc, MRES, PhD; Shrinivas Nivrutti Dighe[1], MSc, PhD; Sophie E Watson[1], BSc, MRES, PhD; Rafael A Badell-Grau[3], BSc, MRES, PhD; Andrew J Weightman[1], BSc, PhD; Davey L Jones[4], BSc, PhD; Peter Kille[1], BSc, PhD

[1]School of Biosciences, Cardiff University, Cardiff, United Kingdom
[2]School of Natural and Environmental Sciences, Newcastle University, Newcastle-upon-Tyne, United Kingdom
[3]Division of Genetics, Department of Paediatrics, University of California, San Diego, La Jolla, CA, United States
[4]School of Natural Sciences, Bangor University, Bangor, United Kingdom

Corresponding Author:
Peter Kille, BSc, PhD
School of Biosciences
Cardiff University
Sir Martin Evans Building
Museum Avenue
Cardiff, CF10 3AX
United Kingdom
Phone: 44 29 2087 4974
Email: Kille@cardiff.ac.uk

Abstract

Background: The COVID-19 pandemic necessitated rapid real-time surveillance of epidemiological data to advise governments and the public, but the accuracy of these data depends on myriad auxiliary assumptions, not least accurate reporting of cases by the public. Wastewater monitoring has emerged internationally as an accurate and objective means for assessing disease prevalence with reduced latency and less dependence on public vigilance, reliability, and engagement. How public interest aligns with COVID-19 personal testing data and wastewater monitoring is, however, very poorly characterized.

Objective: This study aims to assess the associations between internet search volume data relevant to COVID-19, public health care statistics, and national-scale wastewater monitoring of SARS-CoV-2 across South Wales, United Kingdom, over time to investigate how interest in the pandemic may reflect the prevalence of SARS-CoV-2, as detected by national testing and wastewater monitoring, and how these data could be used to predict case numbers.

Methods: Relative search volume data from Google Trends for search terms linked to the COVID-19 pandemic were extracted and compared against government-reported COVID-19 statistics and quantitative reverse transcription polymerase chain reaction (RT-qPCR) SARS-CoV-2 data generated from wastewater in South Wales, United Kingdom, using multivariate linear models, correlation analysis, and predictions from linear models.

Results: Wastewater monitoring, most infoveillance terms, and nationally reported cases significantly correlated, but these relationships changed over time. Wastewater surveillance data and some infoveillance search terms generated predictions of case numbers that correlated with reported case numbers, but the accuracy of these predictions was inconsistent and many of the relationships changed over time.

Conclusions: Wastewater monitoring presents a valuable means for assessing population-level prevalence of SARS-CoV-2 and could be integrated with other data types such as infoveillance for increasingly accurate inference of virus prevalence. The importance of such monitoring is increasingly clear as a means of objectively assessing the prevalence of SARS-CoV-2 to circumvent the dynamic interest and participation of the public. Increased accessibility of wastewater monitoring data to the public, as is the case for other national data, may enhance public engagement with these forms of monitoring.

(JMIR Infodemiology 2023;3:e43891) doi:10.2196/43891

KEYWORDS

COVID-19; Google Trends; infodemiology; quantitative reverse transcription polymerase chain reaction; RT-qPCR; wastewater; infoveillance; public health care; health care statistics; correlation analysis; analysis; public health; online health; eHealth; public interest

Introduction

The COVID-19 pandemic has given rise to a range of public responses that have dynamically driven the cooperation of the public with governmental guidance and public recognition of the need for regular testing. Health care systems have been stretched beyond capacity by sudden, large-volume influxes of patients following sometimes unpredictable waves of the virus [1]. There is a pressing need for local, national, and global adaptability to manage these outbreaks of the disease to minimize the impact on health care systems, the first requirement of which is the stringent collection of reliable and accurate data on viral prevalence [2].

Many strategies have been used to monitor SARS-CoV-2, for example, self-reporting [3] and participatory surveillance [4-6], including through the use of platforms such as accessible phone apps [7]. Surveys and self-reporting, achieved through participatory surveillance and even active crowdsourcing strategies, have proven highly effective in monitoring symptoms such as loss of taste [8]; participatory surveillance platforms such as this have been a crucial component of monitoring in partnership with the public [8,9]. Relying on surveys and personal testing data, however, allows only a reactive approach to mitigating the health care burden imposed by COVID-19, which is often too little, too late to mitigate the heavy case numbers and death tolls. Case data, while sometimes collected by standardized surveys, can otherwise depend on self-reporting by the public, many members of which may not self-test given poor access to tests, may not feel obliged due to asymptomatic cases, or may receive false negative results. Others may unreliably or even dishonestly report the results of tests given the restrictions that a positive test for COVID-19 imposed [10], or they may be disenfranchised with the efforts to reduce the prevalence of the disease given the overwhelming extent of misinformation in circulation [11].

Search engine use has been explored as a means for ascertaining the prevalence of diseases [12,13], but this method is not infallible and its accuracy over time must be assessed in different epidemiological contexts [14,15]. Such data could anecdotally track COVID-19 or specific related symptoms [16-19] but the public searching for particular character strings cannot be directly ascribed to the prevalence of the disease. This “infoveillance” does, however, facilitate analysis of public interest in subjects such as the pandemic [11,20], which can be an important factor in health care management and the pandemic response. Infoveillance can be integrated into interdisciplinary frameworks such as “One Health” [21,22] and, more specifically, “One Digital Health” [23], which aim to view health care matters more holistically, particularly the interaction between human and veterinary health and its implications for zoonotic diseases, but also the environmental dimension of disease occurrence and transmission.

Given the latency of surveys and testing by the public, and the potential inaccuracies of infoveillance approaches, objective means for disease surveillance without the requirement of public participation have become increasingly important throughout the COVID-19 pandemic. The presence of coronaviruses and other human pathogenic viruses in human feces and their subsequent presence in urban wastewater is a long-established tool for assessing disease prevalence within a community [24,25]. This approach provides a noninvasive means for assessing SARS-CoV-2 prevalence across whole populations via wastewater [25-31]. The monitoring of wastewater has provided a robust and accurate means of assessing the population-level prevalence of COVID-19, facilitating some prediction of health care burden before symptoms arise [32]. Wastewater monitoring circumvents several barriers preclusive to accurate testing data such as hesitancy, the availability of testing, asymptomatic patients, and socioeconomic or cultural barriers by passively sampling from whole communities [10,33]. The efficacy of this approach does not depend on public participation, possibly leading to some inconsistencies with national testing statistics. A strong positive correlation between direct testing, wastewater monitoring data, and public interest in the pandemic has been demonstrated [34], but the dynamic relationship between these data and how public interest dictates the accuracy of monitoring data are still poorly characterized.

Here, we compare public interest in the pandemic through search engine use data against wastewater SARS-CoV-2 surveillance data and nationally reported statistics over time to assess how public interest dictated the relationship between disease prevalence and reporting over a year of the COVID-19 pandemic in South Wales, United Kingdom. This study also explores the efficacy of wastewater monitoring and infoveillance as means for assessing the national state of the pandemic, how these relationships change over time, and how they could inform predictions of case numbers for streamlined monitoring.

Methods

Wastewater Monitoring

Since mid-September 2020, wastewater samples were collected every Monday, Wednesday, and Friday from Cardiff Bay, Newport Nash, Llanfoist, Ponthir, Ogmore, Cog Moors, Swansea Bay, and Gowerton wastewater treatment plants, and samples from Carmarthen and Haverfordwest were collected every Wednesday. Samples were transported on ice in a cooler box to designated wastewater processing facilities at Cardiff University. The processing of samples was based on Farkas et al [35]. From each site, 200 mL of wastewater was spun at 3000×g for 30 minutes, and 150 mL of supernatant was neutralized to pH 7-7.4 using 1 M NaOH. The supernatant was incubated with 50 mL of 40% PEG and 8% NaCl overnight. Samples were then spun at 10,000×g for 30 minutes and the pellet was dissolved in 500 µL of PBS (pH 7.4). Of the dissolved pellet, 100 µL was spiked with 10,000 copies of synthetic murine norovirus DNA to check the extraction efficiency. Subsequent nucleic acid extraction and amplification took place in the COVID-19 testing facilities at Cardiff University. Total RNA was extracted using the methodology published by Oberacker et al [36]. Total RNA was eluted in 100 µL of nuclease-free water. For SARS-CoV-2 detection, 4 primer sets published by the US Centers for Disease Control and Prevention (CDC), Charite´, and Hong Kong University [37] were used for quantitative reverse transcription polymerase chain reaction (RT-qPCR). Primer sets N1 and N2 target different regions of the nucleocapsid (N genes); E_Sarbeco and ORF1b target the SARS-CoV-2 E and nsp14 genes, respectively. For the controls, a set of primers that target virus crAssphage [38] (which is present in human fecal material) and murine norovirus [39] (which was used to assess extraction efficiency) were selected (Table 1). Samples were run in triplicate on Fast 384-well plates (Applied Biosystems) using QuantStudio 7 Flex (Applied Biosystems). A 10 µL RT-qPCR reaction was performed containing 4 µL of extracted RNA template, 5 µL of Luna Universal Probe One-step Reaction Mix (2X; NEB), 0.04 µL of each primer set (100 µM), 0.02 µL of fluorescent probe (100 µM), 0.5 µL NEB Luna reverse transcriptase (20X), and 0.4 µL nuclease-free water. The reverse transcription (RT) was carried out at 55 °C for 10 minutes, followed by polymerase activation at 95.0 °C for 1 minute and 40 cycles of denaturation, annealing, and extension at 95.0 °C for 10 seconds and then 60.0 °C for 1 minute, respectively. Serial dilutions of the heat-inactivated SARS-CoV-2 viral standards were run on every PCR plate to generate standard curves used to quantify the copies of SARS-CoV-2 genes. Additionally, RT-qPCR runs were validated by positive (Qnostics, SCV2QC01-QC) and negative controls (nuclease-free water). Resultant data were normalized to account for population size in each area, and to correct for dilution as described by Wilde et al [40].

Table 1. The quantitative polymerase chain reaction (qPCR) primers used for wastewater monitoring.

Assay	Target gene	Sequences (5’-3’)
E_Sarbeco	E	• F- 5’-ACAGGTACGTTAATAGTTAATAGCGT-3’ • R- 5’-ATATTGCAGCAGTACGCACACA-3’ • P- 5’-HEX-ACACTAGCCATCCTTACTGCGCTTCG- IBFQ-3
ORF1b	nsp14	• F-5’-TGGGGYTTTACRGGTAACCT-3’ • R-5’-AACRCGCTTAACAAAGCACTC-3’ • P- 5’-FAM-TAGTTGTGATGCWATCATGACTAG- IBFQ-3’
N1	Nucleocapsid	• F- 5’-GACCCCAAAATCAGCGAAAT-3’ • R- 5’-TCTGGTTACTGCCAGTTGAATCTG-3’ • P-5-HEX- ACCCCGCATTACGTTTGGTGGACC- IBFQ-3’
N2	Nucleocapsid	• F- 5’-TTACAAACATTGGCCGCAAA-3’ • R- 5’-GCGCGACATTCCGAAGAA-3’ • P- 5’-FAM- ACAATTTGCCCCCAGCGCTTCAG-IBFQ-3’
crAssphage	Q56	• F- 5’-CAGAAGTACAAACTCCTAAAAAACGTAGAG-3’ • R- 5’-GATGACCAATAAACAAGCCATTAGC-3’ • P- 5’-HEX- AATAACGATTTACGTGATGTAAC-IBFQ-3’
MNV[a]	—[b]	• F- 5’-CCGCAGGAACGCTCAGCAG-3’ • R- 5’-GGYTGAATGGGGACGGCCTG-3’ • P-5’-FAM- ATGAGTGATGGCGCA- IBFQ-3’

[a]MNV: murine norovirus.

[b]Not applicable.

National Statistics and Search Volume Data Extraction

This study concerns 2 periods: the primary study period (between the weeks of October 11, 2020 and October 31, 2021; the focus of all analyses and visualizations aside from comparison with model-based predictions described below) and the full study period (the primary study period with extension up to July 17, 2022 to facilitate comparison of real-world data with model-based predictions). All data were generated or extracted to encompass the full study period. National statistics on the daily number of COVID-19 cases, deaths, and vaccinations in Wales were extracted from the UK government’s COVID-19 data portal for the full study period [41]. Case data were new cases by publish date (ie, the number of new cases reported since the previous update; API=“newCasesByPublishDate”). Death data were new daily national statistics office deaths by death date (ie, daily numbers of deaths of people whose death certificate mentioned COVID-19 as one of the causes; API=“newDailyNsoDeathsByDeathDate”). Vaccine data were new vaccines given by publish date (ie, daily numbers of new vaccines [all doses] given; API=“newVaccinesGivenByPublishDate”). These data can be downloaded via a permanent download link [41].

Search volume data were extracted from Google Trends. These data provide a proxy for public interest in or response to the extent of the COVID-19 pandemic. The data extracted from

Google Trends are relative search volumes (RSVs) for predetermined search terms, allowing comparison of search rates for different terms via Google, the most widely used internet search engine. These RSVs are presented for each date of a given period within a given country, nation, or region and are normalized relative to the highest search volume peak in that search batch in the time period specified (this peak is represented as a search volume of 100%). Search volumes were releveled so that the highest peak in the primary study period was represented by "100" and any higher peaks across the full study period exceeded 100 to reflect the limitations of making real-time predictions from existing data. Given the representation of numbers less than 1 as "<1" by Google Trends, all RSVs of "<1" were converted to 0 to facilitate quantitative comparison.

Search terms were selected based on their broad relevance throughout the study period and the high volume of searches generated during that period. These included "COVID lockdown," "COVID rules," "COVID symptoms," "COVID test," and "COVID vaccine." "COVID" was included in each search term to ensure relevance to the COVID-19 pandemic; "COVID" was selected over "coronavirus," "SARS-CoV-2," and other variations due to the greater prevalence of searches related to this string, and its inclusion within other search strings such as "COVID-19".

Statistical Analysis

Statistical analyses and plotting of data were carried out using R (version v4.0.3; R Core Team) [42] and all data and code are openly available [43]. Since wastewater sites were sampled weekly, all data were averaged first by site and then by week. Wastewater quantitative polymerase chain reaction (qPCR) data were log-transformed to improve model fit and visualization. Data were processed and aggregated using *tidyverse* packages for reproducibility [44].

Correlations between search volumes; wastewater SARS-CoV-2 prevalence; and nationally reported cases, deaths, and vaccinations were tested using Spearman ρ rank correlation via the *rcor* function of the *Hmisc* package [45]. To facilitate the assessment of correlation, week dates were transformed into successive study weeks (ie, cumulative weeks of the study). The data were identified as nonnormally distributed via Shapiro-Wilk tests, so nonparametric correlation analyses were selected. The output was visualized in a correlogram via the *corrplot* function of the *corrplot* package [46], with colors to denote the strength of correlations assigned via the *viridis* package [47].

To assess how RSV for the selected search terms changed with differences in the number of COVID-19–related cases, deaths, and vaccines and the estimated prevalence of COVID-19 in wastewater, a multivariate linear model (MLM) was built via *manylm* in the *mvabund* package [4]. The dependent variable comprised the RSVs, log-transformed (log[n+1]) to achieve normality, and the independent variables were week; national cases, deaths, and vaccinations; and 2-way interactions between study week and each of the other variables. For visualization via line plots, data were releveled so that their minimum and maximum values were 0 and 100, respectively. These normalized search volume, wastewater, and government data were plotted against time using the *ggplot2* package [48], with colors assigned via the *paired* palette in the *RColorBrewer* package [49] and data lines smoothed using the *loess* method.

Pairwise plots were generated for reported case data, qPCR data, and RSVs from each of the Google Trends search terms separately using *ggpairs* from the *GGAlly* package. Linear models (LMs) were generated with the number of reported cases as the dependent variable and, in a separate model for each, the qPCR and Google Trends data as independent variables. The *predict* function was used to make interpolated predictions of case numbers across the primary study period and extrapolated predictions of case numbers beyond the primary study period for the remainder of the full study period. These predicted case numbers were plotted against the reported case numbers, and a correlation analysis was carried out as described above. A generalized linear model (GLM) with a Gaussian error family was built with reported cases as the dependent variables and predicted case numbers, time, and pairwise interactions between predictions and time as independent variables.

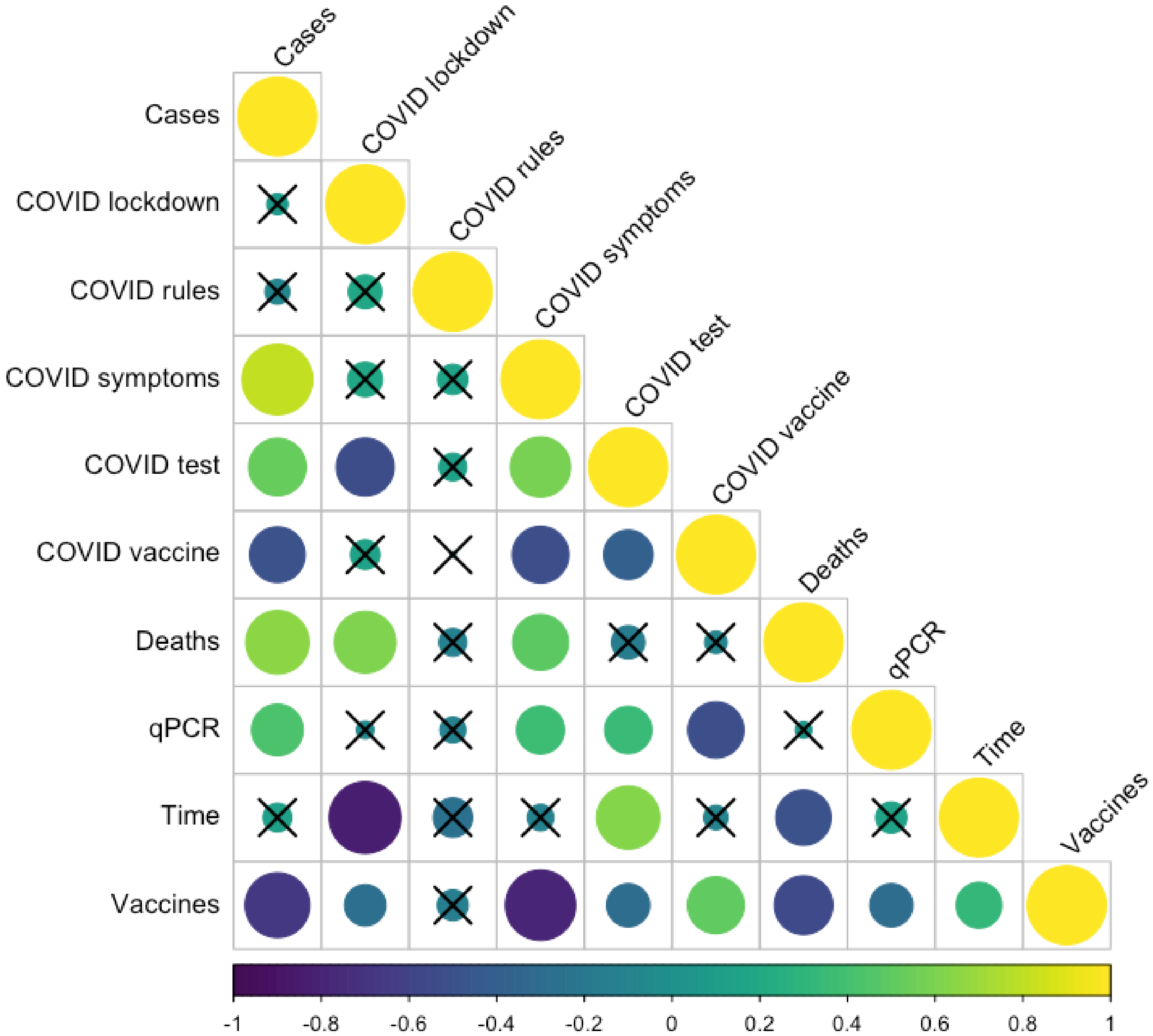
Information Sources and Reliability

Wastewater monitoring data were generated by the authors of this study at Cardiff University as part of the Welsh government–funded WEWASH project. The national statistics on COVID-19 cases, deaths, and vaccinations were extracted from the UK government's COVID-19 data portal [41], which is internationally recognized as a reputable source used for national reporting, scientific research, and public awareness. The Google Trends data should be reliable as indicators of Google use since they are collected by Google based on the input of users of their service.

Results

Overall, significant correlations were identified between many of the variables (Figure 1 and Table S1 in Multimedia Appendix 1). Notably, wastewater SARS-CoV-2 RNA prevalence significantly positively correlated with the number of reported cases (Spearman ρ=0.428; *P*=.001) but did not correlate with the number of reported deaths (Spearman ρ=0.044; *P*=.75). Of the search terms included, wastewater prevalence positively correlated with "COVID symptoms" (Spearman ρ=0.369; *P*=.005) and "COVID test" (Spearman ρ=0.356; *P*=.007) and significantly negatively correlated with "COVID vaccine" (Spearman ρ=–0.504; *P*<.001). The number of reported cases, however, positively correlated with both "COVID symptoms" (Spearman ρ=0.805; *P*<.001) and "COVID test" (Spearman ρ=0.531; *P*<.001) but negatively correlated with "COVID vaccine" (Spearman ρ=–0.495; *P*=.001). All search terms except "COVID rules" significantly negatively correlated with national vaccinations (all *P*<.05; Table S1 in Multimedia Appendix 1).

Figure 1. Correlogram of time (study week, ie, progressive number of weeks into the study period), Google Trends search volumes (variables starting with "COVID"), nationally reported cases, deaths and vaccinations, and qPCR-based wastewater SARS-CoV-2 RNA prevalence. Circle size and color (purple, through teal to yellow—denoting negative through neutral to positive) indicate the extent and directionality of the correlation. Crossed-out circles are those for which correlations were not significant. qPCR: quantitative polymerase chain reaction.

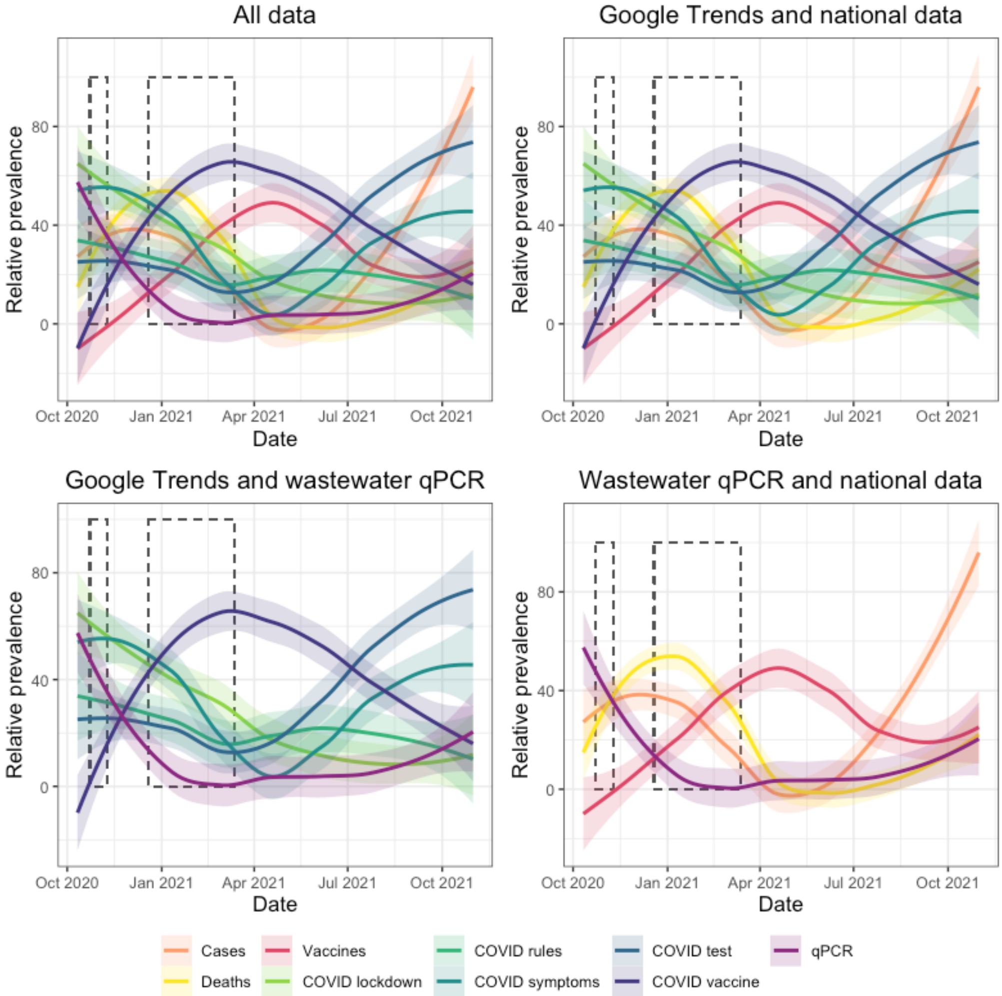


Search volumes were significantly related to several of the independent variables and their interactions (Table 2 and Figure 2), comprising wastewater SARS-CoV-2 prevalence (MLM: $F_{1,54}$=34.89; P=.002); time (MLM: $F_{1,53}$=120.89; P=.002); national COVID-19 cases reported (MLM: $F_{1,52}$=117.77; P=.002); national COVID-19–related deaths reported (MLM: $F_{1,51}$=65.84; P=.002); national COVID-19 vaccines administered (MLM: $F_{1,50}$=54.31; P=.002); and the interactions between time and national COVID-19 cases (MLM: $F_{1,48}$=46.32; P=.002), time and national COVID-19 deaths (MLM: $F_{1,48}$=26.09; P=.004), and time and national vaccinations (MLM: $F_{1,46}$=15.10; P=.02). The interaction between time and wastewater SARS-CoV-2 RNA prevalence (MLM: $F_{1,49}$=0.77; P=.97) was not significantly related to RSVs.

Table 2. Univariate results from the multivariate linear model results for search volume data analyzed against time (progressive study weeks); wastewater SARS-CoV-2 RNA prevalence; nationally reported COVID-19 cases, deaths, and vaccines; and 2-way interactions between time and each other variable.

Independent variable	"COVID symptoms," *F* test (*df*)	*P* value	"COVID test," *F* test (*df*)	*P* value	"COVID vaccine," *F* test (*df*)	*P* value	"COVID rules," *F* test (*df*)	*P* value	"COVID lockdown," *F* test (*df*)	*P* value
Wastewater SARS-CoV-2 prevalence	2.211 (1, 54)	.34	0.418 (1, 54)	.69	28.838 (1, 54)	.002	0.583 (1, 54)	.69	2.834 (1, 54)	.31
Time	0.189 (1, 53)	.88	34.716 (1, 53)	.002	0.120 (1, 53)	.88	4.414 (1, 53)	.12	81.453 (1, 53)	.002
National COVID-19 cases reported	77.157 (1, 52)	.002	28.501 (1, 52)	.002	4.122 (1, 52)	.11	0.677 (1, 52)	.41	7.315 (1, 52)	.03
National COVID-19–related deaths	2.373 (1, 51)	.22	13.42 (1, 51)	.003	18.621 (1, 51)	.003	30.232 (1, 51)	.002	1.193 (1, 51)	.24
Vaccines administered nationally	17.880 (1, 50)	.002	21.308 (1, 50)	.002	8.766 (1, 50)	.02	0.586 (1, 50)	.43	5.770 (1, 50)	.048
Time: wastewater prevalence	0.284 (1, 49)	.98	0.067 (1, 49)	.98	0.011 (1, 49)	.98	0.243 (1, 49)	.98	0.165 (1, 49)	.98
Time: cases	3.349 (1, 48)	.16	15.165 (1, 48)	.002	10.632 (1, 48)	.004	15.869 (1, 48)	.002	1.301 (1, 48)	.27
Time: deaths	3.536 (1, 47)	.18	4.113 (1, 47)	.15	3.04 (1, 47)	.18	0.246 (1, 47)	.59	15.155 (1, 47)	.004
Time: vaccines	0.241 (1, 46)	.81	0.171 (1, 46)	.81	6.898 (1, 46)	.06	1.89 (1, 46)	.37	5.903 (1, 46)	.07

Figure 2. Relative search volumes extracted from Google Trends compared against nationally reported data and qPCR-based estimates of prevalence for SARS-CoV-2 in wastewater. All values are normalized so that the maximum value for each variable is 100. Lines are loess-smoothed curves, thus representing the overall trend, and do not always represent the most extreme (eg, maximum) values. Dashed rectangles represent periods of national lockdown in Wales for reference. Wastewater qPCR-based SARS-CoV-2 prevalence is given in light purple, Google Trends data are given in green or blue, and national data are given in orange or red or purple. A figure containing nonsmoothed trends is presented in Figure S1 in Multimedia Appendix 1. qPCR: quantitative polymerase chain reaction.



National case data significantly related to Google Trends data for “COVID symptoms” (LM: t_{54}=7.248, *P*<.001; Figure S4 in Multimedia Appendix 1), “COVID test” (LM: t_{54}=6.070, *P*<.001; Figure S5 in Multimedia Appendix 1), and “COVID vaccine” (LM: t_{54}=–3.301, *P*=.002; Figure S6 in Multimedia Appendix 1 but not qPCR-based wastewater SARS-CoV-2 prevalence (LM: t_{54}=1.360, *P*=.18 Figures S2-6 in Multimedia Appendix 1) nor Google Trends data for “COVID lockdown” (LM: t_{54}=0.897, *P*=.37; Figure S2 in Multimedia Appendix 1) and “COVID rules” (LM: t_{54}=0.320, *P*=.75; Figure S3 in Multimedia Appendix 1). Notably, wastewater SARS-CoV-2 RNA prevalence-based predictions significantly positively correlated with the number of reported cases (Spearman ρ=0.274; *P*=.008). Of the search terms included, case data correlated with predictions based on “COVID symptoms” (Spearman ρ=0.683; *P*<.001), “COVID test” (Spearman ρ=0.706; *P*<.001), and “COVID rules” (Spearman ρ=0.409; *P*<.001). National case data significantly related to case numbers predicted by “COVID symptoms” (GLM: t_{92}=5.158, *P*<.001) and “COVID test” (GLM: t_{92}=–4.997, *P*<.001) RSVs, but these relationships changed over time (“COVID symptoms”: t_{92}=–5.162, *P*<.001; “COVID test”: t_{92}=5.029, *P*<.001; Figure 4). National case data marginally insignificantly related to case numbers predicted by qPCR wastewater SARS-CoV-2 prevalence (GLM: t_{92}=–1.896, *P*=.02) and “COVID rules” RSVs (GLM: t_{92}=1.853, *P*=.07), but these relationships were marginally insignificantly related to time (qPCR: t_{92}=1.920, *P*=.06; “COVID rules”: t_{92}=–1.866, *P*=.07; Figure 4).

Figure 3. Correlogram of time (study week, ie, progressive number of weeks into the study period), nationally reported cases, and the number of cases predicted based on linear models of cases against Google Trends search volumes and qPCR-based wastewater SARS-CoV-2 prevalence. Circle size and color (purple, through teal to yellow—denoting negative through neutral to positive) indicate the extent and directionality of the correlation. Crossed-out circles are those for which correlations were not significant. qPCR: quantitative polymerase chain reaction.

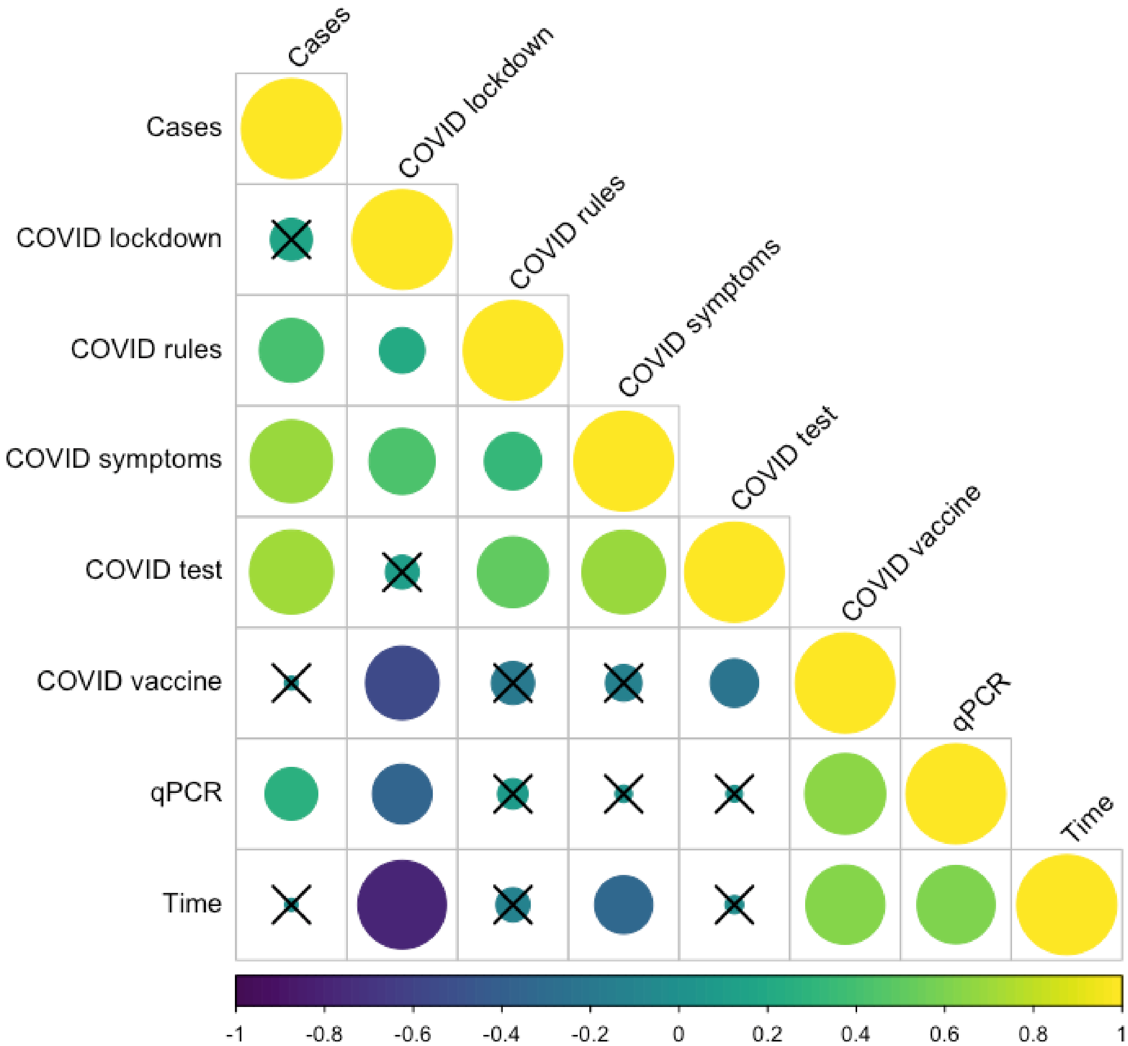
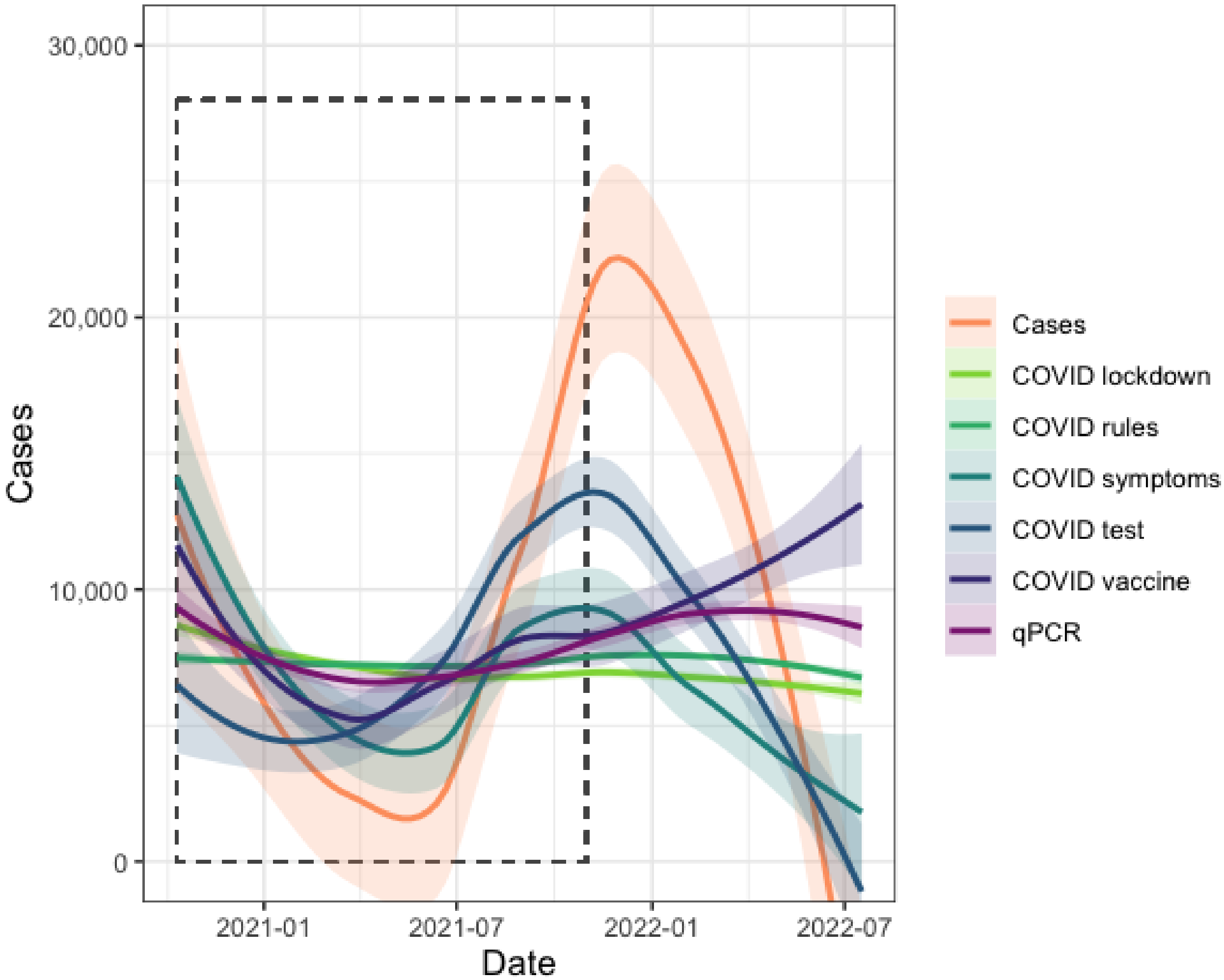


Figure 4. COVID-19 case numbers, and predicted case numbers interpolated and extrapolated based on linear models of case numbers and, separately, each Google Trends search term and qPCR-based SARS-CoV-2 prevalence in wastewater. The dashed rectangle denotes the primary study period, within which data are interpolated. Interpolations are based on data from the primary study period from which models were generated. Extrapolations (outside of the rectangle) are based on data from the following 9 months. Wastewater qPCR-estimated SARS-CoV-2 prevalence is given in light purple, Google Trends data are given in green or blue, and national reported case data are given in orange. Nonsmoothed data are presented in Figure S7 in Multimedia Appendix 1. qPCR: quantitative polymerase chain reaction.



Discussion

Principal Findings

This study provides evidence to suggest that public interest in topics related to the pandemic changed dynamically across the study period, with some relation to the prevalence of the virus in wastewater and the number of reported cases. Both internet search volume and qPCR-based SARS-CoV-2 RNA prevalence data provide some predictive potential for monitoring SARS-CoV-2 and could be applied across other contexts.

During the course of this study, comprising 2 significant waves of the COVID-19 pandemic in Wales, the correlation between reported COVID-19 cases and wastewater-quantified SARS-CoV-2 prevalence was significantly positive overall, as has been demonstrated in previous studies [28,34], but this correlation may have changed over time. Comparing the prevalence of wastewater SARS-CoV-2 estimates and national cases across the full study period shows that wastewater prevalence of SARS-CoV-2 peaked substantially higher in October 2020 than the rest of the study period, whereas case data peaked the following October (Figure S1 in Multimedia Appendix 1). Indications of correlation between SARS-CoV-2 prevalence in wastewater and COVID-19 disease prevalence were recognized at an early stage of the pandemic in other countries [32]. The Google Trends search volume data show web-based searching for some COVID-19–related strings largely reduced over time, although this was highly dependent on the search string. This could indicate reduced public interest, fluctuations that were reported even in the initial months of the pandemic despite the importance of sustained public action to ensure the success of public health measures [50].

In this same period, many of the search volumes, with the intuitive exception of “COVID vaccine,” appear to inversely correlate with increased vaccinations. This suggests that the public may have been seeking vaccine opportunities and otherwise expressed less interest in COVID-19 following mass vaccinations, although additional data would be required to confirm this. Importantly, searches for “COVID vaccine” may also represent those that were concerned with misinformation or conspiracy theories that were commonplace, particularly around the vaccine [11].

The search term "COVID test" was maintained at a relatively constant level throughout the study and, along with "COVID symptoms" and "COVID vaccine," correlated with the wastewater SARS-CoV-2 prevalence just as national case data did. This indicates the potential of carefully selected search terms for estimating the prevalence of the virus, further ratified by the predictions made in this study. The relationship between predictions and case data varied greatly depending on the data used to guide predictions and, importantly, these relationships changed over time. The variable potential of infoveillance to predict epidemiological trends has been recorded in other cases, such as for Google Flu Trends [13,15], and is an important consideration for the use of infoveillance in a monitoring context. The efficacy of infoveillance is contingent on public interest consistently reflecting epidemiology, which is ultimately unlikely for global pandemics given natural spikes and fluctuations in public interest. It is, however, important to contextualize this with the likely reasons for members of the public searching with this particular string. Search volume data could nonetheless provide anecdotal monitoring of disease prevalence, especially since many nations face difficulties in monitoring the virus using molecular methods or population-level testing. Search volume data, while imperfect, may provide a valuable alternative for anecdotal epidemiological monitoring in nations or regions lacking access to alternatives [51], but the search terms must be carefully considered, closely monitored, and interpreted with appropriate skepticism.

The strong positive correlation between national testing, wastewater monitoring data, and Google RSVs has previously been demonstrated in the United States [34]. The relation of search term data to SARS-CoV-2 prevalence in wastewater changed over time, suggesting that such approaches require monitoring and constant evaluation, again suggesting that an approach combining data types may be optimal [34]. Importantly, the predictions made based on qPCR-based wastewater monitoring were marginally insignificantly related to recorded cases. Given the relative objectivity of this molecular monitoring, this is likely to reflect the inconsistent accuracy of national case data reporting as the pandemic progressed, highlighting the need for objective measures of virus prevalence irrespective of public participation. While these different data types dynamically interact and often imperfectly reflect one another, as demonstrated by our univariate predictions, together they could generate models with greater predictive power for forecasting improved above that of univariate approaches [34]. This aligns with the "One Health" perspective of integrating different data types across disciplinary boundaries to monitor health care and epidemiological events more holistically [22,23]. Wastewater monitoring has been integrated into One Health frameworks for pathogen monitoring [52] and emerging concepts such as antimicrobial resistance in the environment [53]. Given that infoveillance similarly aligns with the principles of One Health [23], this presents an ideal opportunity to integrate different data types for sociobiological monitoring of SARS-CoV-2 and other pandemic agents.

Limitations

Regarding infoveillance, this study relied exclusively on Google search volume data; while this represents the most used search engine and thus the greatest single source of such data, other search engines are regularly used that might provide different insights. Web-based search data, while an asset for assessing public responses, is also collected without the context of its users' motives; thus, assumptions cannot reliably be made about the specific interests related to each search string. Even without this context, however, the search volumes presented in this study indicate interest, positive or negative, in those topics. Previous studies have demonstrated that the efficacy of these data in predicting epidemiological trends can be, at best, variable and, at worst, ineffective [13-15]; this can be mitigated to some degree via robust statistical methods to increase the reliability and accuracy of infoveillance for epidemiological "nowcasting" [15], but the integration of these data into more holistic frameworks across disciplinary boundaries could further ameliorate these inaccuracies and provide increasingly accurate predictions [22,23].

While the qPCR data in this study represent a nationwide effort to monitor SARS-CoV-2, they do not comprehensively cover the nation of Wales, which is otherwise fully represented by the Google Trends and national reporting data. Importantly, the qPCR data do account for all of South Wales, which, in turn, accounts for approximately 71% of the national population [54], meaning that these data should accurately reflect the overall national SARS-CoV-2 prevalence. Future studies could investigate how different spatiotemporal resolutions of data affect the accuracy and outcomes of analyses such as these, especially given that this will impact the feasibility of long-term monitoring using most methods.

The progression of COVID-19 as a global pandemic continues to be extremely complicated and unpredictable, and the findings of this study focus on just 1 period in this evolving situation, prior to the emergence of the SARS-CoV-2 Omicron variant and its sublineages. More importantly, the early months of the pandemic are not represented due to the unavailability of qPCR data for that period. While this study relates primarily to those later months of the first year of the pandemic through to the second year, the use of Google Trends data may have been more powerful in the early months of the pandemic when public familiarity was lower and more people were seeking information.

Conclusions

Both molecular monitoring of wastewater and infoveillance approaches demonstrate potential for monitoring and prediction of epidemiological trends. Personal testing and surveys can introduce latency to monitoring, lack randomization, and can receive reduced participation for fear of positive test outcomes [10]; thus, reduced dependency on these data through widespread adoption of wastewater monitoring will likely improve the accuracy of epidemiological data. Wastewater monitoring has previously correlated strongly with national case data [32], but any decrease in this correlation must importantly be viewed with respect to the public interest and how this might impact reported case data. Disease surveillance via wastewater monitoring provides many potential benefits, not least its objectivity. As public interest in the pandemic wanes, widespread molecular analysis of wastewater will become increasingly important as personal testing data become increasingly inaccurate at the population level. Public access to wastewater monitoring data has been facilitated through web-based reporting, including the data used in this study [38], but accessible presentation of these data in interactive dashboards, as has been the case for other national data, may increase public understanding, appreciation, and use of this important data source.

Acknowledgments

This work was supported by a grant from the Welsh government under the Welsh Wastewater Programme (C035/2021/2022). The authors also thank Tony Harrington at Dŵr Cymru-Welsh Water alongside staff at the wastewater treatment facility for their support in collecting the samples.

Authors' Contributions

JPC, SND, and SEW conceptualized the study. SND carried out the quantitative polymerase chain reaction testing of wastewater. JPC and RABG analyzed the data. JPC prepared figures. PK, AJW, and DLJ oversaw the study. All authors contributed to the writing and editing of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 1417 KB - infodemiology_v3i1e43891_app1.docx]

References

1. McLellan A, Godlee F. COVID 19: Christmas relaxation will overwhelm services. BMJ 2020;371:m4847 [FREE Full text] [doi: 10.1136/bmj.m4847] [Medline: 33323367]
2. Lange KW. The prevention of COVID-19 and the need for reliable data. Movement Nutr Health Dis 2020;4:53-63 [FREE Full text] [doi: 10.5283/mnhd.25]
3. McDonald SA, van den Wijngaard CC, Wielders CCH, Friesema IHM, Soetens L, Paolotti D, et al. Risk factors associated with the incidence of self-reported COVID-19-like illness: data from a web-based syndromic surveillance system in the Netherlands. Epidemiol Infect 2021;149:e129 [FREE Full text] [doi: 10.1017/S0950268821001187] [Medline: 34006340]
4. McNeil C, Verlander S, Divi N, Smolinski M. The landscape of participatory surveillance systems across the one health spectrum: systematic review. JMIR Public Health Surveill 2022;8(8):e38551 [FREE Full text] [doi: 10.2196/38551] [Medline: 35930345]
5. 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;7(11):e33576 [FREE Full text] [doi: 10.2196/33576] [Medline: 34727046]
6. Leal-Neto OB, Santos FAS, Lee JY, Albuquerque JO, Souza WV. Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning. Int J Med Inform 2020;143:104263 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104263] [Medline: 32877853]
7. Pandit JA, Radin JM, Quer G, Topol EJ. Smartphone apps in the COVID-19 pandemic. Nat Biotechnol 2022;40(7):1013-1022 [FREE Full text] [doi: 10.1038/s41587-022-01350-x] [Medline: 35726090]
8. Drew DA, Nguyen LH, Steves CJ, Menni C, Freydin M, Varsavsky T, et al. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. Science 2020;368(6497):1362-1367 [FREE Full text] [doi: 10.1126/science.abc0473] [Medline: 32371477]
9. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med 2020;26(7):1037-1040 [FREE Full text] [doi: 10.1038/s41591-020-0916-2] [Medline: 32393804]
10. Murakami M, Hata A, Honda R, Watanabe T. Letter to the editor: wastewater-based epidemiology can overcome representativeness and stigma issues related to COVID-19. Environ Sci Technol 2020;54(9):5311 [FREE Full text] [doi: 10.1021/acs.est.0c02172] [Medline: 32323978]
11. Badell-Grau RA, Cuff JP, Kelly BP, Waller-Evans H, Lloyd-Evans E. Investigating the prevalence of reactive online searching in the COVID-19 pandemic: infoveillance study. J Med Internet Res 2020;22(10):e19791 [FREE Full text] [doi: 10.2196/19791] [Medline: 32915763]

12. Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. Detecting influenza epidemics using search engine query data. Nature 2009;457(7232):1012-1014 [FREE Full text] [doi: 10.1038/nature07634] [Medline: 19020500]
13. Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T, et al. Influenza forecasting with Google Flu Trends. PLoS One 2013;8(2):e56176 [FREE Full text] [doi: 10.1371/journal.pone.0056176] [Medline: 23457520]
14. Lazer D, Kennedy R. What we can learn from the epic failure of Google Flu Trends. Wired. 2015. URL: https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/#:~:text=And%20then%2C%20GFT%20failed%E2%80%94and,the%20foibles%20of%20big%20data [accessed 2023-10-24]
15. Kandula S, Shaman J. Reappraising the utility of Google Flu Trends. PLoS Comput Biol 2019;15(8):e1007258 [FREE Full text] [doi: 10.1371/journal.pcbi.1007258] [Medline: 31374088]
16. Rovetta A, Bhagavathula AS. COVID-19-related web search behaviors and infodemic attitudes in Italy: infodemiological study. JMIR Public Health Surveill 2020;6(2):e19374 [FREE Full text] [doi: 10.2196/19374] [Medline: 32338613]
17. Walker A, Hopkins C, Surda P. Use of Google trends to investigate loss-of-smell-related searches during the COVID-19 outbreak. Int Forum Allergy Rhinol 2020;10(7):839-847 [FREE Full text] [doi: 10.1002/alr.22580] [Medline: 32279437]
18. Ortiz-Martínez Y, Garcia-Robledo JE, Vásquez-Castañeda DL, Bonilla-Aldana DK, Rodriguez-Morales AJ. Can Google® trends predict COVID-19 incidence and help preparedness? the situation in Colombia. Travel Med Infect Dis 2020;37:101703 [FREE Full text] [doi: 10.1016/j.tmaid.2020.101703] [Medline: 32360323]
19. Saegner T, Austys D. Forecasting and surveillance of COVID-19 spread using Google trends: literature review. Int J Environ Res Public Health 2022;19(19):12394 [FREE Full text] [doi: 10.3390/ijerph191912394] [Medline: 36231693]
20. Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with internet search volumes: a Google trends analysis. Int J Infect Dis 2020;95:192-197 [FREE Full text] [doi: 10.1016/j.ijid.2020.04.033] [Medline: 32305520]
21. Zinsstag J, Schelling E, Waltner-Toews D, Tanner M. From "one medicine" to "one health" and systemic approaches to health and well-being. Prev Vet Med 2011;101(3-4):148-156 [FREE Full text] [doi: 10.1016/j.prevetmed.2010.07.003] [Medline: 20832879]
22. Mackenzie JS, Jeggo M. The one health approach-why is it so important? Trop Med Infect Dis 2019;4(2):88 [FREE Full text] [doi: 10.3390/tropicalmed4020088] [Medline: 31159338]
23. Benis A, Tamburis O, Chronaki C, Moen A. One digital health: a unified framework for future health ecosystems. J Med Internet Res 2021;23(2):e22189 [FREE Full text] [doi: 10.2196/22189] [Medline: 33492240]
24. Gundy PM, Gerba CP, Pepper IL. Survival of coronaviruses in water and wastewater. Food Environ Virol 2008;1(1):10-14 [FREE Full text] [doi: 10.1007/s12560-008-9001-6]
25. Wade MJ, Jacomo AL, Armenise E, Brown MR, Bunce JT, Cameron GJ, et al. Understanding and managing uncertainty and variability for wastewater monitoring beyond the pandemic: lessons learned from the United Kingdom national COVID-19 surveillance programmes. J Hazard Mater 2022;424(Pt B):127456 [FREE Full text] [doi: 10.1016/j.jhazmat.2021.127456] [Medline: 34655869]
26. Ni G, Lu J, Maulani N, Tian W, Yang L, Harliwong I, et al. Novel multiplexed amplicon-based sequencing to quantify SARS-CoV-2 RNA from wastewater. Environ Sci Technol Lett 2021;8(8):683-690. [doi: 10.1021/acs.estlett.1c00408] [Medline: 37566375]
27. Bogler A, Packman A, Furman A, Gross A, Kushmaro A, Ronen A, et al. Rethinking wastewater risks and monitoring in light of the COVID-19 pandemic. Nat Sustain 2020;3(12):981-990 [FREE Full text] [doi: 10.1038/s41893-020-00605-2]
28. Hillary LS, Farkas K, Maher KH, Lucaci A, Thorpe J, Distaso MA, et al. Monitoring SARS-CoV-2 in municipal wastewater to evaluate the success of lockdown measures for controlling COVID-19 in the UK. Water Res 2021;200:117214 [FREE Full text] [doi: 10.1016/j.watres.2021.117214] [Medline: 34058486]
29. Xu Y, Li X, Zhu B, Liang H, Fang C, Gong Y, et al. Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. Nat Med 2020;26(4):502-505 [FREE Full text] [doi: 10.1038/s41591-020-0817-4] [Medline: 32284613]
30. Usman M, Farooq M, Hanna K. Existence of SARS-CoV-2 in wastewater: implications for its environmental transmission in developing communities. Environ Sci Technol 2020;54(13):7758-7759 [FREE Full text] [doi: 10.1021/acs.est.0c02777] [Medline: 32525667]
31. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 2020;26(5):672-675 [FREE Full text] [doi: 10.1038/s41591-020-0869-5] [Medline: 32296168]
32. Medema G, Heijnen L, Elsinga G, Italiaander R, Brouwer A. Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in the Netherlands. Environ Sci Technol Lett 2020;7(7):511-516 [FREE Full text] [doi: 10.1021/acs.estlett.0c00357] [Medline: 37566285]
33. Naughton CC, Roman FA, Alvarado AGF, Tariqi AQ, Deeming MA, Kadonsky KF, et al. Show us the data: global COVID-19 wastewater monitoring efforts, equity, and gaps. FEMS Microbes 2023;4:xtad003 [FREE Full text] [doi: 10.1093/femsmc/xtad003] [Medline: 37333436]
34. Liu Z, Jiang Z, Kip G, Snigdha K, Xu J, Wu X, et al. An infodemiological framework for tracking the spread of SARS-CoV-2 using integrated public data. Pattern Recognit Lett 2022;158:133-140 [FREE Full text] [doi: 10.1016/j.patrec.2022.04.030] [Medline: 35496673]

35. Farkas K, Hillary LS, Thorpe J, Walker DI, Lowther JA, McDonald JE, et al. Concentration and quantification of SARS-CoV-2 RNA in wastewater using polyethylene glycol-based concentration and qRT-PCR. Methods Protoc 2021;4(1):17 [FREE Full text] [doi: 10.3390/mps4010017] [Medline: 33672247]
36. Oberacker P, Stepper P, Bond DM, Höhn S, Focken J, Meyer V, et al. Bio-On-Magnetic-Beads (BOMB): open platform for high-throughput nucleic acid extraction and manipulation. PLoS Biol 2019;17(1):e3000107 [FREE Full text] [doi: 10.1371/journal.pbio.3000107] [Medline: 30629605]
37. Vogels CBF, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT-qPCR primer-probe sets. Nat Microbiol 2020;5(10):1299-1305 [FREE Full text] [doi: 10.1038/s41564-020-0761-6] [Medline: 32651556]
38. Stachler E, Kelty C, Sivaganesan M, Li X, Bibby K, Shanks OC. Quantitative CrAssphage PCR assays for human fecal pollution measurement. Environ Sci Technol 2017;51(16):9146-9154 [FREE Full text] [doi: 10.1021/acs.est.7b02703] [Medline: 28700235]
39. Kitajima M, Tohya Y, Matsubara K, Haramoto E, Utagawa E, Katayama H. Chlorine inactivation of human norovirus, murine norovirus and poliovirus in drinking water. Lett Appl Microbiol 2010;51(1):119-121. [doi: 10.1111/j.1472-765X.2010.02869.x] [Medline: 20536708]
40. Wilde H, Perry WB, Jones O, Kille P, Weightman A, Jones DL, et al. Accounting for dilution of SARS-CoV-2 in wastewater samples using physico-chemical markers. Water 2022;14(18):2885 [FREE Full text] [doi: 10.3390/w14182885]
41. Gov.UK Coronavirus (COVID-19) in the UK. UK Government. URL: https://coronavirus.data.gov.uk/details/cases [accessed 2021-12-01]
42. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
43. Cuff JP, Dighe SN, Watson SE, Badell-Grau RA, Weightman AJ, Jones DL, et al. An infoveillance analysis of public interest, national data and wastewater monitoring in Wales, UK. Zenodo. 2022 Oct 28. URL: https://zenodo.org/records/7260269 [accessed 2023-11-06]
44. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. JOSS 2019;4(43):1686 [FREE Full text] [doi: 10.21105/joss.01686]
45. Harrell FE. Hmisc: Harrell Miscellaneous. The Comprehensive R Archive Network. 2023 Sep 12. URL: https://cran.r-project.org/web/packages/Hmisc/index.html [accessed 2023-11-06]
46. Wei T, Simko V, Levy M, Xie Y, Jin Y, Zemla F, et al. R package “corrplot”: visualization of a correlation matrix. The Comprehensive R Archive Network. 2022 Oct 12. URL: https://cran.r-project.org/web/packages/corrplot/corrplot.pdf [accessed 2023-10-24]
47. Garnier S, Ross N, Rudis B, Sciaini M, Scherer C. Package 'viridis': default color maps from 'matplotlib'. The Comprehensive R Archive Network. 2018 Feb 2. URL: http://cran.nexr.com/web/packages/viridis/viridis.pdf [accessed 2023-10-24]
48. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Cham: Springer International Publishing; 2016.
49. Neuwirth E. RColorBrewer: R Color Brewer's palettes. R Graph Gallery. 2014. URL: https://r-graph-gallery.com/38-rcolorbrewers-palettes.html [accessed 2023-10-24]
50. Husain I, Briggs B, Lefebvre C, Cline DM, Stopyra JP, O'Brien MC, et al. Fluctuation of public interest in COVID-19 in the United States: retrospective analysis of Google trends search data. JMIR Public Health Surveill 2020;6(3):e19969 [FREE Full text] [doi: 10.2196/19969] [Medline: 32501806]
51. Nindrea RD, Sari NP, Lazuardi L, Aryandono T. Validation: the use of Google trends as an alternative data source for COVID-19 surveillance in Indonesia. Asia Pac J Public Health 2020;32(6-7):368-369 [FREE Full text] [doi: 10.1177/1010539520940896] [Medline: 32643957]
52. Xiao K, Zhang L. Wastewater pathogen surveillance based on one health approach. Lancet Microbe 2023;4(5):e297 [FREE Full text] [doi: 10.1016/S2666-5247(23)00039-3] [Medline: 36848921]
53. Miłobedzka A, Ferreira C, Vaz-Moreira I, Calderón-Franco D, Gorecki A, Purkrtova S, et al. Monitoring antibiotic resistance genes in wastewater environments: the challenges of filling a gap in the one-health cycle. J Hazard Mater 2022;424(Pt C):127407 [FREE Full text] [doi: 10.1016/j.jhazmat.2021.127407] [Medline: 34629195]
54. Summary statistics by regions of Wales: 2020. Llywodraeth Cymru Welsh Government. 2020. URL: https://www.gov.wales/summary-statistics-regions-wales-2020 [accessed 2023-10-24]

Abbreviations

CDC: Centers for Disease Control and Prevention
GLM: generalized linear model
LM: linear model
MLM: multivariate linear model
qPCR: quantitative polymerase chain reaction
RSV: relative search volume
RT: reverse transcription

RT-qPCR: quantitative reverse transcription polymerase chain reaction

Edited by T Mackey; submitted 28.10.22; peer-reviewed by V Brahmbhatt, M Haupt, O Leal Neto; comments to author 14.06.23; revised version received 15.08.23; accepted 30.09.23; published 23.11.23.

Please cite as:
Cuff JP, Dighe SN, Watson SE, Badell-Grau RA, Weightman AJ, Jones DL, Kille P
Monitoring SARS-CoV-2 Using Infoveillance, National Reporting Data, and Wastewater in Wales, United Kingdom: Mixed Methods Study
JMIR Infodemiology 2023;3:e43891
URL: https://infodemiology.jmir.org/2023/1/e43891
doi:10.2196/43891
PMID:37903300

Original Paper

Attitudes of Swedish Language Twitter Users Toward COVID-19 Vaccination: Exploratory Qualitative Study

Safwat Beirakdar[1], BSc, MSc; Leon Klingborg[1], BSc, MSc; Sibylle Herzig van Wees[1], BA, MSc, PhD

Karolinska Institute, Department of Global Public Health, Stockholm, Sweden

Corresponding Author:
Sibylle Herzig van Wees, BA, MSc, PhD
Karolinska Institute
Department of Global Public Health
K9 Global folkhälsa
Tomtebodavägen 18A
Stockholm, 17177
Sweden
Phone: 46 735351190
Email: sibylle.hvw@ki.se

Abstract

Background: Social media have played an important role in shaping COVID-19 vaccine choices during the pandemic. Understanding people's attitudes toward the vaccine as expressed on social media can help address the concerns of vaccine-hesitant individuals.

Objective: The aim of this study was to understand the attitudes of Swedish-speaking Twitter users toward COVID-19 vaccines.

Methods: This was an exploratory qualitative study that used a social media–listening approach. Between January and March 2022, a total of 2877 publicly available tweets in Swedish were systematically extracted from Twitter. A deductive thematic analysis was conducted using the World Health Organization's 3C model (*confidence*, *complacency*, and *convenience*).

Results: *Confidence* in the safety and effectiveness of the COVID-19 vaccine appeared to be a major concern expressed on Twitter. Unclear governmental strategies in managing the pandemic in Sweden and the belief in conspiracy theories have further influenced negative attitudes toward vaccines. *Complacency*—the perceived risk of COVID-19 was low and booster vaccination was unnecessary; many expressed trust in natural immunity. *Convenience*—in terms of accessing the right information and the vaccine—highlighted a knowledge gap about the benefits and necessity of the vaccine, as well as complaints about the quality of vaccination services.

Conclusions: Swedish-speaking Twitter users in this study had negative attitudes toward COVID-19 vaccines, particularly booster vaccines. We identified attitudes toward vaccines and misinformation, indicating that social media monitoring can help policy makers respond by developing proactive health communication interventions.

(JMIR Infodemiology 2023;3:e42357) doi:10.2196/42357

KEYWORDS

COVID-19; vaccine hesitancy; COVID-19 vaccines; social media; Twitter; qualitative analysis; World Health Organization; WHO's 3C model

Introduction

Background

Sweden was seriously affected by the pandemic compared with some European countries. In relation to neighboring Nordic countries, Sweden has had the highest number of infected and terminally ill patients, with 2,500,000 positive cases and >16,000 deaths as of February 2022 [1,2].

The Public Health Agency of Sweden implemented various interventions and strategies to speed up COVID-19 vaccine uptake, including media campaigns to promote vaccination and facilitate vaccination accessibility across the Stockholm region through mobile vaccination buses [3]. However, vaccine hesitancy among the public arguably slowed down the vaccination process, and a small percentage of the public is still showing reluctance to COVID-19 vaccination in general and boosters in particular [4]. According to the Public Health Agency of Sweden, 86.4% of the population received 2 doses of the COVID-19 vaccine. Uptake of the third vaccine has been lower, and only 66.6% of the Swedish population who are eligible for a third vaccine have been vaccinated [5]. Although these numbers are not alarming in comparison with other countries, Sweden has in the past few years encountered persistent vaccine hesitancy and the circulation of rumors about vaccines in certain migrant communities, in communities that hold fringe political views, and in anthroposophic communities [6].

Globally, skepticism about vaccine effectiveness and safety has been a consistent challenge, and the rise in vaccine hesitancy has become an urgent concern and one of the top 10 threats to global health in 2019, according to the World Health Organization (WHO) [7]. Several factors contribute to the personal decision to take the vaccine, but social media have played an important role in promoting vaccine scarcity. Social media accelerated the spread of misinformation by providing a platform for vaccine-hesitant communities to spread rumors, ultimately shaking public trust in the COVID-19 vaccine [8,9].

Since the beginning of the COVID-19 pandemic, myths and rumors have circulated on social media regarding the virus's origin, spread, symptoms, severity, treatments, and the safety and effectiveness of its vaccines [10]. Some of these rumors include concerns about the safety of COVID-19 vaccines because of their rapid development and the use of the novel concept of messenger RNA, which some claimed causes infertility [11]. Furthermore, conspiracy theories linking the spread of the virus to 5G mobile technology and implanted microchips have been prominent on social media [10]. In Sweden, there is little information about what rumors circulate in the Swedish language on social media.

The WHO defines false information that systematically spreads in time of disease outbreaks as *infodemics* [12]. Infodemics constitute the rapid proliferation of harmful messages through social media platforms, causing confusion and mistrust among the public [12]. For instance, social media platforms played an important role in polarizing the public against the human papillomavirus vaccination in Japan in 2013 [13], where negative media campaigns overtook the scientific evidence provided by local authorities, leading to a decline in vaccine uptake to less than 1% [13]. Another incident was seen in Denmark, where public confidence in the human papillomavirus vaccination dropped significantly after the spread of a documentary based on teenagers' experiences with complications after getting vaccinated [14]. Moreover, increasing evidence suggests that negative vaccine posts on social media contribute to vaccine hesitancy by altering the risk perceptions of individuals [15]. An experimental study by Betsch et al [15] demonstrated that 5-10 minutes of exposure to such materials is sufficient to trigger negative attitudes about vaccination. Similarly, a study on the uptake of the influenza vaccine showed that uptake was lower among people who were exposed to misinformation distributed on the internet [16]. Given the growing concern over fading confidence in the COVID-19 vaccine and as little knowledge is available on rumors and misinformation in Sweden, this study examines potential traces of *infodemics* that are at play in the Swedish Twitter discourse about COVID-19 vaccines.

Objective

The European Center for Disease Prevention and Control published a report in June 2021 that encouraged member states to gain a better understanding of the misinformation landscape on social media [17]. In Sweden, several surveys have been conducted by the Public Health Agency of Sweden in the past 2 years to measure the public acceptance of COVID-19 vaccines. However, to date, no study has been published on the public discourse found on social media platforms, such as Twitter, where people express their opinions without probing from researchers. Previous literature has focused more on immigrants,such as those from the Somali community in Stockholm and the anthroposophic communities [6], as these 2 groups have shown a pattern of vaccine hesitancy [6]. Hence, there is a gap in the literature in exploring vaccination concerns and rumors among people who are active on social media in Sweden, and this study aimed to contribute to this knowledge.

Methods

Study Design

This was an exploratory qualitative study that used a social media–listening approach. Data from Twitter were gathered using Netlytic, a wrapper for the Twitter application programming interface (API; version 1), and Boolean operator search queries. Qualitative deductive thematic data analysis was guided by the WHO's 3C model: confidence, complacency, and convenience. The WHO's 3C model classifies the factors influencing vaccine hesitancy in individuals or groups into 3 main categories: confidence, complacency, and convenience [18,19]. Confidence refers to both trust in the effectiveness and safety of the vaccine and trust in governmental policies and motivation behind recommending the vaccine [7]. Complacency is related to the level of risk that individuals perceive in terms of becoming infected with the disease, thereby shaping their personal belief in the necessity of vaccination [7]. Finally, convenience refers to the availability and accessibility of vaccination and is also related to the quality of vaccination services [7].

Twitter was selected as the main data source because it is an important social media platform for disseminating information and sharing opinions [20]. It is a popular and trusted source used by many governmental agencies, political leaders, and famous influencers to address and interact with the public [21]. In addition, compared with other social media platforms, Twitter provides greater access to data and the ability to retrieve real-time data [22]. Twitter allows users to post pictures, videos, and "tweets" that constitute short texts with a maximum of 280 characters [22], and users interact with each other and engage in conversations using the like, reply, and retweet features [20].

Data Extraction

The study analyzed public attitudes by reviewing tweets posted in Swedish. This was accomplished through "Netlytic," a web-based service that allows for real-time data scraping from various social media platforms that publish publicly available posts [23,24]. Specifically, this study used Netlytic's wrapper and interface for the Twitter API. Netlytic has been used in multiple social media–listening studies [24]. The parameters in Netlytic were set to capture tweets in Swedish using complex search queries linked with Boolean operators (OR and AND), instead of single terms and hashtags, to obtain more relevant tweets when retrieving data from the Twitter API [22].

Boolean search queries were set based on the most-used hashtags and terms in Sweden regarding COVID-19 according to Google Trends, Statista, and Twitter [25]. The final search queries are presented in Textbox 1. Only neutral and general terms were used to avoid skewing the data and influencing results.

To reduce duplicates, Netlytic was set to exclude retweets while importing data. The study did not filter according to geographical location, as many users chose to hide their location for privacy concerns. Data scraping was scheduled to run the search queries weekly to match the Netlytic settings, because tweets older than 1 week would not be captured [23].

The data scraping covered 2 months from January 24, 2022, to March 24, 2022. The timeline reflected an important period of many changes, including the start of the booster shot recommendation [26] and the dominance of "omicron," a new variant that is highly transmissible and less susceptible to vaccines [27,28]. In addition, by February 9, 2022, Sweden entered a new phase of the pandemic, where all restrictions implemented to control the virus were removed [29].

Textbox 1. Boolean search queries used in the study.

(Corona OR covid OR coronaviruset OR coronavirussverige OR coronasverige OR coronavirus OR COVID-19) AND (spruta OR vaccin OR coronavaccin OR vaccinspruta OR coronaspruta OR påfyllnadsdos OR tredjedos)

Sample Size

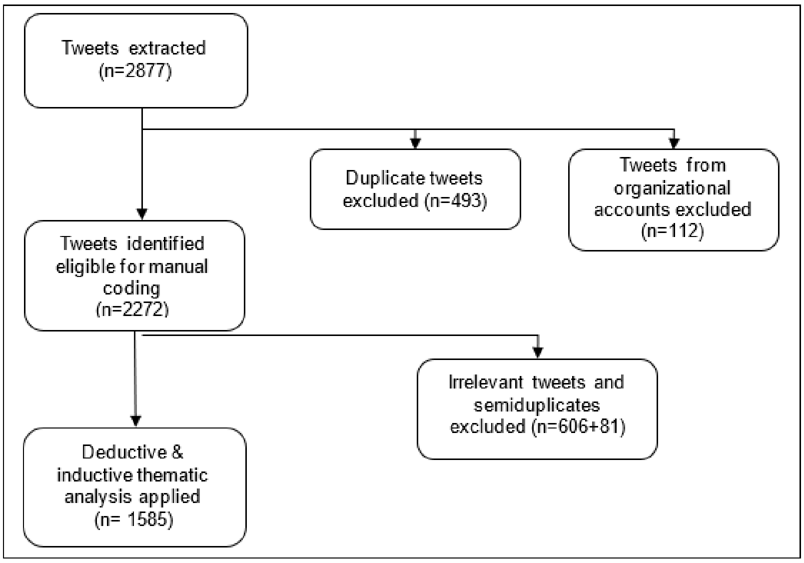
All tweets from scraping iterations were merged into a single data set. The total number of retrieved tweets was 2877, which underwent cleaning and eligibility screening phases (Figure 1).

The master sheet was cleaned from duplicate tweets (n=493), which included copy-pasted text with no changes. In addition, as the study aimed to explore individuals' attitudes, tweets from organizational accounts (n=112), such as RegionStockholm, Krisinformation, WHO, Public Health Agency of Sweden, and Dagensnyheter, were removed from the data set.

Account names were also removed from the data set for ethical considerations.

The eligibility screening phase was conducted using the qualitative data analysis software NVivo (version 12 Pro; QSR International) with 2272 tweets. Tweets that did not present personal opinions about COVID-19 vaccines, were irrelevant to the research topic, or contained unclear statements were coded as irrelevant and excluded from the study (n=606). Irrelevant posts were predominantly posts that did not present personal opinions (including news, posts from organizations, and advertisements), and there were a few posts that were not included because the statement was not legible. In addition, semiduplicated tweets that included changes but did not present additional context compared with their original tweets were also excluded (n=81). As a result, tweets that contained a clear attitude related to the COVID-19 vaccines—whether the tweets were in favor of vaccination or skeptical toward it—were eligible for the qualitative analysis (n=1585). All tweets found eligible (n=1585) were included in the final sample.

Figure 1. Tweet selection flow.



Data Analysis

The data were analyzed using qualitative thematic analysis (TA), inspired by the 2006 guide by Braun and Clarke [30]. TA was chosen because of its flexibility in answering research questions [30]. Moreover, TA is suitable for large amounts of data, as it provides a rich and inclusive analysis by reflecting the nuances within the data [30,31].

The TA was also led by the 3C model: confidence, complacency, and convenience [4]. This model has been used in many vaccine studies to understand the factors that influence vaccination [32,33]. Qualitative coding was conducted blindly by 2 individual coders (SB and SHvW) using NVivo 12. A mixture of deductive and inductive qualitative methods was applied by adopting TA according to Braun and Clarke [30,34]. This hybrid approach allowed for the flexibility of creating categories that emerged from the data rather than relying solely on the 3C framework. The data set was reviewed to gain familiarity with the data. Next, the themes were deductively predetermined, and the 3C of confidence, complacency, and convenience were set in NVivo. Subsequently, the tweets were inductively coded according to their meanings within the corresponding themes. For each theme, tweets were organized within categories according to emerging patterns. Themes and categories were not mutually exclusive; however, a tweet could be coded into one or more themes or categories. The themes' titles were adjusted according to the findings. The resulting coding tree was discussed and agreed between the two coders.

Trustworthiness

To ensure the trustworthiness of the qualitative research process, we applied the following strategies [35]: First, we applied clear criteria for the purposive sampling strategy. Second, the coding was completed by 2 researchers blindly. We also applied the overall peer scrutiny of the research project, whereby the research team regularly discussed emerging research challenges. This included reflection on their own backgrounds, which may lead to bias [35].

Ethical Considerations

No ethics approvals were needed as the study analyzed publicly available data on the internet. All tweets identified in the study were anonymized after screening for eligibility to protect the privacy of users. Aspects of confidentiality and anonymization of data were respected as no data used in the final report can be linked to actual users. One of the measures taken to protect the identity of Twitter users behind the tweets in our data set was to translate all the quotes used in the analysis, ensuring that they cannot be traced back to their author. In addition, no interaction occurred between the study researchers and Twitter users. This approach is consistent with guidelines on the ethical conduct of qualitative research in web-based communities [36,37].

Results

Overview of Themes

The analysis resulted in 3 main themes and 18 categories, guided by the WHO's 3C model (Textbox 2). The themes included confidence—safety and effectiveness concerns and mistrust in authorities; complacency—fading belief in vaccination necessity; and convenience—unappealing vaccination services and unclear information. In this section, each theme is described, and selective quotes are used to present the categories.

Textbox 2. Study results.

Themes and categories

- Confidence: safety and effectiveness concerns and mistrust in authorities
  - Concerns about messenger RNA COVID-19 vaccines safety and side effects
  - COVID-19 vaccines effectiveness is limited
  - The risks from the vaccines outweigh the benefit
  - Concerns about the number of booster shots
  - The spread of rumors and conspiracy theories
  - Lack of transparency from the government and the Public Health Agency of Sweden
  - Limited trust in the authority’s management of the COVID-19 pandemic
  - Against mandatory vaccination and vaccine passes
  - The media presented biased evidence in favor of COVID-19 vaccination
  - Mistrust in scientific experts and pharmaceutical companies
- Complacency: fading belief in vaccination necessity
  - The perceived necessity of vaccination against COVID-19 is low, especially among healthy adults
  - Changes in the perceived risk of COVID-19 infection
  - Natural immunity is superior to vaccines
  - Conflicting opinions toward children’s vaccination against COVID-19
- Convenience: unappealing vaccination services and unclear information
  - Limited availability of COVID-19 vaccination appointments
  - COVID-19 vaccination services are unorganized
  - Contradicting evidence
  - Unanswered questions related to COVID-19 vaccination

Confidence: Safety and Effectiveness Concerns and Mistrust in Authorities

The analysis shows that *confidence* is an important barrier to COVID-19 vaccination uptake.

Safety and Effectiveness Concerns

The analysis demonstrated that there are multiple safety concerns related to COVID-19 vaccines. There is a shared belief that the messenger RNA vaccines are produced too quickly and do not undergo all the testing processes required for approval. This rapid development of vaccines has resulted in side effects:

> *It takes several years to develop and test vaccines before they are released on the market. The C-vax [COVID-19 vaccines] is quickly developed and emergency-approved in just a few months. Therefore, more people get side effects than COVID.*

In addition to the number of side effects, there was a specific concern regarding the severity of side effects and deaths related to vaccination:

> *COVID-19 injections have probably caused 2-50 million deaths and many more disabling injuries worldwide.*

In addition, some tweets compared the number and severity of COVID-19 vaccines’ side effects to the H1N1 pandemic influenza vaccine Pandemrix, for example:

> *Side-effects reported by the Medical Products Agency. Right now, just over 95,000 are reported, just over 17,000 handled and just over 9,000 treated as serious. Compare with Pandemrix which had 300 severe, narcolepsy.*

In terms of effectiveness, the data show that Twitter users in this study perceived vaccines as prophylactic injections, which reduce the severity of the infection. However, they did not consider them as effective as traditional vaccines. Many tweets expressed people’s frustration with becoming infected after being vaccinated:

> *Vaccines usually prevent diseases, right? At least the ones I have taken from birth onwards. The current COVID-19 “vaccine” is useless.*

The skepticism and concerns about the vaccines seen in the data were primarily related to COVID-19 vaccines, as many tweets clearly expressed trust in other vaccines:

> *Being against the COVID-19 vaccine does not mean that you are against all vaccines.*

However, some people expressed mistrust of future vaccines:

> *The failure of the COVID-19 vaccine makes me hesitant about any future vaccine.*

Moreover, many people argued that the risk of vaccination outweighs its benefits:

> *There's no way I'm taking a third Covid-19 syringe! The side effects after both make me give up. In addition, my daughter and her husband were sick, really sick, despite two syringes! So no thanks!*

Several tweets reported safety concerns regarding the number of injections the body can tolerate and the immunity period provided by the vaccines, which are continuously decreasing. There were many sarcastic tweets on the booster shots:

> *You need to take booster shots until you are dead.*

To add to the uncertainty about effectiveness, there were repeated rumors, myths, and misinformation about the COVID-19 vaccine, for example, that COVID-19 vaccines cause AIDS by weakening the immune system of the human body:

> *The more "vaccines" a person receives against the COVID-19 coronavirus, the faster the body will die from the AIDS-like immune loss syndrome!*

Moreover, theories regarding the origin of the virus are rampant. Some argue that the virus was synthesized by political forces and that COVID-19 vaccines are biological weapons used against the public:

> *The virus came from a lab in Wuhan that was sponsored by the US government to conduct "gain-of-function" research that was banned by Obama.*

Mistrust in Authorities

The data show that there was mistrust in the statistics and numbers related to the COVID-19 vaccines' side effects and infection rates published by the authorities. For example, it was highlighted that the Public Health Agency of Sweden did not report the full numbers to encourage vaccination:

> *The problem is generally that the numbers are inflated and unreliable...Why not report who died OF COVID? Why are patients with syringes 1 + 2 reported as unvaccinated? Why not report figures for the unvaccinated?*

The tweets in this study show that Sweden's regulation was inconsistent with those of other countries. In addition, many tweets criticized the government's delay in taking action, which led to serious consequences. For example, they accused the government of not protecting older adults:

> *Sweden's strategy can never be "right." A choice was made, in February 2020, where it was decided that it was ok to let the elderly get sick and die before they knew how to cure COVID-19 or have a vaccine. It is morally indefensible.*

Furthermore, Twitter users in this study expressed skepticism based on governmental recommendations concerning COVID-19 vaccines:

> *On 12 January, the Swedish Public Health Agency stated that the vaccine protects well against serious illness, also against the omicron variant, for more than 6 months. That was less than THREE weeks ago! They currently have no idea what they are doing.*

The enforcement of the vaccine pass was further criticized. The tweets expressed their disapproval of mandatory vaccination, even for people who took the vaccine, as it was perceived as a violation of personal freedom:

> *I have taken two doses of the vaccine and became ill with corona. The vaccine pass does not reduce the spread of infection, it is only a way to control people.*

Moreover, the media agencies were criticized for being biased toward the government, where they blindly supported governmental decisions and undermined space for critical opinions.

The last category within this theme is the mistrust of scientific experts and pharmaceutical companies. The data demonstrate that there was a common belief that pharmaceutical companies benefit the most from the sale of vaccines. Some tweets suggested that scientific experts were pushed to ignore other potential factors to promote vaccination:

> *Everything that was not done to prevent the spread of COVID-19 that could have worked, ivermectin [ivermectin is an antiparasitic drug used by some countries to treat COVID-19], Vitamin D, etc. Instead, they all invested in one card—vaccination with a vaccine they had not tried before.*

Complacency: Fading Belief in Vaccination Necessity

Tweets analyzed under the *complacency* theme showed that young healthy adults felt they were not at risk. Instead, Twitter users in this study believed that only the older adults and people with chronic diseases were at high risk of hospitalization:

> *No one under the age of 50 would have become sicker without a vaccine. It would most likely have been just as mild symptoms anyway. Greater risk of crossing the street than getting seriously ill in COVID-19 if you are healthy and younger.*

Many studies have compared COVID-19 infection with the usual influenza infection. In addition, the analysis of the tweets highlights that the new mild variants negatively affect people's willingness to be vaccinated, even among infected individuals who reported strong symptoms:

> *Now COVID-19 is like a severe cold, I had Omicron now, had pain in the body, a little runny nose, sore throat, headache where I thought the eyes would fall out a little awkwardly with asthma but always so with a cold. I did not need a vaccine for this.*

Moreover, there was a common belief that direct infection provides better protection than the immunity provided by vaccines. Many tweets expressed that individuals would rather be infected with COVID-19 than get vaccinated:

> *No one who has had COVID-19 needs to be vaccinated, natural immunity is superior to the temporary protection that this vaccine provides.*

Swedish Twitter users in this study expressed a strong trust in natural immunity; therefore, many tweets encouraged other people to get infected to be protected:

> *Omicron is just a vaccine without a reservation* [for a time slot for vaccination]

Conflicting opinions were observed among the retrieved tweets on the necessity of vaccinating children against COVID-19, where the same arguments used to demand vaccination for children were used against vaccination but in a different context. Many tweets argued that children were also at risk of contracting COVID-19, contrary to what was previously believed. Tweets in favor of vaccination highlight that it is a child's right to get vaccinated and be provided with the best possible care:

> *COVID-19 is to be spread and children between 5-11 years are deprived of the opportunity to be protected with a vaccine.*

In addition, some Twitter users expressed fear for their children, as they can become severely ill and require hospitalization:

> *Many children are cared for in COVID-19 hospitals. The fact that they are offered vaccines is important to reduce the risk of them being seriously affected.*

By contrast, tweets against vaccinating children expressed disbelief in the other group's evidence, where they insisted that there is certainly no reason to fear COVID-19 infection:

> *COVID-19 is not dangerous for children. This is just propaganda because they want to throw vaccines at everyone when they have invested so much in it.*

Similar to what was found in terms of the low perceived risk of COVID-19 in healthy adults, many people expressed that the risk of becoming seriously ill among healthy children is relatively low:

> *There is still no reason at all to vaccinate healthy children. Children who have risk factors are another matter, but there is in principle no healthy child in the whole world who has died from COVID-19 during the latter part of the pandemic.*

Convenience: Unappealing Vaccination Services and Unclear Information

The *convenience* theme revealed that some tweets discussed unappealing vaccination services and unclear health information. The analysis showed that there were complaints regarding the limited availability of vacant slots:

> *I'm unvaccinated, my wife had COVID-19 last week. It was a severe flu with a high fever for a few days...I was going to get vaccinated, but it was hard to find anything near where I live.*

Available vaccination appointments were especially a problem for booster shots, and those who managed to get vaccinated complained of long waiting queues:

> *Today I took the 3rd vaccine against COVID-19. Cheers to us who stood in line for about an hour.*

Moreover, some tweets revealed dissatisfaction with the vaccination system, as they did not receive invitations for their doses according to the published guidelines:

> *Tested the phone booking and seemed to be free to come forward. Most people wonder why I did not receive an offer...According to 1177 [Swedish health information website and number], those who received the second dose in mid-August will receive an offer today, I received it at the end of July.*

In addition, many people on Twitter felt lost while following contradicting evidence and information related to COVID-19 vaccines distributed on the internet. Many people have highlighted their limited ability to understand scientific reports:

> *I think people have a hard time understanding that what is coming out here is true. There are research articles that claim completely different things, so it is not surprising that people get confused.*

Many Twitter users felt that they needed more clarification regarding the COVID-19 vaccines. Tweets described insufficient information on the effects and immunity provided by the COVID-19 vaccines.

> *Why would it make sense to get vaccinated when you have had COVID-19? Why not highlight the risks of vaccines as well as the benefits?*

In addition, many questions were related to the COVID-19 booster intervals. Twitter users in this study demonstrated a low understanding of the dose guidelines and how they should schedule their boosters after getting infected:

> *Some thoughts about vaccination. How do you do it if you just had COVID-19, do you take booster 3 or should you wait a couple of months?*

Some tweets were very specific in that they asked questions related to certain medical conditions or age groups:

> *Look at the risk to the foetus/mother. These are extremely low if the mother is healthy, not overweight...How much risk should COVID-19 pose to recommend a vaccine where the clinical studies are not complete?*

Discussion

Principal Findings

The overall aim of the study was to understand the attitudes of Swedish Twitter users toward COVID-19 vaccines. The study found that tweets expressing opinions about vaccines and the vaccination process were predominantly negative. The tweets expressed low confidence in the COVID-19 vaccines, policy makers, and scientific experts. Further concerns were related to complacency, which reflected a low understanding of the severity of COVID-19 infection and a low perception of the necessity to vaccinate, particularly with regard to booster shots. Moreover, the study found that convenience was not seen as a major challenge; however, the accessibility of information and the quality and availability of vaccination services were criticized.

Swedish Twitter users in this study had major concerns about the safety of the vaccines. This supports the findings of other studies that highlight the importance of people's confidence in vaccine safety in promoting vaccination uptake [13,14,38]. Moreover, the results indicate the presence of rumors about the vaccine. For example, the fear of acquiring AIDS from vaccines could have undermined people's willingness to be vaccinated, and there have been numerous studies that have highlighted rumors and misinformation about COVID-19 vaccines [5,39].

This study further demonstrates that people's beliefs in the effectiveness of COVID-19 vaccines have decreased over time. The findings show that these arguments were commonly raised against COVID-19 vaccines and booster shot uptakes and might be related to the low booster coverage seen among the Swedish population [5].

The data further show that people express mistrust in authorities and demand more transparency from the government and the responsible authorities, mainly from the Public Health Agency of Sweden and the Swedish Medical Products Agency (SMPA), about the incidence and severity of the side effects caused by the vaccines. The SMPA releases a monthly update of the registered side effects and a list of death cases suspected to be related to COVID-19 vaccines [40,41]. This indicates that the information was available but might not have been effectively shared with the public. The mistrust in the government and health authorities' management of the COVID-19 pandemic found in this study is a new finding and is not consistent with previous studies in the field. A survey from 2020 found that most of the Swedish population supported the government's strategy in managing the pandemic and had strong trust in health authorities [42]. However, these findings do not necessarily reflect the opinion of the population today; in fact, the findings from our study indicated that trust in the authority's management might have been negatively affected by the continuous changes in guidelines for taking booster shots.

Tweets analyzed under the *complacency* theme suggested that the perceived severity of COVID-19 infection was low, and consequently, that the perceived importance of the vaccine has been fading. This may be largely because of rumors and limited knowledge of vaccines. This study shows a widespread belief in the superiority of natural immunity and the low risk of COVID-19 infection in healthy individuals. The data also show that people were actively encouraging others to get infected rather than get vaccinated. These results are consistent with a Portuguese study, which documented that a low perceived risk among healthy adults contributes to their vaccine hesitancy [38]. The SMPA and the Centers for Disease Control and Prevention warned against these beliefs and emphasized that the risks associated with COVID-19 infection are greater than the risks of taking the vaccine, and that the immune response against the infection is not foreseeable; therefore, no groups are protected from becoming seriously ill [43,44]. This study further found conflicting views on the risks and benefits of vaccinating children against COVID-19. Users presented opposing evidence regarding the risk of infection and safety of vaccines for children. These findings are consistent with the Public Health Agency of Sweden's survey results, which highlight the uncertainty among parents regarding their children's vaccination [45].

The *convenience* concerns expressed among the tweets were related to limited access to and availability of vaccination appointments. The findings implied that although drop-in vaccination services were introduced, better organization shortened the queue times. These results suggest that enhancing the efficiency of vaccination services could encourage vaccination. Furthermore, this study reveals that there was a lack of knowledge about COVID-19 and the vaccines, which could arguably have resulted in people turning to social media to seek answers to their unanswered questions. This aligns with a survey conducted in 2020 that showed that part of the Swedish population was not satisfied with the information provided by health authorities on COVID-19 [42]. There is growing evidence of a lack of information causing hesitancy. For instance, a study conducted in the United States before the development of COVID-19 vaccines found that people were willing to get vaccinated if they received adequate information about the vaccines [46], whereas another study on the influenza vaccine found that those with better influenza literacy had higher chances of choosing vaccination [47].

Limitations

This study has some limitations. It is important to note that Twitter API provides free access to only a 1% sample of all Twitter data [48], thus limiting the generalization of the findings. In addition, the study timeline was limited to 2 months, which does not represent the general perspective of the web-based population. Longer studies would strengthen the validity of these results.

In addition, although the number of web-based users in Sweden has increased in recent years, they cannot be considered representative of the entire Swedish population. As it was not possible to capture users' demographics for technical reasons, as such details are not available on Twitter, the transferability of the study is also limited because of the lack of such information. However, studies on the demographics of Twitter users have shown that the web-based population constitutes the younger generation, with females especially overrepresented [49]. Furthermore, the study's results and parameters are specific to the Swedish context; thus, a similar study in other settings could present different concerns and opinions.

Finally, there are limitations to language-restricted searches as they may include Swedish expats who live in a different context than Sweden. It is of course possible that expats are part of the discussion; according to this report, as many as 700,000 Swedes actually live abroad, or around 7% of Swedish people [50]. However, as they are a relatively small proportion of the total population, we doubt that they heavily skew our findings.

Public Health and Practical Implications

This study contributes to ongoing public health efforts to promote COVID-19 vaccination and address vaccine hesitancy, particularly in Sweden. These data were collected in early 2022, when COVID-19 vaccine coverage for 2 doses was nearly 90%

of the eligible Swedish population. Since then, Sweden has experienced a decrease in the uptake of the third dose [5]. Weaning confidence in booster vaccines was observed in this study. This highlights the importance of monitoring and analyzing public sentiments regarding health-related matters, particularly vaccine decision-making, on social media platforms. As the number of social media users is rapidly increasing, the social media landscape has emerged as an important platform that must be considered when working with public health awareness. Moreover, this study provides evidence of the dominance of negative attitudes on social media, which forms a threat to public health and needs to be addressed.

This study indicates that COVID-19 vaccine acceptance is not unchallenged and should be closely monitored to address emerging COVID-19 vaccine hesitancy. It also shows that social media studies provide valuable insights into the factors that shape public attitudes toward vaccination. Furthermore, the evidence illustrates that clear health communication and consistent messages are needed to maintain public trust. Moreover, this study emphasizes the importance of addressing the spread of rumors and misinformation to overcome COVID-19 vaccine hesitancy and its potential implications for future vaccines.

This study contributes to the existing literature on COVID-19 vaccines by exploring the attitudes of Swedish users on the web. Still, further social media studies are needed to explore and quantify attitudes toward COVID-19 vaccines on the entire spectrum of social media platforms.

Public trust in government, experts, and authorities can be reinforced by facilitating open dialogue and channels with the public. Furthermore, innovative approaches, such as internet-based interventions to address the growing web-based community, could be considered to increase public trust in COVID-19 vaccines. The Public Health Agency of Sweden and other health-related authorities should expand their presence on the web to provide accurate information on various social media platforms. Additional resources should be considered to increase the quality of vaccination services and the vaccination support system to provide opportunities for personalized consultations on vaccination.

Conclusions

This study shows that Swedish Twitter users engaged in discussing COVID-19 vaccination expressed safety and effectiveness concerns about COVID-19 vaccines and mistrust in governmental authorities, scientific organizations, and media agencies. The tweets indicated a fading belief in vaccination necessity linked to changes in the perceived severity of COVID-19 infection and belief in the superiority of natural immunity. In comparison, the quality of vaccination services was discussed less frequently; however, some complaints related to the limited availability of vaccination appointments did appear. Moreover, there was an observed information gap on COVID-19 and vaccines related to contradicting evidence and unanswered questions. The study highlights the importance of enhancing health communication, increasing public trust in the government, and countering misinformation.

Conflicts of Interest

None declared.

References

1. Ludvigsson JF. The first eight months of Sweden's COVID-19 strategy and the key actions and actors that were involved. Acta Paediatr 2020 Dec 11;109(12):2459-2471 [FREE Full text] [doi: 10.1111/apa.15582] [Medline: 32951258]
2. WHO coronavirus (COVID-19) dashboard. World Health Organization. URL: https://covid19.who.int [accessed 2022-03-07]
3. Drop-in vaccination buses in Stockholm. 1177. URL: https://www.1177.se/Stockholm/en/other-languages/other-languages/covid-19/vaccination-bus-in-botkyrka-jarva-sodertalje-and-skarholmen/ [accessed 2022-03-13]
4. Vaccinationstäckning per födelseland, inkomst och utbildningsgrad – vecka 21. Folkhälsomyndigheten. URL: https://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/statistikdatabaser-och-visualisering/vaccinationsstatistik/statistik-for-vaccination-mot-covid-19/uppfoljning-av-vaccination/vaccinationstackning-i-undergrupper/ [accessed 2022-04-20]
5. Statistik för vaccination mot covid-19. Folkhälsomyndigheten. URL: https://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/statistikdatabaser-och-visualisering/vaccinationsstatistik/statistik-for-vaccination-mot-covid-19/ [accessed 2022-02-24]
6. Sidan kan inte hittas. Folkhälsomyndigheten. URL: http://www.folkhalsomyndigheten.se/pagefiles/20261/Barriers-motivating-factors-MMR-vaccination-communities-low-coverage-Sweden-15027.pdf [accessed 2022-03-24]
7. Ten health issues WHO will tackle this year. World Health Organization. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2022-03-07]
8. Cu A, Meister S, Lefebvre B, Ridde V. Assessing healthcare access using the Levesque's conceptual framework- a scoping review. Int J Equity Health 2021 May 07;20(1):116 [FREE Full text] [doi: 10.1186/s12939-021-01416-3] [Medline: 33962627]
9. 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]
10. Sidan finns inte. Myndigheten för samhällsskydd och beredskaps. URL: https://www.msb.se/sv/publikationer/conspiracy.-theories-and-covid-19--the-mechanisms-behind-a-rapidly-growing-societal-challenge/ [accessed 2022-03-07]

11. Goodman B. Why COVID vaccines are falsely linked to infertility. WebMD. URL: https://www.webmd.com/vaccines/covid-19-vaccine/news/20210112/why-covid-vaccines-are-falsely-linked-to-infertility [accessed 2022-03-22]
12. Infodemic. World Health Organization. URL: https://www.who.int/westernpacific/health-topics/infodemic [accessed 2022-03-14]
13. Kunitoki K, Funato M, Mitsunami M, Kinoshita T, Reich MR. Access to HPV vaccination in Japan: increasing social trust to regain vaccine confidence. Vaccine 2021 Oct 01;39(41):6104-6110 [FREE Full text] [doi: 10.1016/j.vaccine.2021.08.085] [Medline: 34507858]
14. Suppli CH, Hansen ND, Rasmussen M, Valentiner-Branth P, Krause TG, Mølbak K. Decline in HPV-vaccination uptake in Denmark - the association between HPV-related media coverage and HPV-vaccination. BMC Public Health 2018 Dec 10;18(1):1360 [FREE Full text] [doi: 10.1186/s12889-018-6268-x] [Medline: 30526589]
15. Betsch C, Renkewitz F, Betsch T, Ulshöfer C. The influence of vaccine-critical websites on perceiving vaccination risks. J Health Psychol 2010 Apr 26;15(3):446-455. [doi: 10.1177/1359105309353647] [Medline: 20348365]
16. Ahmed N, Quinn SC, Hancock GR, Freimuth VS, Jamison A. Social media use and influenza vaccine uptake among White and African American adults. Vaccine 2018 Nov 26;36(49):7556-7561. [doi: 10.1016/j.vaccine.2018.10.049] [Medline: 30389192]
17. European. Countering online vaccine misinformation in the EU/EEA. European Centre for Disease Prevention Control. LU: Publications Office URL: https://data.europa.eu/doi/10.2900/329304 [accessed 2022-02-20]
18. Report of the SAGE working group on vaccine hesitancy. 2014. URL: https://www.asset-scienceinsociety.eu/sites/default/files/sage_working_group_revised_report_vaccine_hesitancy.pdf [accessed 2022-03-07]
19. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
20. Griffith J, Marani H, Monkman H. COVID-19 vaccine hesitancy in Canada: content analysis of tweets using the theoretical domains framework. J Med Internet Res 2021 Apr 13;23(4):e26874 [FREE Full text] [doi: 10.2196/26874] [Medline: 33769946]
21. Pascual-Ferrá P, Alperstein N, Barnett DJ. Social network analysis of COVID-19 public discourse on Twitter: implications for risk communication. Disaster Med Public Health Prep 2022 Apr 10;16(2):561-569 [FREE Full text] [doi: 10.1017/dmp.2020.347] [Medline: 32907685]
22. Systematic scoping review on social media monitoring methods and interventions relating to vaccine hesitancy. European Centre for Disease Prevention Control. URL: https://www.ecdc.europa.eu/en/publications-data/systematic-scoping-review-social-media-monitoring-methods-and-interventions [accessed 2022-03-07]
23. About. Netlytic. URL: https://netlytic.org/home/?page_id=10834 [accessed 2021-11-29]
24. Netlytic Making sense of public discourse online. Netlytic. URL: https://netlytic.org/home/?page_id%3d2#cmtoc_anchor_id_5 [accessed 2022-03-07]
25. Leading hashtags on social media in Sweden in March 2020, by number of mentions. Statista. URL: https://www.statista.com/statistics/1103995/leading-hashtags-on-social-media-in-sweden-by-number-of-mentions/ [accessed 2022-02-18]
26. Alla vuxna i Stockholm erbjuds påfyllnadsdosen. Region Stockholm. URL: https://www.regionstockholm.se/verksamhet/halsa-och-vard/nyheter-vaccin-covid-19/2022/01/alla-vuxna-i-stockholm-erbjuds-pafyllnadsdosen/ [accessed 2022-03-07]
27. The Omicron variant BA.2 was first seen in wastewater from Malmö and Kalmar. Swedish University of Agricultural Sciences. URL: https://www.slu.se/en/ew-news/2022/2/the-omicron-variant-ba.2-was-first-seen-in-wastewater-from-malmo-and-kalmar/ [accessed 2022-03-07]
28. The Omicron variant: sorting fact from myth. World Health Organization. URL: https://tinyurl.com/4rszubzw [accessed 2022-03-07]
29. Restriktioner tas bort 9 februari. Krisinformation. URL: https://www.krisinformation.se/nyheter/2022/februari/restriktioner-tas-bort-9-februari [accessed 2022-03-07]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017 Oct 02;16(1):160940691773384. [doi: 10.1177/1609406917733847]
32. Gerretsen P, Kim J, Caravaggio F, Quilty L, Sanches M, Wells S, et al. Individual determinants of COVID-19 vaccine hesitancy. PLoS One 2021 Nov 17;16(11):e0258462 [FREE Full text] [doi: 10.1371/journal.pone.0258462] [Medline: 34788308]
33. Wang J, Zhang Y, Long S, Fu X, Zhang X, Zhao S, et al. Non-EPI vaccine hesitancy among Chinese adults: a cross-sectional study. Vaccines (Basel) 2021 Jul 10;9(7):772 [FREE Full text] [doi: 10.3390/vaccines9070772] [Medline: 34358188]
34. 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 Nov 29;5(1):80-92 [FREE Full text] [doi: 10.1177/160940690600500107]
35. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Educ Inform 2004 Jul 19;22(2):63-75. [doi: 10.3233/efi-2004-22201]

36. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [doi: 10.1136/bmj.323.7321.1103] [Medline: 11701577]
37. Bond CS, Ahmed OH, Hind M, Thomas B, Hewitt-Taylor J. The conceptual and practical ethical dilemmas of using health discussion board posts as research data. J Med Internet Res 2013 Jun 07;15(6):e112 [FREE Full text] [doi: 10.2196/jmir.2435] [Medline: 23748147]
38. Soares P, Rocha JV, Moniz M, Gama A, Laires PA, Pedro AR, et al. Factors associated with COVID-19 vaccine hesitancy. Vaccines (Basel) 2021 Mar 22;9(3):300 [FREE Full text] [doi: 10.3390/vaccines9030300] [Medline: 33810131]
39. Pertwee E, Simas C, Larson HJ. An epidemic of uncertainty: rumors, conspiracy theories and vaccine hesitancy. Nat Med 2022 Mar 10;28(3):456-459. [doi: 10.1038/s41591-022-01728-z] [Medline: 35273403]
40. Angående utdrag ur Läkemedelsverkets databas för spontant rapporterade misstänkta biverkningar. Läkemedelsverket. URL: https://www.lakemedelsverket.se/498f29/globalassets/dokument/behandling-och-forskrivning/coronavirus/misstankta-biverkningar/comirnaty-handlagda-rapporter-om-misstankta-biverkningar-2022-05-05.pdf [accessed 2022-06-01]
41. Rapporterade misstänkta biverkningar - coronavaccin. Läkemedelsverket. URL: https://www.lakemedelsverket.se/sv/coronavirus/coronavaccin/rapporterade-misstankta-biverkningar-coronavaccin [accessed 2022-06-01]
42. Helsingen LM, Refsum E, Gjøstein DK, Løberg M, Bretthauer M, Kalager M, Clinical Effectiveness Research group. The COVID-19 pandemic in Norway and Sweden - threats, trust, and impact on daily life: a comparative survey. BMC Public Health 2020 Oct 23;20(1):1597 [FREE Full text] [doi: 10.1186/s12889-020-09615-3] [Medline: 33097011]
43. 5 reasons it is important for adults to get vaccinated. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/vaccines/adults/reasons-to-vaccinate.html [accessed 2022-02-24]
44. Detta gäller om coronavaccin. Lakemedelsverket. URL: https://www.lakemedelsverket.se/sv/coronavirus/coronavaccin/detta-galler-om-coronavaccin [accessed 2022-02-02]
45. Vårdnadshavares acceptans för vaccination mot covid-19 av barn 5–11 år. Folkhälsomyndigheten. URL: https://tinyurl.com/4c75da98 [accessed 2022-05-06]
46. Fisher KA, Bloomstone SJ, Walder J, Crawford S, Fouayzi H, Mazor KM. Attitudes toward a potential SARS-CoV-2 vaccine. Ann Intern Med 2020 Dec 15;173(12):964-973. [doi: 10.7326/m20-3569]
47. Yeung MP, Lam FL, Coker R. Factors associated with the uptake of seasonal influenza vaccination in adults: a systematic review. J Public Health (Oxf) 2016 Dec 02;38(4):746-753. [doi: 10.1093/pubmed/fdv194] [Medline: 28158550]
48. Boyd D, Crawford K. Critical questions for big data. Inform Commun Soc 2012 Jun;15(5):662-679. [doi: 10.1080/1369118x.2012.678878]
49. Tomeny TS, Vargo CJ, El-Toukhy S. Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. Soc Sci Med 2017 Oct;191:168-175 [FREE Full text] [doi: 10.1016/j.socscimed.2017.08.041] [Medline: 28926775]
50. Svenskar i världen. Kartlaggning 2022. URL: https://www.sviv.se/wp-content/uploads/2022/07/Kartlaggning-2022_Svenskar-i-varlden.pdf [accessed 2022-12-02]

Abbreviations

API: application programming interface
SMPA: Swedish Medical Products Agency
TA: thematic analysis
WHO: World Health Organization

Edited by T Mackey; submitted 07.09.22; peer-reviewed by J Alpert, B Eshrati; comments to author 30.11.22; revised version received 05.12.22; accepted 10.01.23; published 22.02.23.

Please cite as:
Beirakdar S, Klingborg L, Herzig van Wees S
Attitudes of Swedish Language Twitter Users Toward COVID-19 Vaccination: Exploratory Qualitative Study
JMIR Infodemiology 2023;3:e42357
URL: https://infodemiology.jmir.org/2023/1/e42357
doi:10.2196/42357
PMID:37012999

Original Paper

Public Figure Vaccination Rhetoric and Vaccine Hesitancy: Retrospective Twitter Analysis

Vlad Honcharov[1,2], MPH; Jiawei Li[3,4], MS; Maribel Sierra[1,2], MPH; Natalie A Rivadeneira[1,2], MPH; Kristan Olazo[1,2], MPH; Thu T Nguyen[5,6], MSPH, SCD; Tim K Mackey[3,4,7], MAS, PhD; Urmimala Sarkar[1,2], MD, MPH

[1]Division of General Internal Medicine at Zuckerberg San Francisco General Hospital and Trauma Center, University of California San Francisco, San Francisco, CA, United States

[2]Center for Vulnerable Populations, University of California San Francisco, San Francisco, CA, United States

[3]S-3 Research LLC, San Diego, CA, United States

[4]Global Health Policy and Data Institute, San Diego, CA, United States

[5]Department of Family and Community Medicine, University of California San Francisco, San Francisco, CA, United States

[6]Department of Epidemiology & Biostatistics, University of Maryland School of Public Health, College Park, MD, United States

[7]Global Health Program, Department of Anthropology, University of California San Diego, La Jolla, CA, United States

Corresponding Author:
Urmimala Sarkar, MD, MPH
Division of General Internal Medicine at Zuckerberg San Francisco General Hospital and Trauma Center
University of California San Francisco
Department of Medicine, Box 1364
1001 Potrero Ave, Building 10, Floor 3
San Francisco, CA, 94143
United States
Phone: 1 628 206 4273
Email: usarkar@medsfgh.ucsf.edu

Abstract

Background: Social media has emerged as a critical mass communication tool, with both health information and misinformation now spread widely on the web. Prior to the COVID-19 pandemic, some public figures promulgated anti-vaccine attitudes, which spread widely on social media platforms. Although anti-vaccine sentiment has pervaded social media throughout the COVID-19 pandemic, it is unclear to what extent interest in public figures is generating anti-vaccine discourse.

Objective: We examined Twitter messages that included anti-vaccination hashtags and mentions of public figures to assess the connection between interest in these individuals and the possible spread of anti-vaccine messages.

Methods: We used a data set of COVID-19–related Twitter posts collected from the public streaming application programming interface from March to October 2020 and filtered it for anti-vaccination hashtags "antivaxxing," "antivaxx," "antivaxxers," "antivax," "anti-vaxxer," "discredit," "undermine," "confidence," and "immune." Next, we applied the Biterm Topic model (BTM) to output topic clusters associated with the entire corpus. Topic clusters were manually screened by examining the top 10 posts most highly correlated in each of the 20 clusters, from which we identified 5 clusters most relevant to public figures and vaccination attitudes. We extracted all messages from these clusters and conducted inductive content analysis to characterize the discourse.

Results: Our keyword search yielded 118,971 Twitter posts after duplicates were removed, and subsequently, we applied BTM to parse these data into 20 clusters. After removing retweets, we manually screened the top 10 tweets associated with each cluster (200 messages) to identify clusters associated with public figures. Extraction of these clusters yielded 768 posts for inductive analysis. Most messages were either pro-vaccination (n=329, 43%) or neutral about vaccination (n=425, 55%), with only 2% (14/768) including anti-vaccination messages. Three main themes emerged: (1) anti-vaccination accusation, in which the message accused the public figure of holding anti-vaccination beliefs; (2) using "anti-vax" as an epithet; and (3) stating or implying the negative public health impact of anti-vaccination discourse.

Conclusions: Most discussions surrounding public figures in common hashtags labelled as "anti-vax" did not reflect anti-vaccination beliefs. We observed that public figures with known anti-vaccination beliefs face scorn and ridicule on Twitter. Accusing public figures of anti-vaccination attitudes is a means of insulting and discrediting the public figure rather than discrediting vaccines. The majority of posts in our sample condemned public figures expressing anti-vax beliefs by undermining their influence, insulting them, or expressing concerns over public health ramifications. This points to a complex information ecosystem, where anti-vax sentiment may not reside in common anti-vax–related keywords or hashtags, necessitating further assessment of the influence that public figures have on this discourse.

(JMIR Infodemiology 2023;3:e40575) doi:10.2196/40575

KEYWORDS

Twitter; anti-vaccination; Biterm Topic modeling; inductive content analysis; COVID-19; social media; health information; vaccination; vaccine hesitancy; infodemiology; misinformation

Introduction

COVID-19 is a member of a large family of viruses called coronaviruses [1]. The virus is spread from person to person through droplets released when an infected person coughs, sneezes, or talks and is less commonly spread by touching a surface with the virus on it and then touching one's eyes, mouth, or nose [1,2]. COVID-19 was first detected in late December 2019 and was declared a pandemic by the World Health Organization in March 2020 [2,3]. Symptoms typically include fever, malaise, and cough, and some who are infected develop acute respiratory distress syndrome, respiratory failure, organ failure, and even death [2,3].

Even in the context of an ongoing COVID-19 pandemic, anti-vaccination rhetoric persists despite scientific evidence validating the safety and efficacy of vaccines as a critical public health tool [4]. Existing literature indicates that false claims regarding COVID-19 vaccines undermine public trust in ongoing vaccination campaigns, which can lead to greater morbidity and mortality from vaccine-preventable diseases [5]. Social media plays a role in shaping vaccination beliefs, as 7 of every 10 Americans report using a social media platform [6]. For example, Twitter, a popular microblogging platform that allows users to share posts of 280 characters or less, commonly referred to as "tweets," boasted 290.5 million users in 2019. Twitter has also been identified as a source of misinformation and disinformation about COVID-19 and vaccines [7,8]. An investigation into misinformation warnings on Twitter found that rather than dispelling misinformation, moderation often led to the development of reverberations of one's own beliefs regardless of the presence of the disclaimer [9]. This is supported by research indicating that social media users heavily relied on social media platforms for COVID-19 information and were unlikely to fact check the information they obtained with a professional [10].

Researchers have previously examined the role public figures on social media have in shaping the public's health beliefs. Prior studies have found that only a handful of individual accounts can be responsible for disseminating information and misinformation that is then shared or retweeted thousands of times, reaching potentially millions of social media users [11]. Furthermore, a 2013 meta-analysis found that individuals are conditioned to react positively to the advice of celebrities, and that celebrity medical advice can be a contagion that diffuses throughout social networks [12]. Consequently, celebrity anti-vaccination rhetoric can have extensive, deleterious consequences on public health. Notably, public figures may even propagate health misinformation inadvertently. A retrospective Twitter analysis examining the diffusion of misinformation following Hank Aaron's death found an increase in erroneous claims connecting his death to vaccine misinformation [13].

Many Twitter posts about COVID-19 vaccination reference public figures, but it remains unclear how the discourse surrounding vaccination integrates attitudes and opinions about public figures. It is also undetermined whether the conversation about public figures' vaccination attitudes is intended to fuel anti-vaccination sentiments. Therefore, we aimed to study Twitter posts about COVID-19 vaccination that specifically mentioned publicly known individuals or groups, while concurrently investigating the themes and sentiments depicted in these associated posts.

Methods

Ethics Approval

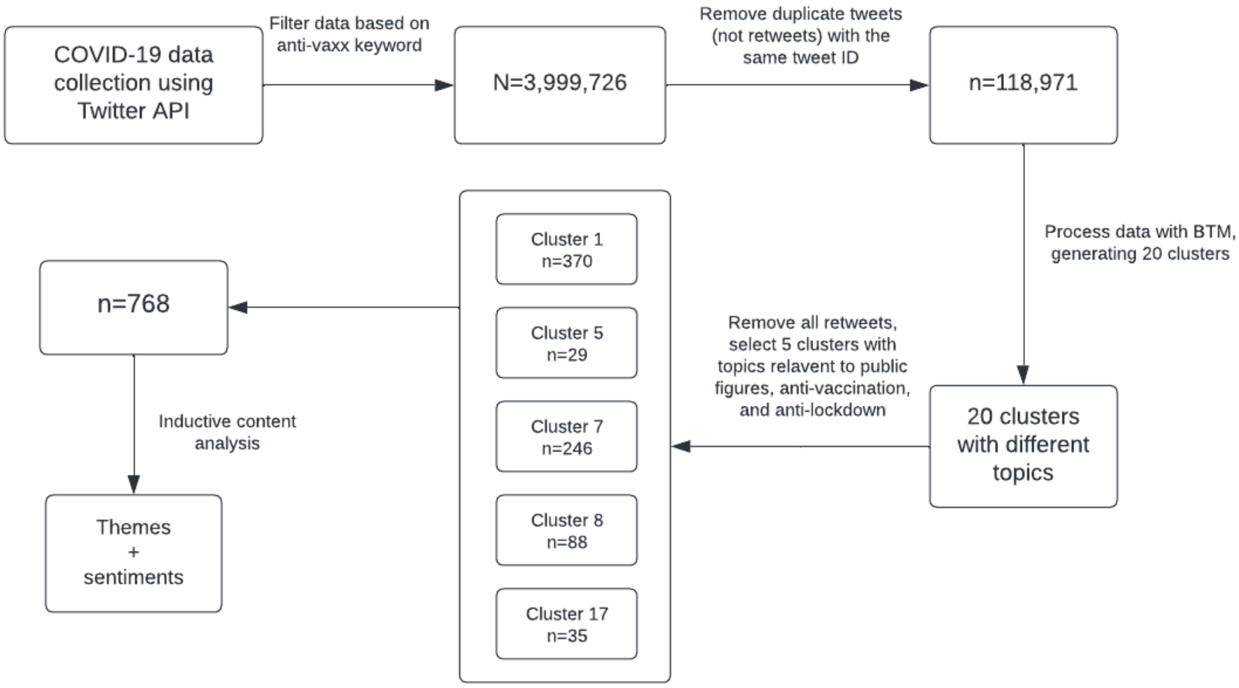
As this study used deidentified, publicly available social media data, the Institutional Review Board of University of California, San Francisco classified our proposal as exempt from review (IRB 13-12815).

Procedure

We collected publicly available data using Twitter's application programming interface (API) as seen in previous social media research (Figure 1) [14]. Specifically, the purpose of this study was to identify tweets associated with anti-vax discussion that also included mentions of public figures. Hence, though data on individual Twitter user accounts or handles were collected, they were subsequently removed from the data set prior to the topic modeling phase of the study and were not analyzed or reported other than in the aggregate.

Next, we removed all duplicate Twitter posts and conducted topic exploration using an unsupervised machine learning approach called Biterm Topic modeling (BTM), which thematically groups related Twitter posts into topic clusters [14,15]. We defined Twitter messages as texts—with 280 characters or less—posted on Twitter, and we used Twitter posts, Twitter messages, and tweets interchangeably. We then removed all retweets, defined as messages that had been shared and circulated by users other than the initial poster. Finally, we used an inductive qualitative coding approach to code Twitter messages from manually selected clusters that contained word groupings related to the study aims.

Figure 1. Study design. API: application programming interface; BTM: Biterm Topic Modeling.



Data Collection and Processing

We collected data from Twitter using the public streaming API over a period of approximately 3 months from March 3, 2020, to Oct 28, 2020, using the Python package Tweepy (version 3.8.0) [16]. This period was selected because it marked the acceleration of the outbreak into a global pandemic and was a crucial period for the establishment of pro- and anti-vaccination sentiment, as vaccine development was widely discussed and debated. The data included the text of the Twitter messages and other metadata associated with the message (eg, geolocation, if available; time stamp information; and user account or handle). We first applied a list of common COVID-19–related keywords used on social media as filters for the Twitter public API. These keywords were chosen on the basis of structured manual searches conducted on Twitter that detected content related to the COVID-19 pandemic as posted by users, and they had also been validated as being able to identify tweets pertaining to general COVID-19 conversations in prior studies [17-21]. These keywords included “coronaoutbreak,” ”corona,” ”anticorona,” “coronavirus,” “covid,” and “pandemic.” We captured and processed all Twitter messages that contained at least one of these keywords or hashtags. The purpose for this first phase of keyword selection was to obtain a broad Twitter corpus that contained general COVID-19–related conversations not specific to any topic, which could then be filtered for more specific hashtags, keywords, and other vocabulary associated with anti-vaccination sentiment, opinions, or statements.

After removing duplicate tweets, we applied a second text filter to isolate tweets with anti-vax–related keywords and conducted BTM. Anti-vax–related keywords included “antivaxxing,” “antivaxx,” “antivaxxers,” “antivax,” “anti-vaxxer,” “discredit,” “undermine,” “confidence,” and “immune.” We chose anti-vax–related keywords, as we were specifically interested in the web-based discourse surrounding these terms, and these terms also appeared as related search terms when conducting testing of related terms associated with “anti-vaccine” on the Google search engine. For the purposes of our analysis, “anti-vax” is equivalent to anti-vaccination, and therefore, the 2 terms are used interchangeably. Our investigation defines vaccine deniers, more commonly referred to as “anti-vaxxers,” as individuals who believe vaccines are dangerous, deny the efficacy of inoculation, or refuse vaccines for themselves and their children, if applicable.

Topic Modeling Using BTM

BTM groups Twitter messages containing the same word-related themes and summarizes the entire corpus of text into distinct highly correlated categories. BTM is best used for short text, and its primary strengths are topic modeling word co-occurrence patterns and identifying such sequences in text that contain few words [15]. The main themes in clusters produced by BTM are considered an aggregation of topics from the text, which are then split into a bag of words, where a discrete probability distribution for all words in each theme is generated. Before running BTM, we cleaned our data set for imbedded hyperlinks, stop words, special characters and punctuation marks, and length using the Natural Language Toolkit package in Python [22]. Specifically, we excluded Twitter posts less than 3 words in length, as they likely do not convey sufficient information for purposes of inductive content coding of themes, which is consistent with prior studies [23]. Using the COVID-19 data set filtered for the anti-vax–related keywords described, we used BTM to parse the data into 20 topic clusters.

We set a total number of 20 different clusters (ie, total number of topics for BTM to output: k=20), resulting in texts with similar themes put into the same clusters. To find the appropriate k value, we used a topic coherence score [21,24]. Coherence score is used to measure the performance of a topic model with different number of clusters and can help differentiate between topics that are semantically interpretable and topics that are artifacts of statistical inference [24,25]. We tested 5 different k values (k=10, 20, 30, 40, and 50) for each data set and found that when k=20, we generated the highest coherence score, and this score did not change significantly with an increase in the k value.

Screening

We manually screened the top 10 tweets that were most highly correlated to the 20 topic cluster word groupings generated following the BTM topic modeling phase. By examining the top 10 tweets, we ensured that we did not miss public figures mentioned in other topics. In BTM, correlation is determined by word co-occurrence patterns in the text, and the outputted clusters were then manually reviewed for relevance. We then manually selected 5 clusters that most closely included messages calling out or making claims about public figures as anti-vaxxers or that called out groups of people such as scientists or political parties. We define a public figure as a Twitter user with a verified Twitter account. Topic clusters that were not included in this study covered topics about government mistrust, conspiracies, mask business promotion, and general statements about anti-vaccination beliefs (not specific to any public figure).

Content Analysis

Our sample included Twitter messages associated with these 5 relevant clusters outputted by BTM and then extracted for all tweets associated with the selected clusters. We applied a grounded theory's inductive coding approach, allowing for themes to emerge while coding rather than prespecifying the content of interest [26]. Grounded theory enables researchers to develop a theory to explain the phenomenon of interest, and as the study progresses, the researcher's initial exploratory question becomes refined until an understanding is reached regarding the topic of investigation [27]. The advantage of this approach is that it allows us to minimize the effect of personal bias surrounding vaccination rhetoric [26] by generating the codes based on the content of the Twitter messages. We conducted qualitative analysis to characterize the discourse (eg, pro-vaccination vs anti-vaccination). 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 Twitter messages in our sample, while iteratively continuing to develop existing codes and definitions as well as new codes.

We also labeled the public figure or group mentioned in each post, where applicable, and calculated the corresponding frequencies and percentages (Table 1). We reached thematic saturation after approximately 200 posts but continued to code the entire data set. Three of the authors (MS, US, and NR) coded the Twitter messages independently and achieved a high intercoder reliability (κ=0.92). For inconsistent results, all coders met and conferred on correct classification and subclassifications to reach consensus. Coders denoted neutral, anti-vaccination, or pro-vaccination sentiments expressed in the messages, along with each theme, throughout 7 rounds of coding.

Table 1. Public figures or groups mentioned in selected sample of Twitter messages sorted by frequency (n=768).

Public Figure[a]	Frequency[b], n (%)[c]
Novak Djokovic	330 (43)
None	149 (19.4)
Kamala Harris	148 (19.3)
Joe Biden	123 (16)
Donald Trump	20 (2.6)
Amy Duncan[d]	13 (1.7)
Isabel Lucas	12 (1.6)
Andrew Cuomo	9 (1.2)
Barack Obama	7 (0.9)
Ammon Bundy	6 (0.8)
Joe Rogan	5 (0.7)
Bill Gates	5 (0.7)
Rand Paul	5 (0.7)
Rebecca Judd	5 (0.7)
Alex Jones	4 (0.5)
Anti-vaxxers[d]	3 (0.4)
Kanye West	3 (0.4)
Jim Carrey	3 (0.4)

[a]We analyzed Twitter messages collected using anti-vaccination hashtags involving public figures or groups. The public figures or groups mentioned in these messages do not all explicitly express anti-vaccination ideology but have been included in analysis for the assessment of Twitter rhetoric surrounding these individuals or entities.

[b]Public figures or groups mentioned ≤2 times were excluded from the table. Excluded public figures or groups are as follows: Washington Post, Jeanine Piro, University of Toronto Scarborough, Carolyn Maloney, Tommy Robinson, Scientists, David Icke, Russian government, Robert Redfield, Qanon, Ian Brown, White House Coronavirus Taskforce, Dejan Lovren, Catholic Archbishops, Jim Acosta, Sebastian Gorka, Mike Lindell, Sharyl Attkisson, Leigh-Allyn Baker, Nikola Jokic, Rita Pala, Lee Zeldin, Bernie Sanders, Judy Mikovits, John Water, Federal Agencies, Democrats, online anti-vax communities, MIA, British Union of Fascists, Glenn Davies, George Stephanopoulos, and Marianne Williamson.

[c]Authors assigned multiple public figures or groups to various Twitter messages, when applicable; therefore, percentages do not add up to 100%.

[d]Amy Duncan is a fictional character, and anti-vaxxers are a group. We manually selected Biterm Topic Modeling clusters based on relevance to public figures, which contained figures and groups with both verified and unverified accounts.

Results

Overview

We collected the initial sample from various anti-vax–related keywords, and the sample contained 3,999,726 Twitter posts. We then removed all duplicate tweets with the same tweet ID that distilled our sample to 118,971 messages. Subsequently, we grouped Twitter posts into topic clusters using BTM, yielding 20 clusters containing various topics. We then selected 5 clusters most closely related to public figures, anti-vaccination, and anti-lockdown. We manually reviewed a cumulative sum of 768 Twitter messages identified in the 5 topic clusters selected.

Of the 768 Twitter messages, 425 (55%) were neutral, 329 (43%) expressed pro-vaccination sentiments, and 14 (2%) expressed anti-vaccination sentiments. Furthermore, 356 (46%) messages called out public figures for their stances or behaviors, 188 (24%) undermined public figures, 157 (20%) expressed concern over the negative public health impact of the actions of certain public figures, 57 (7%) insulted public figures, and 8 (1%) defended anti-vaccination public figures (Table 2). A total of 51 public figures with verified Twitter accounts were identified comprising a mix of athletes, politicians, actors, musicians, radio and political commentators, models, business leaders, anti-government activists, and other personalities. Politicians were some of the most frequently mentioned public figures, along with political commentators.

Table 2. Twitter message themes.

Theme	Definition	Example	Frequency, n (%)
Neutral or none	Absence of expression of a clear judgement even if the message is related to the topic [18].	*Latest on our #globalhealth #COVID19 #vaccine. Note to my antivax friends: this is not Merck & Co the vaccine company, it's Merck Darmstadt in Germany, which owns Sigma Millipore, totally separate*	2 (0.3)
Insults	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.	*Wouldn't it be tragic if anti-vaxxer idiot Novak Djokovic succombs* [sic] *to coronavirus before a vaccine he refuses to take is invented?*	57 (7.4)
Negative public health impact	States or implies that anti-vaxxers and anti-vaccine behaviors have a negative impact on public health. May connect the vaccine to other diseases.	*Opinion | Anti-lockdown and anti-vaxxer protesters have merged. And it could be deadly. via @....*	157 (20.4)
Anti-vax accusations	Accuses or asserts a specific person or groups of people are anti-vaxxers. Subcode of "undermine": accusations intending to undermine or discredit a person or group(s).	*Trump calls Dems anti-vaxxers says their anti-vaccine rhetoric is 'dangerous' after Sen. Harris said she wouldn't trust the him* [sic] *on safety of COVID vaccine before the election.*	188 (24.5)
	Accuses or asserts a specific person or groups of people are anti-vaxxers. Subcode of "call-out": accuses a person or group, call them out, or imply that they are against vaccines.	*Tennis: Novak Djokovic is an anti-vaxxer and won't take coronavirus vaccine #tennis #tennisnews*	356 (46.4)
Defending anti-vax stance	Defends or upholds an anti-vax position.	*Novak Djokovic's wife is shamed with a 'False Information' badge by Instagram for spreading coronavirus 5G conspiracy theories after the tennis world No 1 revealed he's an anti-vaxxer….Stand strong Jelena!*	8 (1)

Theme 1: Anti-Vax Accusations

Twitter users frequently accused public figures of holding anti-vaccination views. Twitter posts that "called out" or accused public figures of harboring anti-vaccination beliefs composed the largest segment of our sample (356/768, 46%). In most cases, these posts refered to public figures who publicly espoused anti-vaccination attitudes, such as Novak Djokovic. Messages intending to undermine or discredit public figures formed 24% (188/768) of our sample, with the majority expressing neutral sentiment toward vaccination. Some of these messages amplified statements made by public figures that undermined or accused other public figures of harboring anti-vax beliefs, which was common among politicians in our sample. Specifically, these messages amplified statements made by the former US president Donald Trump accusing Joe Biden and Kamala Harris of subscribing to anti-vax beliefs. Notably, the majority of undermining messages in our sample mentioned Kamala Harris and Joe Biden.

Theme 2: Insults

Our sample included 57 (7%) Twitter messages insulting public figures. The vast majority of insults were directed toward public figures suspected of being anti-vaxxers or toward public figures providing known anti-vaxxers with a platform to voice their ideologies. Similar to undermining, these messages attempted to discredit anti-vaxxers by degrading their beliefs using derogatory terms; however, these insults differed from undermining messages, as the latter accused or implied an anti-vaccination stance, whereas the former blatantly disrespected public figures with demeaning remarks. Of the 57 messages insulting public figures, 35 (61%) were directed at Novak Djokovic. Overall, Twitter messages containing insults targeted a broad scope of public figures, with the majority either known or suspected to be holding anti-vaccination beliefs.

Theme 3: Negative Public Health Impact

Our sample contained 157 (20%) Twitter messages stating or implying that anti-vaccination behaviors or rhetoric expressed by public figures may have a negative public health impact. These messages typically expressed concern about the effects of the anti-vaccination movement on public health. Of these messages, 88 (56%) expressed a neutral attitude toward vaccination, while 69 (44%) explicitly expressed pro-vaccination sentiments. The majority of Twitter messages (116/157, 74%) characterized as expressing a belief that anti-vaccination rhetoric or behaviors stemming from public figures have a negative public health impact were not directed toward specific public figures, but rather targeted anti-vaxxers in general.

Discussion

Given the prevalence of anti-vaccination attitudes and the known contagion of celebrity beliefs, we expected to see mentions of public figures with anti-vaccination beliefs further espousing vaccine misinformation sentiment and conspiracies on the internet; instead, we found that "anti-vax"–related keywords or hashtags in our corpus of tweets primarily consisted of discourse accusing or insulting public figures of holding an anti-vaccination stance, specifically as a means of publicly calling them out or insulting them on Twitter. Notably, the majority of Twitter messages (56/57, 98%) characterized as insults expressed pro-vaccination sentiment, indicating that insults are frequently sent by supporters of vaccination rather than anti-vaxxers.

There were multiple posts accusing public figures of holding anti-vaccination beliefs (including known vaccine supporters). Undermining messages attempt to discredit public figures by accusing them of holding anti-vax beliefs, even among those known to publicly support vaccines. These posts exemplify the denunciation of suspected anti-vaxxers by Twitter users [28,29].

As expected, an abundance of posts insulted public figures known to be anti-vaxxers, and a recent systematic review examining misinformation found it to be a universal source of stress, fatigue, insomnia, and anger [30]. Further research should focus on identifying common traits among public figures subject to insults from social media users and the effect of this overall rhetoric on other users' attitudes, knowledge, and perceptions about vaccines.

Our study has several limitations. First, we acknowledge that attitudes of Twitter users are unlikely to be representative of the general population or attitudes specifically toward celebrities or public personalities. Second, we sampled based on anti-vaccination keywords and for a specific period of time during the COVID-19 pandemic; this method is a common approach in infodemiology and misinformation studies [13,31,32], but we could have nevertheless missed messages relevant to the study aims that did not include these hashtags or occurred later during the pandemic. Hence, our choice of keywords or hashtags used for this study is not generalizable to all anti-vax posts occurring on Twitter. Third, we performed content analysis with a circumscribed sample of tweets outputted by topic modeling; there may be additional themes linking public figures and vaccination that did not emerge in our sample.

For the majority of our sample, referring to a public figure as an "anti-vaxxer" is a way of condemning public figures, whether or not they espouse anti-vaccination beliefs in their own public communication. Novak Djokovic openly opposes vaccination, but pro-vaccination individuals, including President Biden, have been accused of being "anti-vax" on Twitter. This unexpected finding in the context of user-generated posts associated with anti-vax–related keywords and hashtags (ie, we expected to observe amplification of anti-vax sentiment harbored by known celebrities and public figures) suggests reciprocal influence between public health recommendations and attitudes about public figures rather than the previously described one-way, outsize influence of celebrities on vaccination attitudes. We believe that social media platforms represent a complex information ecosystem, where anti-vax sentiment may not reside in common anti-vax–related keywords or hashtags, but instead in other web-based spaces of discourse that require additional study. Additional research is also needed to fully assess the influence of public figures and users' perception of these individuals on ensuing vaccine discourse, whether positive or negative.

Acknowledgments

US was supported by National Institutes of Health (NIH) and National Cancer Institute (K24CA212294). TTN was supported by the National Institute on Minority Health and Health Disparities under Award Numbers R00MD012615 and R01MD015716. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Data Availability

Deidentified data that contains associated twitter IDs associated with this study are available in a GitHub repository [33].

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*.

References

1. Definition of SARS-CoV-2. National Cancer Institute Dictionary of Cancer Terms. URL: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/sars-cov-2 [accessed 2022-12-12]
2. Ludwig S, Zarbock A. Coronaviruses and SARS-CoV-2: a brief overview. Anesth Analg 2020 Jul;131(1):93-96 [FREE Full text] [doi: 10.1213/ANE.0000000000004845] [Medline: 32243297]
3. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020 Feb 15;395(10223):497-506 [FREE Full text] [doi: 10.1016/S0140-6736(20)30183-5] [Medline: 31986264]
4. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, Türeci, C4591001 Clinical Trial Group. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med 2020 Dec 31;383(27):2603-2615 [FREE Full text] [doi: 10.1056/NEJMoa2034577] [Medline: 33301246]

5. 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 Nov 17;7(11):e30642 [FREE Full text] [doi: 10.2196/30642] [Medline: 34653016]
6. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2022-02-10]
7. Twitter: number of users worldwide 2019-2020. Statista. URL: https://www.statista.com/statistics/303681/twitter-users-worldwide/ [accessed 2021-11-03]
8. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Dec 09;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
9. Sharevski F, Alsaadi R, Jachim P, Pieroni E. Misinformation warnings: Twitter's soft moderation effects on COVID-19 vaccine belief echoes. Comput Secur 2022 Mar;114:102577 [FREE Full text] [doi: 10.1016/j.cose.2021.102577] [Medline: 34934255]
10. 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]
11. Fuentes A, Peterson J. Social media and public perception as core aspect of public health: The cautionary case of @realdonaldtrump and COVID-19. PLoS One 2021;16(5):e0251179 [FREE Full text] [doi: 10.1371/journal.pone.0251179] [Medline: 33951083]
12. Hoffman SJ, Tan C. Following celebrities' medical advice: meta-narrative analysis. BMJ 2013 Dec 17;347(dec17 14):f7151-f7151. [doi: 10.1136/bmj.f7151]
13. 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]
14. Shah N, Nali M, Bardier C, Li J, Maroulis J, Cuomo R, et al. Applying topic modelling and qualitative content analysis to identify and characterise ENDS product promotion and sales on Instagram. Tob Control 2021 Dec 02. [doi: 10.1136/tobaccocontrol-2021-056937] [Medline: 34857646]
15. Yan X, Guo J, Lan Y, Cheng X. A biterm topic model for short texts. In: Proceedings of the 22nd International Conference on World Wide Web. 2013 Presented at: WWW '13; May 13-17; Rio de Janeiro Brazil. [doi: 10.1145/2488388.2488514]
16. Tweepy. GitHub. URL: https://github.com/tweepy/tweepy [accessed 2022-05-23]
17. Cuomo RE, Purushothaman V, Li J, Cai M, Mackey TK. A longitudinal and geospatial analysis of COVID-19 tweets during the early outbreak period in the United States. BMC Public Health 2021 Apr 24;21(1):793 [FREE Full text] [doi: 10.1186/s12889-021-10827-4] [Medline: 33894745]
18. Mackey TK, Li J, Purushothaman V, Nali M, Shah N, Bardier C, et al. Big data, natural language processing, and deep learning to detect and characterize illicit COVID-19 product sales: infoveillance study on Twitter and Instagram. JMIR Public Health Surveill 2020 Aug 25;6(3):e20794 [FREE Full text] [doi: 10.2196/20794] [Medline: 32750006]
19. Cuomo RE, Purushothaman V, Li J, Cai M, Mackey TK. Sub-national longitudinal and geospatial analysis of COVID-19 tweets. PLoS One 2020;15(10):e0241330 [FREE Full text] [doi: 10.1371/journal.pone.0241330] [Medline: 33112922]
20. Cai M, Li J, Nali M, Mackey T. Evaluation of hybrid unsupervised and supervised machine learning approach to detect self-reporting of COVID-19 symptoms on Twitter. 2021 Presented at: 2021 IEEE International Conference on Communications Workshops (ICC Workshops) 2021.?6; June 14-23; Montreal, QC p. 1. [doi: 10.1109/iccworkshops50388.2021.9473830]
21. Mackey T, Purushothaman V, Li J, Shah N, Nali M, Bardier C, et al. Machine learning to detect self-reporting of symptoms, testing access, and recovery associated with COVID-19 on Twitter: retrospective big data infoveillance study. JMIR Public Health Surveill 2020 Jun 08;6(2):e19509 [FREE Full text] [doi: 10.2196/19509] [Medline: 32490846]
22. Natural Language Toolkit. URL: https://www.nltk.org/ [accessed 2022-02-17]
23. Alvarez-Mon MA, Llavero-Valero M, Asunsolo Del Barco A, Zaragozá C, Ortega MA, Lahera G, et al. Areas of interest and attitudes toward antiobesity drugs: thematic and quantitative analysis using Twitter. J Med Internet Res 2021 Oct 26;23(10):e24336 [FREE Full text] [doi: 10.2196/24336] [Medline: 34698653]
24. Xu Q, Nali MC, McMann T, Godinez H, Li J, He Y, et al. Unsupervised machine learning to detect and characterize barriers to pre-exposure prophylaxis therapy: multiplatform social media study. JI 2022 Apr 28;2(1):e35446. [doi: 10.2196/35446]
25. Shan Y, Ji M, Xie W, Lam KY, Chow CY. Public trust in artificial intelligence applications in mental health care: topic modeling analysis. JMIR Hum Factors 2022 Dec 02;9(4):e38799 [FREE Full text] [doi: 10.2196/38799] [Medline: 36459412]
26. Charmaz K. Constructing Grounded Theory: A practical guide through qualitative analysis Kathy Charmaz Constructing Grounded Theory: A practical guide through qualitative analysis Sage 224 £19.99 0761973532 0761973532 [Formula: see text]. In: Nurse Res. Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. Thousand Oaksage Publications; Jul 01, 2006:84.
27. Creswell J. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Qualitative Inquiry & Research Design: Choosing Among Five Approaches Thousand Oaks California: Sage Publications; 2007:-488.

28. Baeza-Rivera MJ, Salazar-Fernández C, Araneda-Leal L, Manríquez-Robles D. To get vaccinated or not? Social psychological factors associated with vaccination intent for COVID-19. J Pac Rim Psychol 2021 Oct 30;15:183449092110517. [doi: 10.1177/18344909211051799]
29. Korn L, Böhm R, Meier NW, Betsch C. Vaccination as a social contract. Proc Natl Acad Sci U S A 2020 Jun 30;117(26):14890-14899 [FREE Full text] [doi: 10.1073/pnas.1919666117] [Medline: 32541033]
30. Rocha YM, de Moura GA, Desidério GA, de Oliveira CH, Lourenço FD, de Figueiredo Nicolete LD. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review. Z Gesundh Wiss 2021 Oct 09:1-10 [FREE Full text] [doi: 10.1007/s10389-021-01658-z] [Medline: 34660175]
31. Quinn EK, Fenton S, Ford-Sahibzada CA, Harper A, Marcon AR, Caulfield T, et al. COVID-19 and vitamin D misinformation on YouTube: content analysis. JMIR Infodemiology 2022 Mar 14;2(1):e32452 [FREE Full text] [doi: 10.2196/32452] [Medline: 35310014]
32. Sidani JE, Hoffman B, Colditz JB, Wolynn R, Hsiao L, Chu K, et al. Discussions and misinformation about electronic nicotine delivery systems and COVID-19: qualitative analysis of twitter content. JMIR Form Res 2022 Apr 13;6(4):e26335 [FREE Full text] [doi: 10.2196/26335] [Medline: 35311684]
33. Coronavirus_vaccine_antivax_research. GitHub. URL: https://github.com/Mathison/coronavirus_vaccine_antivax_research [accessed 2023-03-03]

Abbreviations

BTM: Biterm Topic Modeling
API: application programming interface

Edited by A Mavragani; submitted 27.06.22; peer-reviewed by E Quinn, J White, A Elyashar; comments to author 28.11.22; revised version received 19.12.22; accepted 27.12.22; published 10.03.23.

Please cite as:
Honcharov V, Li J, Sierra M, Rivadeneira NA, Olazo K, Nguyen TT, Mackey TK, Sarkar U
Public Figure Vaccination Rhetoric and Vaccine Hesitancy: Retrospective Twitter Analysis
JMIR Infodemiology 2023;3:e40575
URL: https://infodemiology.jmir.org/2023/1/e40575
doi:10.2196/40575
PMID:37113377

Original Paper

Mining Trends of COVID-19 Vaccine Beliefs on Twitter With Lexical Embeddings: Longitudinal Observational Study

Harshita Chopra[1*], BTech; Aniket Vashishtha[1*], BTech; Ridam Pal[2], BTech; Ashima[2]; Ananya Tyagi[2], BTech, MTech; Tavpritesh Sethi[2,3], MBBS, PhD

[1]Guru Gobind Singh Indraprastha University, New Delhi, India
[2]Indraprastha Institute of Information Technology, New Delhi, India
[3]All India Institute of Medical Sciences, New Delhi, India
[*]these authors contributed equally

Corresponding Author:
Tavpritesh Sethi, MBBS, PhD
Indraprastha Institute of Information Technology
Okhla Industrial Estate
Phase III
New Delhi, 110020
India
Phone: 91 97799 08630
Email: tavpriteshsethi@iiitd.ac.in

Abstract

Background: Social media plays a pivotal role in disseminating news globally and acts as a platform for people to express their opinions on various topics. A wide variety of views accompany COVID-19 vaccination drives across the globe, often colored by emotions that change along with rising cases, approval of vaccines, and multiple factors discussed online.

Objective: This study aims to analyze the temporal evolution of different emotions and the related influencing factors in tweets belonging to 5 countries with vital vaccine rollout programs, namely India, the United States, Brazil, the United Kingdom, and Australia.

Methods: We extracted a corpus of nearly 1.8 million Twitter posts related to COVID-19 vaccination and created 2 classes of lexical categories—emotions and influencing factors. Using cosine distance from selected seed words' embeddings, we expanded the vocabulary of each category and tracked the longitudinal change in their strength from June 2020 to April 2021 in each country. Community detection algorithms were used to find modules in positive correlation networks.

Results: Our findings indicated the varying relationship among emotions and influencing factors across countries. Tweets expressing hesitancy toward vaccines represented the highest mentions of health-related effects in all countries, which reduced from 41% to 39% in India. We also observed a significant change (P<.001) in the linear trends of categories like hesitation and contentment before and after approval of vaccines. After the vaccine approval, 42% of tweets coming from India and 45% of tweets from the United States represented the "vaccine_rollout" category. Negative emotions like rage and sorrow gained the highest importance in the alluvial diagram and formed a significant module with all the influencing factors in April 2021, when India observed the second wave of COVID-19 cases.

Conclusions: By extracting and visualizing these tweets, we propose that such a framework may help guide the design of effective vaccine campaigns and be used by policy makers to model vaccine uptake and targeted interventions.

(JMIR Infodemiology 2023;3:e34315) doi:10.2196/34315

KEYWORDS

COVID-19; COVID-19 vaccination; vaccine hesitancy; public health; unsupervised word embeddings; natural language preprocessing; social media; Twitter

Introduction

The unprecedented spread of COVID-19 has created massive turmoil in public health around the world [1]. The development of vaccines has played a pivotal role in eradicating and mitigating significant outbreaks of infectious diseases like smallpox, tuberculosis, measles, and similar contagious diseases [2]. Major pharmaceutical companies located across the globe are in the phase of developing vaccines, with only a handful of the vaccines authorized for clinical trials [3,4]. As the distribution of vaccines and associated campaigns expand, people continue to express their opinions and personal incidents on social media platforms.

Social media plays a decisive role in propagating information, leading to the emergence of varying perceptions related to the pandemic [5]. During the initial phase of national lockdown in several countries, Twitter had reported an increase of 24% in daily active users due to the increased usage of social media, the highest year-over-year growth rate reported by the company to date [6].

Mass media strongly influences vaccine uptake and vaccination rates, as shown previously for influenza [7,8]. Although some studies have also shown a positive impact of mass media on improving vaccine uptake and mitigating hesitancy [9], its role in the spread of vaccine misinformation and conspiracy theories has been widespread [10]. Recent studies such as "The 'Pandemic' of Disinformation in COVID-19" [11] reported several events for which mass media channels have misinformed the public by sharing incomplete or unverified updates on new treatments, myths about usage of masks, and errors of some hospital organizations that resulted in higher reluctance from patients to go to hospitals or medical centres. The surge in consumption of COVID-19 updates from mass media channels has impacted different age groups by inducing panic and anxiety [12].

The COVID-19 pandemic has been studied in multidisciplinary aspects, and the analysis of Twitter posts remains a widely explored area in public health research [13-15], primarily because of the rapidly evolving nature of the content. Over the last decade, researchers have used multiple methods such as sentiment classification [16], social network analysis [17], and topic identification [18] to study the presence of provaccine and antivaccine communities on social media. It has been observed that vaccine uptake is affected by multiple factors, including rising adverse effect reporting, socioeconomic inequities, and quantitative allocation [19]. In addition, the spread of misinformation online has been a concerning issue, and prior survey-based studies suggest that it is linked with vaccine hesitancy and effects on public health [20,21]. On the other hand, certain marginalized groups continue to face inaccessibility to vaccines [22].

This paper presents a temporal and demographic analysis of lexical categories mined from Twitter conversations around vaccines. We further subdivided these categories into 2 subtypes: emotions and their influencing factors. We examined the relationships between emotions such as hesitancy, rage, contentment, sorrow, faith, and anticipation with influencing factors such as conspiracy theories around vaccines, social inequities, and health effects using unsupervised word embeddings trained on the curated corpus of tweets during an 11-month period. Further, we created correlation-based networks of these categories and performed clustering using the Infomap algorithm. The alluvial diagrams generated by these networks demonstrate the flow of importance of each factor from one month to another. We performed a granular analysis of the temporal-based trends of various outlooks toward COVID-19 vaccine activities. We analyzed their correlation with prominent factors for 5 countries (India, the United States, Brazil, the United Kingdom, and Australia) located on 5 different continents to demonstrate the comparative results among them.

Recent research work has analyzed vaccine hesitancy or sentiment analysis to determine the overall general perception among people toward COVID-19 vaccines. Our work provides a more detailed insight into the variety of outlooks people had toward the emergence of continuous vaccine updates and possible correlations with reasons for these outlooks. Major analysis work on survey data in specific regions or a cohort of the population has helped understand people's opinions toward vaccine uptake or resistance. Still, we have worked on a large corpus of tweets (more than 1.8 million) from different countries. As the meteoric rise in the use of social media has become a substantial influencing source for formulating different perceptions in millions of users, working with such a data source helps gain a broader and better sense of various factors that might be associated with fueling vaccine resistance. We have also analyzed our findings with vaccine developments and news in each country during the specific time periods to support our results.

Methods

Design and Data Set

We performed an observational study by curating a longitudinal data set by scraping more than 1.8 million tweets using the Snscrape library [23] from June 2020 to April 2021. The query used to extract the tweets was created using an "OR" combination of hashtags and words related to vaccines and the names of the vaccines administered in the respective countries. Detailed queries for each country are mentioned in Table 1.

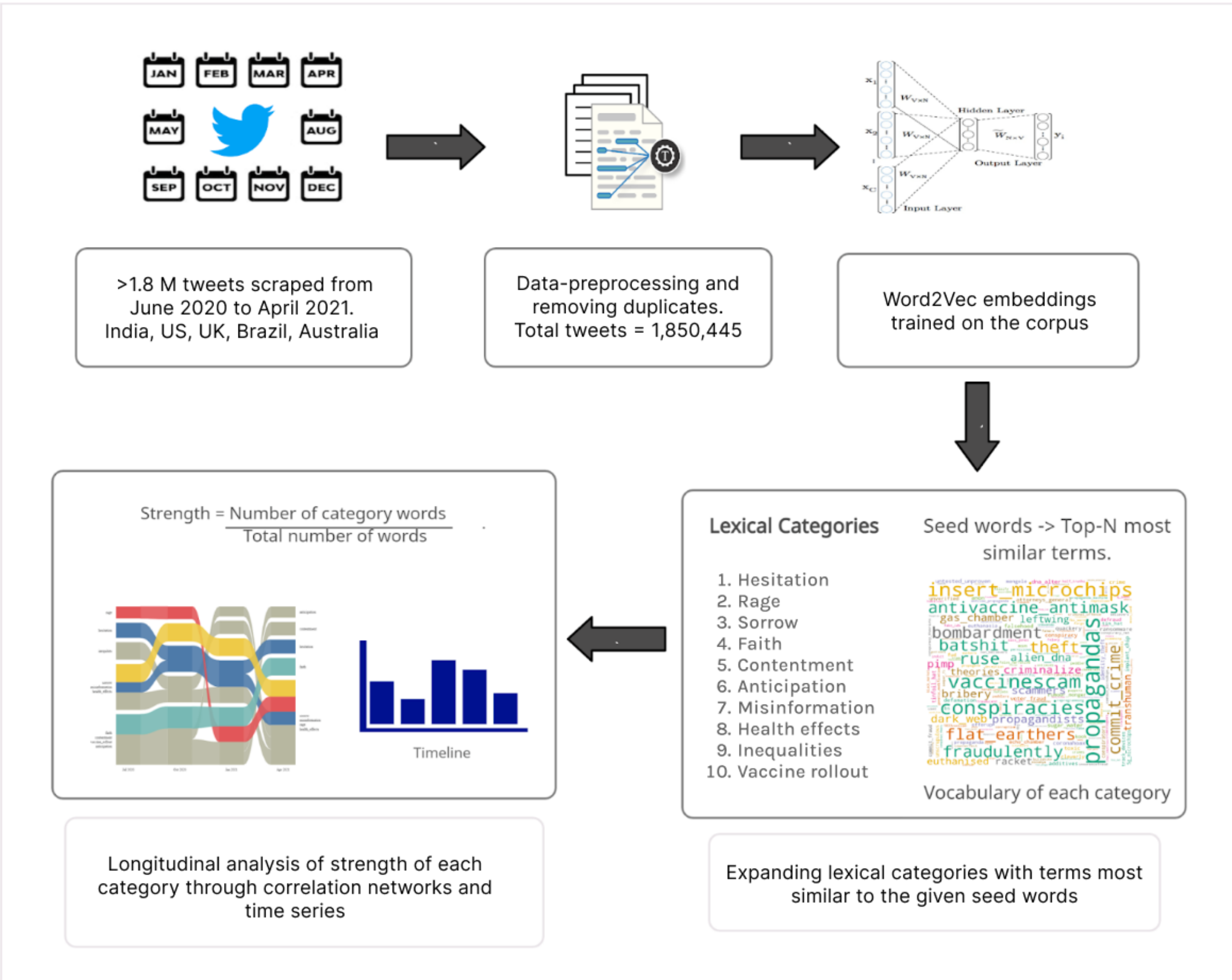
Preprocessing of tweets was carried out on lowercase-converted text by removing white spaces, punctuation, hashtags, mentions, digits, stop words, URLs, and HTML characters. The verbs present in the text were lemmatized using WordNet Lemmatizer from the Natural Language Toolkit (NLTK) package [24]. Duplicate tweets were removed based on identical username, time, and location. Figure 1 illustrates an abstract view of the study design. We list all the software and packages used in further analysis along with the corresponding versions and sources in Multimedia Appendix 1.

Table 1. Queries used for scraping tweets from each country and number of tweets used after preprocessing.

Country	Query[a]	Tweets, n
United States	(General keywords) OR (moderna OR pfizer OR biontech OR astrazeneca OR inovio OR novavax OR #pfizerbiontech)	1,121,216
United Kingdom	(General keywords) OR (pfizer OR biontech OR oxfordvaccine OR astrazeneca OR moderna OR #pfizerbiontech)	432,271
India	(General keywords) OR (covishield OR covaxin)	229,127
Australia	(General keywords) OR (pfizer OR biontech OR oxfordvaccine OR astrazeneca OR moderna OR novavax OR #pfizerbiontech)	50,224
Brazil	(General keywords) OR (coronavac OR Sinovac OR AstraZeneca OR Pfizer OR BioNTech OR #pfizerbiontech OR oxfordvaccine)	17,608

[a]General keywords: (vaccine OR vaccination OR vaccinate OR covax OR #covidvaccine OR #coronavaccine OR #covidvaccination).

Figure 1. Overview of the pipeline followed to create and analyze the strength of lexical categories.



Ethics Approval

Publicly available Twitter data were used, and an aggregated analysis was performed without any attempt to re-identify or link any personal information. The study received institutional review board approval (IIITD/IEC/08/2021-6) and was conducted under the oversight of the associated protocol.

Curating Categories Using Unsupervised Word Embeddings

We created 10 lexical categories for a psychometric evaluation of the tweet content in an approach similar to that by Empath [25]. The categories formed can be broken down into 2 classes: "emotions" and "influencing factors." Emotions consist of the affective processes that help us understand how reactions, feelings, thoughts, and behavior of people evolve in a given situation. We selected 6 COVID-19–related emotions, namely hesitation, rage, sorrow, faith, contentment, and anticipation, along with their putative influencing factors such as misinformation, vaccine rollout, inequities, and health effects in contrast to the COVID-19 vaccines. We specified a set of seed words corresponding to these categories, as shown in Table 2.

We trained a low dimensional representation (d=100) as word embeddings for the unigrams and frequently occurring bigrams (co-occurring at least 5 times with the bigram scoring function [26] greater than a threshold of 50) present in our corpus using the skip-gram algorithm of the Word2Vec model [27] with a sliding window size of 5. We defined lexical categories as sets of words most similar to the assigned seed words. Each seed word, ensured to be present in the model's vocabulary, was mapped to a word vector. We used cosine similarity to measure proximity to find the top N(=50) words in the nearby vector space. Following this approach, k seed words were expanded to a list of maximum k×N words. A category was defined as the union set of seed words and their closest similar words (Table 2). Seed words used for the health effects category were taken from the adverse events mentioned in the Vaccine Adverse Event Reporting System (VAERS) database [28], which occurred in our data set's vocabulary. The resulting set of words in each lexical category was manually verified.

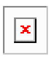
Table 2. Curated categories (emotions and influencing factors), their description, and seed words.

Category	Description	Seed words
Emotions		
1. Hesitation	Sceptic attitude and reluctance toward being vaccinated due to multiple negative factors affecting an individual's opinions	Anxious, nervous, fear, consequences, uncertain, hesitation, suspicion, harm
2. Sorrow	Dissatisfaction and disapproval toward the different phases of COVID-19 vaccine production and distribution	Sad, hopeless, worst, disappointment, setback
3. Faith	Signifies strong belief and confidence in vaccines along with optimistic behavior toward the success of vaccines	Faith, optimism, vaccines work, assurance, grateful
4. Contentment	Signifies a state of happiness, appreciation, and acceptance of the COVID-19 vaccines	Satisfy, glad, proud, gratitude, great, joy
5. Anticipation	State of urgent demand and necessity of vaccines	Anticipate, urgently, priority, quick, await
6. Rage	Anger or aggression is associated with conflict arising from a particular situation	Angry, annoyance, hate, mad, pathetic
Influencing factors		
7. Misinformation	Propagation of false information such as misinterpreted agendas and conceiving vaccines as conspiracy or scam	Propaganda, conspiracy, fraud, fake, poison
8. Vaccine rollout	Availability and distribution of vaccines through campaigns and mass vaccination drives	Vaccinate, distribution, supply, mass, dose, vaccination drive
9. Inequities	Socioeconomic disparities are based on societal norms such as caste, race, religion	Socioeconomic, deprive, racial injustice, racism, underrepresented
10. Health effects	Mentions of health-related adverse events caused by or affected by vaccines, including diseases, symptoms, and pre-existing conditions	From the VAERS[a] database (eg, headache, fatigue, inflammation)

[a]VAERS: Vaccine Adverse Event Reporting System.

Temporal Analysis of Lexical Categories

To measure each category's strength in a given text, we used the word count approach, similar to that by Empath [25] and other lexicon-based tools like Linguistic Inquiry and Word Count (LIWC) [29]. To obtain an unbiased value that is independent of the length of text, we divided the frequency by the total number of words using the following formula:

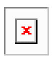


We appended the preprocessed text of all tweets monthly to calculate the strength. The time series of the strength of emotion categories and influencing factors was helpful in analyzing the evolution of perceptions and opinions expressed by the public and how they vary with crucial time stamps like the news of the country's first vaccine approval.

Analysis of Change Before and After Approval

To understand the variation of emotions among social media users in the aftermath of the approval of vaccines, we conducted a before-after change analysis for each lexical category based on the date when the country's government approved the first COVID-19 vaccine.

We created a day-wise time series of the strength of each category from June 2020 to April 2021 and smoothened it using the Moving Average algorithm. The linear nature of the trend was captured using an ordinary linear regression model fit on the strength of a category in the 2 time periods preceding and succeeding the approval date. To calculate the significance of the change, we used the *z* test to compare the regression coefficients [30]:



where b_1 and b_2 denote the slopes and and are the standard errors of the regression lines and before and after the approval, respectively.

Further, we used a change-point detection method based on dynamic programming using the Ruptures package [31] in Python3. The “Dynp” model was used with the “l1” cost function to detect one change point. This was done to verify if the date of approval was close to the change point.

To understand the Influencing factors co-occurring with hesitation, we resampled the tweets with a positive strength of hesitation (n=1000) and calculated the percentage of tweets that also had positive strength of anticipation, rage, misinformation, health effects, and inequities. The resampling was repeated for 100 iterations, and the mean and standard errors were plotted. The percentages of tweets from each of these categories that changed before and after the approval were recorded and tested for significance.

Longitudinal Correlation-Based Networks

The correlation between any 2 categories represents the degree to which they are linearly related. Daily strengths were calculated for each category followed by pairwise Pearson correlation [32]. Weighted networks of categories (nodes) and edge strengths (correlation coefficients) were constructed to evaluate the positive associations among classes ($\rho \geq 0$). Community detection on these networks was carried out using the Infomap algorithm [33], and the dynamic change in these associations was visualized as an alluvial diagram [34]. The use of the Pearson correlation typically requires the verification of some assumptions. We verified the assumption of outliers by plotting box plots of the samples and observed very few or no outliers. To check for a normal distribution, we used the Shapiro-Wilk test (used for n_samples<50), which was satisfied for most but not all months. Hence, we also present the analyses using Spearman correlation, a nonparametric measure, to construct the alluvial diagrams, as shown in Figure S1 in Multimedia Appendix 2.

Results

Analysis of Lexical Categories

Unsupervised word embeddings capture the context of words in the latent space based on their distribution and patterns of co-occurrence [35]. Given the noisy nature of social media data, it becomes difficult to implement a predefined lexicon-based approach with appropriate semantic inclusion. In this paper, we used unsupervised word embeddings trained on our corpus of tweets to find the words most similar to a given set of seed words, hence expanding the vocabulary of a lexical category. Table 3 shows the words belonging to the categories of hesitation and misinformation. The lexical category of hesitation represents words such as “skeptical,” “disillusionment,” “needle-phobic,” “dissonance,” and “consequence,” which demonstrate the uncertainty and doubt regarding vaccines and their effects. Some of the words most similar to “conspiracy” were found to be “implant_microchips” ($\cos\theta=0.844$), “qanon_conspiracy” ($\cos\theta=0.820$), “tinfoil_hat” ($\cos\theta=0.808$), and “echo_chamber” ($\cos\theta=0.806$). These terms denote how people link vaccines to unconventional concepts and propaganda.

Table 3. Words belonging to the lexical categories of hesitation and misinformation, representing the vocabulary expanded from the seed words of the respective categories.

Category	Category words
Hesitation	Confusions, trade_off, shortterm_longterm, frustrate, damage, popularize, apprehension, notions, tire, harmful
Misinformation	Frenzy, propaganda, lethal_injection, false_narratives, black_ market, insert_microchips, euthanised, unsafe_untested, non_believers, conspiracy_theory

Change in Trends Before and After Approval

The difference in slopes of the linear trends of the before and after periods for each category demonstrate 2 significant inferences: the magnitude of change and the direction of change. Figure 2A shows the trends for hesitation in India. A significant change in the direction of the slope is evident (z=10.37, P<.001), which depicts a decrease in its strength after the approval. There was a significant increase (z=–7.65, P<.001) in the magnitude of tweets expressing contentment during the vaccination phase in the United States as shown in Figure 2B. The detected change point was found to be lying within the ranges of 6 days (Figure 2A) and 10 days (Figure 2B) of the date of approval.

The percentage of tweets belonging to different categories was analyzed from the sample of tweets before and after the approval of vaccines in each country. Figure 3A shows that faith and contentment were both significantly higher (both P<.001) before the approval of the first vaccine in India on January 01, 2021 [36]. The factors co-occurring with hesitation were analyzed by calculating the percentage of tweets of 5 other categories (Figures 3C and 3D). Our findings suggest that mentions of health effects contributed the most in tweets with a positive hesitation score. Rage and discussions on misinformation became significantly higher (both P<.001) in the vaccination phase in India (Figure 3C), while an opposite trend was observed in the United States after approval on December 10, 2020 (Figure 3D) [37]. Similar analysis for the United Kingdom, Brazil, and Australia is shown in Figure S2 in Multimedia Appendix 2.

Figure 2. Linear variation in the strength of (A) hesitation in India and (B) contentment in the United States. The dotted line represents the date of approval, and the light blue line depicts the detected change point.

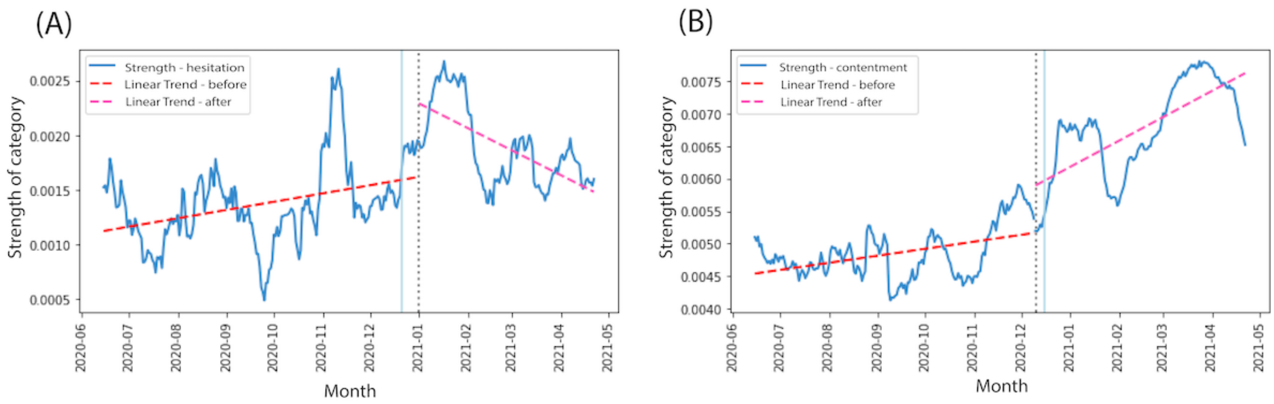
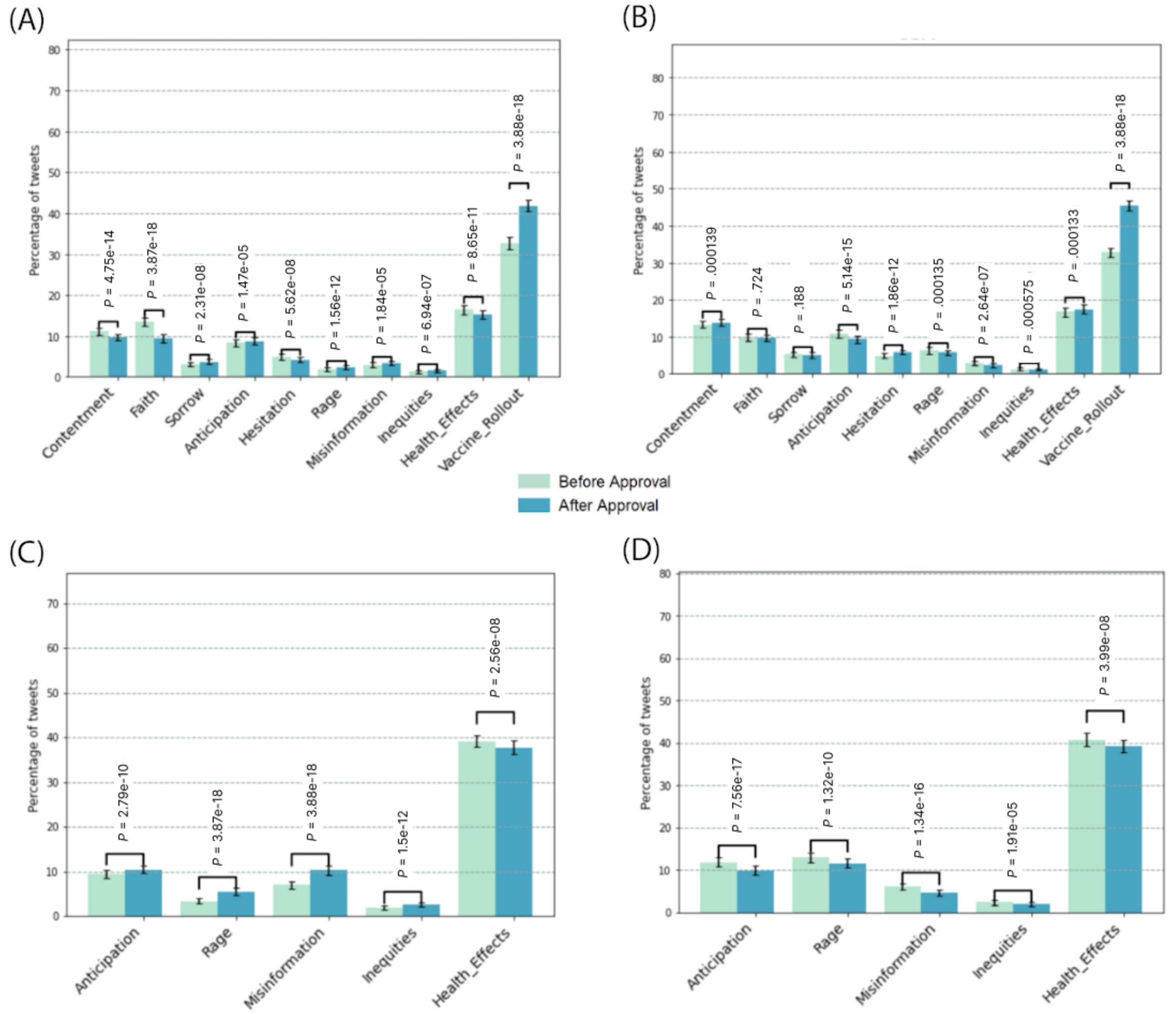


Figure 3. Percentage of tweets with a positive strength in each lexical category before and after approval of COVID-19 vaccine in (A) India (January 1, 2021) and (B) the United States (December 10, 2020) and the percentage of anticipation, rage, misinformation, inequities, and health effects in positive "hesitancy" tweets in (C) India and (D) the United States.



Longitudinal Analysis Using an Alluvial Diagram

Inferences from the alluvial diagrams (Figure 4A) based on Infomap clustering on Pearson correlation networks demonstrated that all the influencing factors (ie, misinformation, health effects, inequities, and vaccine rollout) formed a primary module with emotions of sorrow and rage, which gained the highest PageRank in April 2021, the time when India saw the second wave of COVID-19 cases while the vaccine rollout continued. This articulates the stern sentiment of disappointment due to rising issues and the nonavailability of vaccines for people under the age of 45 years. It also had a high correlation with tweets mentioning the spread of misinformation. Faith, contentment, and anticipation, which were found to be highly associated in the early months of July 2020 and October 2020, were found to be relatively less important and unrelated in April 2021.

On the contrary, lexical categories representing positive sentiment in the United States evolved to a significant module. Faith, contentment, and anticipation toward the vaccine were found to have a positive correlation with each other (Figure 4B). Hesitation was the emotion influenced by mentions of health effects and inequities, whereas rage, sorrow, and misinformation were seen as less central factors in the United States.

Analysis of the temporal trend of misinformation, hesitation, and rage in the 5 countries is depicted in Figure 5. Updates regarding vaccinations started increasing near the end of 2020, which led to changing trends for hesitation expressed on Twitter. A notable inference from the line plots is that hesitation started rising from the beginning of 2021 when primary vaccination drives were initiated. In addition to this, rage is highly expressed in the tweets from the United States, while mentions of misinformation-related terms represented more significant proportions in India and the United Kingdom. Lexical categories of hesitation and rage were found to have similar trends, suggesting a tentative association between the 2 categories.

Figure 4. Alluvial diagram for correlation-based networks showing the evolution of categories from July 2020 to April 2021 at an interval of 3 months in (A) India and (B) the United States.

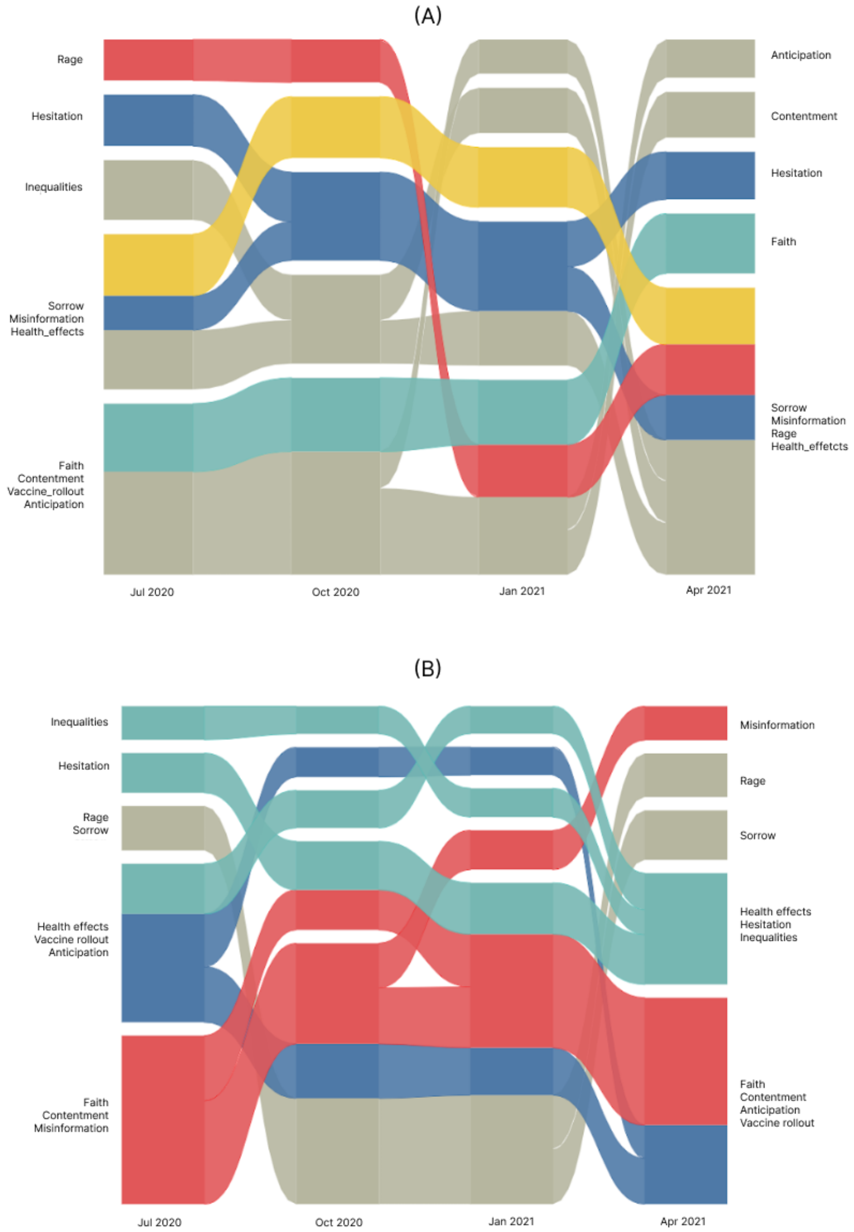


Figure 5. Comparing the temporal flow of strength of 3 categories (misinformation, hesitation, rage) for 5 countries: (A) United States, (B) India, (C) the United Kingdom, (D) Brazil, and (E) Australia.

Discussion

The rise in social media platforms, such as Twitter, has resulted in a valuable source to understand temporal variation in multiple affective and social categories. Influencing factors represented by word embedding–based lexical categories, namely misinformation, vaccine rollout, inequities, and health effects, significantly assisted in studying public perceptions toward emerging vaccine updates from initial approvals to rollout and administration.

Principal Findings

Widespread misinformation being articulated through social media creates panic among users [38]. The misinformation category contains terms similar to "scam" and "conspiracy" from our data set that helped capture references of such words in the context of COVID-19 vaccines. High reporting of adverse effects and severe symptoms in rare cases leading to death [39] becomes a significant factor in increasing vaccination hesitation. The seed words given in the health effects category from the VAERS database led to the formation of its vocabulary containing "restless_sleep," "skin_sensitivity," "hot_flash," "flulike_symptoms," "complications," and more. The semantic similarity-based approach allowed customization of categories according to our data set while ensuring the inclusion of rather noisy words like "feverish" and "achiness," which cannot precisely be found in medical databases.

Inequalities based on socioeconomic status, religion, race, or demographics are standard in different countries, which can lead to inconsistencies while distributing vaccines. The inequities category encapsulated terms related to socioeconomic disparities and helped us identify the impact on other emotions. Based on inspection of our data set of tweets, we found words like "bigotry," "underprivileged," "financial_hardship," and "institutional_racism" were occurring in a highly similar context toward vaccine distribution. Expression of inequities in April 2020 was found to be significantly anticorrelated with faith (P=.03) in India. Inaccessibility to vaccines in marginalized groups has led to lower gratification and higher anxiety among these groups [40].

We analyzed tweets from 5 countries belonging to different continents to get the generalized outlook toward vaccines and how they affect the global immunization process. Figure S3 in Multimedia Appendix 2 depicts sorrow, rage, and misinformation during April 2021 in the United Kingdom as the central module, with the highest PageRank. The Medicines and Healthcare products Regulatory Agency of the United Kingdom issued a new advisory during that period, concluding a possible link between AstraZeneca's COVID-19 vaccine and extremely rare, unlikely occurrences of blood clots [41]. Upon a high-level investigation of the tweets from this period in the United Kingdom, we noticed that this press release had prompted multiple users to talk about blood clots due to the AstraZeneca vaccine. This could have been a potential contributing factor to the high strength of negative emotions expressed on social media platforms. Figure S4 in Multimedia Appendix 2 shows the alluvial diagram for Brazil. The category of rage, which was a relatively less important and independent module in the early months, had associations with sorrow and misinformation in April 2021 in Brazil. It aligned with a major peak in the numbers of cases and deaths during that period of the pandemic in Brazil [42]. In Figure S5 in Multimedia Appendix 2, we can see that faith, contentment, and vaccine rollout were relatively lower than other categories during July 2020, but later in April 2021, they formed a module with anticipation and gained the highest relative importance in the alluvial diagram. The announcement by the Australian government of securing an additional 20 million doses of the Pfizer-BioNTech COVID-19 vaccines overnight [43] happened in April 2021, and multiple tweets expressing optimism possibly contributed to the observed trend. Australia entered into 4 separate agreements with Pfizer, AstraZeneca, Novavax, and COVAX for the supply of COVID-19 vaccines, which resulted in a total number of approximately 170 million vaccine doses, as announced by the Prime Minister.

Related Work

Existing literature on understanding vaccine hesitancy primarily focuses on defined questions from a part of the population belonging to a specific country [44-46]. Although such studies using surveys can help understand the explicit reasoning provided by the individuals, they still pose a limitation on inculcating the variation in outlooks of a larger population over a long period of time. We aimed to fill these gaps by studying important events, such as vaccine trials, highest reported deaths, or import and export of new vaccines, that fueled different populations' emotions, as social media platforms are highly influential due to their comprehensive access and popularity. Our psychometric analysis considers important time stamps and a broader category of emotions to understand the before-after change and the factors with which they associate.

Identification of psychological processes that distinguish between vaccine-hesitant and receptive groups has been carried out in recent research [47]. This helps broadcast public health advisories on social media platforms by strategically taking into account the user's perspective. Effective public health interventions encouraging the uptake of COVID-19 vaccines have benefitted from psychologically oriented approaches [48,49].

Research around understanding the themes and general sentiments toward vaccination programs by analyzing social media posts has also been conducted [50,51]. Although their work provides an overview of positive, negative, or neutral sentiment around other important global developments affiliated with COVID-19 vaccine trials, our analysis provides intricate granularity in understanding the nature of emotions, temporal trends, and the influencing factors that have the highest correlations. Our pipeline effectively clusters the emotion categories and influencing factors around important time stamps based on vaccine approval with categories ranging from negative emotions like hesitation, rage, and sorrow to positive categories like contentment and faith. We further provide a framework to establish lexical categories for understanding the influencing factor correlation and its strength across crucial events. Identification of conspiracy theories related to COVID-19 vaccines has also been carried out [52], which can further be leveraged in addition to our work for improving the understanding of the underlying dynamics of social media posts and disrupting the spread of such content for improving vaccine uptake and tackling hesitancy.

Limitations

Our study has some limitations. We extracted the tweets based on an empirical search of keywords and hashtags relevant to our study in "OR" combination with names of vaccines in the respective countries. Although this approach casts a wide net to retrieve tweets representing discourse around these vaccines, it does not guarantee that all posts were related to COVID-19 vaccine conversations specifically. The chosen keywords for the queries also might not include all relevant terms for capturing tweets specific to our objective. Our framework scores the emotions and influencing factors based on a normalized word count criteria and may miss nuanced language such as sarcasm. However, we interpreted our scores as the amount of discussion happening related to that category, such as hesitancy. Further, the selected categories for our framework are commonly identified emotions that indicate people's perception toward vaccines. Our framework is designed to capture new categories and can be easily expanded and updated periodically to include relevant factors and emotion categories guided by contemporary patterns. Finally, a limitation of our study pertains to the representation bias inherent to social media–based analytics. However, considering that misinformation spreads the fastest through social media and we are considering trends, instead of absolute values, the results are expected to be fairly reliable. Future work may include segmentation of the trends by user demographics, and this information can help in developing tailored solutions for promoting inclusion of minority communities in campaigns. Vaccination drives and policies are targeted heavily toward older populations and minority groups that might not be an active part of such social media platforms. Therefore, for a better understanding of people's opinions toward vaccines, further exploration via other mediums targeting various communities is essential.

Conclusion

Our study provides research and practical implications for public policy making and research on vaccine hesitancy. Our findings offer insights into how the different stages of a pandemic and vaccination process influence emotions and crucial factors like misinformation, health discussions, and socioeconomic disparities on Twitter. This can help decision makers to navigate better solutions in future waves of COVID-19 or similar outbreaks and design appropriate interventions. Our approach can also be utilized to understand the general perception of people during such situations and what preventive measures should be implemented, taking the various influencing factors into account.

Future work can take the direction of local region-level analysis for a specific country to understand the granular emotions within different sections of people and the contributing factors behind them. Providing some weight to the number of reshares and likes the social media post gets can also play an essential role in including the influence the post had in calculating overall strength. Our approach has high adaptability and can be utilized for any online forum, news, or survey data to extract various insights. Designing categories and performing temporal analysis on social media data can also be used to identify multiple ongoing issues like the unavailability of medical resources like oxygen concentrators, intensive care unit beds, and drugs during the second wave of COVID-19. Such analysis can be taken into account while formulating quality allocation of scarce resources based on various factors and their strength. Better information extraction and understanding of such data can be facilitated through our work.

Acknowledgments

TS acknowledges support from the Department of Biotechnology (DBT) Project BT/PR/34245/AI/133/9/2019 and Delhi Knowledge Development Foundation (DKDF) Center of Excellence in Healthcare Project SRP-206 and Delhi Cluster - Delhi Research Implementation and Innovation (DRIIV) Project supported by the Principal Scientific Advisor Office.

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. *JMIR Publications* provided APF support for the publication of this article.

Authors' Contributions

HC and AV designed and implemented the computational framework, interpreted the results, and wrote the paper. RP contributed to writing and provided feedback on statistical methods. Authors A and AT scraped the tweets and curated the data set. TS designed the study, analyzed the results, and contributed to writing. All authors read and approved the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of software and packages used for our study with their sources and identifiers for the reproducibility of this study.

[DOCX File , 18 KB - infodemiology_v3i1e34315_app1.docx]

Multimedia Appendix 2

Supplementary figures.

[DOCX File , 1044 KB - infodemiology_v3i1e34315_app2.docx]

References

1. Timeline: WHO's COVID-19 response. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline [accessed 2023-03-18]
2. Greenwood B. The contribution of vaccination to global health: past, present and future. Philos Trans R Soc Lond B Biol Sci 2014 Jun 19;369(1645):20130433 [FREE Full text] [doi: 10.1098/rstb.2013.0433] [Medline: 24821919]
3. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416. [doi: 10.1056/nejmoa2035389]
4. Pfizer and BioNTech Conclude Phase 3 Study of COVID-19 Vaccine Candidate, Meeting All Primary Efficacy Endpoints. Pfizer. 2020 Nov 18. URL: https://www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-conclude-phase-3-study-covid-19-vaccine [accessed 2023-03-18]
5. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
6. Q1 2020 Letter to Shareholders. Twitter. 2020 Apr 30. URL: http://q4live.s22.clientfiles.s3-website-us-east-1.amazonaws.com/826641620/files/doc_financials/2020/q1/Q1-2020-Shareholder-Letter.pdf [accessed 2023-03-18]
7. Chen W, Stoecker C. Mass media coverage and influenza vaccine uptake. Vaccine 2020 Jan 10;38(2):271-277. [doi: 10.1016/j.vaccine.2019.10.019] [Medline: 31699506]
8. Yoo B, Holland M, Bhattacharya J, Phelps C, Szilagyi P. Effects of mass media coverage on timing and annual receipt of influenza vaccination among Medicare elderly. Health Serv Res 2010 Oct;45(5 Pt 1):1287-1309 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01127.x] [Medline: 20579128]
9. 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 May 12;16(5):e0251095 [FREE Full text] [doi: 10.1371/journal.pone.0251095] [Medline: 33979370]
10. Rovetta A. The impact of COVID-19 on conspiracy hypotheses and risk perception in Italy: infodemiological survey study using Google Trends. JMIR Infodemiology 2021 Aug 6;1(1):e29929 [FREE Full text] [doi: 10.2196/29929] [Medline: 34447925]

11. Tagliabue F, Galassi L, Mariani P. The "Pandemic" of disinformation in COVID-19. SN Compr Clin Med 2020;2(9):1287-1289 [FREE Full text] [doi: 10.1007/s42399-020-00439-1] [Medline: 32838179]
12. Dhanashree, Garg H, Chauhan A, Bhatia M, Sethi G, Chauhan G. Role of mass media and it's impact on general public during coronavirus disease 2019 pandemic in North India: An online assessment. Indian Journal of Medical Sciences 2021 May 29;73(1):21-25. [doi: 10.25259/IJMS_312_2020]
13. Tsao S, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. The Lancet Digital Health 2021 Mar;3(3):e175-e194. [doi: 10.1016/s2589-7500(20)30315-0]
14. Arora A, Chakraborty P, Bhatia MPS, Mittal P. Role of emotion in excessive use of Twitter during COVID-19 imposed lockdown in India. J Technol Behav Sci 2021 Oct 20;6(2):370-377 [FREE Full text] [doi: 10.1007/s41347-020-00174-3] [Medline: 33102690]
15. Aggrawal P, Jolly BLK, Gulati A, Sethi A, Kumaraguru P, Sethi T. Psychometric analysis and coupling of emotions between state bulletins and Twitter in India during COVID-19 infodemic. Front. Commun 2021 Sep 24;6:2005 [FREE Full text] [doi: 10.3389/fcomm.2021.695913]
16. Kim EH, Jeong YK, Kim Y, Kang KY, Song M. Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. Journal of Information Science 2016 Jul 11;42(6):763-781. [doi: 10.1177/0165551515608733]
17. Nagarajan K, Muniyandi M, Palani B, Sellappan S. Social network analysis methods for exploring SARS-CoV-2 contact tracing data. BMC Med Res Methodol 2020 Sep 17;20(1):233 [FREE Full text] [doi: 10.1186/s12874-020-01119-3] [Medline: 32942988]
18. 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]
19. Rodrigues CMC, Plotkin SA. Impact of vaccines; health, economic and social perspectives. Front Microbiol 2020 Jul 14;11:1526 [FREE Full text] [doi: 10.3389/fmicb.2020.01526] [Medline: 32760367]
20. 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 05;5(3):337-348. [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
21. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020 Oct 14;7(10):201199 [FREE Full text] [doi: 10.1098/rsos.201199] [Medline: 33204475]
22. Kantamneni N. The impact of the COVID-19 pandemic on marginalized populations in the United States: A research agenda. J Vocat Behav 2020 Jun;119:103439 [FREE Full text] [doi: 10.1016/j.jvb.2020.103439] [Medline: 32390658]
23. JustAnotherArchivist / snscrape. GitHub. URL: https://github.com/JustAnotherArchivist/snscrape [accessed 2023-03-18]
24. Bird S, Loper E. NLTK: The Natural Language Toolkit. Proceedings of the ACL Interactive Poster and Demonstration Sessions 2004:214-217 [FREE Full text] [doi: 10.3115/1219044.1219075]
25. Fast E, Chen B, Bernstein MS. Empath: Understanding Topic Signals in Large-Scale Text. 2016 Presented at: CHI Conference on Human Factors in Computing Systems; May 7, 2016; San Jose, CA. [doi: 10.1145/2858036.2858535]
26. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed Representations of Words and Phrases and their Compositionality. 2013 Presented at: Advances in Neural Information Processing Systems 26 (NIPS 2013); December 5-10, 2013; Lake Tahoe, NV URL: https://arxiv.org/abs/1310.4546
27. Ma L, Zhang Y. Using Word2Vec to process big text data. 2015 Presented at: IEEE International Conference on Big Data (Big Data); October 29-November 1, 2015; Santa Clara, CA. [doi: 10.1109/bigdata.2015.7364114]
28. VAERS Data Sets. Vaccine Adverse Event Reporting System. URL: https://vaers.hhs.gov/data/datasets.html [accessed 2023-03-18]
29. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927X09351676]
30. Paternoster R, Brame R, Mazerolle P, Piquero A. Using the correct statistical test for the equality of regression coefficients. Criminology 1998 Nov;36(4):859-866. [doi: 10.1111/j.1745-9125.1998.tb01268.x]
31. Truong C, Oudre L, Vayatis N. Selective review of offline change point detection methods. Signal Processing 2020 Feb;167:107299. [doi: 10.1016/j.sigpro.2019.107299]
32. Benesty J, Chen J, Huang Y, Cohen I. Pearson Correlation Coefficient. In: Noise Reduction in Speech Processing. Springer Topics in Signal Processing, vol 2. Berlin, Heidelberg: Springer; 2009:1-4.
33. Bohlin L, Edler D, Lancichinetti A, Rosvall M. Community Detection and Visualization of Networks with the Map Equation Framework. In: Ding Y, Rousseau R, Wolfram D, editors. Measuring Scholarly Impact. Cham, Switzerland: Springer; 2014:3-34.
34. Rosvall M, Bergstrom CT. Mapping change in large networks. PLoS One 2010 Jan 27;5(1):e8694 [FREE Full text] [doi: 10.1371/journal.pone.0008694] [Medline: 20111700]
35. Tshitoyan V, Dagdelen J, Weston L, Dunn A, Rong Z, Kononova O, et al. Unsupervised word embeddings capture latent knowledge from materials science literature. Nature 2019 Jul 3;571(7763):95-98 [FREE Full text] [doi: 10.1038/s41586-019-1335-8] [Medline: 31270483]

36. Coronavirus: India approves vaccines from Bharat Biotech and Oxford/AstraZeneca. BBC News. 2021 Jan 03. URL: https://www.bbc.com/news/world-asia-india-55520658 [accessed 2023-03-18]
37. Thomas K, Weiland N, LaFraniere S. F.D.A. Advisory Panel Gives Green Light to Pfizer Vaccine. The New York Times. 2020 Dec 17. URL: https://www.nytimes.com/2020/12/10/health/covid-vaccine-pfizer-fda.html [accessed 2023-03-18]
38. Burki T. Vaccine misinformation and social media. The Lancet Digital Health 2019 Oct;1(6):e258-e259. [doi: 10.1016/S2589-7500(19)30136-0]
39. Torjesen I. Covid-19: Norway investigates 23 deaths in frail elderly patients after vaccination. BMJ 2021 Jan 15;372:n149. [doi: 10.1136/bmj.n149] [Medline: 33451975]
40. Larson HJ, Cooper LZ, Eskola J, Katz SL, Ratzan S. Addressing the vaccine confidence gap. The Lancet 2011 Aug;378(9790):526-535. [doi: 10.1016/s0140-6736(11)60678-8]
41. Medicines and Healthcare products Regulatory Agency. MHRA issues new advice, concluding a possible link between COVID-19 Vaccine AstraZeneca and extremely rare, unlikely to occur blood clots. GOV.UK. 2021 Apr 07. URL: https://tinyurl.com/4d2ykmyv [accessed 2023-03-18]
42. Roser M, Ritchie H. Coronavirus Pandemic (COVID-19). Our World in Data. URL: https://ourworldindata.org/coronavirus [accessed 2023-03-18]
43. Press Conference - Australian Parliament House. Prime Minister of Australia [Internet]. URL: https://www.pm.gov.au/media/press-conference-australian-parliament-house-act-09april21 [accessed 2021-08-03]
44. Bendau A, Plag J, Petzold MB, Ströhle A. COVID-19 vaccine hesitancy and related fears and anxiety. Int Immunopharmacol 2021 Aug;97:107724 [FREE Full text] [doi: 10.1016/j.intimp.2021.107724] [Medline: 33951558]
45. Larson HJ, Jarrett C, Eckersberger E, Smith DM, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012. Vaccine 2014 Apr 17;32(19):2150-2159 [FREE Full text] [doi: 10.1016/j.vaccine.2014.01.081] [Medline: 24598724]
46. Marti M, de Cola M, MacDonald NE, Dumolard L, Duclos P. Assessments of global drivers of vaccine hesitancy in 2014-Looking beyond safety concerns. PLoS One 2017;12(3):e0172310 [FREE Full text] [doi: 10.1371/journal.pone.0172310] [Medline: 28249006]
47. Murphy J, Vallières F, Bentall RP, Shevlin M, McBride O, Hartman TK, et al. Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. Nat Commun 2021 Jan 04;12(1):29 [FREE Full text] [doi: 10.1038/s41467-020-20226-9] [Medline: 33397962]
48. Cameron L, Leventhal H. The self-regulation of health and illness behaviour. New York, NY: Routledge; 2003.
49. Salovey P, Williams-Piehota P. Field experiments in social psychology. American Behavioral Scientist 2016 Jul 27;47(5):488-505. [doi: 10.1177/0002764203259293]
50. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. J Med Internet Res 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: 10.2196/26627] [Medline: 33724919]
51. 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 Nov 17;7(11):e30642 [FREE Full text] [doi: 10.2196/30642] [Medline: 34653016]
52. Shahsavari S, Holur P, Wang T, Tangherlini TR, Roychowdhury V. Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news. J Comput Soc Sci 2020;3(2):279-317 [FREE Full text] [doi: 10.1007/s42001-020-00086-5] [Medline: 33134595]

Abbreviations

LIWC: Linguistic Inquiry and Word Count
NLTK: Natural Language Toolkit
VAERS: Vaccine Adverse Event Reporting System

Edited by T Mackey; submitted 17.10.21; peer-reviewed by A Rovetta, K Dashtipour; comments to author 11.01.22; revised version received 09.03.22; accepted 31.03.22; published 02.05.23.

Please cite as:
Chopra H, Vashishtha A, Pal R, Ashima, Tyagi A, Sethi T
Mining Trends of COVID-19 Vaccine Beliefs on Twitter With Lexical Embeddings: Longitudinal Observational Study
JMIR Infodemiology 2023;3:e34315
URL: https://infodemiology.jmir.org/2023/1/e34315
doi:10.2196/34315
PMID:37192952

Original Paper

Public Officials' Engagement on Social Media During the Rollout of the COVID-19 Vaccine: Content Analysis of Tweets

Husayn Marani[1,2], MSc, PhD; Melodie Yunju Song[2,3], MSc, PhD; Margaret Jamieson[1,2], BIE; Monika Roerig[1,2], MA; Sara Allin[1,2], MSc, PhD

[1]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[2]North American Observatory on Health Systems and Policies, University of Toronto, Toronto, ON, Canada
[3]Centre for Vaccine Preventable Diseases, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Husayn Marani, MSc, PhD
Institute of Health Policy, Management and Evaluation
Dalla Lana School of Public Health
University of Toronto
155 College Street
Toronto, ON, M5T 3M6
Canada
Phone: 1 (416) 978 432
Email: husayn.marani@mail.utoronto.ca

Abstract

Background: Social media is an important way for governments to communicate with the public. This is particularly true in times of crisis, such as the COVID-19 pandemic, during which government officials played a strong role in promoting public health measures such as vaccines.

Objective: In Canada, provincial COVID-19 vaccine rollout was delivered in 3 phases aligned with federal government COVID-19 vaccine guidance for priority populations. In this study, we examined how Canadian public officials used Twitter to engage with the public about vaccine rollout and how this engagement has shaped public response to vaccines across jurisdictions.

Methods: We conducted a content analysis of tweets posted between December 28, 2020, and August 31, 2021. Leveraging the social media artificial intelligence tool Brandwatch Analytics, we constructed a list of public officials in 3 jurisdictions (Ontario, Alberta, and British Columbia) organized across 6 public official types and then conducted an English and French keyword search for tweets about vaccine rollout and delivery that mentioned, retweeted, or replied to the public officials. We identified the top 30 tweets with the highest impressions in each jurisdiction in each of the 3 phases (approximately a 26-day window) of the vaccine rollout. The metrics of engagement (impressions, retweets, likes, and replies) from the top 30 tweets per phase in each jurisdiction were extracted for additional annotation. We specifically annotated sentiment toward public officials' vaccine responses (ie, positive, negative, and neutral) in each tweet and annotated the type of social media engagement. A thematic analysis of tweets was then conducted to add nuance to extracted data characterizing sentiment and interaction type.

Results: Among the 6 categories of public officials, 142 prominent accounts were included from Ontario, Alberta, and British Columbia. In total, 270 tweets were included in the content analysis and 212 tweets were direct tweets by public officials. Public officials mostly used Twitter for information provision (139/212, 65.6%), followed by horizontal engagement (37/212, 17.5%), citizen engagement (24/212, 11.3%), and public service announcements (12/212, 5.7%). Information provision by government bodies (eg, provincial government and public health authorities) or municipal leaders is more prominent than tweets by other public official groups. Neutral sentiment accounted for 51.5% (139/270) of all the tweets, whereas positive sentiment was the second most common sentiment (117/270, 43.3%). In Ontario, 60% (54/90) of the tweets were positive. Negative sentiment (eg, public officials criticizing vaccine rollout) accounted for 12% (11/90) of all the tweets.

Conclusions: As governments continue to promote the uptake of the COVID-19 booster doses, findings from this study are useful in informing how governments can best use social media to engage with the public to achieve democratic goals.

(JMIR Infodemiology 2023;3:e41582) doi:10.2196/41582

KEYWORDS

Twitter; COVID-19; vaccines; sentiment analysis; public officials

Introduction

Background

With the global usership of popular social media platforms such as Twitter, Facebook, and Instagram in the billions, it is no longer a choice but a necessity for government institutions and public officials therein to have a social media presence [1]. For better or worse, the proliferation of social media in recent decades has transformed how governments achieve various public policy goals [2]. This has given rise to e-governance, whereby communication technologies such as social media are leveraged by government institutions to improve information exchange (improving transparency), enhance citizen engagement in democratic processes (increasing participation), and foster collaboration between government institutions and constituents to improve government-related activities (strengthening collaboration) [3].

Numerous frameworks exist to typify the ways in which government institutions interact with the public over social media to achieve various democratic goals. The seminal framework by Mergel [4] organizes the abovementioned goals into a 3-category framework called the *open government framework of interpreting the impact of social media interactions* (herein *government interaction framework*); the framework describes public sector's tactical use of social media to push information, pull information, and network. Each of these tactics relate to improving transparency, increasing participation, and strengthening collaboration, respectively [4]. Additional goals, such as public service delivery [3], self-presentation and marketing [5], and facilitating local social transactions [6], have yielded adaptations to the framework by Mergel [4]. Broadly, these frameworks are useful to answer important questions about how government institutions may interact with the public over social media in specific contexts, for example, by responding, retweeting—which could imply engagement on Twitter, or mentioning others [7]. These frameworks are also useful for evaluating the effectiveness of public response to government social media communication. For example, indicators such as user sentiment [8], audience engagement (volume of *likes*, *comments*, and *shares* by the public) [9], and emotion in textual responses [10] indicate public receptivity toward policy rollout and government responses [11].

Recently, we have seen government institutions leveraging social media to engage the public during times of crisis, such as weather catastrophes and viral epidemics over short periods [12-14]. Studies on these contexts are largely oriented toward local levels of e-governance, as local governments are the closest to citizens and can more effectively achieve interaction goals, such as participation and public service delivery [3,15]. However, there is a relative dearth of the literature concerning how social media is used by government institutions at the national, regional, and local levels during times of crisis. Filling this gap is important in settings where crucial decision-making concerning crises is derived across multiple levels of government—such as in the case of Canada.

Objectives

The COVID-19 pandemic has seen a significant use of social media by the government to achieve various goals, including mitigating the transmission of COVID-19 through information dissemination, encouraging behavioral changes, and promoting the availability of community-based supports and resources [16,17]. Given the decentralized structure of the Canadian health care system, national (federal) and subnational (provincial and territorial) governments played both independent and collaborative roles in responding to the COVID-19 pandemic. Indeed, recent research from Canada has shown that approaches to health communication over social media regarding the risks of COVID-19 differ between jurisdictions and depend in part on the local burden of COVID-19 [18]. This suggests an important responsibility of regional and local governments (provinces, territories, and municipalities) in delivering messaging tailored to the levels of risk audiences are encountering in each geographic context.

Of particular interest to our study are the activities and engagement of decision makers and policy makers, specifically elected public officials, in achieving the goal of population uptake of the COVID-19 vaccine. The vaccine rollout has generated significant policy and public debate in Canada. Achieving high vaccination coverage was seen as the pathway to normalcy, starting with the end of COVID-19 *lockdowns* across the country [19]. Although the federal government is responsible for vaccine procurement and distribution to the provinces and territories, provincial and territorial governments have been responsible for vaccine rollout. In December 2020, all 10 provinces and 3 territories took a phased approach to first-dose vaccine rollout that generally aligned with the advice from the National Advisory Committee on Immunization [20]. Rollout started with health care workers (HCWs), persons living in high-risk congregate settings such as residential long-term care homes, older adults (aged >75 years), and Indigenous communities. This was generally followed by a phased plan based on age cohorts and the presence of health-related risk factors [21,22]. The second dose rollout followed a similar structure, but a more concerted effort was made to prioritize COVID-19 *hotspots* calculated by case positivity rates in postal districts; for instance, the Ministry of Health in Ontario devised an age-based vaccine rollout strategy that prioritized those residing in communal *hotspots* [23,24].

Emerging scholarship from the United States shows that leaders across the political spectrum have used social media, in particular Twitter, to promote uptake of the COVID-19 vaccine [25]. However, little is known about how Canadian public officials engaged with the public through social media during the COVID-19 pandemic. With 77.6% of the Canadians above the age of 15 years regularly using social media and 25.2% of the Canadians on at least 3 social media platforms in 2018 [26], exposure to vaccine-related misinformation and disinformation has threatened public confidence in the vaccine and overall uptake [27]. Accordingly, Canadian public officials have played an important role not only in combating a historical propensity among Canadian populations to refuse vaccines [28], which was documented near the start of the COVID-19 vaccine rollout [29], but also in combating antivaccine misinformation and disinformation on social media, which has been shown to be predictive of vaccine hesitancy [30]. The purpose of this exploratory study was to gain insight into how public officials across federal and provincial governments leveraged social media to communicate with the public about the COVID-19 vaccine rollout, particularly across the 3 phases of rollout when clear public communication was especially important.

Methods

Overview

To identify social media posts, we used Brandwatch Analytics (henceforth Brandwatch), a social media intelligence tool that uses proprietary artificial intelligence to extract and analyze social media data from various social media platforms. Previously used to conduct textual analyses of Twitter data in the context of the COVID-19 pandemic [31], its enterprise Twitter application programming interface has 2 advantages. First, it makes it possible to retrieve retrospective data. Second, it retrieves a higher percentage of tweets within a given time interval, presenting a more representative set of tweets over time. We focused on Twitter for consistency with previous scholarship on health messaging and crisis communication by public officials [18,32] and because Twitter is the most used social media platform by public officials and departments across the Government of Canada to communicate and engage with the public [33].

Data Collection

To obtain tweets about vaccine rollout, we developed an English and French keyword query using a combination of truncation and Boolean operators and wildcards (* and ?), which were used to identify posts containing common root words and letter substitutions, respectively. This query was used to capture conversations about vaccine rollout in Canada between December 28, 2020, and August 31, 2021 (Textbox 1). This query was run twice: first, to canvas Twitter participation of public officials in Canada and second, to derive a sample of tweets for content analysis.

Textbox 1. Keyword query to capture tweets about vaccine rollout.

(((vaccin* OR vax* OR immuniz* OR immunis*) AND (distribu* OR allocat* OR roll-out OR “roll out” OR deliver* OR provid* OR provision* OR administer* OR administr* OR livraison OR apporter OR alloue*)))

Canvassing Twitter Participation of Public Officials in Canada

A first run of the query showed that most of the conversation driving vaccine rollout across Canada (106,834/124,081, 86.1%) occurred in 3 provinces: Ontario (80,156/124,081, 64.6% of the total tweets), Alberta (15,137/124,081, 12.2%), and British Columbia (BC; 11,539/124,081, 9.3%). Unsurprisingly, these are the most Twitter-engaged provinces [34] and are among the most populous provinces in Canada. To narrow down tweets about vaccine rollout posted by, or mentioning, public officials across Canada, we used results from this query to identify the top 20 public officials, irrespective of being *verified* by Twitter, with the highest cumulative engagement (which we estimated based on total impressions) in each of the 3 provinces. This list aimed to supplement an a priori list of public officials generated by the study team of any public officials, including the organizations or provincial government and public health authorities to which they belong, across Canada, who are on Twitter and who have been involved in broad decision-making related to COVID-19 vaccine procurement and rollout. Public officials were organized across 6 categories of public official types (Textbox 2) inspired by another study with a similar organizational framework [18]. In total, 142 user accounts of public officials were included in our second run of the query (Multimedia Appendix 1).

Textbox 2. Six categories of public officials involved in vaccine rollout decision-making.

1. *First ministers (premiers; n=15)*: this includes publicly elected federal and provincial and territorial heads of government, including the Prime Minister, Deputy Prime Minister, and provincial and territorial Premiers.
2. *Ministers of Health (n=16)*: this includes the official, acting, interim, and deputy Ministers of Health in every province and territory.
3. *Chief Medical Officers of Health (n=6)*: this includes verified Twitter accounts of the Chief Public Health Officer of Canada and provincial Chief Public Health and Medical Officers of Health.
4. *Government bodies (n=53)*: this includes official organizational accounts of the Federal Government (ie, Government of Canada); federal public health authorities (ie, Health Canada and Public Health Agency of Canada); provincial governments (ie, Government of New Brunswick) and provincial public health authorities (ie, Saskatchewan Health Authority) who are involved in vaccine rollout and decision-making.
5. *Municipal officials (n=14)*: this includes publicly elected officials at the municipal level, including mayors of capital cities (eg, Toronto, Ottawa, Vancouver, and Calgary).
6. *Other key public officials (n=38)*: this includes elected members of parliament and PT legislative assemblies, which captures Ministers with any form of engagement in vaccine rollout (eg, Minister of Public Services and Procurement) who are not Ministers of Health.

Identifying Public Official Participation on Social Media During the Initiation of Vaccine Rollout Phase Changes in Alberta, BC, and Ontario

We ran our query (Textbox 1) again, this time in combination with the usernames of all 142 public officials identified (Multimedia Appendix 2). This search yielded 133,155 tweets geotagged by Brandwatch as having originated in Canada that mentioned, retweeted, and replied to our list of public officials. Similar to our first query, most tweets (106,834/133,155, 80.23%) came from 3 provinces: Ontario (80,201/106,834, 75.07%), Alberta (15,096/106,834, 14.13%), and BC (11,537/106,834, 10.8%). We selected these 3 provinces for content analysis because these are the 3 most populous provinces in Canada, bar Quebec, which has a low Twitter participation from public officials based on our canvassing activity above (see the *Canvassing Twitter Participation of Public Officials in Canada* section)

To identify the tweets that were driving the conversation during each phase of the vaccine rollout, we sorted all tweets in each province by phase (Multimedia Appendix 3 [35]), and then sorted tweets by impressions (highest to lowest). To ensure that the tweets were temporally consistent with the phase changes, we only extracted tweets that were tweeted 5 days from the onset of the phase change announcement up to 3 weeks following the announcement, for a total of 26 days. For each 26-day period, we extracted the top 30 tweets per phase in each province, resulting in 90 tweets per province and 270 tweets in total. A complete data collection flowchart is presented in Multimedia Appendix 4. The 270 tweets represent original tweets by public officials, quoted tweets (eg, a retweet embedded with personal commentary above the @publicofficial's original tweet), retweets (*RT@publicofficial*), and tweets that mention or reply to public officials (*@publicofficial*). As filtering ad-based tweets was not a function made available to us in Brandwatch, and sponsored content does not require public disclosure, the collected tweets driving conversations about vaccine rollout may include ad-based or sponsored tweets.

Data Extraction and Analysis

We conducted a content analysis of 270 Canadian-geotagged tweets with the highest impressions posted by or engaging with (ie, tagging, mentioning and replying to, or quoting) the 142 public officials accounts. Tweets were divided evenly among the 4 authors (HM, MYS, MJ, and MR) for extraction, sentiment analysis, and content analysis. In summary, the extracted criteria included the text of each tweet, any URLs in the tweet, metrics of engagement (impressions, retweets, likes, replies, and quoted retweets), province from which the tweet was derived, date of tweet, and interaction type. We also manually conducted sentiment analysis of the 270 tweets to determine whether a tweet expressed a positive, negative, or neutral stance toward the process and delivery of the vaccine rollout. A total of 4 coders (HM, MYS, MJ, and MR) participated in the coding; 3 coders were assigned to code the same tweet for all 270 tweets for both sentiment analysis and content analysis, and any discrepancies were resolved by discussion by all authors (HM, MYS, MJ, MR, and SA) if the 3 coders could not reach an agreement.

The seminal framework that Mergel [4] proposed organizes public sector social media interaction types (*push*, *pull*, and *network*) based on observations of the US federal government interacting with the public in response to the Obamacare website crisis. The *public service social media interaction framework* by Criado and Villodre [3] modified the framework by Mergel [4] after testing the framework on tweets collected in localized city councils in 4 European countries. To represent the bulk of public sector interaction, this framework used the term *public service delivery* corresponding to *networking*, *information provision* corresponding to *push*, and *citizeninteraction* corresponding to *pull* by Mergel [4].

Our content analysis of public officials' tweets regarding the COVID-19 vaccine generated an adapted *push and pull* framework that builds on Mergel [4] and Criado and Villodre [3]. In our framework, push-1 (*information provision*) and push-2 (*publicserviceannouncement*) interaction types complement the framework by Criado and Villodre [3] to differentiate the provision of critical updates (push-1: information provision) versus announcements about public service availability, including vaccine eligibility (push-2: public service announcement). Pull-1 (*citizen engagement*) and pull-2 (*public officialengagement*) interaction types reflect nuances across provinces and the need to elevate communication between public officials with each other over society into its own category. Taken together, the modified framework captures all 4 public sector interaction types in large Canadian provinces.

We then assigned a code (ie, *push 1*, *push 2*, *pull 1*, and *pull 2*) to each tweet based on the public service social media interaction framework [3] inspired by Mergel [4] (Textbox 3) that organizes how public officials interact with members of the public over social media to achieve public policy objectives. To understand how well tweets by each category of public officials were endorsed, we calculated an endorsement ratio derived from the number of likes received divided by the number of impressions (views) for each tweet. The endorsement ratio is a value between 0 and 1, with a higher ratio indicating a higher content-specific endorsement on Twitter.

Textbox 3. Types of public officials' social media interaction that may increase vaccine uptake (modified based on the public service social media interaction frameworks of Criado and Villodre and Mergel).

Push 1: information provision

- Refers to one-way social media posting activities that disseminate data, and have broad aims to increase transparency, accountability, and citizen trust.

Push 2: public service announcement

- Refers to one-way social media posting activities that use these platforms for public service transactions, including call for action related to sharing a location or website to sign up for vaccination.

Pull 1: citizen engagement

- Refers to social media posting activities that aim to have a bidirectional engagement between public officials and the public. Interactions include authorities replying, retweeting posts by the public, mentioning accounts of the public, and relaying information from the public.

Pull 2: public official engagement

- Refers to social media posting activities that aim to have a bidirectional engagement among public officials. Interactions include authorities replying, retweeting posts by other authorities, mentioning, and quoting posts from other authorities.

Next, we described characteristics of tweets to guide our coding of sentiment (positive, negative, or neutral) and interaction type (Multimedia Appendix 5). Positive sentiment tweets use emojis, images, and adjectives that support vaccines and vaccine rollout; negative tweets express disagreement and anger toward the vaccine rollout; and neutral tweets are fact-based information with no emotional cues concerning vaccine rollout. Finally, we conducted a thematic analysis of the tweets to add details to our findings on sentiment, interaction type, and content.

Ethical Considerations

Ethics approval was not required for this study as we conducted a secondary analysis of publicly available data.

Results

Overview

The analysis yielded 2 sets of findings. First, we present descriptive results derived from Brandwatch related to the volume of tweets from the 142 public officials accounts, either tweeted by these public officials or by others who have tagged, mentioned or replied to, or quoted these accounts. Next, we present a content analysis of the 270 included tweets, namely engagement (impressions, retweets, likes, and replies), sentiment (with examples of tweets), and interaction type (as described in Textbox 3). To enhance the findings from our content analysis, we present specific tweets that related to themes on sentiment and interaction type. For the 270 tweets coded, we calculated the interrater reliability between the 3 coders using Krippendorff α for nominal data, achieving an α coefficient of .811 for sentiment analysis and .784 for the content analysis of interactions (push-pull dynamics) of tweets, indicating very good and good agreement among coders, respectively.

Description of Tweet Volume by Public Officials Across Phases and Within Included Provinces

A total of 602,050 tweets from 153,200 unique Canadian users (identified by user geotags) were downloaded (Multimedia Appendix 6). During our extraction period, mention volumes were elevated from the end of December to mid-January 2021 (phase 1) across the 3 provinces (Figure 1). The second peak of conversation occurred at the beginning of phase 2 in April 2021 for Ontario (Figure 2) and Alberta (Figure 3), but similar patterns were not observed in BC (Figure 4). Ontario had the most mentions of public health officials (440,013 tweets by 53,431 unique users), BC had 53,472 tweets by 8653 unique users, and Alberta had 108,580 tweets by 13,274 unique users. Multimedia Appendix 7 provides the context for these phase changes, showing the number of vaccines administered in addition to the newly confirmed COVID-19 cases per day (*case positivity*) calculated by the 7-day moving average. The surge in tweet volume in Ontario in phase 2 (on March 15, 2021) coincided with the province's third wave of COVID-19 transmission (Figure 1 and Multimedia Appendix 7). The highest volume of tweets in BC closely preceded BC's phase 3 vaccine rollout and the province's third wave of COVID-19 transmission. Finally, in Alberta, the highest volume of tweets matched the beginning of Alberta's phase 2 vaccine rollout at the height of the province's third wave. In contrast to the expectations, mention volume decreased to varying degrees in all provinces at phase 3, where all provinces witnessed peak COVID-19 transmission (Multimedia Appendix 7).

Among the 142 public official accounts, government bodies (eg, regional health authorities, provincial government, and public health agencies; n=48) were the most prevalent in the vaccine rollout conversation. This pattern was observed in all 3 provinces and across all periods, though it was most evident at the start of phase 2 in Ontario (around March to April 2021) at the beginning of the third wave, during which the public first became eligible for the first dose of the COVID-19 vaccine (Figure 5). Within provinces, government officials had the highest proportion of the mention volume in Ontario (54/90, 60%), likely explained by the number of federal ministers in the data set who reside or tweet from Ottawa. Government officials are somewhat less represented in BC (45/90, 50%) and represent less than half (35/90, 39%) of the mention volume in Alberta. In Alberta and BC, Ministers of Health and first ministers (premiers) engaged more than their counterparts in Ontario (Figure 6).

Figure 1. Mention volume of tweets from or to public officials during the 2020-2021 period, stratified by province. BC: British Columbia.

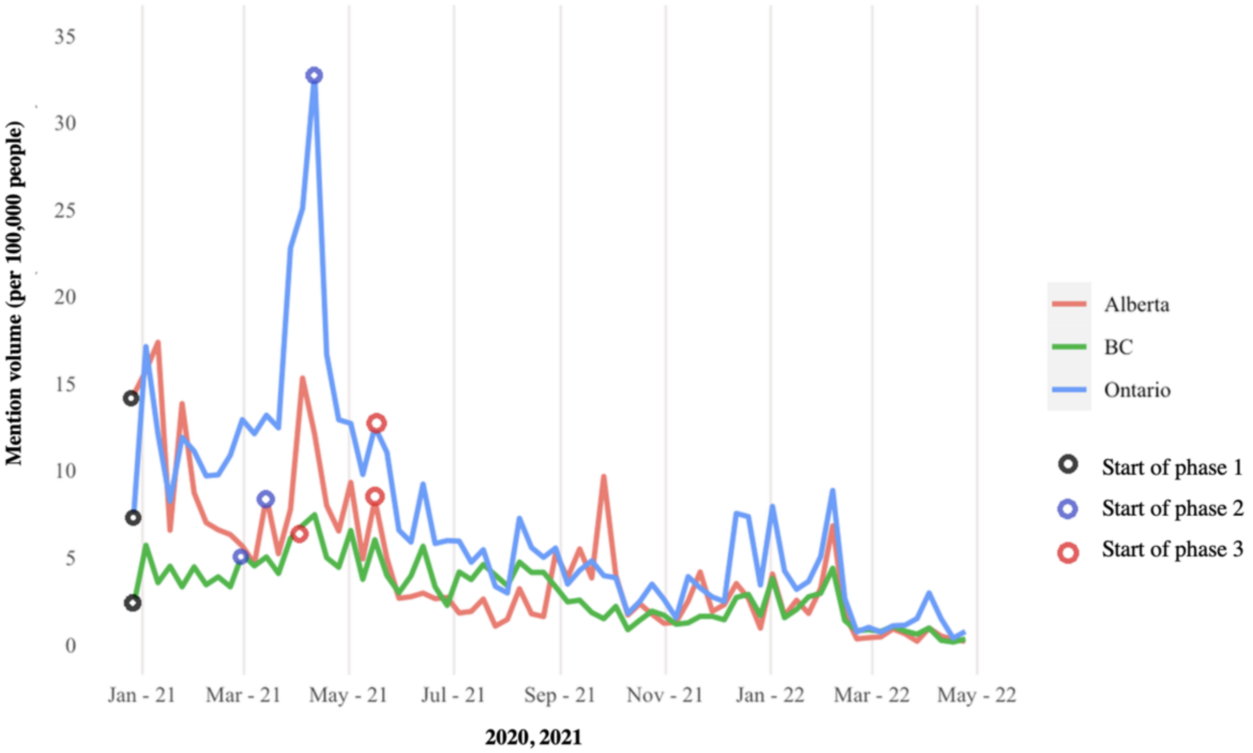


Figure 2. Mention volume of tweets from or to public officials in Ontario (7-day rolling average). Phase 1, phase 2, and phase 3 (gray lines) indicate the starting date of the vaccine rollout phase change.

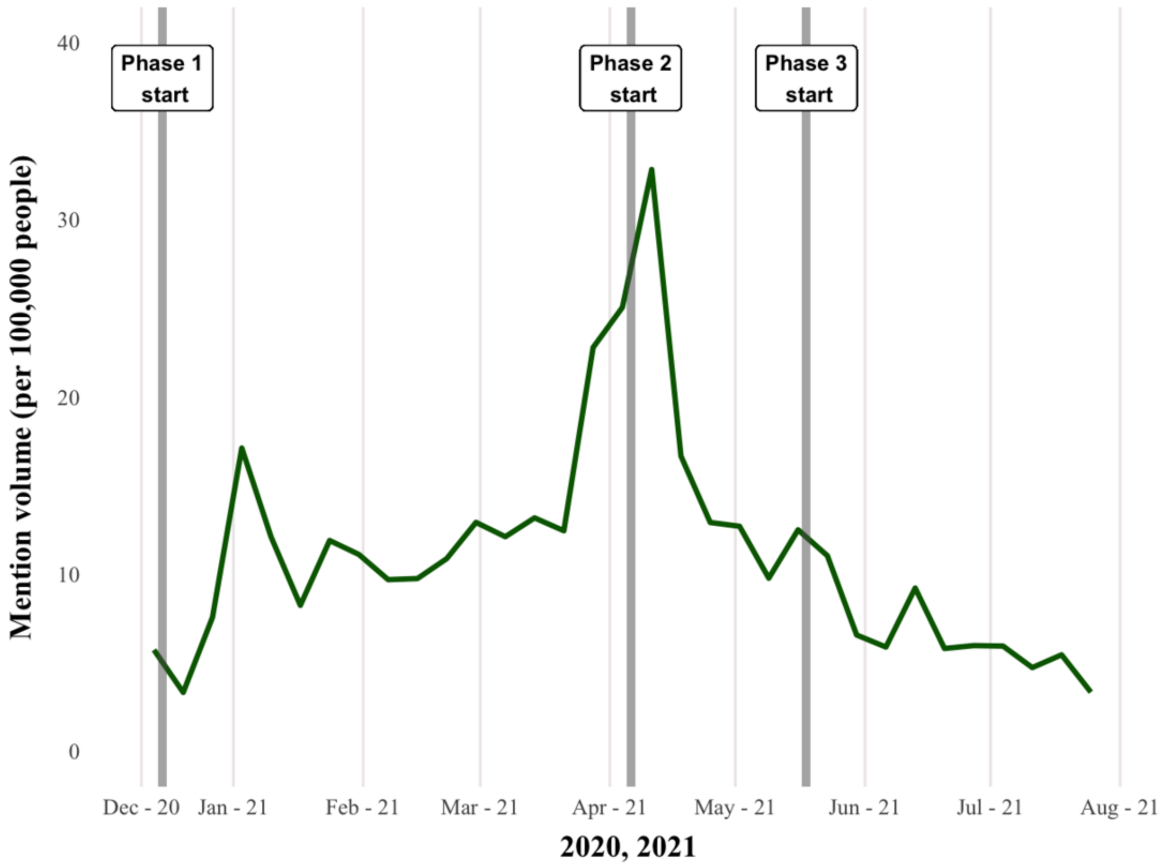


Figure 3. Mention volume of tweets from or to public officials in Alberta (7-day rolling average). Phase 1, phase 2, and phase 3 (gray lines) indicate the starting date of the vaccine rollout phase change.

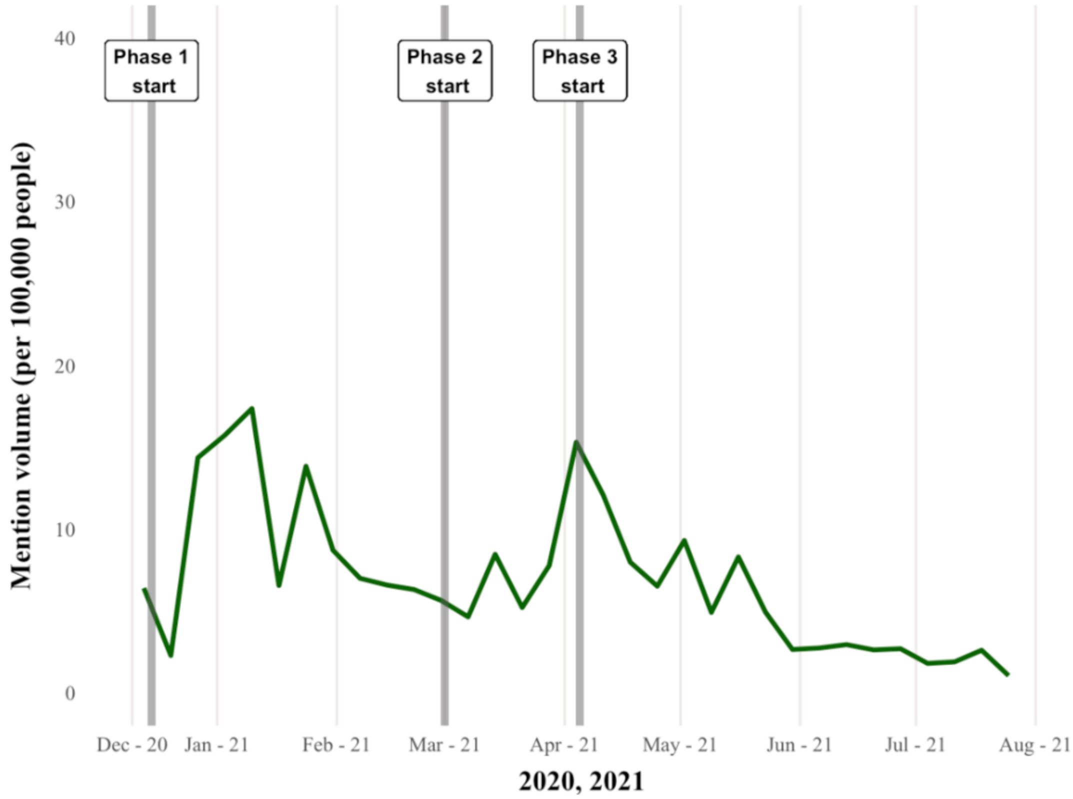


Figure 4. Mention volume of tweets from or to public officials in British Columbia (7-day rolling average). Phase 1, phase 2, and phase 3 (gray lines) indicate the starting date of the vaccine rollout phase change.

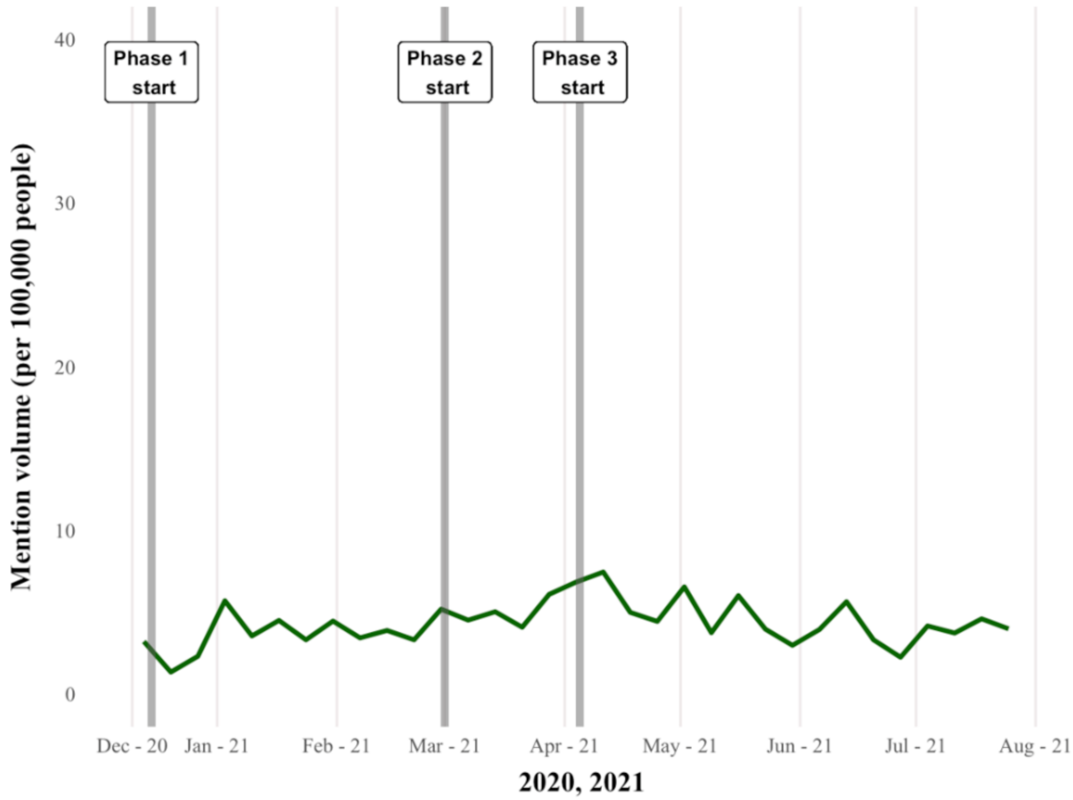


Figure 5. Mention volume of tweets from or to public officials during the 2020-2021 period, stratified by category of public official.

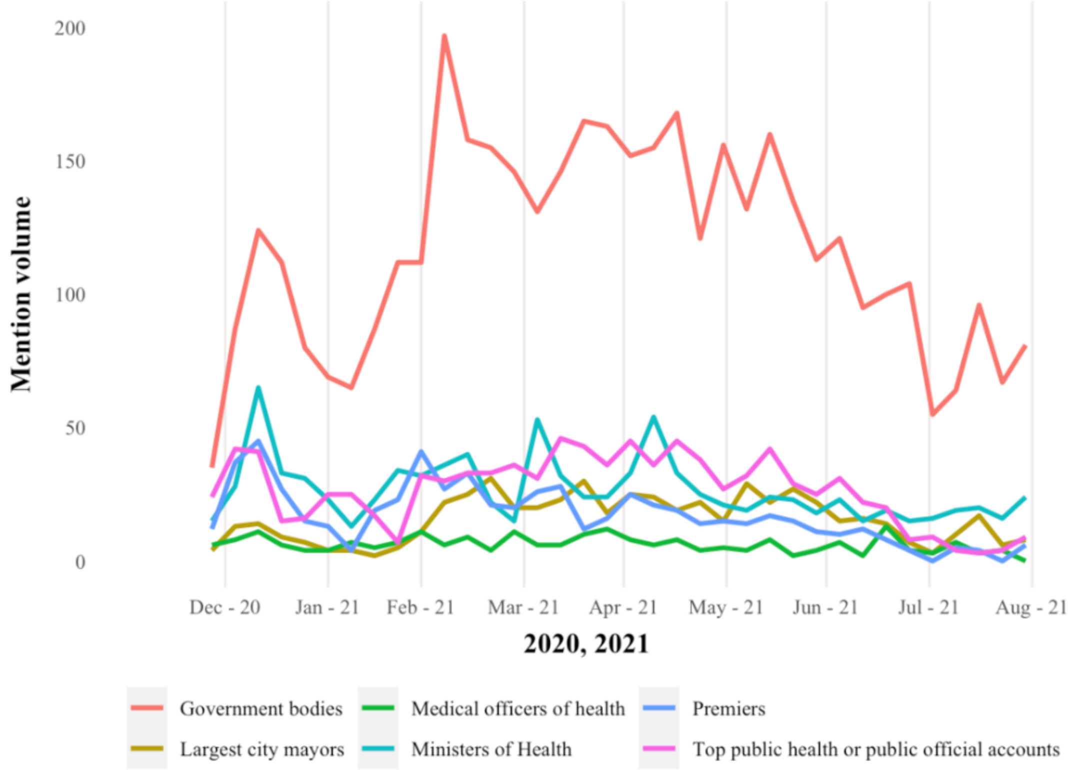
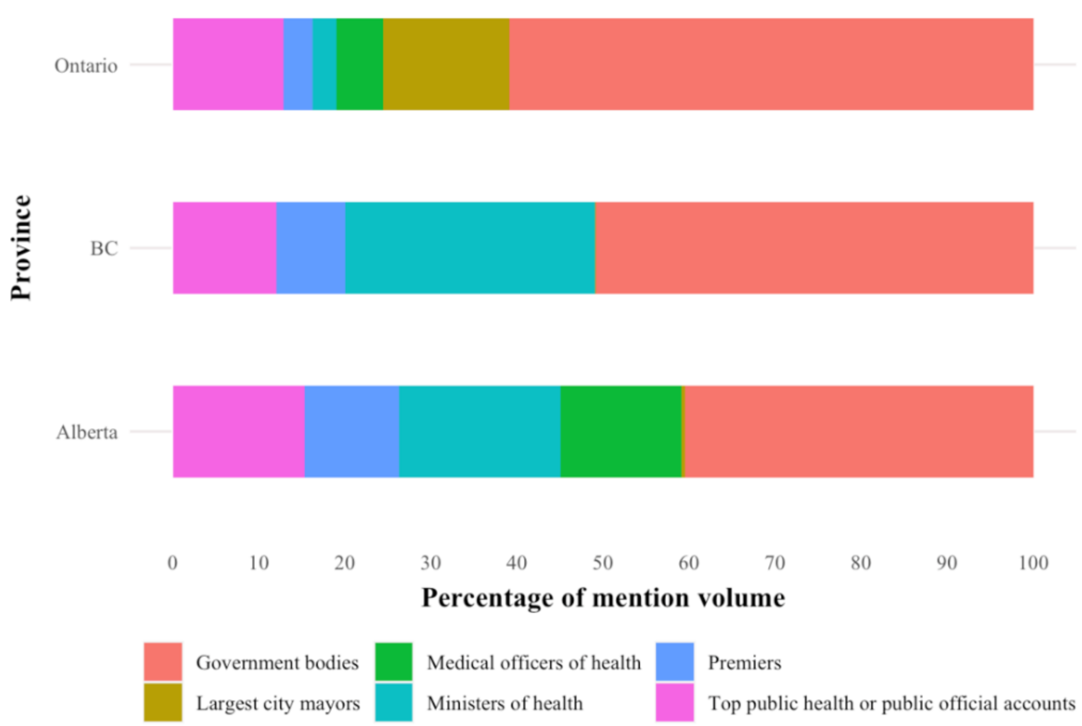


Figure 6. Relative percentage share of mention volume of tweets for each category of public official during the 2020-2021 period, stratified by province. BC: British Columbia.



Content Analysis of Tweets With the Highest Impression

Multimedia Appendix 8 presents engagement metrics in Alberta, BC, and Ontario of tweets with the highest impression (viewership) in each province for each of the 3 phases of vaccine rollout (n=270). Tweets from public officials (212/270, 78.5%) and members of the public (25/270, 9.3%) and media (33/270, 12.2%) who retweeted, replied to, or mentioned public officials were included (definition provided in Multimedia Appendix 5). Table 1 presents the interaction types per our adapted public sector social media interaction framework (Textbox 3).

Table 2 presents a summary of themes for each interaction type. During the 26-day period around each of the 3 phases of rollout across Alberta, BC, and Ontario, 78.5% (212/270) of the most engaged public official tweets were tweeted by public officials, and 21.5% (58/270) of the tweets were from the media and the public who mentioned, quoted, and retweeted public officials. Information provision tweets (*push 1*) by public officials accounted for 65.6% (139/212) of the sampled tweets; public officials interacting with other public officials (37/212, 17.5%), or horizontal engagement (*pull 2*), were more common than public officials interacting with nonpublic officials, or vertical engagement (*pull1*; 24/212, 11.3%). The least common type of engagement was public service announcement tweets (*push 2*), which were only observed in Alberta (eg, tweets that promoted vaccine booking sites or where and how to claim vaccine passports), accounting for 5.7% (12/212) of the tweets sampled.

Table 1. Social media interaction across public officials for each province (n=212).

Interaction type	Alberta, n (%)	British Columbia, n (%)	Ontario, n (%)	Total[a], n (%)
Push 1: information provision	33 (15.6)	63 (29.7)	43 (20.3)	139 (65.6)
Push 2: public service announcement	12 (5.7)	0 (0)	0 (0)	12 (5.7)
Pull 1: citizen engagement[b]	7 (3.3)	6 (2.8)	11 (5.2)	24 (11.3)
Pull 2: public official engagement	23 (10.8)	3 (1.4)	11 (5.2)	37 (17.4)

[a]The percentage of interaction of 212 included tweets by public officials.

[b]Only tweets *by* public officials were analyzed. Tweets engaged with public officials by the public and the media are not considered tweets by public officials.

Table 2. Summary of themes by the modified social media public service interaction framework.

Interaction type	Themes of tweets

Push 1: information provision	• Tweets concerning provincial vaccine rollout policy:
	  • How many and what types of COVID-19 vaccines were received (phases 1 and 2)
	  • How many vaccines were administered (phases 1-3)
	  • Communicating to the public about who should receive the vaccine first; for example, frontline workers, older adults, immunocompromised, hot spots, susceptible populations, or underserved communities (phases 1-3)
	  • Celebratory tweets about vaccine rollout being on schedule (phase 3: April 2021 in British Columbia and May 2021 in Alberta and Ontario)
	• Tweets concerning Federal vaccine rollout policy:
	  • Progress on vaccine distribution to Canada (phase 1)
	  • Progress on vaccine distribution to the Provinces (phase 2)
	  • Progress on vaccine administration in Canada (phase 3)
	• Tweets concerning COVID-19 vaccine safety and efficacy for the Canadian population:
	  • Vaccine approval, safety, efficacy, and contraindications by age groups, gender, pregnancy status, and underlying medical conditions (phase 1)
	  • Updates on recommended dosages, interval, and booster requirements (phase 2 and 3)

Push 2: public service announcement	• Provincial and municipal announcements and sharing of web-based vaccine booking sites, or rollout timeline changes with a corresponding URL to a newsletter subscription invitation (phases 1-3)
Pull 1: citizen engagement	• Provincial public officials engaging with citizens; for example, by retweeting, replying, mentioning, and tagging contents and comments from the public concerning vaccine rollout (phase 2 and 3)
Pull 2: public official engagement	• Provincial government and public health authorities retweeting updates posted by provincial ministers of health (phase 2 and 3) • Federal ministers mentioning other federal provincial government and public health authorities (Federal to Federal) • Members of the legislative assembly or provincial parliament (categorized as top public health officials or public officials) criticizing federal and provincial vaccine rollout (provincial to Federal, provincial, or other provinces)

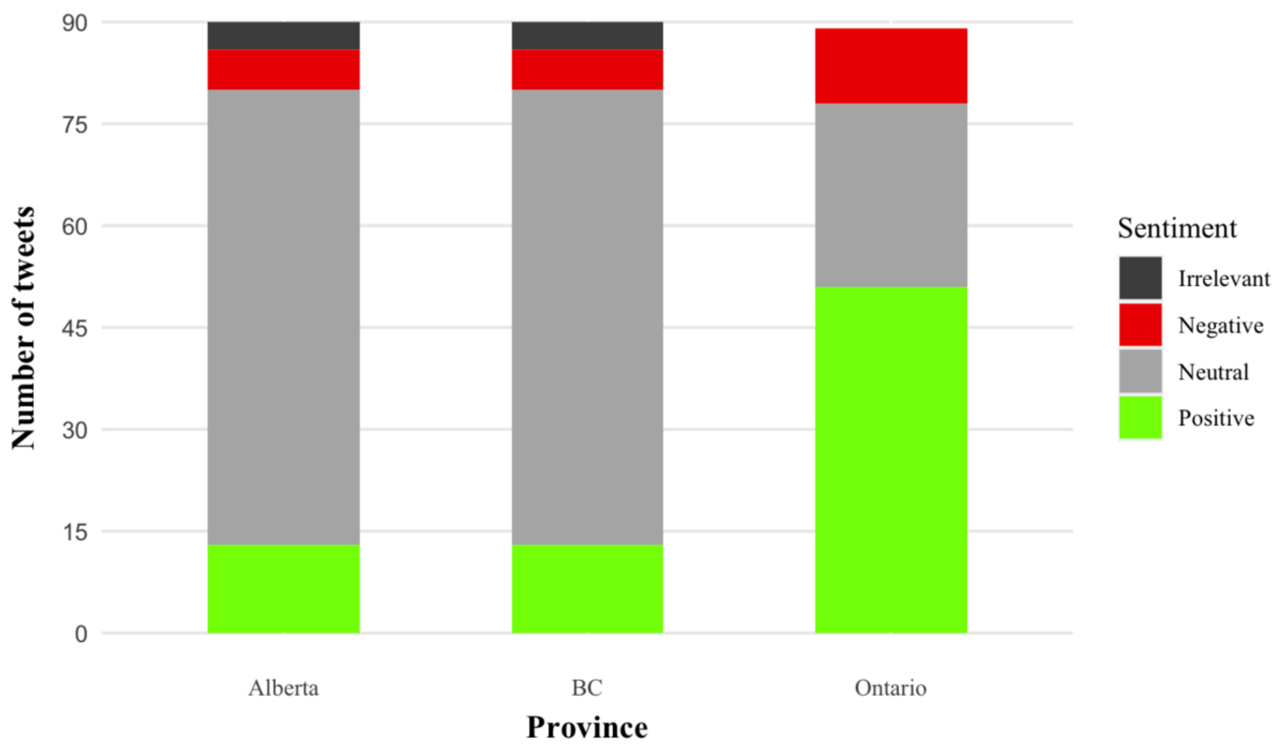
Tweets categorized as information provision (*push 1*) in provinces were generally about the status on the arrival of vaccines, administration of vaccines (eg, number administered), and vaccine eligibility (eg, age and location) updates. Provincial and federal members of parliament (eg, Elizabeth May) with a large social following or influence were observed to retweet information provisional tweets, further driving the overall impression and reach of the original tweets by public officials.

Across the 3 provinces, public officials in Alberta had the highest prevalence of horizontal engagement (*pull 2*; Table 2). These *pull-2* interactions mainly involved praising or criticizing members of the cabinet responsible for the vaccine rollout. The highest prevalence of *pull-1* tweets was observed in Ontario, Canada. These interactions involved public officials mentioning, replying to, or quoting members of the public and the media. This interaction between public officials and influential nonpublic officials (>1000 Twitter followers) generated high viewership and were better endorsed (described in the *Sentiment Analysis* section) than *push-1* and *push-2* type tweets, such as those who tagged the Minister of Health offering to volunteer as vaccinators; tweets from journalists using their personal accounts to broadcast the latest government decisions on vaccine procurement, delivery, and eligibility; and express opinions on issues pertinent to the phase changes. We did not observe the same phenomenon for tweets with high impressions in BC or Alberta.

Sentiment Analysis

Across the provinces, most tweets (138/270, 51.1%) conveyed neutral sentiment across all phases. In BC, 74% (67/90) of the public officials' tweets were neutral, followed by 48% (43/90) in Alberta, and 31% (28/90) in Ontario, likely attributable, in part, to the abundance of *push-1* type tweets that focused on information provision. Negative sentiment accounted for 12.6% (34/270) of all the tweets. Remaining tweets conveyed positive sentiment; Ontario had the highest proportion of positive sentiment (54/90, 60%), followed by Alberta (30/90, 33%) and BC (13/90, 14%; Figure 7). Positive sentiment is demonstrated by public officials to invoke the public for their ongoing commitment to get vaccinated. In Ontario, for example, the Mayor of Toronto expressed gratitude to frontline HCWs for receiving their vaccination, whereas other public officials used a positive tone to thank the public for collaborating with vaccination efforts during the pandemic. In addition, public officials invoked other arms-length provincial government organizations and public health authorities (eg, Canadian Armed Forces) for helping to procure and distribute vaccines, particularly in hard-to-reach or priority areas (Northern Canada and residential long-term care facilities), as well as HCWs and facilities, including physicians and pharmacies. Often, such tweets were associated with dynamic and compelling images with smiling frontline workers, emojis that convey excitement, the use of exclamation points, and a positive overall tone of the tweet (Multimedia Appendix 5). For example, in BC, physicians and executives in the health system praised and expressed positive sentiment toward public officials and quoted tweets from public official accounts to endorse vaccine efforts in phase 1 (December 2020). More examples of positive and negative sentiment tweets are shown in Multimedia Appendix 9.

Figure 7. Sentiment of extracted tweets during the 2020-2021 period, stratified by province. BC: British Columbia.



Across all provinces, negative sentiment (Ontario: 6/90, 7%; Alberta: 17/90, 19%; BC: 12/90, 13%) was tied to feelings of anger and frustration toward federal and provincial public officials for not meeting stated vaccine rollout goals. For example, members of the public in Alberta and Ontario invoked provincial public officials in criticisms around vaccine rollout. Out of the sampled tweets and their responses, no public officials responded to these criticisms directly.

In Multimedia Appendix 8, endorsement ratios (numbers of likes/number of impressions) indicate how receptive Twitter users are toward the vaccine rollout tweet posted by a certain user; the higher likes or impressions, the better endorsed a tweet is. Across the 3 provinces, tweets by and to public officials received a median endorsement ratio of 0.0002. Comparing between provinces, First Ministers in Ontario had the highest endorsement ratio (0.0009) compared with their counterparts in BC (0.0002) and Alberta (0.0007). The low endorsement ratio in BC can be explained by BC's First Minister only tweeting one popular tweet during the peaks of provincial vaccine rollout.

The most popular Minister of Health across the 3 provinces is BC's Adrian Dix (endorsement ratio: 0.0008). In terms of Chief Medical Officers of Health (CMOH), Alberta's CMOH tweeted consistently (18/90, 20%) and received a higher endorsement ratio (0.0006) compared with the CMOH in Ontario (1/90, 1%; 0.0002) and BC (n=0), as BC's CMOH did not have a Twitter presence. Albertan public officials engaged in all 4 methods of engagement (ie, push 1, push 2, pull 1, and pull 2). In Ontario, 3 methods of engagement were used. In BC, 2 methods of engagement were used. Across categories, government bodies in Alberta received the highest endorsement ratio across all provinces (Alberta: 0.0007, BC: 0.0001, and Ontario: 0.0002). The most active municipal officials were from Ontario (23/90, 26%; endorsement ratio=0.0002), who had the highest endorsement ratio compared with BC (n=0; 0.0000) and Alberta (1/90, 1%; 0.0001).

Media endorsement across the provinces was low in Alberta and BC, indicating large viewership and low engagement (Alberta: 0.0000 and BC: 0.0000), whereas tweets from the public received higher endorsement (Alberta: 0.0025 and BC: 0.118). This observation contrasts with Ontario (19/90, 21%; 0.0004), where the media received higher endorsement ratios than tweets from the public (7/90, 8%; 0.0032).

Discussion

Principal Findings

We note 2 salient findings from our results concerning the use of Twitter by public officials to communicate about COVID-19 vaccine rollout in Canada. First, out of all 10 provinces, public officials in 3 provinces—Alberta, BC, and Ontario—use Twitter the most. Out of all 142 sampled public officials' accounts, Twitter was mainly used for unidirectional information provision (*push 1*) to update the public on numbers of vaccines administered. In Ontario and Alberta, we observed a pattern around tweet volume and phase of rollout. An increase in public official interactions on Twitter coincided with the onset of the third wave of the COVID-19 pandemic and matched the start of phase 2 vaccine rollout (phase 2 in Alberta and Ontario: March to April 2021). This can be explained by the governments' readiness to expand vaccine eligibility after the vaccine shortages were resolved. BC, in contrast, had a relatively steady volume of social media interactions by and to public officials across all account categories, which can be explained by the relatively lower daily cases recorded and the absence of public officials from BC on Twitter compared with Ontario and Alberta. BC's CMOH is not on Twitter, which could help explain the lack of any discernable patterns around the use of other interaction types beyond information provision in this province. Social media presence among public officials as a determinant of engagement is therefore a unique area of future investigation. In contrast to our expectations, we also observed a decrease in mention volume across all provinces at phase 3, despite this phase coinciding with peak COVID-19 transmission. We surmise that other COVID-19 conversations overtook vaccine discourse in the public domain by this point.

Second, out of the top viewed tweets, much of the information provided about vaccination rollout on Twitter came from 2 categories of public officials: government bodies (including public health authorities at the federal and provincial levels) and the largest city mayors. Despite accounting for the highest mention volume of tweets, which we attributed to their overrepresentation in our sample, government bodies yielded the lowest endorsement ratios (based on likes and impressions) across all provinces. In comparison, mayors who embedded images and animations in tweets expressing appreciation for frontline workers and the public's vaccination efforts received greater endorsement. This observation is supported by another Canadian study that showed that accounts that tweet frequently per day experience lower engagement per tweet, especially when those tweets do not involve hashtags or multimedia such as animated *gifs* or videos [36]. In contrast, across all 3 provinces, popular tweets by other key public officials not directly responsible for vaccine rollout across received higher than average endorsement, likely attributable to presenting views endorsed by the public (eg, voicing concerns about, or praising, vaccine rollout). Accordingly, we note an opportunity for public officials to engage with other public officials (*pull 2*) to explore bidirectional engagement and its effect on public endorsement during crisis communication.

Furthermore, an interplay of factors explains why a tweet receives many views and many likes (thereby resulting in a high endorsement ratio). For example, the reader may agree with the tweet or show support for the tweet [37]. This was observed during the initial phases of the COVID-19 pandemic when the public liked the tweets of National Health Service to express gratitude [37]. In addition, reading tweets from familiar celebrities correlated with higher endorsement of vaccinations according to a nationwide Twitter experiment that recruited celebrities to endorse vaccination on Twitter [38]. Similarly, there can be many reasons for a tweet to receive many views but a low volume of likes: as impressions are presented as one opens their Twitter feed, tweets that are paid promotions without images or links, such as social marketing campaigns to promote COVID-19 vaccine–related services and information, will likely result in a low endorsement ratio [39]. In addition, viewers may not *like* a tweet when the tweeted content does not align with their beliefs (eg, vaccine beliefs or vaccine eligibility criteria). In a recent large-scale Twitter study, there was substantial empirical evidence pointing toward Twitter's algorithmic amplification of politically right-winged beliefs in Canada [40]; these right-wing beliefs tend to correlate with weaker COVID-19 risk perceptions [41].

Regarding the differences in the sentiment of public officials among the 3 provinces, higher positive sentiment can be explained by the high prevalence of government bodies and mayoral accounts that drove the most views. In particular, Ontario's positive sentiment came from a disproportionately high percentage of government bodies and largest city mayors who used affirming words (eg, "great" and "thanks"), emotive punctuation (eg, exclamation points) to emphasize excitement toward the public's uptake of the COVID-19 vaccine, and emojis implying celebration and strength in numbers (eg, clapping hands and flexing arms). Our findings on positive emoji use echo recent scholarship that notes emoji use by the public during COVID-19 overwhelmingly conveyed positive sentiment [42].

The findings in Ontario on social media behavior driving positive sentiment contrast with Alberta and BC, where government bodies used neutral words and expressions and did not use emojis when conveying information about vaccines and engaging with citizens.

As information about tweet sponsorship by public officials for social marketing campaigns is not publicly available, it is inconclusive why specific public officials' tweets are endorsed more than others. Regardless, it has been noted that highly endorsed tweets are correlated with perceived credibility, which in turn draws more likes [43]. In Ontario, that the highest endorsement ratio was observed for tweets by the Premier may suggest that tweets about vaccine rollout from the Premier, a controversial figure with right-leaning ideology, are well liked. This finding may be surprising to those with opposing ideologies who have been critical of the Premier's pandemic response; however, this may reflect a phenomenon known as *majority illusion* on social media, which suggests that the opinion of a few, amplified in respective echo chambers on social media for which we perceive as a dominant opinion, may in fact be the minority opinion [44].

Comparison With Prior Work

Effective engagement with the public over social media is critical during times of crisis, particularly to protect public safety; to maintain open, clear, and transparent communication of complex issues and risk calculations; and to maintain support for ongoing public health measures and trust in governments. Indeed, recent scholarship from Canada notes the importance of communication strategies by all orders of government to shape change during the COVID-19 pandemic. In particular, transparency is critical in sustaining public trust [45]. Our study found that public officials did promote transparency through the use of Twitter to provide information to the public during the initial rollout of the COVID-19 vaccine, but there was limited engagement and dialogue with the public during this time. Previous studies suggest that public officials at local levels of government have the closest connection to the public, but, as we also found, have used social media in an unengaging and fragmented manner [46-48]. Earlier work has shown a general reluctance by government officials to use social media to engage with the public in times of crisis [49], although this appears to be changing during COVID-19 as noted in our study and elsewhere.

Given the important role of social media communication by government officials in times of crisis, evidence is emerging regarding the use of Twitter by public officials during the COVID-19 pandemic. Zeemering [50] described fragmented communication across municipal public sectors in 3 states in the United States in the early months of the COVID-19 pandemic. They noted that mitigating challenges in communicating public health messaging about COVID-19 requires coordination across all public sectors to ensure better amplification about pandemic responses. In our findings on vaccine communication, this fragmentation was not evident, in part because we focused on multiple layers of government providing uncoordinated messages. Our research focused on communication about vaccine rollout specifically, which was largely the responsibility of government bodies (eg, provincial government and public health authorities), a first minister (health), or a CMOH for provincial updates, and local mayors for municipal updates.

Our findings are consistent with previous research that observes that information provision is the most common type of interaction on social media by public officials [47]. Interestingly, this is inconsistent with a recent study from Poland that also categorized social media communication by public officials during the COVID-19 pandemic using the framework by Mergel [11]. Their analysis of Facebook, Instagram, and Twitter use during the pandemic by local government officials suggests pushing information to be the least used type of interaction. Further work could explore whether public officials in other jurisdictions and at different levels of government use social media in different ways.

Limitations

This study has several limitations. From a methodological perspective, we analyzed a small sample of tweets across select provinces to inform Brandwatch's social media intelligence platform and categorize and track public officials' engagement activities across larger geographic contexts (eg, all Canadian provinces and territories). Along with limited resources and this serving as a pilot study, we were limited in the volume of sentiment and content analysis that we could perform. Via Brandwatch's retrospective database, there were no means to differentiate whether a tweet was promoted (paid) to generate more views and audiences [51], and only geotagged tweets were included.

From an analytic perspective, our study focuses on public officials' engagement on Twitter around the vaccine rollout. Accordingly, we do not analyze other popular social media platforms that may target different audiences, such as Facebook and Instagram, which have been studied in the context of government engagement during crises such as COVID-19 in other countries [11,52,53]. It is possible that several elected public officials do not use Twitter, have Twitter but are inactive, are represented by an organizational user account, or do not have a substantive or engaged following, but are highly engaged on other social media platforms [36]. Furthermore, given the small sample of tweets from which we extracted content and the labor-intensive process of manually coding engagement types (per the public sector social media interaction framework by Mergel [4]), we could not compare how engagement changed or remained consistent across phases of vaccine rollout within each province. In addition, our measures of engagement did not account for the public officials' follower count. Finally, we did not look at the impact of the different types of users or interaction types on vaccine uptake, which could be a focus of future research. To narrow the scope of this study, we did not analyze public resonance to public officials' tweets, but this represents another focus of future research.

Conclusions

The COVID-19 pandemic is an objective lesson in the importance of communicating timely information about vaccine availability to reduce COVID-19 spread. Findings from our study conducted through a Canadian lens advance a growing body of literature on how public officials use social media, particularly Twitter, to communicate with the public during the COVID-19 pandemic. We found a predominant use of information provision (*push-1* interaction) and a reliance on official government accounts to communicate information, which may not be as effective at engaging the public. Our findings leave room for further research, particularly around developing a set of best practices that public officials can lean into when developing communication strategies in times of crisis, COVID-19 related or otherwise.

Acknowledgments

The authors acknowledge the funding support from IQVIA. They are also grateful to Rohit Shome, Pratiksha Vijay Shetty, Jonathan Wiersma, and Tanmay Saraykar for their support in data collection and for their thoughtful review of earlier versions of this manuscript. The statements, findings, conclusions, views, and opinions contained and expressed in this brief report are based in part on data obtained under license from the following IQVIA Solutions Canada Inc information services: Brandwatch database. The statements, findings, conclusions, views, and opinions contained and expressed herein are not necessarily those of IQVIA Solutions Canada Inc or any of its affiliated or subsidiary entities.

Data Availability

Elements of the data sets generated and analyzed during this study are not publicly available because of the proprietary nature of the analytic services provided by Brandwatch Analytics. A data set that has been stripped of the proprietary information will be made available upon reasonable request from the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Number of public officials included in each jurisdiction (federal, provincial, or territorial).
[DOCX File , 20 KB - infodemiology_v3i1e41582_app1.docx]

Multimedia Appendix 2
Full query.
[DOCX File , 21 KB - infodemiology_v3i1e41582_app2.docx]

Multimedia Appendix 3
Vaccine rollout phase changes and milestones in Alberta, British Columbia (BC), and Ontario for the first and second doses.
[DOCX File , 16 KB - infodemiology_v3i1e41582_app3.docx]

Multimedia Appendix 4
Data collection flowchart.
[DOCX File , 81 KB - infodemiology_v3i1e41582_app4.docx]

Multimedia Appendix 5
Definitions of variables extracted for 270 tweets.
[DOCX File , 20 KB - infodemiology_v3i1e41582_app5.docx]

Multimedia Appendix 6
Flow diagram of tweet selection.
[DOCX File , 34 KB - infodemiology_v3i1e41582_app6.docx]

Multimedia Appendix 7
COVID-19 vaccine administration and case positivity in Alberta, British Columbia, and Ontario (December 28, 2020, to August 31, 2021).
[DOCX File , 522 KB - infodemiology_v3i1e41582_app7.docx]

Multimedia Appendix 8
Public officials' Twitter engagement metrics.

[DOCX File , 25 KB - infodemiology_v3i1e41582_app8.docx]

Multimedia Appendix 9
Examples of tweets conveying positive and negative sentiment.
[DOCX File , 19 KB - infodemiology_v3i1e41582_app9.docx]

References

1. Mickoleit A. Social media use by governments a policy primer to discuss trends, identify policy opportunities and guide decision makers. OECD Working Papers on Public Governance No. 26. 2014. URL: https://www.oecd-ilibrary.org/docserver/5jxrcmghmk0s-en.pdf?expires=1687242940&id=id&accname=guest&checksum=6AFA646B41740A4213CE83D533245C13 [accessed 2022-04-14]
2. Bryer TA, Zavattaro SM. Social media and public administration. Adm Theory Prax 2014 Dec 07;33(3):325-340. [doi: 10.2753/atp1084-1806330301]
3. Criado JI, Villodre J. Delivering public services through social media in European local governments. An interpretative framework using semantic algorithms. Local Gov Stud 2020 Mar 04;47(2):253-275. [doi: 10.1080/03003930.2020.1729750]
4. Mergel I. A framework for interpreting social media interactions in the public sector. Gov Inf Q 2013 Oct;30(4):327-334. [doi: 10.1016/j.giq.2013.05.015]
5. DePaula N, Dincelli E, Harrison TM. Toward a typology of government social media communication: democratic goals, symbolic acts and self-presentation. Gov Inf Q 2018 Jan;35(1):98-108. [doi: 10.1016/j.giq.2017.10.003]
6. Khan GF. The Government 2.0 utilization model and implementation scenarios. Inf Dev 2013 Nov 07;31(2):135-149. [doi: 10.1177/0266666913502061]
7. Guerrero-Solé F. Interactive behavior in political discussions on Twitter: politicians, media, and citizens' patterns of interaction in the 2015 and 2016 electoral campaigns in Spain. Soc Media Soc 2018 Nov 11;4(4):205630511880877. [doi: 10.1177/2056305118808776]
8. De Souza Bermejo PH, Pereira JR, Barbosa DA, Oliveira DJ. The application of the sentiment analysis technique in social media as a tool for social management practices at the governmental level. Rev Adm Pública 2019 Feb;53(1):235-251. [doi: 10.1590/0034-7612174204]
9. Nadzir MM. Proposed e-government 2.0 engagement model based on social media use in government agencies. In: Proceedings of the IEEE Conference on e-Learning, e-Management & e- Services (IC3e). 2019 Presented at: 2019 IEEE Conference on e-Learning, e-Management & e- Services (IC3e); Nov 19-21, 2019; Pulau Pinang, Malaysia. [doi: 10.1109/ic3e47558.2019.8971778]
10. Stone JA, Can SH. Linguistic analysis of municipal twitter feeds: factors influencing frequency and engagement. Gov Inf Q 2020 Oct;37(4):101468. [doi: 10.1016/j.giq.2020.101468]
11. Górska A, Dobija D, Grossi G, Staniszewska Z. Getting through COVID-19 together: understanding local governments' social media communication. Cities 2022 Feb;121:103453 [FREE Full text] [doi: 10.1016/j.cities.2021.103453] [Medline: 34566232]
12. Graham MW, Avery EJ, Park S. The role of social media in local government crisis communications. Public Relat Rev 2015 Sep;41(3):386-394. [doi: 10.1016/j.pubrev.2015.02.001]
13. Hagen L, Keller T, Neely S, DePaula N, Robert-Cooperman C. Crisis communications in the age of social media. Soc Sci Comput Rev 2017 Aug 21;36(5):523-541. [doi: 10.1177/0894439317721985]
14. Wendling C, Radisch J, Jacobzone S. The use of social media in risk and crisis communication. OECD Working Papers on Public Governance 24. 2013. URL: https://ideas.repec.org/p/oec/govaaa/24-en.html [accessed 2022-04-14]
15. Cegarra-Navarro JG, Pachón JR, Cegarra JL. E-government and citizen's engagement with local affairs through e-websites: the case of Spanish municipalities. Int J Inf Manag 2012 Oct;32(5):469-478. [doi: 10.1016/j.ijinfomgt.2012.02.008]
16. Hyland-Wood B, Gardner J, Leask J, Ecker UK. Toward effective government communication strategies in the era of COVID-19. Humanit Soc Sci Commun 2021 Jan 27;8(1):30. [doi: 10.1057/s41599-020-00701-w]
17. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194 [FREE Full text] [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
18. Slavik CE, Darlington JC, Buttle C, Sturrock SL, Yiannakoulias N. Has public health messaging during the COVID-19 pandemic reflected local risks to health?: a content analysis of tweeting practices across Canadian geographies. Health Place 2021 May;69:102568 [FREE Full text] [doi: 10.1016/j.healthplace.2021.102568] [Medline: 33865003]
19. Vilches TN, Zhang K, Van Exan R, Langley JM, Moghadas SM. Projecting the impact of a two-dose COVID-19 vaccination campaign in Ontario, Canada. Vaccine 2021 Apr 22;39(17):2360-2365 [FREE Full text] [doi: 10.1016/j.vaccine.2021.03.058] [Medline: 33812742]
20. Archived: Guidance on the prioritization of key populations for COVID-19 immunization . Government of Canada. URL: https://www.canada.ca/en/public-health/services/immunization/national-advisory-committee-on-immunization-naci/guidance-prioritization-key-populations-covid-19-vaccination.html [accessed 2022-06-01]

21. Here's the COVID-19 vaccine rollout plan, province by province. CBC News. 2020 Dec 10. URL: https://www.cbc.ca/news/canada/covid19-vaccine-rollout-plans-canada-1.5836262 [accessed 2022-05-22]
22. Ontario ready to rollout phase two of COVID-19 vaccine distribution plan. Government of Ontario. 2021 Mar 5. URL: https://news.ontario.ca/en/release/60568/ontario-ready-to-rollout-phase-two-of-covid-19-vaccine-distribution-plan [accessed 2022-05-22]
23. Fitzpatrick T, Allin S, Camillo CA, Habbick M, Mauer-Vakil D, Muhajarine N, et al. COVID-19 vaccination rollout: Ontario. CoVaRR-Net and North American Observatory on Health Systems and Policies. 2022 Feb 7. URL: https://covarrnet.ca/covid-19-vaccination-rollout-ontario/ [accessed 2022-06-01]
24. Chung H, Azimaee M, Bronskill SE, Cartagena R, Guttmann A, Ho MM, et al. Pivoting data and analytic capacity to support Ontario's COVID-19 response. Int J Popul Data Sci 2020 Jan 18;5(3):1682 [FREE Full text] [doi: 10.23889/ijpds.v5i4.1682] [Medline: 35141430]
25. Engel-Rebitzer E, Stokes DC, Buttenheim A, Purtle J, Meisel ZF. Changes in legislator vaccine-engagement on Twitter before and after the arrival of the COVID-19 pandemic. Hum Vaccin Immunother 2021 Sep 02;17(9):2868-2872 [FREE Full text] [doi: 10.1080/21645515.2021.1911216] [Medline: 33970786]
26. Schimmele C, Fonberg J, Schellenberg G. Canadians' assessments of social media in their lives. Statistics Canada. 2021 Mar 24. URL: https://www150.statcan.gc.ca/n1/pub/36-28-0001/2021003/article/00004-eng.htm [accessed 2021-09-07]
27. Barua Z, Barua S, Aktar S, Kabir N, Li M. Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. Prog Disaster Sci 2020 Dec;8:100119 [FREE Full text] [doi: 10.1016/j.pdisas.2020.100119] [Medline: 34173443]
28. Dubé E, Bettinger JA, Fisher WA, Naus M, Mahmud SM, Hilderman T. Vaccine acceptance, hesitancy and refusal in Canada: challenges and potential approaches. Can Commun Dis Rep 2016 Dec 01;42(12):246-251 [FREE Full text] [doi: 10.14745/ccdr.v42i12a02] [Medline: 29769995]
29. Griffith J, Marani H, Monkman H. COVID-19 vaccine hesitancy in Canada: content analysis of tweets using the theoretical domains framework. J Med Internet Res 2021 Apr 13;23(4):e26874. [doi: 10.2196/26874] [Medline: 33769946]
30. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. BMJ Glob Health 2020 Oct;5(10):e004206 [FREE Full text] [doi: 10.1136/bmjgh-2020-004206] [Medline: 33097547]
31. Anderson JT, Bouchacourt LM, Sussman KL, Bright LF, Wilcox GB. Telehealth adoption during the COVID-19 pandemic: a social media textual and network analysis. Digit Health 2022 Mar 31;8:20552076221090041 [FREE Full text] [doi: 10.1177/20552076221090041] [Medline: 35392254]
32. Kaur M, Verma R, Otoo FN. Emotions in leader's crisis communication: Twitter sentiment analysis during COVID-19 outbreak. J Hum Behav Soc Environ 2021 Feb 19;31(1-4):362-372. [doi: 10.1080/10911359.2020.1829239]
33. Social media. Government of Canada. URL: https://www.canada.ca/en/social.html [accessed 2022-04-14]
34. Canada top social media used by province 2021. Statista. URL: https://www.statista.com/statistics/262804/social-networks-used-by-internet-users-in-canada-provinces/ [accessed 2022-04-14]
35. Canadian COVID-19 intervention timeline. Canadian Institute for Health information. 2022 Oct 13. URL: https://www.cihi.ca/en/covid-19-intervention-timeline-in-canada [accessed 2023-07-04]
36. Teichmann L, Nossek S, Bridgman A, Loewen PJ, Owen T, Ruths D, et al. Public health communication and engagement on social media during the COVID-19 pandemic. OSF Preprints. Preprint posted online on July 27, 2020 . [doi: 10.31219/osf.io/7hypj]
37. Day G, Robert G, Leedham-Green K, Rafferty AM. An outbreak of appreciation: a discursive analysis of tweets of gratitude expressed to the National Health Service at the outset of the COVID-19 pandemic. Health Expect 2022 Feb 20;25(1):149-162 [FREE Full text] [doi: 10.1111/hex.13359] [Medline: 34543519]
38. Alatas V, Chandrasekhar AG, Mobius M, Olken BA, Paladines C. When celebrities speak: a nationwide Twitter experiment promoting vaccination in Indonesia. NBER Working Paper No. 25589. 2019 Feb. URL: https://www.nber.org/system/files/working_papers/w25589/w25589.pdf [accessed 2022-04-14]
39. McLachlan S, Martin N. Experiment: what kind of promoted tweet gets a higher click-through rate? Hootsuite. 2021 May 27. URL: https://blog.hootsuite.com/experiment-promoted-tweets/ [accessed 2022-07-08]
40. Huszár F, Ktena SI, O'Brien C, Belli L, Schlaikjer A, Hardt M. Algorithmic amplification of politics on Twitter. Proc Natl Acad Sci U S A 2022 Jan 04;119(1):e2025334119 [FREE Full text] [doi: 10.1073/pnas.2025334119] [Medline: 34934011]
41. Pennycook G, McPhetres J, Bago B, Rand DG. Beliefs about COVID-19 in Canada, the United Kingdom, and the United States: a novel test of political polarization and motivated reasoning. Pers Soc Psychol Bull 2022 May;48(5):750-765 [FREE Full text] [doi: 10.1177/01461672211023652] [Medline: 34180276]
42. Al-Rawi A, Siddiqi M, Morgan R, Vandan N, Smith J, Wenham C. COVID-19 and the gendered use of emojis on Twitter: infodemiology study. J Med Internet Res 2020 Nov 05;22(11):e21646 [FREE Full text] [doi: 10.2196/21646] [Medline: 33052871]
43. Borah P, Xiao X. The importance of 'likes': the interplay of message framing, source, and social endorsement on credibility perceptions of health information on Facebook. J Health Commun 2018;23(4):399-411. [doi: 10.1080/10810730.2018.1455770] [Medline: 29601271]

44. Lerman K, Yan X, Wu XZ. The "majority illusion" in social networks. PLoS One 2016 Feb 17;11(2):e0147617 [FREE Full text] [doi: 10.1371/journal.pone.0147617] [Medline: 26886112]
45. Palmer KS, Bélanger-Hardy E, Kokorelias K, Law S. How language shapes change: perspectives on the most and least effective communication strategies and tactics during the COVID-19 pandemic. CanCOVID. 2023 Feb 9. URL: https://cancovid.ca/wp-content/uploads/2023/02/CanCOVID-Brief-Communications-During-COVID-EN.pdf [accessed 2023-03-27]
46. Bonsón E, Royo S, Ratkai M. Citizens' engagement on local governments' Facebook sites. An empirical analysis: the impact of different media and content types in Western Europe. Gov Inf Q 2015 Jan;32(1):52-62. [doi: 10.1016/j.giq.2014.11.001]
47. Trencher G. Towards the smart city 2.0: empirical evidence of using smartness as a tool for tackling social challenges. Technol Forecast Soc Change 2019 May;142:117-128. [doi: 10.1016/j.techfore.2018.07.033]
48. Zavattaro SM, Sementelli AJ. A critical examination of social media adoption in government: introducing omnipresence. Gov Inf Q 2014 Apr;31(2):257-264. [doi: 10.1016/j.giq.2013.10.007]
49. Medina RZ, Diaz JC. Social media use in crisis communication management: an opportunity for local communities? In: Sobaci M, editor. Social Media and Local Governments. Cham: Springer; 2016.
50. Zeemering ES. Functional fragmentation in city hall and twitter communication during the COVID-19 pandemic: evidence from Atlanta, San Francisco, and Washington, DC. Gov Inf Q 2021 Jan;38(1):101539. [doi: 10.1016/j.giq.2020.101539]
51. Data restrictions. Brandwatch Developers. URL: https://developers.brandwatch.com/docs/data-restrictions [accessed 2022-05-22]
52. El Baradei L, Kadry M, Ahmed G. Governmental social media communication strategies during the COVID-19 pandemic: the case of Egypt. Int J Public Adm 2021 May 03;44(11-12):907-919. [doi: 10.1080/01900692.2021.1915729]
53. Wukich C. Government social media engagement strategies and public roles. Public Perform Manag Rev 2020 Dec 17;44(1):187-215. [doi: 10.1080/15309576.2020.1851266]

Abbreviations

BC: British Columbia
CMOH: Chief Medical Officers of Health
HCW: health care worker

Edited by T Mackey; submitted 05.12.22; peer-reviewed by A Lundberg, Z Han; comments to author 06.03.23; revised version received 27.03.23; accepted 30.05.23; published 20.07.23.

Please cite as:
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;3:e41582
URL: https://infodemiology.jmir.org/2023/1/e41582
doi:10.2196/41582
PMID:37315194

Original Paper

Using COVID-19 Vaccine Attitudes on Twitter to Improve Vaccine Uptake Forecast Models in the United States: Infodemiology Study of Tweets

Nekabari Sigalo[1], BSc, MPH, PhD; Naman Awasthi[2], MSc; Saad Mohammad Abrar[2], MSc; Vanessa Frias-Martinez[1], BSc, MSc, PhD

[1]College of Information Studies, University of Maryland, College Park, MD, United States

[2]Department of Computer Science, University of Maryland, College Park, MD, United States

Corresponding Author:
Nekabari Sigalo, BSc, MPH, PhD
College of Information Studies
University of Maryland
4130 Campus Dr
College Park, MD, 20742-4345
United States
Phone: 1 301 405 2033
Email: nsigalo@terpmail.umd.edu

Abstract

Background: Since the onset of the COVID-19 pandemic, there has been a global effort to develop vaccines that protect against COVID-19. Individuals who are fully vaccinated are far less likely to contract and therefore transmit the virus to others. Researchers have found that the internet and social media both play a role in shaping personal choices about vaccinations.

Objective: This study aims to determine whether supplementing COVID-19 vaccine uptake forecast models with the attitudes found in tweets improves over baseline models that only use historical vaccination data.

Methods: Daily COVID-19 vaccination data at the county level was collected for the January 2021 to May 2021 study period. Twitter's streaming application programming interface was used to collect COVID-19 vaccine tweets during this same period. Several autoregressive integrated moving average models were executed to predict the vaccine uptake rate using only historical data (baseline autoregressive integrated moving average) and individual Twitter-derived features (autoregressive integrated moving average exogenous variable model).

Results: In this study, we found that supplementing baseline forecast models with both historical vaccination data and COVID-19 vaccine attitudes found in tweets reduced root mean square error by as much as 83%.

Conclusions: Developing a predictive tool for vaccination uptake in the United States will empower public health researchers and decisionmakers to design targeted vaccination campaigns in hopes of achieving the vaccination threshold required for the United States to reach widespread population protection.

(JMIR Infodemiology 2023;3:e43703) doi:10.2196/43703

KEYWORDS

social media; Twitter; COVID-19; vaccine; surveys; SARS-CoV-2; vaccinations; hesitancy; vaccine hesitancy; forecast model; vaccine uptake; health promotion; infodemiology; health information; misinformation

Introduction

Background

Since the onset of the COVID-19 pandemic, there has been a global effort to develop vaccines that protect against COVID-19. Individuals who are fully vaccinated are far less likely to contract and therefore transmit the virus to others [1]. Up until recently, public health experts have stressed the importance of achieving a numerical threshold of herd immunity, but this is only possible if a significant proportion of the population is fully vaccinated. More recent research suggests that the traditional concept of herd immunity may not apply to COVID-19 [2]. Instead, the goal is to increase vaccination uptake to optimize population protection without prohibitive restrictions on our daily lives [3]. Accurately forecasting vaccination uptake allows policymakers and researchers to evaluate how close we are to achieving normalcy again.

Researchers have turned to traditional methods for forecasting COVID-19 infection and vaccination rates [4-6]. For example, one of the most common forecasting methods used, univariate time series, involves predicting future vaccination rates using historical vaccination rates. While this method can be useful in many cases, it fails to account for other time-dependent factors that may also influence vaccinations. For example, the COVID-19 vaccine conversation on social media has been deemed an infodemic, with antivaccination misinformation spreading across social media platforms [7]. Researchers have found that the internet and social media both play a role in shaping personal or parental choices about vaccinations [8,9]. Additionally, previous research showed a positive relationship between positive sentiment scores in COVID-19 vaccine–related tweets and an increase in vaccination rates [10]. These findings suggest it is important to consider the daily conversations on social media when developing vaccine uptake forecast models.

Forecasting COVID-19–Related Measures Using Social Media

There is no shortage of studies that sought to forecast COVID-19-related measures using information from social media. Researchers Yousefinaghani et al [11] conducted a study using COVID-19–related terms mentioned in tweets and Google searches to predict COVID-19 waves in the United States. Researchers found that tweets that mentioned COVID-19 symptoms predicted 100% of first waves of COVID-19 days sooner than other data sources. Another study used data from Google searches, tweets, and Wikipedia page views to predict COVID-19 cases and deaths in the United States [12]. Researchers found models that included features from all 3 sources performed better than baseline models that did not include these features. Researchers also found that Google searches were a leading indicator of the number of cases and deaths across the United States. Another study [13] examined the relationship between daily COVID-19 cases and COVID-19–related tweets and Google Trends. In a study conducted by Shen et al [14], researchers used reports of symptoms and diagnoses on Weibo, a popular social media platform in China, in order to predict COVID-19 case counts in mainland China. Researchers found reports of symptoms and diagnoses on the social media platform to be highly predictive of daily case counts. Although each of these studies forecast COVID-19 cases and deaths, none of these studies forecast COVID-19 vaccination rates.

Forecasting Vaccinations

Very few studies have conducted time series forecasting of the COVID-19 vaccinated population in the United States. In a study conducted by Sattar and Arifuzzaman [15], researchers developed a time series model to predict the percentage of the US population that would get at least 1 dose of the COVID-19 vaccine or be fully vaccinated. Researchers projected that by the end of July 2021, 62.44% and 48% of the US population would get at least 1 dose of the COVID-19 vaccine or be fully vaccinated, respectively. Although this paper also included a separate tweet sentiment analysis, researchers did not include Twitter-related features in the forecast model. Additionally, researchers used aggregated vaccination data for the entire United States, rather than a more granular geographic level.

Another study aimed to evaluate if and when the world would reach a vaccination rate sufficient enough for herd immunity by forecasting the number of people fully vaccinated against COVID-19 in various countries, including the United States [16]. In this study, researchers used a common univariate time series forecasting method, autoregressive integrated moving average (ARIMA), to forecast the future number of fully vaccinated people using only historical vaccination data. Based on the resulting projections, researchers concluded that countries were nowhere near the necessary herd immunity threshold needed to end the COVID-19 pandemic.

A study conducted by Cheong et al [17] sought to predict COVID-19 vaccine uptake using various sociodemographic factors. Although not a time series forecasting model, the results of this study showed that geographic location, education level, and online access were highly predictive of vaccination uptake in the United States. The model predicted vaccine uptake with 62% accuracy.

Although there are very few studies related to COVID-19 vaccination forecasting, other studies have been conducted to predict immunizations for other illnesses. For example, 1 study analyzed electronic medical records of a cohort of 250,000 individuals over the course of 10 years [18]. Researchers developed a model to predict vaccination uptake of individuals in the upcoming influenza season based on previous personal and social behavioral patterns. Another study developed a tool for leveraging immunization related content from Twitter and Google Trends to develop a model for predicting whether a child would receive immunizations [19]. Researchers were able to predict child immunization statuses with 76% accuracy.

Study Objectives

Although previous studies have developed forecast models for COVID-19 vaccination rates in the United States, to our knowledge, there are no studies that aim to factor in the real-time vaccination attitudes present on Twitter. The vaccine attitudes on Twitter change daily, as do vaccination rates, so analyzing vaccine attitudes on social media might contribute to the performance of vaccine forecast models. Additionally, previous studies developed forecast models that focused on the entire United States as a whole. These forecast models fail to appreciate the differences in vaccination roll out, behaviors, and attitudes across different geographic regions. This study seeks to fill this gap by examining vaccine uptake at the metropolitan level.

The purpose of this study is to develop a time series forecasting algorithm that can predict future vaccination rates across US metropolitan areas. Specifically, this study aims to determine whether supplementing forecast models with real-time vaccine attitudes found in tweets—measured via sentiments and emotions—improves over baseline models that only use historical vaccination data. Developing a predictive tool for vaccination uptake in the United States will empower public health researchers and decision makers to design targeted vaccination campaigns in hopes of achieving the vaccination threshold required for us to reach herd immunity.

Methods

Data Collection and Preprocessing

Twitter Data

The Twitter streaming application programming interface, which provides access to a random sample of 1% of publicly available tweets, was used to collect tweets from 8 of the most populated metropolitan areas in the United States from January 2021 to May 2021 (Textbox 1) [20]. We chose to focus on large metropolitan areas to gather a sufficient number of tweets for the analysis. Additionally, larger metropolitan areas also tend to have users who enable the location feature when tweeting [21,22]. All tweets had "place" information (usually city and state). The place information found in tweets was used to determine the metropolitan area associated with each tweet. Next, to extract tweets related to COVID-19 vaccines, tweets were further filtered by matching variations of vaccine-related keywords, such as *vaccine, pfizer, moderna, johnson & johnson, and dose*. Additional vaccine keywords can be found in Multimedia Appendix 1. A language filter was then applied to identify tweets written in the English language. The tweets sample was further preprocessed to minimize "noise" resulting from tweets that matched our vaccine-related keywords but did not necessarily reflect the thoughts and opinions of individual Twitter users. For example, companies often promote job postings and advertisements on Twitter using targeted hashtags in hopes of reaching their target audience. To prevent these tweets from adding noise to the sample, tweets related to job postings and advertisements were removed by excluding tweets with hashtags and keywords, including "jobs," "hiring," "advertisement," "apply," and "ad."

Textbox 1. Targeted metropolitan areas for Twitter data collection, January 1, 2021, to May 20, 2021.

- Phoenix-Mesa-Chandler, AZ
- Miami–Fort Lauderdale–Pompano Beach, FL
- Atlanta–Sandy Springs–Alpharetta, GA
- New York–Newark–Jersey City, NY-NJ-PA
- Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
- Washington-Arlington-Alexandria, DC-VA-MD-WV
- Chicago-Naperville-Elgin, IL-IN-WI
- Los Angeles–Long Beach–Anaheim, CA

COVID-19 Vaccination Data

Daily COVID-19 vaccination data at the county level was collected for the January 2021 to May 2021 study period from the Centers for Disease Control and Prevention's publicly available vaccination data set [23]. This data set includes daily vaccination data from clinics, pharmacies, long-term care facilities, dialysis centers, Federal Emergency Management Agency and Health Resources and Services Administration partner sites, and federal entity facilities. Vaccination administration data are reported to the Centers for Disease Control and Prevention via immunization information systems, the vaccine administration management system, and data submissions directly to the COVID-19 Data Clearinghouse [23]. Each county was linked to its respective metropolitan area according to the US Census delineation file [24]. Next, the data were aggregated to the daily-metropolitan level and the 7-day rolling average of the percentage of individuals who have been administered at least 1 vaccine dose was calculated.

Data Analysis

Sentiment and Emotion Analysis of Tweets

For the purposes of this study, we measure COVID-19 vaccine attitudes via sentiment and emotion analyses of tweets. We evaluated both sentiments and emotions because both methods offer different levels of granularity. Sentiment analysis focuses on determining the overall sentiment or polarity of a text, such as positive, negative, or neutral. It provides a high-level understanding of the sentiment expressed. Emotion analysis, on the other hand, aims to identify specific emotions within the text, such as joy, anger, and sadness. It offers a more detailed and nuanced understanding of the emotional states. By utilizing both sentiment and emotion analysis, we gain a comprehensive understanding of the text, covering both the overall sentiment and the specific emotions expressed.

To capture the sentiments and emotions found in COVID-19 vaccine-related tweets, a sentiment and emotion analysis of all tweets was conducted using bidirectional encoder representation from transformer (BERT) [25], a pretrained language model trained using bidirectional (left to right and right to left) context training to learn joint probability distributions of text. We leveraged the fine-tuned BERT models in the *TweetNLP* package in Python (Python Software Foundation) [26] to calculate the valence of 8 different emotions (fear, joy, anticipation, anger, disgust, sadness, surprise, trust), along with overall neutral, positive, and negative sentiment of tweets in our analysis sample. The sentiment analysis and emotion recognition BERT models were fine-tuned with the TweetEval benchmark [27].

The outputs from BERT are softmax of logits, one corresponding to each of the emotions or sentiments. For each tweet, we performed argmax over the probability distribution for each tweet, to get the most likely emotion and sentiment.

Next, we found the percentage of tweets classified as each of the emotions and sentiments for each day and metro area combination. For example, the count of anger tweets on January 1 for the New York–Newark–Jersey City, NY-NJ-PA metropolitan area divided by the total number of tweets on January 1 for the New York–Newark–Jersey City, NY-NJ-PA metropolitan area gives percentage of anger tweets for January 1 in the New York–Newark–Jersey City, NY-NJ-PA metropolitan area.

The total number of COVID-19 vaccine related tweets and users per 100,000 population was also calculated for each day of data collection, at the metropolitan level. Finally, user engagement metrics, including the average number of retweets and favorites, were calculated for each day of data collection, at the metropolitan level. Retweets and favorites suggest, after processing the information, that a user resonates with an idea expressed in a tweet [28,29]. Therefore, we believe these engagement metrics might also reflect vaccine attitudes.

Time Series Model

The data were divided into training and test data sets, where the time series analysis was trained using the data set created from the January 1, 2021, to April 12, 2021, time period, and tested on the data set created from the April 13, 2021, to May 20, 2021, time period. ARIMA models were executed for forecasting the proportion of individuals who have been administered at least 1 vaccine dose. Autoregressive integrated moving average exogenous variable model (ARIMAX) models, which are extensions of ARIMA models that include independent predictors called exogenous variables, were also executed. The ARIMA method has been widely used in time series forecasting and public health surveillance [30-32]. An ARIMA model typically consists of three components: (1) auto-regression, notated in the model as *p*; (2) differencing, notated in the model as *d*; and (3) moving average, notated in the model as *q* [33]. In an ARIMA model, the present value of the time-series is a linear function of random noise and its previous values; the present value is also a linear function of both present and past values of the residuals in the model; and the auto-regressive moving average model includes both the auto-regressive and moving average models, in addition to the historical values in the time series and its residuals [30].

Stationarity of a time series is a key assumption when making predictions based on past observations of a variable [34]. Stationarity requires the properties (mean and variance) of a time series to remain constant over time, thus making future values easier to predict [35]. Otherwise, the results are spurious and analyses are not valid [30]. The stationarity of all variables included in the time series was assessed using the Dickey-Fuller (dfuller) test. If the null hypothesis is rejected, stationarity is satisfied. If stationarity is not satisfied, variables must undergo differencing, a process that removes any trend in the times series that is not of interest [35]. All differencing and model selection was performed by the auto_arima function from the *pmdarima* package in Python [36], which is a function that selects the optimal order of the model based on the Hyndman-Khandakar algorithm for automatic ARIMA modeling [37]. A combination of unit root tests and minimization of the Akaike information criterion and Bayesian information criterion allows this algorithm to select the best preforming model order by fitting several variations of model components *p*, *d*, and *q* [38]. By including a penalty that is an increasing function of the number of estimated parameters, the information criteria scores maximize the goodness of fit while minimizing the number of model parameters, effectively dealing with both the risk of overfitting and the risk of underfitting [39,40].

For each metropolitan area, a baseline ARIMA model with no exogenous variables was constructed to forecast the 7-day rolling average of the number of individuals who have been administered at least 1 vaccine dose, using only past values of this outcome. To assess the ability of vaccine attitudes on Twitter to improve COVID-19 vaccination forecasts, multiple ARIMAX models were executed, each with individual Twitter-derived features included as exogenous variables. Additionally, we executed a multivariate ARIMAX model that included those Twitter attitudes that showed improvement over the ARIMA baseline across all metro areas. A final ARIMAX model that contained all Twitter features regardless of performance was attempted but did not converge. A complete list of the constructed time series models can be found in Table 1.

Table 1. Time series models predicting COVID-19 vaccine uptake, January 1, 2021, to May 20, 2021.

Model type	Exogenous variables
ARIMA[a]	None (baseline)
ARIMAX[b]	Number of users per 100,000 population
ARIMAX	Number of tweets per 100,000 population
ARIMAX	Average favorites
ARIMAX	Average retweets
ARIMAX	% Positive sentiment
ARIMAX	% Negative sentiment
ARIMAX	% Neutral sentiment
ARIMAX	% Trust
ARIMAX	% Surprise
ARIMAX	% Sadness
ARIMAX	% Joy
ARIMAX	% Fear
ARIMAX	% Disgust
ARIMAX	% Anticipation
ARIMAX	% Anger
ARIMAX	Best predictors (predictors that show improvement over baseline across all metro areas)

[a]ARIMA: autoregressive integrated moving average.

[b]ARIMAX: autoregressive integrated moving average exogenous variable model.

Ethical Considerations

This project does not meet the definition of human participant research under the purview of the University of Maryland Institutional Review Board according to federal regulations, section 45CFR46.102(e) [41].

Results

Twitter Data

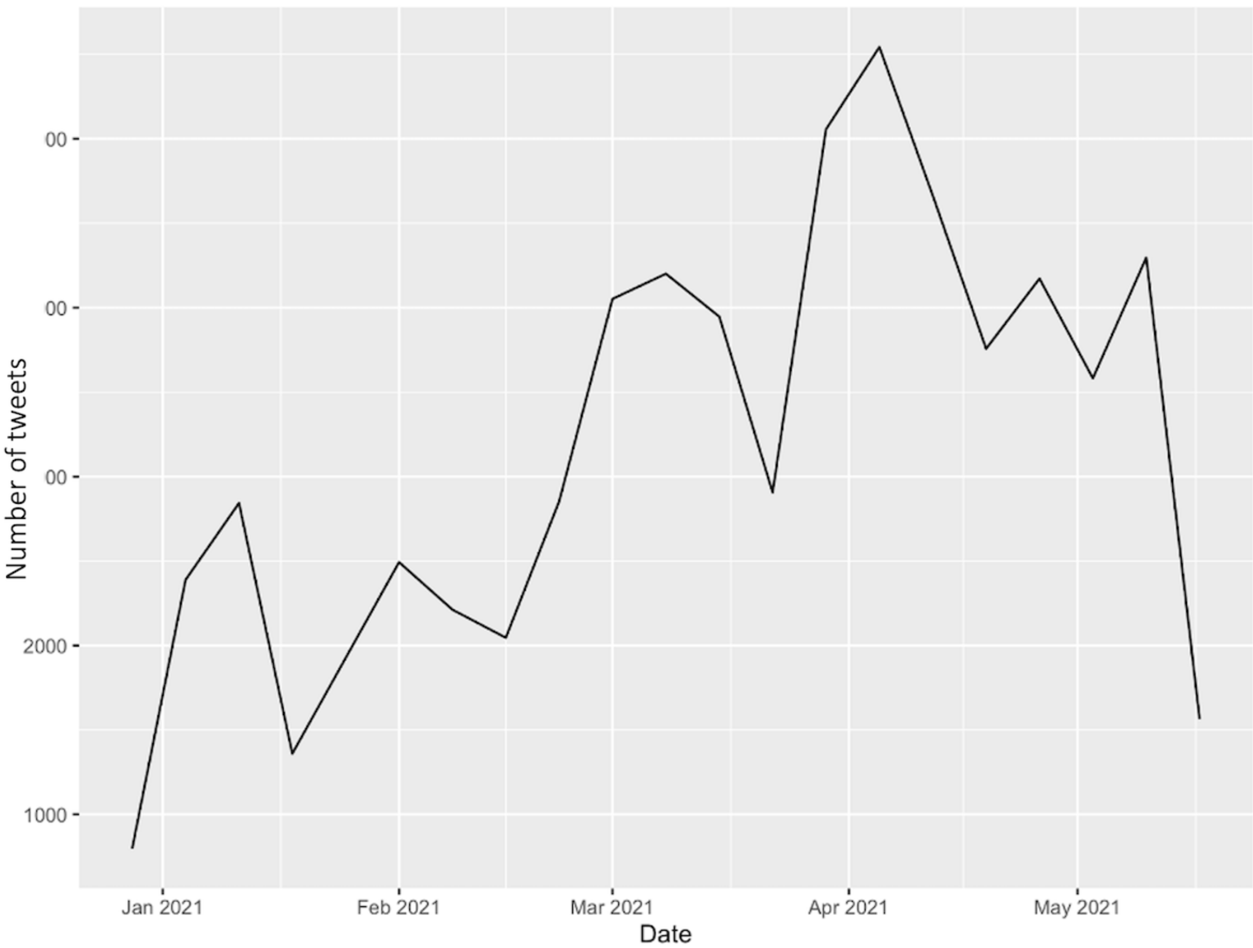
A total of 59,687 COVID-19 vaccine-related tweets were collected during the data collection period, across 23,878 users (Table 2). The Los Angeles–Long Beach–Anaheim metropolitan area had the largest representation of tweets (13,125/59,687, 21.99%) as well as the largest representation of users (5620/23,878, 23.54%). The Miami–Fort Lauderdale–Pompano Beach metropolitan area had the smallest representation of tweets (1631/59,687, 2.73%) as well as the smallest representation of users (625/23,878, 2.62%). The maximum number of tweets by a single individual was 228 (from a user in the Washington-Arlington-Alexandria metropolitan area).

The temporal trends for the number of COVID-19 vaccine–related tweets from January to May 2021 are presented in Figure 1. The number of COVID-19 vaccine–related tweets fluctuated over time; however, a peak in the number of tweets was observed during the week of April 5, 2021, to April 11, 2021. This was the week that President Joe Biden announced that every adult in the United States would be eligible to receive a COVID-19 vaccine starting April 19, 2021 [42].

Table 2. Number of COVID-19 vaccine tweets (n=59,687) and users (n=23,878) by city, January 1, 2021, to May 20, 2021.

Metropolitan area	Tweets, n, %	Users, n, %	Average retweets, mean (SD)	Average favorites, mean (SD)
Atlanta–Sandy Springs–Alpharetta, GA	12,623 (21.1)	5431 (22.7)	438 (5140)	10 (178)
Chicago-Naperville-Elgin, IL-IN-WI	6857 (11.5)	2847 (11.9)	543 (9579)	11 (118)
Los Angeles–Long Beach–Anaheim, CA	12,387 (20.8)	4858 (20.3)	351 (4209)	13 (224)
Miami–Fort Lauderdale–Pompano Beach, FL	4345 (7.3)	1558 (6.5)	267 (3187)	131 (2389)
New York–Newark–Jersey City, NY-NJ-PA	2231 (3.7)	914 (3.8)	169 (1704)	6 (20)
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	6488 (10.9)	2025 (8.5)	304 (3952)	13 (124)
Phoenix-Mesa-Chandler, AZ	12,623 (21.1)	5431 (22.7)	438 (5140)	10 (178)
Washington-Arlington-Alexandria, DC-VA-MD-WV	6857 (11.5)	2847 (11.9)	543 (9579)	11 (118)

Figure 1. Number of COVID-19 vaccine tweets over time, across all metropolitan areas, January 1, 2021, to May 20, 2021.



Sentiment and Emotion Analysis

A sentiment analysis classified most tweets across all metropolitan areas as having neutral sentiment, with joy as the predominantly expressed emotion (Table 3). The Phoenix-Mesa-Chandler metropolitan area had the largest proportion of tweets with positive sentiment (3875/12,623, 30.7%), while the Miami–Fort Lauderdale–Pompano Beach metropolitan area had the lowest proportion of tweets with positive sentiment (1065/4345, 24.5%). Anger and disgust were the most perceived negative emotions. The Atlanta–Sandy Springs–Alpharetta, GA metropolitan area had the largest proportion of tweets with negative sentiment (3888/12,623, 30.8%), while the Miami–Fort Lauderdale–Pompano Beach metropolitan area had the lowest proportion of tweets with negative sentiment (1060/4345, 24.4%).

Table 3. Distribution of sentiments and emotions among COVID-19 vaccine tweets collected from January 1, 2021, to May 20, 2021 (N=59,687).

Metropolitan area	Anger, %	Anticipation, %	Disgust, %	Fear, %	Joy, %	Sadness, %	Surprise, %	Trust, %	Negative, %	Neutral, %	Positive, %
Atlanta–Sandy Springs–Alpharetta, GA	18.1	24	14.2	7	30.4	6	0.2	0	30.8	42	27.2
Chicago-Naperville-Elgin, IL-IN-WI	16.9	23.4	13.9	6.3	33.7	5.7	0.1	0.1	29.1	40.4	30.6
Los Angeles–Long Beach–Anaheim, CA	17.4	22.4	12.7	7.1	33.4	6.9	0.2	0	30.4	39.9	29.7
Miami–Fort Lauderdale–Pompano Beach, FL	14.2	28.1	13.9	5.6	32.4	5.5	0.1	0.1	24.4	51.1	24.5
New York–Newark–Jersey City, NY-NJ-PA	17.5	24.9	13.7	6.5	31.8	5.3	0.2	0.1	28.8	42.3	28.9
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	18.5	26	13.4	6.9	29.2	5.8	0.1	0	28.6	44.1	27.4
Phoenix-Mesa-Chandler, AZ	18	23	13.9	7.1	32.2	5.7	0.1	0	28.5	40.8	30.7
Washington-Arlington-Alexandria, DC-VA-MD-WV	15.3	28.6	14.3	6.8	29.6	5.3	0.1	0	28.2	44.7	27.1

Time Series Forecast

Multiple time series models were constructed to forecast the vaccine uptake rate (7-day rolling average). The results of the Dickey-Fuller (dfuller) test for stationarity revealed that across all metropolitan areas, stationarity did not hold for several of the variables (Tables 4 and 5). However, the necessary differencing was automatically applied via the auto_arima function.

The performance of the optimal models across all regions, as determined by the auto_arima function, can be found in Tables 6 and 7. The best-performing model for each metropolitan area is marked by an asterisk. Models that performed better than the baseline model are bolded. Model performance for the “out-sample” forecasts was evaluated using the root mean square error (RMSE) instead of Akaike information criterion because RMSE measures how close the data are around the line of best fit [43]. This measure is commonly used in time series forecasting to evaluate how close the forecasted values are to the actual values [44]. When evaluating model performance using RMSE, across all metropolitan areas, the addition of a Twitter-derived feature related to COVID-19 vaccination attitudes improved model performance by up to 83%. For example, across all metropolitan areas, adding the *percentage of vaccine tweets expressing joy, negative sentiment, surprise, or trust* individually as exogenous variables resulted in a lower RMSE compared to the baseline ARIMA model. Additionally, across all metropolitan areas, most of the ARIMAX models, which each had 1 Twitter-derived feature related to COVID-19 vaccination attitudes, showed improvement over the baseline ARIMA model that did not factor in Twitter-derived features. A final model that contained the 3 features that consistently showed improvement over baseline across all metro areas (negative sentiment [%], surprise [%], joy [%], trust [%]) showed improvement over the baseline ARIMA when combined into 1 model (ARIMAX with multiple exogenous variables) across all metropolitan areas except for Philadelphia-Camden-Wilmington, PA-NJ-DE-MD and Phoenix-Mesa-Chandler, AZ.

Table 4. Dickey-Fuller (dfuller) test for stationarity in Atlanta–Sandy Springs–Alpharetta, GA, Chicago-Naperville-Elgin, IL-IN-WI Los Angeles-–Long Beach–Anaheim, CA, and Miami–Fort Lauderdale-–Pompano Beach, FL.

Variable	Atlanta–Sandy Springs–Alpharetta, GA		Chicago-Naperville-Elgin, IL-IN-WI		Los Angeles–Long Beach–Anaheim, CA		Miami–Fort Lauderdale–Pompano Beach, FL	
	Test statistic	*P* value	Test statistic	*P* value	Test statistic	*P* value	Test statistic	*P* value
% Anger	−1.237	.66[a]	−3.091	.03	−2.512	.11[a]	−3.776	.003
% Anticipation	−1.879	.34[a]	−2.579	.10[a]	−1.594	.49[a]	−3.375	.01
Average favorites	−4.154	.001	−3.073	.03	−4.080	.001	−2.956	.04
Average retweets	−2.882	.047	1.632	.99[a]	−3.526	.007	−3.239	.02
% Disgust	−2.915	.04	−2.435	.13[a]	−1.711	.42[a]	−3.414	.01
% Fear	−2.908	.04	−3.698	.004	−3.195	.02	−2.707	.07[a]
% Joy	−1.548	.51[a]	−2.500	.12[a]	−1.354	.60[a]	−2.264	.18[a]
% Negative sentiment	−1.666	.45[a]	−2.198	.21[a]	−1.425	.57[a]	−2.142	.23[a]
% Neutral sentiment	−2.223	.20[a]	−3.820	.003	−1.655	.46[a]	−2.841	.05[a]
Number of tweets per 100,000 population	−1.521	.52[a]	−1.333	.61[a]	−1.221	.66[a]	−2.001	.29[a]
Number of users per 100,000 population	−1.450	.56[a]	−1.334	.61[a]	−1.241	.66[a]	−1.947	.31[a]
% Positive sentiment	−1.281	.64[a]	−2.626	.09[a]	−1.256	.65[a]	−2.847	.05[a]
Percentage of individuals who have been administered at least 1 vaccine dose (7-day rolling average)	−0.569	.88[a]	−0.814	.82[a]	−0.048	.95[a]	0.057	.96[a]
% Sadness	−3.817	.003	−2.619	.09[a]	−3.157	.02	−3.249	.02
% Surprise	−4.030	.001	−2.349	.16[a]	−3.883	.002	−1.658	.45[a]
% Trust	−2.739	.07[a]	−3.128	.02	−2.120	.24[a]	−5.039	<.001

[a]Nonstationary variable results.

Table 5. Dickey-Fuller (dfuller) test for stationarity in New York–Newark–Jersey City, NY-NJ-PA, Philadelphia-Camden-Wilmington, PA-NJ-DE-MD, Phoenix-Mesa-Chandler, AZ, and Washington-Arlington-Alexandria, DC-VA-MD-WV.

	New York–Newark–Jersey City, NY-NJ-PA		Philadelphia-Camden-Wilmington, PA-NJ-DE-MD		Phoenix-Mesa-Chandler, AZ		Washington-Arlington-Alexandria, DC-VA-MD-WV	
Variable	Test statistic	*P* value	Test statistic	*P* value	Test statistic	*P* value	Test statistic	*P* value
% Anger	−3.084	.03	−4.880	<.001	−3.275	.02	−2.233	.20[a]
% Anticipation	−3.336	.01	−3.586	.006	−3.400	.01	−2.111	.24[a]
Average favorites	−2.786	.06[a]	−3.001	.04	−3.367	.01	−2.507	.11[a]
Average retweets	−2.724	.07[a]	−1.647	.46[a]	−2.596	.10[a]	−2.451	.13[a]
% Disgust	−2.218	.20[a]	−2.307	.17[a]	−1.678	.44[a]	−2.330	.16[a]
% Fear	−2.730	.07[a]	−3.129	.02	−3.625	.005	−2.960	.043
% Joy	−2.383	.15[a]	−2.702	.07[a]	−1.826	.37[a]	−1.493	.54[a]
% Negative sentiment	−1.432	.57[a]	−2.897	.046	−1.846	.36[a]	−1.747	.41[a]
% Neutral sentiment	−1.993	.29[a]	−3.635	.005	−1.738	.41[a]	−1.998	.29[a]
Number of tweets per 100,000 population	−1.402	.58[a]	−1.617	.47[a]	−1.205	.67[a]	−1.890	.34[a]
Number of users per 100,000 population	−1.461	.55[a]	−1.697	.43[a]	−1.116	.71[a]	−1.796	.38[a]
% Positive sentiment	−1.702	.43[a]	−3.793	.003	−2.080	.25[a]	−1.572	.498[a]
Percentage of individuals who have been administered at least 1 vaccine dose (7 day rolling average)	−0.792	.82[a]	−1.064	.73[a]	−1.483	.54[a]	−0.085	.95[a]
% Sadness	−2.862	.05	−2.263	.18[a]	−3.206	.02	−1.954	.31[a]
% Surprise	−2.893	.046	−2.599	.09[a]	−3.082	.03	−1.544	.51[a]
% Trust	−2.733	.069[a]	−2.463	.06[a]	−2.854	.05[a]	−3.078	.03

[a]Nonstationary variable results.

Table 6. ARIMA[a]/ARIMAX[b] model performance (RMSE[c]) for Atlanta–Sandy Springs–Alpharetta, GA, Chicago-Naperville-Elgin, IL-IN-WI, Los Angeles–Long Beach–Anaheim, CA, and Miami–Fort Lauderdale–Pompano Beach, FL. Models that performed better than the baseline ARIMA are shown in italics.

Variables	Atlanta–Sandy Springs–Alpharetta, GA, RMSE	Chicago-Naperville-Elgin, IL-IN-WI, RMSE	Los Angeles–Long Beach–Anaheim, CA, RMSE	Miami–Fort Lauderdale–Pompano Beach, FL, RMSE
(Baseline) percentage of individuals who have been administered at least 1 vaccine dose (7 day rolling average)	4.0855	4.2182	4.1473	3.7109
Number of users per 100,000 population	*1.6356*	*1.3900*	*0.7198*	*1.2971*
Number of tweets per 100,000 population	*1.6420*	*1.3700*	*0.7131*	*1.3492*
Average favorites	*2.1176*	*4.2100*	*0.6878*	4.0045
Average retweets	5.3545	*1.2414*	*4.1356*	*1.4062*
% Positive sentiment,	*1.6182*	*1.3000*	*0.7051*	*1.1691*
% Negative sentiment	*1.6238*	*1.3300*	*0.6915*	*1.2168*
% Neutral sentiment	*1.6236*	*1.1600* [d]	*0.7213*	3.7217
% Trust	*4.0854*	4.2183	*4.1471*	*1.1407*
% Surprise	*1.6522*	*1.3371*	*0.7078*	*1.1314*
% Sadness	*1.5826* [d]	*1.2400*	*0.7077*	3.7117
% Joy	*1.6243*	*1.3600*	*0.6865* [d]	*1.2322*
% Fear	*1.6751*	*4.1900*	*0.6973*	*3.7028*
% Disgust	*1.6401*	*1.3200*	*0.7054*	3.7670
% Anger	4.6909	*4.2000*	*0.7037*	*1.1006* [d]
% Anticipation	*1.6589*	*1.2800*	*0.7079*	3.7115
Best predictors: joy (%), negative sentiment (%), surprise (%), trust (%)	*1.7324*	*1.2878*	*0.6921*	*1.2901*

[a]ARIMA: autoregressive integrated moving average.

[b]ARIMAX: autoregressive integrated moving average exogenous variable model.

[c]RMSE: root mean square error.

[d]Best-performing model for each metropolitan area.

Table 7. ARIMA[a]/ARIMAX[b] model performance (RMSE[c]) for New York–Newark–Jersey City, NY-NJ-PA, Philadelphia-Camden-Wilmington, PA-NJ-DE-MD, Phoenix-Mesa-Chandler, AZ, and Washington-Arlington-Alexandria, DC-VA-MD-WV. Models that performed better than the baseline ARIMA are shown in italics.

	New York–Newark–Jersey City, NY-NJ-PA, RMSE	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD, RMSE	Phoenix-Mesa-Chandler, AZ, RMSE	Washington-Arlington-Alexandria, DC-VA-MD-WV, RMSE
(Baseline) percentage of individuals who have been administered at least 1 vaccine dose (7 day rolling average)	4.7686	4.4737	2.7743	2.5829
Number of users per 100,000 population	*1.8473*	*1.2524*	*1.9983*	*1.1554*
Number of tweets per 100,000 population	*1.8329*	*1.2737*	*1.9769*	*1.1259*
Average favorites	*4.7564*	5.6078	*1.9406*	*0.7570*
Average retweets	*1.8838*	*0.9200*[d]	*2.7728*	*0.7570*
% Positive sentiment,	*1.8682*	4.8339	*1.9452*	*1.1168*
% Negative sentiment	*1.8757*	*1.2486*	*1.9118*	*1.1206*
% Neutral sentiment	4.7722	*1.2392*	*1.8932*	2.5825
% Trust	*1.8659*	*1.2503*	*1.9372*	*1.1327*
% Surprise	*4.7668*	*1.2279*	*1.9374*	*1.1210*
% Sadness	*4.4896*	*1.1615*	*1.9355*	*2.4392*
% Joy	*1.8397*	*1.1956*	*1.9424*	*1.1450*
% Fear	4.7720	4.5114	*1.9371*	*1.0632*
% Disgust	*1.8207*[d]	*1.2506*	*1.9520*	*1.1380*
% Anger	*1.9003*	4.6179	*1.8858*[d]	*0.6834*
% Anticipation	*1.9060*	*1.2348*	*1.9454*	*1.1088*
Best predictors: joy (%), negative sentiment (%), surprise (%), trust (%)	2.7323	33.5446	5.1538	*0.6816*[d]

[a]ARIMA: autoregressive integrated moving average.

[b]ARIMAX: autoregressive integrated moving average exogenous variable model.

[c]RMSE: root mean square error.

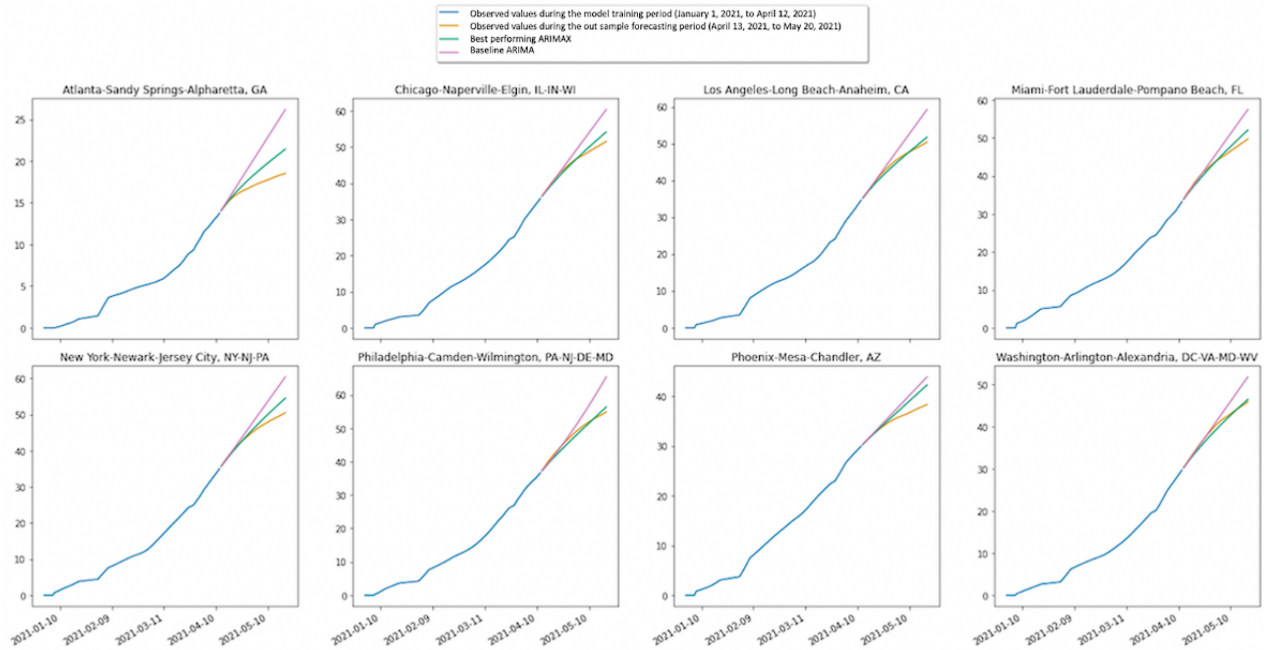
[d]Best-performing model for each metropolitan area.

Effect of Models on Performance

To understand the effect of modeling choices on the usefulness of Twitter-derived features to improve COVID-19 vaccination rate predictions, we evaluated 2 additional models: one that used the *Syuzhet* package [45]—instead of BERT—to extract the same set of sentiments and emotions from tweets and then ARIMA/ARIMAX to predict COVID-19 vaccination rates; and another model that used BERT to extract sentiments and emotions from tweets and deep learning—a Temporal Fusion Transformer Model [46]—to predict COVID-19 vaccination rates, instead of ARIMA/ARIMAX. We confirmed that independently of the model selected, the same findings hold—the results of these models show that adding Twitter-based features to COVID-19 vaccination rates in predictive models improves most baselines, independently of the model and the city, albeit with higher RMSE than the ones shown in Tables 6 and 7. We have included descriptions, results, and a discussion of these other 2 models in Multimedia Appendix 2.

Figure 2 illustrates the performance of the baseline ARIMA models and the best-performing ARIMAX models, compared to the observed values of the outcome variable during the “out-sample” forecasting period (April 13, 2021, to May 20, 2021). Across all metropolitan areas, the ARIMAX time series models with Twitter-derived features aligned more closely with the actual values of the vaccination rates compared to the baseline ARIMA model that relied on past historical vaccination data alone.

Figure 2. Predicted versus observed COVID-19 vaccination rates, January 1, 2021, to May 20, 2021. ARIMA: autoregressive integrated moving average; ARIMAX: autoregressive integrated moving average exogenous variable model.



Discussion

Principal Findings

In this study, we sought to determine whether supplementing forecast models with COVID-19 vaccine attitudes found in tweets—modeled via sentiments and emotions—improves over baseline models that only use historical vaccination data. When evaluating model performance across all metropolitan areas, the addition of COVID-19 vaccine attitudes found in tweets resulted in improved model performance, as reflected by RMSE, when compared to baseline forecast models that did not include these features. Specifically, compared with the traditional ARIMA model with vaccination data alone, ARIMAX models with the predictions of both historical vaccination data and COVID-19 vaccine attitudes found in tweets reduced RMSE by as much as 83%. We were able to replicate similar findings across various modeling choices, including the *Syuzhet* package to extract sentiments and emotions, instead of BERT, and deep learning (temporal fusion transformer model) to predict COVID-19 vaccination rates, instead of ARIMA/ARIMAX.

Study Findings in Context

The ongoing COVID-19 pandemic emphasizes the need for innovative approaches to public health surveillance. The global public health community has monitored the COVID-19 pandemic by tracking case counts, hospitalizations, deaths, and vaccinations. For the United States, these data sets are publicly available. Forecasting case counts and vaccination rates using existing historical data has been a key approach in COVID-19 surveillance efforts [47]. Previous forecast models for predicting vaccine uptake rate relied on traditional ARIMA methods, where historical data were used to predict future rates [48]. However, social media data sources, such as Twitter, reveal society's attitudes toward the pandemic and current vaccination efforts on a real-time basis. This provides an opportunity for a large volume of raw and uncensored data related to vaccine attitudes, across various geographic locations, to be leveraged for disease surveillance, which can subsequently be used to supplement and improve existing models.

The findings of this study suggest that attitudes extracted from Twitter data can be added to existing forecast models for monitoring vaccination uptake across various metropolitan areas. In certain metropolitan areas, the mere volume of tweets and users engaged in vaccine-related conversations improved model performance when compared to baseline models. These results echo the findings in the study by Maugeri et al [33], which revealed another social media source, Google Trends data, improved the prediction of COVID-19 vaccination uptake in Italy when compared to baseline models. In this study, Google Trends data were represented as the relative search volume for each vaccine-related keyword. Another similar study developed a framework for predicting vaccination rates in the United States based on traditional clinical data and web search queries [49]. The results of this study also revealed the ability for online networks to predict societal willingness to receive vaccinations. Specifically, the authors similarly found improvement in model performance as in this study—with a reduction in RMSE of 9.1%.

Although few studies sought to supplement current vaccine models with social media data, to our knowledge, there are no studies that go beyond the mere volume of relevant Twitter data and factor in the sentiment and emotion of vaccine-related conversations. Over the course of the pandemic, some states experienced low vaccination rates despite comprehensive vaccine roll out programs. In these cases, it is important to consider the public's emotions and sentiments toward vaccines. This study contributes to the literature by evaluating the ability for sentiments and emotions related to the COVID-19 vaccine to predict vaccine uptake. Specifically, the results show an improvement in model performance across metropolitan areas when models were supplemented with the percentage of tweets expressing anger, fear, joy, positive sentiment, or neutral sentiment. A study conducted by Alegado and Tumibay [48] examined the association between sentiments and emotions found in tweets and vaccine uptake via regression coefficient analysis. This study showed similar insights—tweets expressing fear, sadness, and anger appeared to be significantly associated with vaccination rates.

The results of this study have several implications for the present COVID-19 response. Public health experts now argue that the traditional concept of herd immunity may not apply to COVID-19 [2]. Instead, the focus is to increase vaccination uptake to substantially control community spread, without the societal disruptions caused by the virus [3]. Accurately forecasting vaccination uptake allows policy makers and researchers to evaluate how close we are to achieving normalcy again. Additionally, similar algorithms allow public health practitioners to better anticipate vaccine uptake behaviors and therefore develop targeted policies. As the global community builds toward achieving herd immunity, researchers should also "listen" to the vaccine conversation on social media—monitoring misconceptions and misinformation and implementing targeted vaccine education campaigns that address these misconceptions. Although the COVID-19 pandemic appears to be improving, the present framework can also be used to improve vaccine forecast models for future pandemics.

Limitations and Future Work

It is important to note that this study has some limitations. The study period was limited to the first half of 2021. However, vaccines were not yet available to most of the US adult population until April 2021. Therefore, the study period did not capture the height of vaccination efforts. Another limitation is that as the COVID-19 pandemic evolves, vaccine related keywords may change, requiring frequent updating of the model. Future work may involve the use of topic modeling to capture the general themes surrounding the COVID-19 pandemic.

Another limitation is related to the geographic scope of this study. This study only focused on forecasting vaccine uptake in the United States. However, it is important to note that vaccination efforts must be addressed on a global scale, not just domestically, for normalcy to be attained. Future work should consider collecting tweets and vaccination data from other countries to see if similar models improve vaccine forecasts globally. Additionally, this study only examined tweets posted in the English language. Limiting the study to the collection of Tweets only in the English language poses a limitation as it may overlook valuable insights and perspectives expressed in other languages. This exclusion could lead to a biased understanding of sentiments and emotions, potentially missing out on crucial data from non–English-speaking populations. Language barriers may hinder the study's generalizability and restrict the representation of diverse cultural contexts. Future work should involve the use of sentiment and emotion classifiers that include lexicons in other languages.

Conclusions

Researchers have found that the internet and social media both play a role in shaping personal or parental choices about vaccinations. Although few previous studies have developed forecast models for COVID-19 vaccination rates in the United States, to our knowledge, there are no studies that aim to factor in the real-time vaccination attitudes present on Twitter. This study suggests the benefits of using the linguistic constructs found in tweets to improve predictions of the COVID-19 vaccination rate. In this study, we found that supplementing baseline forecast models with both historical vaccination data and COVID-19 vaccine attitudes found in tweets reduced RMSE by as much as 83%. Developing a predictive tool for vaccination uptake in the United States will empower public health researchers and decision makers to design targeted vaccination campaigns in hopes of achieving the vaccination threshold required for widespread population protection.

Conflicts of Interest

None declared.

Multimedia Appendix 1

COVID-19 vaccine keywords.

[DOCX File , 15 KB - infodemiology_v3i1e43703_app1.docx]

Multimedia Appendix 2

Alternative prediction models.

[DOCX File , 292 KB - infodemiology_v3i1e43703_app2.docx]

References

1. Koyama T, Weeraratne D, Snowdon JL, Parida L. Emergence of drift variants that may affect COVID-19 vaccine development and antibody treatment. Pathogens 2020;9(5):324 [FREE Full text] [doi: 10.3390/pathogens9050324] [Medline: 32357545]
2. Morens DM, Folkers GK, Fauci AS. The concept of classical herd immunity may not apply to COVID-19. J Infect Dis 2022;226(2):195-198 [FREE Full text] [doi: 10.1093/infdis/jiac109] [Medline: 35356987]

3. NIH experts discuss controlling COVID-19 in commentary on herd immunity. National Institutes of Health. 2022. URL: https://www.nih.gov/news-events/news-releases/nih-experts-discuss-controlling-covid-19-commentary-herd-immunity [accessed 2022-09-12]
4. Bhattacharjee A, Vishwakarma GK, Gajare N, Singh N. Time series analysis using different forecast methods and case fatality rate for COVID-19 pandemic. Region Sci Policy Practice 2022;15(3):506-519 [FREE Full text] [doi: 10.1111/rsp3.12555]
5. Kim M. Prediction of COVID-19 confirmed cases after vaccination: based on statistical and deep learning models. SciMed J 2021;3(2):153-165 [FREE Full text] [doi: 10.28991/scimedj-2021-0302-7]
6. Davahli MR, Karwowski W, Fiok K. Optimizing COVID-19 vaccine distribution across the United States using deterministic and stochastic recurrent neural networks. PLoS One 2021;16(7):e0253925 [FREE Full text] [doi: 10.1371/journal.pone.0253925] [Medline: 34228740]
7. Germani F, Biller-Andorno N. The anti-vaccination infodemic on social media: a behavioral analysis. PLoS One 2021;16(3):e0247642 [FREE Full text] [doi: 10.1371/journal.pone.0247642] [Medline: 33657152]
8. 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]
9. Smith N, Graham T. Mapping the anti-vaccination movement on Facebook. Inf Commun Soc 2017;22(9):1310-1327. [doi: 10.1080/1369118x.2017.1418406]
10. Bari A, Heymann M, Cohen RJ, Zhao R, Szabo L, Apas Vasandani S, et al. Exploring coronavirus disease 2019 vaccine hesitancy on Twitter using sentiment analysis and natural language processing algorithms. Clin Infect Dis 2022;74(Suppl_3):e4-e9 [FREE Full text] [doi: 10.1093/cid/ciac141] [Medline: 35568473]
11. Yousefinaghani S, Dara R, Mubareka S, Sharif S. Prediction of COVID-19 waves using social media and Google search: a case study of the US and Canada. Front Public Health 2021;9:656635 [FREE Full text] [doi: 10.3389/fpubh.2021.656635] [Medline: 33937179]
12. O'Leary DE, Storey VC. A Google–Wikipedia–Twitter model as a leading indicator of the numbers of coronavirus deaths. Intell Sys Acc Fin Mgmt 2020;27(3):151-158 [FREE Full text] [doi: 10.1002/isaf.1482]
13. Sun J, Gloor PA. Assessing the predictive power of online social media to analyze COVID-19 outbreaks in the 50 U.S. States. Future Internet 2021;13(7):184 [FREE Full text] [doi: 10.3390/fi13070184]
14. Shen C, Chen A, Luo C, Zhang J, Feng B, Liao W. Using reports of symptoms and diagnoses on social media to predict COVID-19 case counts in Mainland China: observational infoveillance study. J Med Internet Res 2020;22(5):e19421 [FREE Full text] [doi: 10.2196/19421] [Medline: 32452804]
15. Sattar NS, Arifuzzaman S. COVID-19 vaccination awareness and aftermath: public sentiment analysis on Twitter data and vaccinated population prediction in the USA. Appl Sci 2021;11(13):6128 [FREE Full text] [doi: 10.3390/app11136128]
16. Cihan P. Forecasting fully vaccinated people against COVID-19 and examining future vaccination rate for herd immunity in the US, Asia, Europe, Africa, South America, and the World. Appl Soft Comput 2021;111:107708 [FREE Full text] [doi: 10.1016/j.asoc.2021.107708] [Medline: 34305491]
17. Cheong Q, Au-Yeung M, Quon S, Concepcion K, Kong JD. Predictive modeling of vaccination uptake in US Counties: a machine learning-based approach. J Med Internet Res 2021;23(11):e33231 [FREE Full text] [doi: 10.2196/33231] [Medline: 34751650]
18. Shaham A, Chodick G, Shalev V, Yamin D. Personal and social patterns predict influenza vaccination decision. BMC Public Health 2020;20(1):222 [FREE Full text] [doi: 10.1186/s12889-020-8327-3] [Medline: 32050948]
19. Mannion N. Predictions of changes in child immunization rates using an automated approach: USA [dissertation]. National College of Ireland. URL: https://norma.ncirl.ie/4368/1/niallmannion.pdf [accessed 2023-07-07]
20. Metropolitan and micropolitan statistical areas population totals: 2020-2022. United States Census Bureau. 2022. URL: https://www.census.gov/data/tables/time-series/demo/popest/2020s-total-metro-and-micro-statistical-areas.html [accessed 2022-08-07]
21. Melotte S, Kejriwal M. Predicting zip code-level vaccine hesitancy in US Metropolitan Areas using machine learning models on public tweets. PLoS Digit Health 2022;1(4):e0000021 [FREE Full text] [doi: 10.1371/journal.pdig.0000021] [Medline: 36812517]
22. Malik M, Lamba H, Nakos C, Pfeffer J. Population bias in geotagged Tweets. In: Proceedings of the International AAAI Conference on Web and Social Media. 2021 Aug 03 Presented at: International AAAI Conference on Web and Social Media; May 26-29, 2015; Oxford, UK p. 18-27. [doi: 10.1609/icwsm.v9i4.14688]
23. COVID-19 vaccinations in the United States, County. US Centers for Disease Control and Prevention. 2021. URL: https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh [accessed 2022-01-13]
24. Delineation Files: core based statistical areas (CBSAs), metropolitan divisions, and combined statistical areas (CSAs). United States Census Bureau. URL: https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html [accessed 2022-08-07]
25. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. ArXiv Preprint posted online on May 24, 2019 [FREE Full text] [doi: 10.48550/arXiv.1810.04805]

26. Camacho-Collados J, Rezaee K, Riahi T, Ushio A, Loureiro D, Antypas D, et al. TweetNLP: cutting-edge natural language processing for social media. ArXiv Preprint posted online on October 25, 2022 [FREE Full text]
27. Barbieri F, Camacho-Collados J, Neves L, Espinosa-Anke L. TweetEval: unified benchmark and comparative evaluation for tweet classification. ArXiv Preprint posted online on October 26, 2020 [FREE Full text] [doi: 10.18653/v1/2020.findings-emnlp.148]
28. Liu Z, Liu L, Li H. Determinants of information retweeting in microblogging. Internet Res 2012;22(4):443-466 [FREE Full text] [doi: 10.1108/10662241211250980]
29. Gittelman S, Lange V, Gotway Crawford CA, Okoro CA, Lieb E, Dhingra SS, et al. A new source of data for public health surveillance: Facebook likes. J Med Internet Res 2015;17(4):e98 [FREE Full text] [doi: 10.2196/jmir.3970] [Medline: 25895907]
30. Talaei-Khoei A, Wilson JM, Kazemi SF. Period of measurement in time-series predictions of disease counts from 2007 to 2017 in Northern Nevada: analytics experiment. JMIR Public Health Surveill 2019;5(1):e11357 [FREE Full text] [doi: 10.2196/11357] [Medline: 30664479]
31. Alzahrani SI, Aljamaan IA, Al-Fakih EA. Forecasting the spread of the COVID-19 pandemic in Saudi Arabia using ARIMA prediction model under current public health interventions. J Infect Public Health 2020;13(7):914-919 [FREE Full text] [doi: 10.1016/j.jiph.2020.06.001] [Medline: 32546438]
32. Zheng A, Fang Q, Zhu Y, Jiang C, Jin F, Wang X. An application of ARIMA model for predicting total health expenditure in China from 1978-2022. J Glob Health 2020;10(1):010803 [FREE Full text] [doi: 10.7189/jogh.10.010803] [Medline: 32257167]
33. Maugeri A, Barchitta M, Agodi A. Using Google trends to predict COVID-19 vaccinations and monitor search behaviours about vaccines: a retrospective analysis of Italian data. Vaccines (Basel) 2022;10(1):119 [FREE Full text] [doi: 10.3390/vaccines10010119] [Medline: 35062780]
34. Shumway RH, Stoffer DS. Springer Texts in Statistics: Time Series Analysis and its Applications: With R Examples. Cham, Switzerland: Springer International Publishing; 2017.
35. Jebb AT, Tay L, Wang W, Huang Q. Time series analysis for psychological research: examining and forecasting change. Front Psychol 2015;6:727 [FREE Full text] [doi: 10.3389/fpsyg.2015.00727] [Medline: 26106341]
36. pmdarima: ARIMA estimators for Python. pmdarima. URL: https://alkaline-ml.com/pmdarima/ [accessed 2023-06-05]
37. Forecast package. RDocumentation. URL: https://www.rdocumentation.org/packages/forecast/versions/8.16 [accessed 2022-08-07]
38. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. J Stat Soft 2008;27(3):1-22. [doi: 10.18637/jss.v027.i03]
39. Burnham KP, Anderson DR. Practical use of the information-theoretic approach. In: Model Selection and Inference: A Practical Information-Theoretic Approach. New York, NY: Springer; 1998:75-117.
40. Yang S, Berdine G. Model selection and model over-fitting. Southwest Respir Crit Care Chron 2015;3(12):52-55. [doi: 10.12746/swrccc2015.0312.160]
41. IRB Definitions. Division of Research - University of Maryland. URL: https://research.umd.edu/resources/research-compliance/institutional-review-board-irb/irb-definitions [accessed 2023-08-18]
42. Remarks by President Biden Marking the 150 millionth COVID-19 vaccine shot. The White House. 2021. URL: https://www.whitehouse.gov/briefing-room/speeches-remarks/2021/04/06/remarks-by-president-biden-marking-the-150-millionth-covid-19-vaccine-shot/ [accessed 2022-08-08]
43. Zhang J, Shang R, Rittenhouse C, Witharana C, Zhu Z. Evaluating the impacts of models, data density and irregularity on reconstructing and forecasting dense Landsat time series. Sci Remote Sensing 2021;4:100023 [FREE Full text] [doi: 10.1016/j.srs.2021.100023]
44. Jiang B, Liang S, Wang J, Xiao Z. Modeling MODIS LAI time series using three statistical methods. Remote Sens Environ 2010;114(7):1432-1444. [doi: 10.1016/j.rse.2010.01.026]
45. Silge J, Robinson D. tidytext: text mining and analysis using tidy data principles in R. J Open Source Softw 2016;1(3):37 [FREE Full text] [doi: 10.21105/joss.00037]
46. Lim B, Arık S, Loeff N, Pfister T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. Int J Forecast 2021;37(4):1748-1764 [FREE Full text] [doi: 10.1016/j.ijforecast.2021.03.012]
47. Alabdulrazzaq H, Alenezi MN, Rawajfih Y, Alghannam BA, Al-Hassan AA, Al-Anzi FS. On the accuracy of ARIMA based prediction of COVID-19 spread. Results Phys 2021;27:104509 [FREE Full text] [doi: 10.1016/j.rinp.2021.104509] [Medline: 34307005]
48. Alegado RT, Tumibay GM. Forecasting measles immunization coverage using ARIMA model. J Comput Commun 2019;7(10):157-168 [FREE Full text] [doi: 10.4236/jcc.2019.710015]
49. Zhou X, Li Y. Forecasting the COVID-19 vaccine uptake rate: an infodemiological study in the US. Hum Vaccin Immunother 2022;18(1):2017216 [FREE Full text] [doi: 10.1080/21645515.2021.2017216] [Medline: 35050825]

Abbreviations

ARIMA: autoregressive integrated moving average
ARIMAX: autoregressive integrated moving average exogenous variable model
BERT: bidirectional encoder representation from transformer
RMSE: root mean square error

Edited by R Cuomo; submitted 20.10.22; peer-reviewed by PA Gloor, SF Tsao, W Ahmed; comments to author 07.03.23; revised version received 08.06.23; accepted 26.06.23; published 21.08.23.

Please cite as:
Sigalo N, Awasthi N, Abrar SM, Frias-Martinez V
Using COVID-19 Vaccine Attitudes on Twitter to Improve Vaccine Uptake Forecast Models in the United States: Infodemiology Study of Tweets
JMIR Infodemiology 2023;3:e43703
URL: https://infodemiology.jmir.org/2023/1/e43703
doi:10.2196/43703
PMID:37390402

Original Paper

Addressing Antivaccine Sentiment on Public Social Media Forums Through Web-Based Conversations Based on Motivational Interviewing Techniques: Observational Study

David Scales[1,2], MPhil, MD, PhD; Lindsay Hurth[2], MPP; Wenna Xi[1], PhD; Sara Gorman[2], MPH, PhD; Malavika Radhakrishnan[2]; Savannah Windham[2], MS; Azubuike Akunne[2]; Julia Florman[2]; Lindsey Leininger[3], MA, PhD; Jack Gorman[2], MD

[1]Weill Cornell Medicine, New York City, NY, United States

[2]Critica, Bronx, NY, United States

[3]Tuck School of Business, Dartmouth College, Hannover, NH, United States

Corresponding Author:
David Scales, MPhil, MD, PhD
Weill Cornell Medicine
525 East 68th Street
Box 331
New York City, NY, 10065
United States
Phone: 1 2127464071
Email: das9289@med.cornell.edu

Abstract

Background: Health misinformation shared on social media can have negative health consequences; yet, there is a dearth of field research testing interventions to address health misinformation in real time, digitally, and in situ on social media.

Objective: We describe a field study of a pilot program of "infodemiologists" trained with evidence-informed intervention techniques heavily influenced by principles of motivational interviewing. Here we provide a detailed description of the nature of infodemiologists' interventions on posts sharing misinformation about COVID-19 vaccines, present an initial evaluation framework for such field research, and use available engagement metrics to quantify the impact of these in-group messengers on the web-based threads on which they are intervening.

Methods: We monitored Facebook (Meta Platforms, Inc) profiles of news organizations marketing to 3 geographic regions (Newark, New Jersey; Chicago, Illinois; and central Texas). Between December 2020 and April 2021, infodemiologists intervened in 145 Facebook news posts that generated comments containing either false or misleading information about vaccines or overt antivaccine sentiment. Engagement (emojis plus replies) data were collected on Facebook news posts, the initial comment containing misinformation (level 1 comment), and the infodemiologist's reply (level 2 reply comment). A comparison-group evaluation design was used, with numbers of replies, emoji reactions, and engagements for level 1 comments compared with the median metrics of matched comments using the Wilcoxon signed rank test. Level 2 reply comments (intervention) were also benchmarked against the corresponding metric of matched reply comments (control) using the Wilcoxon signed rank test (paired at the level 1 comment level). Infodemiologists' level 2 reply comments (intervention) and matched reply comments (control) were further compared using 3 Poisson regression models.

Results: In total, 145 interventions were conducted on 132 Facebook news posts. The level 1 comments received a median of 3 replies, 3 reactions, and 7 engagements. The matched comments received a median of 1.5 (median of IQRs 3.75) engagements. Infodemiologists made 322 level 2 reply comments, precipitating 189 emoji reactions and a median of 0.5 (median of IQRs IQR 0) engagements. The matched reply comments received a median of 1 (median of IQRs 2.5) engagement. Compared to matched comments, level 1 comments received more replies, emoji reactions, and engagements. Compared to matched reply comments, level 2 reply comments received fewer and narrower ranges of replies, reactions, and engagements, except for the median comparison for replies.

Conclusions: Overall, empathy-first communication strategies based on motivational interviewing garnered less engagement relative to matched controls. One possible explanation is that our interventions quieted contentious, misinformation-laden threads about vaccines on social media. This work reinforces research on accuracy nudges and cyberbullying interventions that also reduce engagement. More research leveraging field studies of real-time interventions is needed, yet data transparency by technology platforms will be essential to facilitate such experiments.

(JMIR Infodemiology 2023;3:e50138) doi:10.2196/50138

KEYWORDS

anti-vaccine; digital environment; engagement; health misinformation; infodemic; infodemiology; information environment; medical misinformation; misinformation; observational study; social media engagement metrics; social media

Introduction

Extensive research has shown that health misinformation has real, negative consequences. It can influence people to hold misperceptions and adopt unhealthy behaviors [1-8]. This led the US Surgeon General to issue a special advisory on the topic [9], in which health misinformation was defined as "information that is false, inaccurate, or misleading according to the best available evidence at the time" [10,11]. We also know that although supplying facts is often necessary to counteract misinformation, it is usually not sufficient to change opinions or behavior [12,13]. Belief in misinformation can be deeply ingrained, reinforced by psychological and social pressures, and difficult to dislodge [14]. This is especially the case with information on the internet and social media presented in misleading contexts and subjected to repeated sharing, reposting, and commenting. Some such information, whether true or false, can be spread with the intent to deliberately create misperceptions or sway public opinion [15].

From the onset of the COVID-19 pandemic, the World Health Organization (WHO) described the situation of a parallel "infodemic" [16], defined as "excess information, including false or misleading information, in digital and physical environments during an acute public health event" [17]. This infodemic focus has renewed interest in "infodemiology," the epidemiological study of these digitally enabled flows of information [18], and the need for professionals equipped to assess and respond to misinformation of public health importance as a core function of public health [19,20]. While an epidemic metaphor has its limits and externalities [21], it offers a framework with which to marshal resources toward understanding and mitigating the problem. Put another way, the lens of an infodemic suggests the need to develop field epidemiologists to deploy in public health and infodemic emergencies for rapid support of public health communications and interventions [22].

While public health institutions such as the Centers for Disease Control and Prevention (CDC) or the WHO issue messages on social media, these public health broadcasts are often at the periphery of web-based discussions about vaccines, Ebola, and the Zika virus [23-25]. Similarly, during the COVID-19 pandemic, information with public health relevance was decreasingly reliant on top-down recommendations from doctors and public health institutions (eg, the CDC) and more reliant on socially contextualized, decentralized, interpersonal, horizontal, and networked communication like that found on social media [26]. In contrast, antivaccine advocates are often leveraging the affordances of digital platforms to communicate in a coordinated, networked fashion [25,27]. We therefore hypothesized that best practice public health recommendations would not speak for themselves but would require trusted, community-linked advocates to communicate and interpret them within the value frames of those networks. We therefore developed a protocol based on principles of motivational interviewing and other evidence-based approaches, including inoculation, use of narratives, and promoting critical thinking, to address misinformation in web-based contexts and used it to intervene on Facebook (Meta Platforms, Inc) when misinformation about COVID-19 vaccines appeared [22].

Our approach is based on a menu of tactics derived from 3 main strategies: assessing how receptive the person posting health misinformation may be to an intervention; increasing high-quality, science-based messages across the web-based communication network; and reducing misinformation across that network. A critical principle underlying the protocol is derived from motivational interviewing (MI) techniques [28], which have shown efficacy in addressing vaccine hesitancy [29]. Using MI principles in this setting meant that the interventionist attempted to establish common ground with the person who posted misinformation and expressed empathy and an interest in understanding their point of view before responding directly to misinformed comments. Open-ended questioning and reflective listening in the spirit of MI are used throughout. Work to fully adapt MI to this setting, which we term community-oriented motivational interviewing, is ongoing [30].

As the person posting misinformation on social media is often committed to the misinformed point of view and unlikely to be immediately persuaded to consider an alternative perspective, the infodemiologists also consider the perspective of "bystanders" to the conversation, those observing but not necessarily engaging or commenting [31]. Such "bystanders" are hypothesized to be part of the "moveable middle" [32] and more persuadable about issues such as COVID-19 vaccines than the initial commentator. As misinformation can be perceived as true through repetition, infodemiologists seek to disrupt that "illusory truth effect" [33] while also role modeling how community members can make decisions commensurate with their values despite scientific uncertainty.

The purpose of this study is to provide a detailed description of infodemiologists' interventions. We also present an outline of an initial evaluation framework for such work, highlighting major gaps in the lack of accessibility of social media data that hinder researchers' ability to tie their work to more concrete outcomes, like behaviors. Finally, we quantify the impact these in-group messengers have on the web-based threads on which they are intervening.

Methods

Overview

Infodemiologists were drawn from the communities in which they intervene to help ensure trust through shared identity and values [34-37]. Details on the recruitment, credentials, training, and supervision of the 4 infodemiologists involved in this report have been published previously, along with descriptions of the intervention process [22,30]. Briefly, infodemiologists underwent a skills-based training consisting of practice interventions and weekly supervision sessions with one of the authors (DS or JMG) for feedback, totaling approximately 20-30 hours of training, practice, and supervision. They were first assigned independent reading to provide guidance on the evidence behind different communication techniques and then conducted a series of web-based training interventions with supervision, reflection, and feedback with DS and JMG [13,38,39]. The instructions on how to conduct interventions were broad, emphasizing that they needed to be tailored to the context. After the initial training, supervision continued at weekly group reflection sessions with all infodemiologists throughout the course of the study. All infodemiologist interventions were included in data collection and analysis, and none were excluded. A total of 145 pilot interventions were conducted between December 2020 and April 2021.

Full details on our misinformation monitoring and identification process are available from Gorman and Scales [22] and Scales et al [30]. In short, we monitored web-based Facebook profiles of news organizations marketing to 3 geographic regions: Newark, New Jersey; Chicago, Illinois; and central Texas. Regions were chosen for demographic, geographic, and urban or rural diversity. Infodemiologists were trained to select local media postings on Facebook that had generated comments containing either misinformation about vaccines or antivaccine sentiment within several hours of their posting. We defined misinformation about COVID-19 vaccines practically as any post that contained factually incorrect material or overtly negative sentiment about the vaccines, regardless of the motive of the person posting. This was a subjective assessment based on the infodemiologist's perception of what could be considered negative from the perspective of their community. For more details on how threads were chosen for interventions, see [30]. Infodemiologists recorded deidentified transcripts of the conversations (including ancillary comments from bystanders) as well as native engagement metrics (likes, shares, etc). Information on matched comments and replies for benchmarking was collected later, but sensitivity analysis did not find significant changes in conversation metrics over time. Infodemiologists were supported in their work through a process of written reflection after each intervention, direct written feedback on their interventions, and weekly group supervision sessions. Moreover, to protect them from harassment, infodemiologists were instructed to exit conversations that became emotionally heated or where they felt unsafe. To minimize web-based harassment, infodemiologists used the Critica's Facebook account and only identified themselves by their first names. The structure of the comments can be found in Figure 1, and an example of an intervention can be seen in Figure 2, paraphrased to protect the privacy of participants [40].

Figure 1. Visual description of comment and reply levels.

The infodemiologist intervened on a local news article from Texas describing how employers could require COVID-19 vaccines. To ensure the original post cannot be retroactively identified, the engagement and response numbers are rounded, and the transcript has been paraphrased.

In any given infodemiology session, infodemiologists deployed various evidence-informed communication techniques depending on the context (Table 1). Since there is little evidence to guide what communication technique should be used at any given time, we developed a 2-pronged approach to guide how and when to apply different intervention techniques. We engaged in discursive reflection among our team members, sometimes in real time through email or Slack (Slack Technologies) channels and at our weekly reflection meetings, to assess what techniques appeared to curtail conversations and engagement or promote reflection or resistance (ie, "change talk" or "sustain talk" per motivational interviewing language). Additionally, we paid particular attention to whether interventions elicited backfire effects, or psychological reactance, defined as escalating negative emotions through the course of an interaction with the infodemiologist [41]. Immediately after every initial infodemiologist intervention, they posted a disclaimer identifying themselves as researchers and a web link to further information about the research study, including options to request data be removed from our database. Of note, no requests for data removal were received. Table 1 provides a glimpse into the range of communication techniques that infodemiologists may use and the evidence behind them. It is not a comprehensive compilation of such techniques or the supporting evidence. A full review of this literature is beyond the scope of this study.

Table 1. Infodemiologist menu of techniques, protocols, and corresponding evidence.

Principle or approach and goal	Explanation or example	Reference
Receptivity to finding misinformation credible		
Infodemiologists should be from and within the communities and networks in which they will be intervening	Example cues (language, register, and slang) as markers of in-group identity	[35,42,43]
Assess readiness for change	Precontemplation, contemplation, preparation, action, and maintenance	[44,45]
Apply relevant principles of motivational interviewing	Open-ended questions, affirmations, reflective listening, summarizing, and promoting self-efficacy	[38,46,47]
Focus on the “fencers” or the “moveable middle”	People with heavily committed beliefs are unlikely to change their views quickly.	[31,32,48]
Promote critical thinking	“Please tell me more about that? Maybe give an example?”	[49]
Inoculation	“Misinformation will use various methods to make you doubt vaccines, like saying, “vaccines will make you infertile forever!”	[35,50-53]
Increase high-quality, science-based messages across the network		
Targeting highly visible Facebook news sources with little comment moderation	Facebook is often slow to implement its own misinformation takedown policies.	[54-56]
Reframing negative comments according to in-group cultural values regarding uncertainty, freedom to choose, etc	“I understand masks feel to you like they restrict freedom, but I’m proud to wear a mask with traditional designs that also helps protect our elders.”	[57-59]
Use personal narratives or anecdotes	“I was so relieved when I got vaccinated. I stopped worrying that I’d die if I got COVID”	[60]
Respond as quickly as possible after comments are posted	Ensures misinformation does not become entrenched or misinterpreted.	[61,62]
Detailed rebuttals, if needed	Rebuttals without explanations are less effective.	[49,63,64]
Link to sources likely to be trusted by commenters	Example: not citing the CDC[a] or the FDA[b] as sources if stakeholders are antigovernment	[34,61,65]
Repetition	People attribute more accuracy to repeated information.	[66]
Reduce misinformation across the network		
Reframing uncertainty as congruent with values	Unknown risks psychologically loom large; reframing them around known benefits provides balance.	[57,58]
Offering alternate, more plausible explanations	Plausible explanations accompanying warnings or rebuttals increase effectiveness.	[67]
Appeal to expert consensus	“97% of climate scientists agree that human-caused climate change is happening.”	[35,68-70]
Appeal to accuracy	Asking, “how accurate is that headline?”	[71,72]
Recontextualizing information taken out of context	“While true, these data are unverified user reports, not official statistics”	[73]

To build an evaluation strategy, we used the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework from implementation science [74], which provided a strategic guide for ideal evaluation of infodemiology interventions. Originally designed to incentivize scientists to be transparent and reflect on internal and external validity across the continuum of translational research from pilot to effectiveness studies, the RE-AIM framework was chosen here for 2 main reasons. First, it is familiar, being one of the most widely used and cited implementation science frameworks. Second, it has been successfully adapted to multiple and diverse contexts, suggesting it could also be applied to the web-based setting in which this research was done [75].

While implementation science is often used for interventions whose efficacy and effectiveness have already been established and which require further intervention to ensure their uptake into a specific context of interest, we believe it offers useful frameworks for work in digital spaces to counter misinformation, even though the evidence base is still emerging. In that context, we note that the communication techniques listed in Table 1 have demonstrated varying degrees of efficacy, and not all have demonstrated effectiveness in real-world, web-based settings. However, due to the constantly changing nature of web-based platforms, we recognized that effectiveness data either would not be forthcoming quickly or would essentially be outdated by publication due to platform changes (algorithmic, graphical user interface, or other). We therefore took the approach outlined by Lane-Fall et al [76] that hybrid implementation-effectiveness studies may be most appropriate when the urgency of the situation requires it, coupled with strong indirect evidence of potential effectiveness in the context of interest. In this study, we focus on the effectiveness results. Refer to Gorman and Scales [22] for an implementation discussion.

We collected data on the Facebook news post, the initial comment containing misinformation, which we will refer to as a "level 1 comment," and the infodemiologist reply comment, which is the start of the intervention, referred to as a "level 2 reply comment" (Figures 1 and 2 contain visual descriptions and an example of different comments and replies described here). Facebook organizes comments into threads, with level 2 comments branching off level 1 comments. Engagement was defined as the total number of comments and emoji reactions (like, love, hug, mad, haha, wow, and sad). In the context of this topic, emoji reactions "like," "love," and "hug" are interpreted as positive reactions, and "mad" and "haha" are negative reactions, with the latter interpreted as sarcasm. "Wow" and "sad" are considered neutral reactions.

Figure 2. Example transcript of an infodemiologist intervention on a local news article from Texas describing how employers could require COVID-19 vaccines. To prevent retroactive identification, engagement and response numbers are rounded and the transcript paraphrased.

Reactions: Like 200; Mad 40; Haha 30; Love 20, Wow 20; Sad 0, Hug 0;
Comments: 200
Shares: 30

Level 1 Initial Comment - Stakeholder A: This is a major violation of my rights. I totally support anyone who wants to get the vaccine. But I personally don't trust it. It's not political, but it's my right not to get the shot. I shouldn't be forced, that is tyranny. (Like 10; Haha 10)

Level 2 Reply Comment - Infodemiologist: @A I've heard a few people say they don't trust it, whether or not their job will require it. What are the reasons you don't trust it? (DISCLAIMER)

Stakeholder B: @A [Flashing GIF reading "idiot"]

Stakeholder C: @A Even Fauci said the mRNA vax was totally experimental. There's not enough research yet.

Stakeholder A: @B I'm an idiot because you disagree with me? Sheesh, at least now I know what kind of person you are.

Stakeholder D: @A But this is your employer "forcing" you to take the vaccine, not the government. How can you call it tyranny when it isn't the government that's being mean and oppressive? Since it's your choice to work for that employer NO ONE is "forcing" you at all. You may not be dumb, but I'd certainly say you're ignorant.

Infodemiologist: @C It's true that this is the first time mRNA vaccines have been widely used, but years of research led to the vaccines being ready for prime time. I like this article because it helps explain the history behind the vaccines and how they could be produced both quickly yet safely. [link to peer-reviewed paper in high-impact journal]

We implemented an innovative comparison-group evaluation design, building upon existing designs that measure engagement without a comparative benchmark [77-79]. Specifically, we collected data on comparison comments and replies adjacent to the intervention threads. Comments gathered for benchmarking were the five level 1 comments immediately above and below the level 1 comment to which the infodemiologist replied at the time of subsequent data collection, which, due to Facebook algorithms that are not transparent, may have changed from the time of the intervention. In rare circumstances where the level 1 comment subject to an infodemiologist's intervention could not be found (eg, absorbed into "Relevant" by Facebook), data were collected on 10 comments from the middle of the comments thread. Reply comments gathered for benchmarking were collected in the same way: the five level 2 comments above and below the infodemiologist's level 2 reply comment.

Collecting the comments data for benchmarking allowed us to consider the question, "How does engagement on the level 1 comment the infodemiologist chose to respond to compare to the engagement on the typical level 1 comment on the same Facebook news post?" Benchmarking with reply comments data allowed us to compare the engagement the infodemiologist interventions (ie, level 2 reply comments) received to the engagement on the typical level 2 reply comment on the same Facebook news post. For each level 1 comment, we summarized the infodemiologists' level 2 reply comments into 6 metrics: the median and IQR of the respective numbers of replies, emoji reactions, and engagements infodemiologists' reply comments received. For comparative purposes, we also summarized the numbers of comments and reply comments and the corresponding 6 metrics for the comments used for benchmarking. The numbers of replies, emoji reactions, and engagements level 1 comments received were compared with the median metrics of matched comments using the Wilcoxon signed rank test. Each metric of infodemiologists' level 2 reply comments (intervention) was benchmarked against the corresponding metric of matched reply comments (control) using the Wilcoxon signed rank test (paired at the level 1 comment level). The number of replies, emoji reactions, and engagements between infodemiologists' level 2 reply comments (intervention) and matched reply comments (control) were further compared using 3 Poisson regression models: treating intervention as a fixed effect with the Huber-White robust SE estimates, as a random effect (nested within location), and as a random effect (nested within location) with the number of page followers/1,000,000 as an offset. The fixed effect design is particularly strong as an internal validity test, as it controls for any confounding, both observed and unobserved, across level 1 posts. For example, level 1 posts differed by geography, timing, and news organization. Limiting comparisons to level 2 intervention and treatment posts nested within the same level 1 post sweeps away any of these concerns. The tradeoff to fixed effects estimation is a relative loss of statistical precision and an inability to characterize level 1 influences; as such, we estimated random effects models as a robustness exercise. The significance level was set at =.05, and no correction was made for multiple testing as this was an exploratory and hypothesis-generating analysis.

Ethical Considerations

This study was deemed exempt from institutional review board (IRB) review by the Salus IRB (#2014) and approved by the Weill Cornell Medicine IRB (20-10022858). To comply with Facebook Terms of Service, all data were manually collected by infodemiologists and manually reviewed for accuracy during a follow-up assessment approximately 10 months later.

Results

A total of 145 interventions were conducted on 132 Facebook news posts, of which 55 interventions (38%) focused on Illinois news sources, 8 (6%) in New Jersey, 45 (31%) in Texas, and 37 (26%) in other states or were nationally oriented. Two-thirds (93/145, 64.14%) of infodemiologist interventions precipitated some form of engagement (either an emoji reaction or a reply comment) from commenters or activated bystanders. In keeping with related literature, comment engagements were right-skewed. Accordingly, we calculated medians and IQRs. The Facebook page for the news organizations on which infodemiologist interventions (level 2 reply comments) were posted had a median of 915,860 (Q1-Q3 range 634,473-2,689,864) page followers. The Facebook news posts received a median of 19 (Q1-Q3 range 7-86) shares, 119 (Q1-Q3 range 37-352) comments, 190 (Q1-Q3 range 73-510) emoji reactions, and 354 (Q1-Q3 range 119-918) engagements, which are the sum of the numbers of comments and emoji reactions received (Table 2).

Table 2. Metrics for Facebook news page and Facebook news post.

	Page followers, median (Q1-Q3)	Shares, median (Q1-Q3)	Comments, median (Q1-Q3)	Reactions, median (Q1-Q3)	Engagements[a], median (Q1-Q3)
Facebook Page	915,860 (634,473-2,689,864)	—[b]	—	—	—
Facebook news post	—	19 (7-86)	119 (37-352)	190 (73-510)	354 (119-918)

[a]Engagements = comments + emoji reactions for Facebook news post.

[b]Not available.

The level 1 comments received a median of 3 replies, 3 reactions, and 7 engagements. The matched comments received a median of 0 replies with a median IQR of 0.75, 1 emoji reaction with a median IQR of 2, and therefore 1.5 engagements with a median IQR of 3.75. Compared to the matched comments, the level 1 comments received more replies, emoji reactions, and engagements (Table 3).

Table 3. Metrics for level 1 comment and matched comment.

	Level 1 comment, median (Q1-Q3)	Matched comment, median (Q1-Q3)	*P* value[a]
Replies			
Median	3 (2-7)	0 (0-0.5)	<.001
IQR	—[b]	0.75 (0-2)	—
Emoji reactions			
Median	3 (1-6)	1 (0-2)	<.001
IQR	—	2 (1-4)	—
Engagements[c]			
Median	7 (3-12)	1.5 (0.5-2.5)	<.001
IQR	—	3.75 (1.75-6)	—

[a]The Wilcoxon signed rank test was used to examine the difference between level 1 comments and matched comments and revealed that level 1 comments received more replies, emoji reactions, and engagements.

[b]Not available.

[c]Engagements = replies + emoji reactions.

In total, infodemiologists made 322 level 2 reply comments, precipitating 189 emoji reactions, of which 151 (79.9%) were positive (141 like, 10 love, and 0 hug emojis), 37 (19.6%) were negative (37 haha and 0 mad emojis), and 1 (0.5%) was neutral (1 wow and 0 sad emojis). The level 2 reply comments received a median of 0 replies with a median IQR of 0, 0 emoji reactions with a median IQR of 0, and 0.5 engagements with a median IQR of 0. The matched reply comments received a median of 0 replies with a median IQR of 0.75, 0.5 emoji reactions with a median IQR of 1, and 1 engagement with a median IQR of 2.5. Compared to the matched reply comments, the level 2 reply comments received fewer and narrower ranges of replies, reactions, and engagements, except for the median comparison for replies (Table 4).

Table 4. Metrics for infodemiologist intervention (level 2 reply comment) and matched reply comment.

	Level 2 reply comment, median (Q1-Q3)	Matched reply comment, median (Q1-Q3)	*P* value[a]	Poisson fixed-effects with robust SE estimates	Poisson random effects	Poisson random effects with the number of page followers/1,000,000 as an offset
Replies				–0.98[b]	–0.98[b]	–0.99[b]
Median	0 (0-0.5)	0 (0-0.25)	.94			
IQR	0 (0-0.5)	0.75 (0-1.75)	<.001			
Emoji reactions				–1.19[b]	–1.21[b]	–1.20[b]
Median	0 (0-0.5)	0.5 (0-1)	<.001			
IQR	0 (0-0.3125)	1 (1-2.5)	<.001			
Engagements [c]				–1.10[b]	–1.11[b]	–1.10[b]
Median	0.5 (0-1)	1 (1-2)	<.001			
IQR	0 (0-0.75)	2.5 (1-4.25)	<.001			

[a]Wilcoxon signed rank test.

[b]*P*<.001.

[c]Engagements = replies + emoji reactions.

The median number of individuals involved in conversation threads with the infodemiologist was 2 with a median IQR of 3. Qualitative evidence of psychological reactance or a backfire effect (assessed by observing if discussion with an infodemiologist appeared to immediately lead to a commenter leaving more extreme comments) was rare, appearing in 1% (2/145) of interventions.

Discussion

The primary purpose of this research was to see if infodemiologist interventions would receive attention in web-based settings. We found that, by and large, they do; however, the evidence for such attention is limited due to Facebook's data collection limitations, which require active engagement and do not provide data on the number of viewers that do not either comment or react using emojis. We also sought to develop a basis for assessing the extent of that attention by comparing engagement with our comments to engagement with comments made by others to the same post (matched comments). More specifically, on average, infodemiologists' interventions (level 2 reply comments) received fewer replies and less overall sentiment (positive, neutral, or negative), as evidenced through native metrics such as "likes," than matched reply comments. We found in this case that we had statistically significantly less engagement than the matched reply comments. Moreover, according to the IQR comparisons, we also observed that the infodemiologists' comments received a narrower range of reactions, replies, and engagements than matched benchmark comments.

This could suggest that our impact was less than that of antivaccination comments, but another, more nuanced interpretation of our results, based on the context in which the interventions were made, is that our reply comments led to a quieting of the conversation rather than stimulating more antivaccination comments. On a highly charged political topic, such as the discourse surrounding COVID-19 vaccines, reducing engagement may be one effective way to reduce the spread of misinformation, even if the impact on participants' and bystanders' beliefs and behavior remains unclear. For example, numerous studies of accuracy nudges demonstrate such interventions reduce intentions to share misleading or false content [80-82].

We recognize that reduced engagement does not necessarily imply agreement. Moreover, there are several potential explanations for why infodemiologists comments received less engagement, including participant boredom, inattention, apathy, undetected or silent backfire effects, or algorithmic downregulation. However, Facebook data limitations preclude us from tracking the comments of either participants or bystanders after infodemiologist interventions to assess whether interventions changed their attitudes about vaccines or led to anyone's decision to subsequently receive a vaccine.

While some engagement is useful or even necessary to algorithmically drive attention to infodemiologist interventions and therefore increase overall views, attempts to drive too much engagement—that is, through the strong emotional reactions of outrage or fear that may be required to drive such metrics—could be detrimental to the tenor of conversations infodemiologists are seeking to have. In addition, high degrees of engagement and the emotional valence such conversations are likely to bring may expose infodemiologists to other risks, such as harassment or doxing. The optimal amount of engagement that balances these 2 competing priorities is not clear.

Considering the optimal amount of engagement alludes to a larger issue about the metrics being used to assess such digital interventions overall. On social media, the metrics most convenient to use are designed for monitoring the impact of brand marketing [83,84]. Such metrics are not conducive to public health evaluations of the dynamics of misinformation [20]. Moreover, as has been previously noted, engagement itself does not necessarily align with efforts to prevent or mitigate the spread of misinformation about science and health. Efforts to stymie the production or spread of misinformation then face a strategic dilemma: maximize engagement through the native metrics made available by social media platforms (and incur the subsequent externalities) or engage in time- and labor-intensive practices of data collection to generate alternate metrics. The former approach implies that the solution to misinformation about science or medicine is simply to make science more engaging; however, as implied by Brandolini's law that the time, effort, and cost of addressing falsehoods are orders of magnitude larger than the resources required to produce them [85], that approach ignores the diversity and ease with which misinformation about science spreads [86].

This study has several limitations. First and foremost are the limits on data accessibility that curtail efforts to fully understand the impact of infodemiology interventions on the digital information environment and other actors in this space. Indeed, various features of technology platforms' algorithms and user interfaces made it challenging to maintain the same intervention strategy over time, follow ongoing interventions, and collect sufficient data at scale. This manifested as well in that Facebook approximates counts of comments and emojis once the number becomes large, leading to similar approximates in our reported data, especially for the number of page followers. Similarly, Facebook algorithms are constantly reordering comments in active threads. Therefore, the benchmarking data, which were collected after the original posting, may represent slightly different results than if the benchmarking were collected at the same time as the intervention itself. However, as the benchmarking is sufficiently broad, covering the comments both above and below the comment to which the infodemiologist replied, we believe it provides a representative "control" against which we can weigh the infodemiologists' interventions.

Moreover, while third-party applications offer insight into "social media marketing metrics," such services are structured to provide data to Facebook landing pages, not to assess engagement metrics of individuals doing multiple interventions across pages hosted by others. Specifically, because we did not own the pages on which we intervened, we could not access metrics on the views of those bystanders who neither commented nor provided emoji reactions through commercial software that permits such monitoring. Additionally, because engagement is defined as the combination of 2 nonmutually exclusive events (ie, emoji reactions and comments), it is possible that it is an overstatement of the number of people engaging in the discussion if someone both commented and expressed an emoji reaction.

Furthermore, it is impossible to separate the effect of the interventions completed by the infodemiologists from the potential impact of how they were viewed by other stakeholders in the discussion. For example, the perception that they may be trusted sources, researchers, or central nodes in a social network could influence how the interventions are received by participants. It is possible that infodemiologists were responding to, or the engagement numbers were inflated by, fake Facebook accounts. Infodemiologists sought to minimize this risk through the selection of which threads to intervene in and by examining a poster's public profile. Additionally, this work does not address the myriad individual, community, and structural factors that lead to disparities in information sharing on social networks [87]. It also does not effectively address how the structure of the social network affects how health information diffuses [88]. Future work will seek to collect sufficient data to assess these potential confounders.

An additional challenge lies in how misinformation is defined. While our infodemiologists sought to respond to posts with factually incorrect material or overtly negative sentiment about the vaccines, those determinations were made based on context and therefore subjective. Moreover, the sheer volume of such material on social media meant infodemiologists could not respond to all such instances in their communities but only a subset.

Our sample was geographically skewed with relatively few New Jersey-focused interventions, we believe due to the overlap of New York and New Jersey media markets. Finally, our pilot prioritized infodemiologist safety and well-being both in the choice of the site of interventions (Facebook posts of news stories) and in the interventions themselves, instructing infodemiologists to avoid highly contentious forums and exit conversations if they felt threatened. It is likely that an explicit focus on engagement would lead to higher engagement statistics, though such approaches would need to be weighed against the externalities (eg, more contentious, emotionally laden content generating fear or outrage or the emotional safety or doxing of the infodemiologist).

Here we described an approach to addressing health-related misinformation derived from evidence on methods to intervene against misinformation about various topics in web-based and offline spaces that were associated with less engagement relative to comparable comments and replies in the same comment threads. More research building on some of the toolkits and frameworks presented here will be needed to further guide research on addressing misinformation in digital communities. We attempted to show that the RE-AIM framework is an effective schema to guide evaluations in this space, even if direct evidence of real-world efficacy is lacking. However, the gap between ideal evaluation metrics and the available data through social media platforms remains wide.

With this in mind, it raises the question of whether the inexorable drive for engagement—a metric prioritized by social media companies, not public health—is a solution to the problem of misinformation or further exacerbates it by not addressing the underlying mechanisms, incentives, and logic by which misinformation spreads. The infodemiology work described here and its impact on reducing the temperature of web-based conversations and avoiding backfire effects raise questions about how much engagement is optimal to improve science communication. More research is needed, of course, to correlate this approach with the effect on participants' and bystanders' subsequent beliefs and behaviors.

In previous reports, we examined our intervention protocol from several perspectives. In a study [30], we found a clear tension between using principles of motivational interviewing and the imperative to limit the amount of misinformation that remains unchecked by facts. Separately, we discussed that infodemiologists adopt several informal roles in web-based discussions, serving both as hosts and translators [22]. In this study, we quantitatively evaluated the impact of these interventions, drawing inspiration from implementation science frameworks as a guide, with the intention of understanding to what extent these interventions attract attention from bystanders. Viewed in combination, our qualitative analyses plus this quantitative assessment provide a novel mixed methods approach to evaluating interventions to address web-based antivaccine sentiment specifically and digital misinformation in general. Such approaches provide a more complete picture of the extent to which interventions based on a blend of motivational interviewing principles and evidence-based interventions focusing on bystanders can be useful in counteracting networked misinformation on web-based platforms. While labor-intensive, such interventions can be one part of a comprehensive strategy to address medical misinformation in digital spaces, along with other evidence-based strategies.

Acknowledgments

This study was supported by the Robert Wood Johnson Foundation (grant numbers 76935 and 78084), Weill Cornell Medicine's JumpStart program, and the Cornell Center for Social Science. The authors acknowledge the helpful insights of Kathleen Hall Jamieson, Oktawia Wojcik, and Nancy Barrand. The opinions expressed in this study do not necessarily reflect those of either the Robert Wood Johnson Foundation or Weill Cornell Medicine.

Conflicts of Interest

None declared.

References

1. Simonov A, Sacher SK, Dubé JPH, Biswas S. The persuasive effect of Fox News: non-compliance with social distancing during the Covid-19 pandemic. National Bureau of Economic Research. 2020. URL: https://www.nber.org/papers/w27237 [accessed 2023-10-05]
2. Hornik R, Kikut A, Jesch E, Woko C, Siegel L, Kim K. Association of COVID-19 misinformation with face mask wearing and social distancing in a nationally representative US sample. Health Commun 2021;36(1):6-14. [doi: 10.1080/10410236.2020.1847437] [Medline: 33225745]

3. 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 [FREE Full text] [doi: 10.1038/s41562-021-01056-1] [Medline: 33547453]
4. Pertwee E, Simas C, Larson HJ. An epidemic of uncertainty: rumors, conspiracy theories and vaccine hesitancy. Nat Med 2022;28(3):456-459 [FREE Full text] [doi: 10.1038/s41591-022-01728-z] [Medline: 35273403]
5. 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;12(1):5966 [FREE Full text] [doi: 10.1038/s41598-022-10070-w] [Medline: 35474313]
6. Enders AM, Uscinski JE, Klofstad C, Stoler J. The different forms of COVID-19 misinformation and their consequences. Harvard Kennedy School (HKS) Misinformation Review 2020:1-21 [FREE Full text] [doi: 10.37016/mr-2020-48]
7. Jolley D, Douglas KM. The effects of anti-vaccine conspiracy theories on vaccination intentions. PLoS One 2014;9(2):e89177 [FREE Full text] [doi: 10.1371/journal.pone.0089177] [Medline: 24586574]
8. Lee SJ, Lee CJ, Hwang H. The impact of COVID-19 misinformation and trust in institutions on preventive behaviors. Health Educ Res 2023;38(1):95-105 [FREE Full text] [doi: 10.1093/her/cyac038] [Medline: 36564938]
9. Murthy VH. Confronting Health Misinformation: The U.S. Surgeon General's Advisory on Building a Healthy Information Environment. Washington, D.C: Office of the US Surgeon General; 2021.
10. Vraga EK, Bode L. Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. Political Commun 2020;37(1):136-144. [doi: 10.1080/10584609.2020.1716500]
11. Chou WYS, Gaysynsky A, Cappella JN. Where we go from here: health misinformation on social media. Am J Public Health 2020;110(S3):S273-S275 [FREE Full text] [doi: 10.2105/AJPH.2020.305905] [Medline: 33001722]
12. Kahan DM. Misconceptions, misinformation, and the logic of identity-protective cognition. SSRN J 2017 [FREE Full text] [doi: 10.2139/ssrn.2973067]
13. Gorman GE, Gorman JM. Denying to the Grave: Why We Ignore the Science That Will Save Us, Revised and Updated Edition. New York, NY: Oxford University Press; 2021.
14. 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 [FREE Full text] [doi: 10.1038/s44159-021-00006-y]
15. Wardle C, Derakhshan H. Information disorder: toward an interdisciplinary framework for research and policy making. Council of Europe Report. 2017. URL: https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html [accessed 2023-10-05]
16. Zarocostas J. How to fight an infodemic. Lancet 2020;395(10225):676 [FREE Full text] [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
17. Purnat TD, Vacca P, Czerniak C, Ball S, Burzo S, Zecchin T, Dubé, et al. Infodemic signal detection during the COVID-19 pandemic: development of a methodology for identifying potential information voids in online conversations. JMIR Infodemiology 2021;1(1):e30971 [FREE Full text] [doi: 10.2196/30971] [Medline: 34447926]
18. Eysenbach G. Infodemiology: the epidemiology of (mis)information. Am J Med 2002;113(9):763-765 [FREE Full text] [doi: 10.1016/s0002-9343(02)01473-0] [Medline: 12517369]
19. 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;4(2) [FREE Full text] [doi: 10.1056/cat.22.0198]
20. Chiou H, Voegeli C, Wilhelm E, Kolis J, Brookmeyer K, Prybylski D. The future of infodemic surveillance as public health surveillance. Emerg Infect Dis 2022;28(13):S121-S128 [FREE Full text] [doi: 10.3201/eid2813.220696] [Medline: 36502389]
21. Simon FM, Camargo CQ. Autopsy of a metaphor: the origins, use and blind spots of the ‘infodemic’. New Media Soc 2021;25(8):2219-2240 [FREE Full text] [doi: 10.1177/14614448211031908]
22. Gorman JM, Scales DA. Leveraging infodemiologists to counteract online misinformation: experience with COVID-19 vaccines. Harvard Kennedy School (HKS) Misinformation Review. 2022. URL: https://misinforeview.hks.harvard.edu/article/leveraging-infodemiologists-to-counteract-online-misinformation-experience-with-covid-19-vaccines/ [accessed 2023-10-05]
23. Seymour B, Gyenes N, Roberts H, Fish SA, Bermejo F, Zuckerman E. Public health, social networks, and the digital media ecosystem: emerging hypotheses. In: Substance Abuse Library and Information Studies. 2017 Presented at: Proceedings of the 39th Annual SALIS/AMHL Conference; May 3-6, 2017; Worcester URL: https://salis.org/wp-content/uploads/2020/08/SALISJournal-vol4_2017.pdf#page=8
24. Getman R, Helmi M, Roberts H, Yansane A, Cutler D, Seymour B. Vaccine hesitancy and online information: the influence of digital networks. Health Educ Behav 2018;45(4):599-606. [doi: 10.1177/1090198117739673] [Medline: 29267129]
25. DiResta R, Lotan G. Anti-vaxxers are using Twitter to manipulate a vaccine bill. Wired. 2015 Jun 08. URL: https://www.wired.com/2015/06/antivaxxers-influencing-legislation/ [accessed 2023-10-05]
26. Young DG, Miller JM. Political communication. In: Sears DO, Levy JS, Jerit J, Huddy L, editors. The Oxford Handbook of Political Psychology, 3rd Edition. New York, NY: Oxford University Press; 2023.

27. Carpiano RM, Callaghan T, DiResta R, Brewer NT, Clinton C, Galvani AP, et al. Confronting the evolution and expansion of anti-vaccine activism in the USA in the COVID-19 era. Lancet 2023;401(10380):967-970 [FREE Full text] [doi: 10.1016/S0140-6736(23)00136-8] [Medline: 36871571]
28. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change. London: Guilford Press; 2012.
29. Gagneur A, Battista MC, Boucher FD, Tapiero B, Quach C, De Wals P, et al. Promoting vaccination in maternity wards—motivational interview technique reduces hesitancy and enhances intention to vaccinate, results from a multicentre non-controlled pre- and post-intervention RCT-nested study, Quebec, march 2014 to february 2015. Euro Surveill 2019;24(36):1800641 [FREE Full text] [doi: 10.2807/1560-7917.ES.2019.24.36.1800641] [Medline: 31507265]
30. Scales D, Gorman JM, DiCaprio P, Hurth L, Radhakrishnan M, Windham S, et al. Community-oriented Motivational Interviewing (MI): a novel framework extending MI to address COVID-19 vaccine misinformation in online social media platforms. Comput Human Behav 2023;141:107609 [FREE Full text] [doi: 10.1016/j.chb.2022.107609] [Medline: 36531901]
31. How to respond to vocal vaccine deniers in public: best practice guidance. World Health Organization, European Regional Office. 2017. URL: https://www.who.int/europe/publications/i/item/WHO-EURO-2017-2899-42657-59427 [accessed 2023-10-05]
32. Omari A, Boone KD, Zhou T, Lu PJ, Kriss JL, Hung MC, et al. Characteristics of the moveable middle: opportunities among adults open to COVID-19 vaccination. Am J Prev Med 2023;64(5):734-741 [FREE Full text] [doi: 10.1016/j.amepre.2022.11.003] [Medline: 36690543]
33. Hassan A, Barber SJ. The effects of repetition frequency on the illusory truth effect. Cogn Res Princ Implic 2021;6(1):38 [FREE Full text] [doi: 10.1186/s41235-021-00301-5] [Medline: 33983553]
34. Vraga EK, Bode L. Using expert sources to correct health misinformation in social media. Sci Commun 2017;39(5):621-645. [doi: 10.1177/1075547017731776]
35. Lewandowsky S. Climate change disinformation and how to combat it. Annu Rev Public Health 2021;42:1-21 [FREE Full text] [doi: 10.1146/annurev-publhealth-090419-102409] [Medline: 33355475]
36. Dong L, Bogart LM, Gandhi P, Aboagye JB, Ryan S, Serwanga R, et al. A qualitative study of COVID-19 vaccine intentions and mistrust in Black Americans: recommendations for vaccine dissemination and uptake. PLoS One 2022;17(5):e0268020 [FREE Full text] [doi: 10.1371/journal.pone.0268020] [Medline: 35503797]
37. DiRusso C, Stansberry K. Unvaxxed: a cultural study of the online anti-vaccination movement. Qual Health Res 2022;32(2):317-329. [doi: 10.1177/10497323211056050] [Medline: 34847821]
38. Gagneur A, Gosselin V, Dubé È. Motivational interviewing: a promising tool to address vaccine hesitancy. Vaccine 2018;36(44):6553-6555. [doi: 10.1016/j.vaccine.2017.10.049] [Medline: 30301523]
39. Scales D, Gorman J, Leff C, Gorman S. Effective ways to combat online medical and scientific misinformation: a hermeneutic narrative review and analysis. mediArXiv. Preprint posted online on February 16, 2021 2021 [FREE Full text] [doi: 10.33767/osf.io/rcugs]
40. Townsend L, Wallace C. Social media research: a guide to ethics. University of Aberdeen. 2016. URL: https://www.gla.ac.uk/media/Media_487729_smxx.pdf [accessed 2023-10-05]
41. Swire-Thompson B, DeGutis J, Lazer D. Searching for the backfire effect: measurement and design considerations. J Appl Res Mem Cogn 2020;9(3):286-299 [FREE Full text] [doi: 10.1016/j.jarmac.2020.06.006] [Medline: 32905023]
42. Sillence E. Seeking out very like-minded others: exploring trust and advice issues in an online health support group. Int J Web Based Communities 2010;6(4):376-394. [doi: 10.1504/ijwbc.2010.035840]
43. Vraga EK, Bode L. I do not believe you: how providing a source corrects health misperceptions across social media platforms. Inf Commun Soc 2017;21(10):1337-1353. [doi: 10.1080/1369118x.2017.1313883]
44. Prochaska JO, Johnson S, Lee P. The transtheoretical model of behavior change. In: Shumaker SA, Ockene JK, Riekert KA, editors. The Handbook of Health Behavior Change, 3rd Edition. New York, NY: Springer Publishing Company; 2009:59-83.
45. Philip JK, Kerr G, Don ES, McColl R, Pals H. The elaboration likelihood model: review, critique and research agenda. Eur J Mark 2014;48(11/12):2033-2050. [doi: 10.1108/ejm-12-2011-0776]
46. Reno JE, O'Leary S, Garrett K, Pyrzanowski J, Lockhart S, Campagna E, et al. Improving provider communication about HPV vaccines for vaccine-hesitant parents through the use of motivational interviewing. J Health Commun 2018;23(4):313-320. [doi: 10.1080/10810730.2018.1442530] [Medline: 29474117]
47. Lemaitre T, Carrier N, Farrands A, Gosselin V, Petit G, Gagneur A. Impact of a vaccination promotion intervention using motivational interview techniques on long-term vaccine coverage: the PromoVac strategy. Hum Vaccin Immunother 2019;15(3):732-739 [FREE Full text] [doi: 10.1080/21645515.2018.1549451] [Medline: 30457421]
48. Vaccine Persona Explainer: How do we get America vaccinated? Surgo Ventures. 2020. URL: https://surgoventures.org/vaccine-persona-explainer [accessed 2021-12-09]
49. Chan MPS, Jones CR, Jamieson KH, 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]

50. Roozenbeek J, van der Linden S, Goldberg B, Rathje S, Lewandowsky S. Psychological inoculation improves resilience against misinformation on social media. Sci Adv 2022;8(34):eabo6254 [FREE Full text] [doi: 10.1126/sciadv.abo6254] [Medline: 36001675]
51. Schmid P, Schwarzer M, Betsch C. Weight-of-evidence strategies to mitigate the influence of messages of science denialism in public discussions. J Cogn 2020;3(1):36 [FREE Full text] [doi: 10.5334/joc.125] [Medline: 33043246]
52. Cook J, Lewandowsky S, Ecker UKH. Neutralizing misinformation through inoculation: exposing misleading argumentation techniques reduces their influence. PLoS One 2017;12(5):e0175799 [FREE Full text] [doi: 10.1371/journal.pone.0175799] [Medline: 28475576]
53. Traberg CS, Roozenbeek J, van der Linden S. Psychological inoculation against misinformation: current evidence and future directions. Ann Am Acad Pol Soc Sci 2022;700(1):136-151 [FREE Full text] [doi: 10.1177/00027162221087936]
54. Stanford Internet Observatory. Memes, magnets and microchips: narrative dynamics around COVID-19 vaccines. Virality Project. 2022. URL: https://fsi.stanford.edu/publication/memes-magnets-and-microchips-narrative-dynamics-around-covid-19-vaccines [accessed 2023-10-05]
55. Stocking G, Matsa KE, Khuzam M. News organizations were most prominent source in COVID-19 content shared in public Facebook spaces. Pew Research Center. 2020. URL: https://www.pewresearch.org/journalism/2020/06/24/news-organizations-were-most-prominent-source-in-covid-19-content-shared-in-public-facebook-spaces/ [accessed 2023-10-05]
56. Stocking G, Matsa KE, Khuzam M. As COVID-19 emerged in U.S., Facebook posts about it appeared in a wide range of public pages, groups. Pew Research Center. 2020. URL: https://www.pewresearch.org/journalism/2020/06/24/as-covid-19-emerged-in-u-s-facebook-posts-about-it-appeared-in-a-wide-range-of-public-pages-groups/ [accessed 2023-10-05]
57. Chou WYS, Budenz A. Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence. Health Commun 2020;35(14):1718-1722. [doi: 10.1080/10410236.2020.1838096] [Medline: 33124475]
58. Penţa MA, Băban A. Message framing in vaccine communication: a systematic review of published literature. Health Commun 2018;33(3):299-314. [doi: 10.1080/10410236.2016.1266574] [Medline: 28059557]
59. Ilgen JS, Eva KW, de Bruin A, Cook DA, Regehr G. Comfort with uncertainty: reframing our conceptions of how clinicians navigate complex clinical situations. Adv Health Sci Educ Theory Pract 2019;24(4):797-809. [doi: 10.1007/s10459-018-9859-5] [Medline: 30390181]
60. Sillence E, Briggs P, Harris P, Fishwick L. A framework for understanding trust factors in web-based health advice. Int J Hum Comput Stud 2006;64(8):697-713 [FREE Full text] [doi: 10.1016/j.ijhcs.2006.02.007]
61. Walter N, Tukachinsky R. A meta-analytic examination of the continued influence of misinformation in the face of correction: how powerful is it, why does it happen, and how to stop it? Commun Res 2019;47(2):155-177. [doi: 10.1177/0093650219854600]
62. Brashier NM, Pennycook G, Berinsky AJ, Rand DG. Timing matters when correcting fake news. Proc Natl Acad Sci U S A 2021;118(5):e2020043118 [FREE Full text] [doi: 10.1073/pnas.2020043118] [Medline: 33495336]
63. Walter N, Murphy ST. 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]
64. Schmid P, Betsch C. Effective strategies for rebutting science denialism in public discussions. Nat Hum Behav 2019;3(9):931-939. [doi: 10.1038/s41562-019-0632-4] [Medline: 31235861]
65. Kahan DM, Jenkins ‐ Smith H, Braman D. Cultural cognition of scientific consensus. J Risk Res 2011;14(2):147-174. [doi: 10.1080/13669877.2010.511246]
66. Pennycook G, Cannon TD, Rand DG. Prior exposure increases perceived accuracy of fake news. J Exp Psychol Gen 2018;147(12):1865-1880 [FREE Full text] [doi: 10.1037/xge0000465] [Medline: 30247057]
67. Ecker UKH, Lewandowsky S, Cheung CSC, Maybery MT. He did it! She did it! No, she did not! Multiple causal explanations and the continued influence of misinformation. J Mem Lang 2015;85:101-115 [FREE Full text] [doi: 10.1016/j.jml.2015.09.002]
68. van der Linden SL, Clarke CE, Maibach EW. Highlighting consensus among medical scientists increases public support for vaccines: evidence from a randomized experiment. BMC Public Health 2015;15:1207 [FREE Full text] [doi: 10.1186/s12889-015-2541-4] [Medline: 26635296]
69. Maibach EW, van der Linden SL. The importance of assessing and communicating scientific consensus. Environ Res Lett 2016;11(9):091003 [FREE Full text] [doi: 10.1088/1748-9326/11/9/091003]
70. Bolsen T, Druckman JN. Counteracting the politicization of science. J Commun 2015;65(5):745-769. [doi: 10.1111/jcom.12171]
71. Epstein Z, Berinsky AJ, Cole R, Gully A, Pennycook G, Rand DG. Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online. Harvard Kennedy School Misinformation Review 2021:1-12 [FREE Full text] [doi: 10.37016/mr-2020-71]
72. Pennycook G, Rand DG. Nudging social media toward accuracy. Ann Am Acad Pol Soc Sci 2022;700(1):152-164 [FREE Full text] [doi: 10.1177/00027162221092342] [Medline: 35558818]

73. Orpin D. Chapter 9. #Vaccineswork: recontextualizing the content of epidemiology reports on Twitter. In: Pérez-Llantada C, Luzón MJ, editors. Science Communication on the Internet: Old Genres Meet New Genres. Amsterdam, the Netherlands: John Benjamins Publishing Company; 2019:173-194.
74. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999;89(9):1322-1327 [FREE Full text] [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
75. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
76. Lane-Fall MB, Curran GM, Beidas RS. Scoping implementation science for the beginner: locating yourself on the "subway line" of translational research. BMC Med Res Methodol 2019;19(1):133 [FREE Full text] [doi: 10.1186/s12874-019-0783-z] [Medline: 31253099]
77. Platt T, Platt J, Thiel DB, Kardia SLR. Facebook advertising across an engagement spectrum: a case example for public health communication. JMIR Public Health Surveill 2016;2(1):e27 [FREE Full text] [doi: 10.2196/publichealth.5623] [Medline: 27244774]
78. Kite J, Grunseit A, Li V, Vineburg J, Berton N, Bauman A, et al. Generating engagement on the make healthy normal campaign Facebook page: analysis of Facebook analytics. JMIR Public Health Surveill 2019;5(1):e11132 [FREE Full text] [doi: 10.2196/11132] [Medline: 31344679]
79. Benis A, Khodos A, Ran S, Levner E, Ashkenazi S. Social media engagement and influenza vaccination during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2021;23(3):e25977 [FREE Full text] [doi: 10.2196/25977] [Medline: 33651709]
80. 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;31(7):770-780 [FREE Full text] [doi: 10.1177/0956797620939054] [Medline: 32603243]
81. Pennycook G, Epstein Z, Mosleh M, Arechar AA, Eckles D, Rand DG. Shifting attention to accuracy can reduce misinformation online. Nature 2021;592(7855):590-595 [FREE Full text] [doi: 10.1038/s41586-021-03344-2] [Medline: 33731933]
82. Fazio L. Pausing to consider why a headline is true or false can help reduce the sharing of false news. Harvard Kennedy School Misinformation Review. 2020. URL: https://misinforeview.hks.harvard.edu/article/pausing-reduce-false-news/ [accessed 2023-10-05]
83. Waszak PM, Kasprzycka-Waszak W, Kubanek A. The spread of medical fake news in social media—the pilot quantitative study. Health Policy Technol 2018;7(2):115-118 [FREE Full text] [doi: 10.1016/j.hlpt.2018.03.002]
84. Liadeli G, Sotgiu F, Verlegh PWJ. A meta-analysis of the effects of brands‘ owned social media on social media engagement and sales. J Mark 2022;87(3):406-427 [FREE Full text] [doi: 10.1177/00222429221123250]
85. Williamson P. Take the time and effort to correct misinformation. Nature 2016;540(7632):171-171 [FREE Full text] [doi: 10.1038/540171a]
86. Allchin D. Ten competencies for the science misinformation crisis. Sc Edu 2022;107(2):261-274. [doi: 10.1002/sce.21746]
87. Southwell BG. Social Networks and Popular Understanding of Science and Health: Sharing Disparities. Baltimore, MD: Johns Hopkins University Press; 2013.
88. Zhang J, Centola D. Social networks and health: new developments in diffusion, online and offline. Annu Rev Sociol 2019;45(1):91-109 [FREE Full text] [doi: 10.1146/annurev-soc-073117-041421]

Abbreviations

CDC: Centers for Disease Control and Prevention
IRB: institutional review board
MI: motivational interviewing
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance
WHO: World Health Organization

Edited by T Mackey; submitted 20.06.23; peer-reviewed by B Southwell, E Wilhelm, F Medina; comments to author 10.08.23; revised version received 16.09.23; accepted 30.09.23; published 14.11.23.

Please cite as:
Scales D, Hurth L, Xi W, Gorman S, Radhakrishnan M, Windham S, Akunne A, Florman J, Leininger L, Gorman J
Addressing Antivaccine Sentiment on Public Social Media Forums Through Web-Based Conversations Based on Motivational Interviewing Techniques: Observational Study
JMIR Infodemiology 2023;3:e50138
URL: https://infodemiology.jmir.org/2023/1/e50138
doi:10.2196/50138
PMID:37962940

Original Paper

Using COVID-19 Vaccine Attitudes Found in Tweets to Predict Vaccine Perceptions in Traditional Surveys: Infodemiology Study

Nekabari Sigalo[1], BSc, MPH, PhD; Vanessa Frias-Martinez[1], BSc, MSc, PhD

College of Information Studies, University of Maryland, College Park, MD, United States

Corresponding Author:
Nekabari Sigalo, BSc, MPH, PhD
College of Information Studies
University of Maryland
4130 Campus Dr
College Park, MD, 20742
United States
Phone: 1 301 405 2033
Email: nsigalo@terpmail.umd.edu

Abstract

Background: Traditionally, surveys are conducted to answer questions related to public health but can be costly to execute. However, the information that researchers aim to extract from surveys could potentially be retrieved from social media, which possesses data that are highly accessible and lower in cost to collect.

Objective: This study aims to evaluate whether attitudes toward COVID-19 vaccines collected from the Household Pulse Survey (HPS) could be predicted using attitudes extracted from Twitter (subsequently rebranded X). Ultimately, this study aimed to determine whether Twitter can provide us with similar information to that observed in traditional surveys or whether saving money comes at the cost of losing rich data.

Methods: COVID-19 vaccine attitudes were extracted from the HPS conducted between January 6 and May 25, 2021. Twitter's streaming application programming interface was used to collect COVID-19 vaccine tweets during the same period. A sentiment and emotion analysis of tweets was conducted to examine attitudes toward the COVID-19 vaccine on Twitter. Generalized linear models and generalized linear mixed models were used to evaluate the ability of COVID-19 vaccine attitudes on Twitter to predict vaccine attitudes in the HPS.

Results: The results revealed that vaccine perceptions expressed on Twitter performed well in predicting vaccine perceptions in the survey.

Conclusions: These findings suggest that the information researchers aim to extract from surveys could potentially also be retrieved from a more accessible data source, such as Twitter. Leveraging Twitter data alongside traditional surveys can provide a more comprehensive and nuanced understanding of COVID-19 vaccine perceptions, facilitating evidence-based decision-making and tailored public health strategies.

(JMIR Infodemiology 2023;3:e43700) doi:10.2196/43700

KEYWORDS

social media; Twitter; COVID-19; vaccine; surveys; SARS-CoV-2; vaccinations; hesitancy

Introduction

Background

The implementation of successful COVID-19 vaccine rollouts is essential for COVID-19 to remain under control globally. Although vaccines are essential in the global battle against COVID-19, vaccine hesitancy continues to be a barrier to effective and consistent vaccine rollout programs. According to the US Census Bureau's Household Pulse Survey (HPS), individuals who reported being hesitant about receiving a COVID-19 vaccine cited concerns about side effects, safety, and a lack of trust in the vaccine or the government [1]. Although the number of vaccine-hesitant individuals continues to decline, the fact that vaccine hesitancy still exists interferes with infection control through vaccination.

Vaccine hesitancy has been fueled in part by the spread of vaccine misinformation, both in the media and on the web. In fact, the COVID-19 vaccine became a popular topic of discussion among social media users, with many individuals expressing their concerns about taking the vaccine on social media platforms [2]. Amid the new normal of self-quarantine and lockdown, Twitter (subsequently rebranded X) quickly emerged as an important means of COVID-19 communications and discussion [3]. This is in part due to the real-time availability of social media messaging, compared with traditional news reporting methods [4,5]. Twitter users often not only take to the platform to announce their own experiences and opinions about the pandemic but also see Twitter as a source of up-to-date information about the pandemic [6].

The COVID-19 vaccine conversation on social media platforms has been both beneficial and detrimental to vaccination efforts across the world [7]. Although the exact effect of social media on this unprecedented pandemic is difficult to quantify, there has been a constant battle between facts and misinformation, trust and fearmongering, and hope and anger [8]. Research has shown that social media use plays a role in the low acceptance of vaccines [9,10]. Therefore, studying the public COVID-19 vaccine–related discussion on social media can help researchers better understand attitudes related to the vaccine [9].

Traditionally, surveys are conducted to understand attitudes related to public health. For example, many studies leveraged surveys to examine COVID-19 vaccine hesitancy and compliance. In April 2020, Ward et al [11] administered 4 web-based, nationally representative surveys to adults in France to identify the reasons why individuals would or would not take the COVID-19 vaccine once it became available. Researchers found that nearly a quarter of the respondents refused to take the COVID-19 vaccine once it was made available to them, citing reasons such as not trusting vaccines in general; concerns about the expedited vaccine development process; and a lack of fear of COVID-19, deeming the vaccine unnecessary.

In the study by Wang et al [12], a cross-sectional, self-administered survey was conducted to evaluate the COVID-19 vaccine intent among nurses in Hong Kong, China. Researchers found higher rates of vaccine hesitancy compared with vaccine acceptance, with nurses citing concerns about the safety and efficacy of the vaccines. A web-based survey conducted in the study by Ruiz and Bell [13] attempted to identify the predictors of the intent to vaccinate against COVID-19 among Americans. Nearly 40% of the survey respondents refused to be vaccinated against COVID-19. Among the survey respondents, male, older, White, and married individuals and individuals with higher socioeconomic status were more likely to be vaccinated against COVID-19. Researchers also found that Republicans and Fox News viewers were less likely to get vaccinated, whereas individuals who were previously vaccinated for influenza were more likely to be vaccinated for COVID-19.

Although several studies have examined COVID-19 vaccine attitudes through surveys, to our knowledge, no study has evaluated the ability of Twitter, a newer data source, to predict the attitudes reflected in traditionally collected surveys, such as the HPS. In recent years, researchers have looked at social media as a data source, citing the availability of more readily available data and no- or low-cost data collection efforts [14,15]. Traditional mail, field, and telephone interviewer surveys come with high costs of administration, and even though web-based surveys eliminate the costs of postage, paper, printing, and data entry, the newer web-based survey services may still cost up to thousands of dollars for 1 survey [16]. Although relatively inexpensive compared with traditional surveys, web-based surveys are not always cost-effective [16]. Evaluating the ability of information extracted from social media to predict information found in traditional surveys would suggest whether researchers may use this more cost-effective data source to provide similar rich information to that seen in traditional surveys or whether saving money comes at the cost of losing rich data.

Study Overview

The main objective of this study was to examine whether aggregate attitudes extracted from social media can predict vaccine attitudes collected via surveys. We hypothesized that social media may contain attitudes similar to those found in traditional surveys, with the added benefits of more readily available data and no- or low-cost data collection efforts. Predictive models of vaccine attitudes at the metropolitan level can be useful for 2 purposes. First, predictions can be used to identify metropolitan areas where vaccine hesitancy is high and create targeted campaigns to increase vaccination. Second, the relationships between sentiments and emotions and vaccine attitudes can be used to understand human perceptions of vaccines and create effective social media messages for vaccination campaigns. Specifically, we hypothesize that there is a direct, positive relationship between (1) positive sentiments and emotions found in Twitter data and the HPS and (2) negative sentiments and emotions found in Twitter data and the HPS survey.

Methods

Data Collection and Preprocessing

HPS Data

In April 2020, the US Census Bureau began releasing a cross-sectional nationally representative survey, the HPS, in an effort to assess the social and economic impacts of the COVID-19 pandemic on American households [17]. The data from this survey were made publicly available in near real time with the purpose of informing federal and state response and recovery planning [18]. The HPS sample was selected using a stratified random sampling method [19]. Data were collected via computer-assisted telephone interviewing, a data collection method in which surveyors use computer software to conduct telephone interviews with respondents [19].

On January 6, 2021, the US Census Bureau added COVID-19 vaccine–related questions to the HPS with the goal of understanding the factors contributing to vaccine hesitancy and compliance among Americans [20] (Table 1). These questions assessed COVID-19 vaccine receipt, whether respondents received or planned to receive all required doses, intentions to get vaccinated, and reasons why respondents refused to get vaccinated.

Table 1. COVID-19–related Household Pulse Survey questions.

Condition	Question	Responses
Age >18 years	Have you received a COVID-19 vaccine?	(1) Yes (2) No
Answered "yes" to "have you received a COVID-19 vaccine?"	Did you receive (or do you plan to receive) all required doses?	(1) Yes (2) No
Answered "no" to "have you received a COVID-19 vaccine?"	Once a vaccine to prevent COVID-19 is available to you, would you...	(1) Definitely get a vaccine (2) Probably get a vaccine (3) Be unsure about getting a vaccine (4) Probably not get a vaccine (5) Definitely not get a vaccine

Measures of vaccine compliance and hesitancy were assessed for each survey wave overall and in the metropolitan areas in Textbox 1. At the start of the survey period, January 2021, vaccine rollout in the United States had just begun, and with most people unvaccinated at that point, the intent to vaccinate was the only option. For the purposes of this analysis, individuals who answered that they would "definitely get a vaccine" or "probably get a vaccine" once available were considered *vaccine compliant*, and individuals who answered that they would "probably not get a vaccine," or "definitely not get a vaccine" once available were considered *vaccine hesitant*.

The HPS refers to the data collection cycles as *weeks* for consistency with earlier phases, even though the cycles actually span a 2-week collection period. For this study, we used the HPS microdata from weeks 22 to 30, which were collected between January 6 and May 25, 2021, with response rates ranging from 6.4% to 7.5% (Table 2).

Textbox 1. Targeted metropolitan areas for data collection (January to May 2021).

- New York–Newark–Jersey City, New York–New Jersey–Pennsylvania metropolitan area
- Los Angeles–Long Beach–Anaheim, California metropolitan area
- Chicago-Naperville-Elgin, Illinois-Indiana-Wisconsin metropolitan area
- Dallas–Fort Worth–Arlington, Texas metropolitan area
- Houston–The Woodlands–Sugar Land, Texas metropolitan area
- Washington-Arlington-Alexandria, District of Columbia–Virginia–Maryland–West Virginia metropolitan area
- Miami–Fort Lauderdale–Pompano Beach, FL metropolitan area
- Philadelphia-Camden-Wilmington, Pennsylvania–New Jersey–Delaware–Maryland metropolitan area
- Atlanta–Sandy Springs–Alpharetta, Georgia metropolitan area
- Phoenix-Mesa-Chandler, Arizona metropolitan area
- Boston-Cambridge-Newton, Massachusetts–New Hampshire metropolitan area
- San Francisco–Oakland–Berkeley, California metropolitan area
- Detroit-Warren-Dearborn, Michigan metropolitan area
- Seattle-Tacoma-Bellevue, Washington metropolitan area

Table 2. Household Pulse Survey data collection schedule.

Collection dates	Week	Response rate (%)
January 6-January 18, 2021	22	6.4
January 20-February 1, 2021	23	7.5
February 3-February 15, 2021	24	7.3
February 17-March 1, 2021	25	7.3
March 3-March 15, 2021	26	7.4
March 17-March 29, 2021	27	7.2
April 14-April 26, 2021	28	6.6
April 28-May 10, 2021	29	7.4
May 12-May 24, 2021	30	6.8

Twitter Data

To align with the HPS data collection period outlined in Table 2, the Twitter Streaming application programming interface, which provides access to a random sample of 1% of publicly available tweets, was used to collect tweets from the metropolitan areas represented in the HPS (Textbox 1) from January to May 2021. All tweets had *place* information (usually city and state). The place information found in tweets was used to determine the metropolitan area associated with each tweet. Next, to extract tweets related to COVID-19 vaccines, tweets were further filtered by matching variations of vaccine-related keywords, such as *vaccine*, *pfizer*, *moderna*, *johnson & johnson*, and *dose*. The tweet sample was further preprocessed to minimize *noise* resulting from tweets that matched our vaccine-related keywords but did not necessarily reflect the thoughts and opinions of individual Twitter users. For example, companies often promote job postings and advertisements on Twitter using targeted hashtags in hopes of reaching their target audience. To prevent these tweets from adding noise to the sample, tweets related to job postings and advertisements were removed by excluding tweets with hashtags and keywords such as *#jobs*, *#hiring*, and *#ad*.

Sentiment and Emotion Analysis of Tweets

To capture the attitudes found in COVID-19 vaccine–related tweets, a sentiment and emotion analysis of all tweets was conducted using the Natural Language Understanding Research Consortium (NRC) lexicon from the *Syuzhet* package in R (R Core Team) [21]. The NRC lexicon, developed by Saif Mohammad, contains a list of manually labeled English words and their associations with negative and positive sentiments and common human emotions, such as trust, fear, sadness, surprise, and disgust [22]. The *Syuzhet* package applies the NRC lexicon by independently evaluating and rating each word or expression within a tweet [23]. The *get_nrc_sentiment* function was applied to all tweets to calculate the valence of 8 different emotions (fear, joy, anticipation, anger, disgust, sadness, surprise, and trust), along with the overall positive and negative sentiments, toward the COVID-19 vaccine. To assess the accuracy of the sentiment classifier, a random sample of 1000 tweets was selected for manual classification as having a positive or negative sentiment. Among the 1000 tweets in the random sample, 734 (73.4%) were accurately classified by the automated sentiment classifier.

The percentage of tweets expressing the 8 emotions, along with the percentage of tweets expressing a positive or negative sentiment, was calculated at the metropolitan level. For the purposes of this analysis, we used the proportion of tweets with a positive sentiment and positive emotions toward vaccines as a proxy to capture vaccine compliance among Twitter users, and the proportion of tweets with a negative sentiment and negative emotions toward vaccines was used as a proxy to capture vaccine hesitancy among Twitter users.

Data Analysis

Statistical analysis was conducted using the R software packages *betareg* and *GLMMadaptive* [24,25]. To determine whether COVID-19 vaccine attitudes on Twitter can predict the proportion of COVID-19 vaccine perceptions ultimately expressed in the HPS (unweighted), both generalized linear models (GLMs) and generalized linear mixed models (GLMMs) were constructed (Table 3). The models were developed using a total of 126 data points, including proportional vaccine compliance and hesitancy proxies from 14 metropolitan areas across 9 survey waves.

Table 3. Regression models evaluating the relationship between Twitter sentiments and emotions and HPS[a] vaccine hesitancy and compliance.

Model	Features	Outcome
Model 1a (GLM[b])	• Percentage of positive sentiment • Percentage of joy • Percentage of surprise • Percentage of trust • Percentage of anticipation • Percentage of survey week (fixed effect) • Percentage of metropolitan area (fixed effect)	Percentage of vaccine-compliant HPS respondents
Model 1b (GLMM[c])	• Percentage of positive sentiment • Percentage of joy • Percentage of surprise • Percentage of trust • Percentage of anticipation • Percentage of percentage of survey week (fixed effect) • Percentage of metropolitan area (random effect)	Percentage of vaccine-compliant HPS respondents
Model 2a (GLM)	• Percentage of negative sentiment • Percentage of anger • Percentage of disgust • Percentage of sadness • Percentage of fear • Percentage of anticipation • Survey week (fixed effect) • Metropolitan area (fixed effect)	Percentage of vaccine-hesitant HPS respondents
Model 2b (GLMM)	• Percentage of negative sentiment • Percentage of anger • Percentage of disgust • Percentage of sadness • Percentage of fear • Percentage of anticipation • Survey week (fixed effect) • Metropolitan area (random effect)	Percentage of vaccine-hesitant HPS respondents

[a]HPS: Household Pulse Survey.

[b]GLM: generalized linear model.

[c]GLMM: generalized linear mixed model.

GLMs were implemented with both time and geographic location as fixed effects, whereas the GLMMs were a multilevel approach with time as a fixed effect and metropolitan area as a random effect. The main reason behind this dual modeling choice is that random effects can capture the latent variation in the data that cannot be explained by fixed effects or the error term [26]. Random effects represent factors with multiple levels, such as geographic location, and possess distinct components that vary across these levels. Random effects prove especially valuable when dealing with hierarchical or nested data structures, where observations are not independent but grouped at a higher level, and enable us to account for similarities within these groups and prevent overfitting [25]. However, if there is limited variability across locations, including location as a random effect may lead to unstable estimates or unreliable inferences [27]. In such cases, it is often better to treat location as a fixed effect or aggregate the data at a higher level. Thus, we constructed both GLMs and GLMMs to assess the prediction power of using metropolitan areas as either fixed or random effects. We fit beta regression models with the logit link, which is the most appropriate for modeling proportional data [28]. In beta regression, the outcome variable is assumed to follow a beta distribution. Prior to evaluating the models, we conducted assumption checks and checked for multicollinearity and outliers to determine whether the necessary conditions were met. These model diagnostics are presented in Multimedia Appendix 1.

As shown in Table 3, we constructed 2 models to predict vaccine compliance (models 1a, GLM, and 1b, GLMM) and 2 models to predict vaccine hesitancy (models 2a, GLM, and 2b, GLMM). In model 1a, we fit a GLM in which the predictor variables were each of the 5 positive Twitter-derived sentiment and emotion features and the outcome variable was the proportion of vaccine-compliant HPS respondents. This model controlled for survey week (time) and metropolitan area as fixed effects. In model 1b, to account for variations in time and location, we fit a GLMM with each of the 5 positive Twitter-derived sentiment and emotion features and survey week (time) as fixed effects and metropolitan area as a random effect.

By contrast, in model 2a, we fit a GLM in which the predictor variables were each of the 6 negative Twitter-derived sentiment and emotion features and the outcome variable was the proportion of vaccine-hesitant HPS respondents. This model controlled for survey week (time) and metropolitan area as fixed effects. In model 2b, we fit a GLMM with each of the 6 negative Twitter-derived sentiment and emotion features and survey week (time) as fixed effects and metropolitan area as a random effect. As anticipation can be perceived as both positive and negative, this emotion was included as a feature in all models.

Ethical Considerations

This project does not meet the definition of human participant research under the purview of the University of Maryland Institutional Review Board according to federal regulations, section 45CFR46.102(e) [29].

Results

Descriptive Statistics

There were a total of 92,453 tweets from 32,645 users across the 14 metropolitan areas in this study (Table 4). The Los Angeles–Long Beach–Anaheim metropolitan area had the largest representation of tweets (21,500/92,453, 23.26%), whereas the New York–Newark–New Jersey metropolitan area had the largest representation of users (18,400/32,645, 56.36%). The maximum number of tweets by a single individual was 274 (from a user in the New York–Newark–New Jersey metropolitan area). There were a total of 240,242 respondents to the HPS across the 14 metropolitan areas and 9 waves in this study, with the largest sample being the sample from the Washington-Arlington-Alexandria metropolitan area (Table 5).

Table 4. Number of tweets (N=92,453) and users (N=32,645) by metropolitan area (January to May 2021).

Metropolitan area, state	Tweets, n (%)	Users, n (%)	Weekly number of tweets, mean (SD)	Weekly number of users, mean (SD)
Atlanta–Sandy Springs–Alpharetta, Georgia	4234 (4.58)	1542 (4.72)	470 (186)	254 (106)
Boston-Cambridge-Newton, Massachusetts–New Hampshire	3019 (3.27)	1298 (3.98)	335 (133)	218 (83)
Chicago-Naperville-Elgin, Illinois-Indiana-Wisconsin	5821 (6.3)	2561 (7.84)	647 (252)	426 (160)
Dallas–Fort Worth–Arlington, Texas	6203 (6.71)	2299 (7.04)	689 (265)	371 (133)
Detroit-Warren-Dearborn-Michigan	1082 (1.17)	518 (1.59)	120 (56)	84 (40)
Houston–The Woodlands–Sugar Land, Texas	5125 (5.54)	2421 (7.42)	569 (234)	388 (145)
Los Angeles–Long Beach–Anaheim, California	21,500 (23.26)	5429 (16.63)	2389 (983)	891 (344)
Miami–Fort Lauderdale–Pompano Beach, Florida	1954 (2.11)	849 (2.6)	217 (74)	131 (40)
New York–Newark–Jersey City, New York–New Jersey–Pennsylvania	18,400 (19.9)	7259 (22.24)	2044 (683)	1272 (400)
Philadelphia-Camden-Wilmington, Pennsylvania–New Jersey–Delaware-Maryland	3652 (3.95)	1406 (4.31)	406 (156)	250 (88)
Phoenix-Mesa-Chandler, Arizona	4778 (5.17)	1573 (4.82)	531 (183)	260 (81)
San Francisco–Oakland–Berkeley, California	6376 (6.9)	2008 (6.15)	708 (261)	347 (116)
Seattle-Tacoma-Bellevue, Washington	3089 (3.34)	1333 (4.08)	343 (157)	227 (103)
Washington-Arlington-Alexandria, District of Columbia–Virginia–Maryland–West Virginia	7220 (7.81)	2419 (7.41)	802 (313)	436 (155)

Table 5. Number of survey respondents (N=240,242) by city.

Metropolitan area, state	Respondents, n (%)	Weekly number of respondents, mean (SD)
Atlanta–Sandy Springs–Alpharetta, Georgia	12,611 (5.25)	1261 (48)
Boston-Cambridge-Newton, Massachusetts–New Hampshire	20,078 (8.36)	2008 (121)
Chicago-Naperville-Elgin, Illinois-Indiana-Wisconsin	16,044 (6.68)	1604 (89)
Dallas–Fort Worth–Arlington, Texas	15,859 (6.6)	1586 (88)
Detroit-Warren-Dearborn-Michigan	12,149 (5.06)	1215 (88)
Houston–The Woodlands–Sugar Land, Texas	14,179 (5.9)	1418 (125)
Los Angeles–Long Beach–Anaheim, California	17,006 (7.08)	1701 (101)
Miami–Fort Lauderdale–Pompano Beach, Florida	11,641 (4.85)	1164 (67)
New York–Newark–Jersey City, New York–New Jersey–Pennsylvania	19,730 (8.21)	1973 (124)
Philadelphia-Camden-Wilmington, Pennsylvania–New Jersey–Delaware–Maryland	20,240 (8.42)	2024 (162)
Phoenix-Mesa-Chandler, Arizona	14,027 (5.84)	1403 (106)
San Francisco–Oakland–Berkeley, California	17,787 (7.4)	1779 (78)
Seattle-Tacoma-Bellevue, Washington	18,615 (7.75)	1862 (106)
Washington-Arlington-Alexandria, District of Columbia–Virginia–Maryland–West Virginia	30,276 (12.6)	3028 (227)

Attitudes Toward COVID-19 Vaccines in Twitter Data

A sentiment analysis classified most tweets (50,415/92,453, 54.53% of tweets overall) across all metropolitan areas as having a positive sentiment (Table 6). The Washington-Arlington-Alexandria metropolitan area had the largest proportion of tweets with a positive sentiment (53,715/92,453, 58.1%), whereas the Miami–Fort Lauderdale–Pompano Beach metropolitan area had the lowest proportion of tweets with a positive sentiment (47,059/92,453, 50.9%). Tweets with a negative sentiment held the smallest proportions across all metropolitan areas (13,970/92,453, 15.11% of tweets overall). The Los Angeles–Long Beach–Anaheim metropolitan area had the largest proportion of tweets with a negative sentiment (15,162/92,453, 16.4%), whereas the Miami–Fort Lauderdale–Pompano Beach metropolitan area had the lowest proportion of tweets with a negative sentiment (11,926/92,453, 12.9%).

Table 6. Distribution of sentiments and emotions found in COVID-19 vaccine tweets (N=92,453; January to May 2021).

Sentiment or emotion	Tweets, n (%)
Positive sentiment	50,415 (54.53)
Trust	41,317 (44.69)
Anticipation	32,127 (34.75)
Fear	27,227 (29.45)
Sadness	24,935 (26.97)
Joy	24,241 (26.22)
Anger	21,671 (23.44)
Surprise	20,562 (22.24)
Disgust	14,746 (15.95)
Negative	13,970 (15.11)

The emotion analysis revealed trust as the predominantly expressed emotion in COVID-19 vaccine tweets across all metropolitan areas (41,317/92,453, 44.69%). The most perceived negative emotion across all metropolitan areas was fear (27,227/92,453, 29.45%). The least perceived positive emotions were joy (24,241/92,453, 26.22%) and surprise (20,562/92,453, 22.24%), whereas the least perceived negative emotions were anger (21,671/92,453, 23.44%) and disgust (14,746/92,453, 15.95%). Examples of tweets expressing positive, neutral, and negative sentiments are presented in Textbox 2.

Textbox 2. Examples of tweets expressing a positive or negative sentiment toward COVID-19 vaccines.

Positive sentiments

- "Feeling blessed to be healthy this birthday. My two biggest presents are coming in the next week: Inauguration and my second vaccine."
- "Hubby received his first vaccine does this morning-the sense of relief is for real, folks. #vaccinated"
- "With my granddaughter Aurora, Andy, and Elliot. I can see them again and give them a hug now that I am fully Covid 19 vaccinated. I have had both shots plus over 2 weeks since shot two. Thank you President Biden."
- "My mom gets her second dose Sunday, big relief!"
- "I love so much that I got vaccinated today."
- "Proud to work for you @bswhealth-my parents received their COVID vaccines this week at BUMC and said it was so quick and easy and the staff were so friendly! Thank you for taking care of them."

Negative sentiments

- "This is from the Pfizer v-a-c-c-i-n-e. Please understand these shots cause harm. Injury is REAL & not rare. It's a shame these poor people are being gaslighted, & media giants are censoring them."
- "They way my people been bugging me about this d*mn vaccine, I'm not getting that s*it."
- "No way!! No more lockdowns!! No vaccines!!! Oh and if your so concerned about the virus how about no illegals!!! Thank goodness for New Hampshire and Florida!! Go out."
- Clearly you are ignorant of the fact that they said even if you get the vaccine you still have to wear a mask, social distance & deal with all the same bull shit draconian orders. Even after blatant evidence you still want to get it. Heres 100% evidence of brain wash mind control."
- "I am 80. You can have my vaccine. I refuse to get one. I take 2 grams of vitamin C hourly. That makes me IMMUNE. Read: Linus Pauling. No mask. I am out every day working & walking in the park. Paul Kangas 4 Governor."
- "F the stupid vaccine."

Attitudes Toward COVID-19 Vaccines in the HPS Data

Most survey respondents (127,833/240,242, 53.21%) across all metropolitan areas indicated that they received a COVID-19 vaccine, ranging from 50.2% (7041/14,027) of the survey respondents in the Phoenix-Mesa-Chandler metropolitan area to 56.4% (10,032/17,787) of the survey respondents in the San Francisco–Oakland–Berkeley metropolitan area (Table 7). Among the respondents who indicated that they received a COVID-19 vaccine, the majority (65,195/127,833, 51%) also indicated that they received or planned to receive all required doses. Among the respondents who indicated that they had not received a COVID-19 vaccine, the majority (89,759/112,409, 79.85% combined) indicated that they *probably* or *definitely* would get vaccinated), ranging from 48% (6733/14,027) of the survey respondents in the Phoenix-Mesa-Chandler metropolitan area to 75.2% (13,376/17,787) of the survey respondents in the San Francisco–Oakland–Berkeley metropolitan area. For the purposes of this analysis, individuals who answered they that would "definitely get a vaccine" or "probably get a vaccine" once available were considered *vaccine compliant*, and individuals who answered that they would "probably not get a vaccine" or "definitely not get a vaccine" once available were considered *vaccine hesitant*.

Table 7. Distribution of survey responses (unweighted; January to May 2021).

Question and responses	Respondents, n (%)
Have you received a COVID-19 vaccine? (N= 240,242)	
Yes	127,833 (53.21)
No or did not answer	112,409 (46.79)
Did you receive (or do you plan to receive) all required doses? (n=127,833)	
Yes	65,195 (51)
No or did not answer	62,638 (49)
Once a vaccine to prevent COVID-19 is available to you, would you... (n=112,409)	
Definitely get a vaccine	69,233 (61.59)
Probably get a vaccine	20,526 (18.26)
Be unsure about getting a vaccine	3114 (2.77)
Probably not get a vaccine	10,836 (9.64)
Definitely not get a vaccine	8700 (7.74)

Predicting HPS Vaccine Attitudes Using Twitter-Based Attitudes

We evaluated the performance of each GLM in terms of R-squared value and root mean square error (RMSE). Model 1a revealed significant associations (*P*<.001) between the percentage of vaccine-compliant HPS respondents and the percentage of tweets expressing a positive sentiment and trust (Table 8). The R-squared value for the vaccine-compliant GLM (model 1a) was 94.11%, and the RMSE was 0.053, which suggests that we can predict vaccine compliance in the HPS fairly well using positive sentiments and emotions found on Twitter. The GLM coefficients showed that an increase in the percentage of tweets expressing a positive sentiment (*P*<.001) was significantly associated with an increase in the percentage of vaccine-compliant HPS respondents. By contrast, an increase in the percentage of tweets expressing trust (*P*<.001) was significantly associated with a decrease in the percentage of vaccine-compliant HPS respondents.

Table 8. Model results.

Model and features	β coefficient (SE)	*P* value	R-squared (%)	Root mean square error
Model 1a (GLM[a])			94.1	0.053
Percentage of positive sentiment	5.007 (.865)	<.001[b]		
Percentage of joy	.043 (1.482)	.98		
Percentage of surprise	1.084 (1.163)	.35		
Percentage of trust	−4.696 (.865)	<.001[b]		
Percentage of anticipation	.930 (1.197)	.44		
Model 1b (GLMM[c])			82.5	0.062
Percentage of positive sentiment	4.791 (.865)	<.001[b]		
Percentage of joy	.271 (1.483)	.86		
Percentage of surprise	.942 (1.157)	.42		
Percentage of trust	−4.529 (.906)	<.001[b]		
Percentage of anticipation	1.239 (1.107)	.26		
Model 2a (GLM)			93.2	0.01
Percentage of negative sentiment	−1.340 (.581)	.02[b]		
Percentage of anger	.382 (.556)	.49		
Percentage of disgust	−.356 (.612)	.56		
Percentage of sadness	−1.011 (.625)	.11		
Percentage of fear	.715 (.532)	.18		
Percentage of anticipation	−.382 (.369)	.30		
Model 2b (GLMM)			9.4	0.032
Percentage of negative sentiment	−1.334 (.617)	.03[b]		
Percentage of anger	.425 (.587)	.47		
Percentage of disgust	−.312 (.657)	.64		
Percentage of sadness	−1.015 (.662)	.13		
Percentage of fear	.643 (.567)	.26		
Percentage of anticipation	−.456 (.388)	.24		

[a]GLM: generalized linear model.

[b]Statistically significant results (α=.05).

[c]GLMM: generalized linear mixed model.

When we compared the vaccine-compliant GLM (model 1a) with the vaccine-compliant GLMM with metropolitan area as a random effect (model 1b), we observed a lower R-squared value (82.5%) and higher RMSE (0.062). Model 1b results also showed that an increase in the percentage of tweets expressing a positive sentiment (*P*<.001) was significantly associated with an increase in the percentage of vaccine-compliant HPS respondents. By contrast, an increase in the percentage of tweets expressing trust (*P*<.001) was significantly associated with a decrease in the percentage of vaccine-compliant HPS respondents.

Model 2a revealed significant associations (*P*<.05) between the percentage of vaccine-hesitant HPS respondents and the percentage of tweets expressing a negative sentiment. The R-squared value for the vaccine-hesitant GLM (model 2a) was similar to that of the vaccine-compliant GLM (93.17%). However, the vaccine-hesitant GLMM showed a much lower R-squared value (9.4%) and slightly higher RMSE (0.032). When compared with the vaccine-hesitant GLM (model 2a), whose only difference from the GLMM was the use of metropolitan area as a random effect, these results revealed that metropolitan area, as a fixed effect, and negative tweet sentiment (statistically significant in both model 2a and 2b) contributed to the majority of the variation in the percentage of vaccine-hesitant HPS respondents. When looking at the regressors, the results of both the vaccine-hesitant models, model 2a and 2b, showed that an increase in the percentage of tweets expressing a negative sentiment (*P*=.02 and *P*=.03, respectively) was associated with a decrease in the percentage of vaccine-hesitant HPS respondents.

Discussion

Principal Findings

In this study, we sought to determine whether the sentiments and emotions found in COVID-19 vaccine tweets can predict the vaccine hesitancy and compliance expressed in the US Census Bureau's HPS. Depending on the model, GLMs and GLMMs showed significant relationships between (1) the percentage of vaccine-compliant HPS respondents and percentage of tweets expressing a positive sentiment and trust and (2) the percentage of vaccine-hesitant HPS respondents and percentage of tweets expressing a negative sentiment. Positive perceptions expressed on Twitter performed well in predicting positive perceptions in the survey for both GLMs and GLMMs, whereas negative perceptions expressed on Twitter performed well in predicting negative perceptions in the survey only for the GLM.

Study Findings in Context

The main objective of this study was to examine whether aggregate attitudes extracted from social media can predict vaccine attitudes collected via surveys. Specifically, we hypothesized that there is a direct, positive relationship between (1) positive sentiments found in Twitter and the HPS survey and (2) negative sentiments found in Twitter and the HPS survey. We expected to see a positive relationship between positive sentiments and emotions on Twitter and vaccine compliance in the HPS, as suggested in a previous study that showed a positive relationship between positive sentiment scores in COVID-19 vaccine–related tweets and an increase in vaccination rates [30]. The results of both vaccine-compliant models revealed, as expected, significant positive relationships between the percentage of vaccine-compliant HPS respondents and percentage of tweets expressing a positive sentiment. However, in both vaccine-compliant models, the direction of one of the statistically significant relationships that were revealed was not what we expected. Both vaccine-compliant models revealed a significant inverse relationship between the vaccine-compliant measure in the HPS and percentage of tweets expressing trust.

We also expected to see a positive relationship between negative sentiments and emotions on Twitter and vaccine hesitancy in the HPS. Although not significant, the percentages of tweets expressing anger and fear were, as expected, positively related to increases in vaccine hesitancy for both GLM and GLMM. However, the vaccine-hesitant GLM and GLMM both revealed a significant inverse relationship between the vaccine-hesitant measure in the HPS and percentage of tweets expressing a negative sentiment.

The unexpected direction of the relationship between some vaccine perceptions in the survey and vaccine perceptions found on Twitter might be explained via ecological fallacy. An ecological fallacy occurs when an inference is made about the relationship between variables at an aggregate or a group level (eg, the level of a neighborhood, city, or country) based solely on the relationship between the same variables at an individual level [31]. This can be problematic because the relationship between variables at the group level may differ from the relationship between the same variables at the individual level because of factors such as aggregation bias and other contextual factors.

Another possible explanation for the unexpected findings on some of the relationships between vaccine perceptions in the survey and vaccine perceptions found on Twitter could be the possibility that some individuals' web-based personas may not match their reality. For example, an individual might be obligated to get a vaccine because of their job or an upcoming travel, making them vaccine compliant, but rant about it on the web. In our sample data, this type of person would be classified as *provaccine* instead of *antivaccine* in the HPS but would also contribute to the negative perceptions found on Twitter. These findings also align with prior research that suggested an individual's web-based persona may differ from their offline identity [32-34]. This offline identity is often limited by physical, emotional, and financial circumstances that may be beyond an individual's control [33,35-37]. However, individuals have complete control over the identity they choose to present on the web [32-34]. The inverse relationship between the vaccine-hesitant measure in the HPS and percentage of tweets expressing a negative sentiment may have also been due to the use of sarcasm in tweets, where the text itself contradicts what is actually meant by the user [38].

The findings of this study contribute to the literature in 2 ways. First, although many studies have examined COVID-19 vaccine acceptance by extracting information from either surveys or social media, to our knowledge, no study has evaluated the relationship between these vastly different data sources. Unlike social media data collection, surveys come with postage, paper, printing, interviewer, and data entry costs, making them costly to administer [39]. Evaluating the relationship between the attitudes found in surveys and those found on social media allows researchers to determine whether social media data can be trusted to reveal the same information that can be extracted from traditional surveys or whether there is a risk of losing important information in exchange for cutting costs. In this study, we found that COVID-19 vaccine attitudes in the HPS, measured as vaccine compliance and hesitancy, can be predicted using social media attitudes toward vaccines, measured via sentiments and emotions toward vaccines. The results of this study support the efforts of researchers, who over the past few years have looked at social media as a data source, citing the availability of readily available data and no- or low-cost data collection efforts [40,41].

This study makes further contributions by revealing the sentiments and emotions found in tweets across different metropolitan areas. This builds upon several other studies that leveraged natural language processing (NLP) methods, such as sentiment analysis, emotion analysis, and topic modeling, to examine vaccine-related perceptions [42-44]. In this study, we found that most tweets expressed a provaccine sentiment, across all metropolitan areas. However, many tweets also expressed negative feelings and anticipations. This supports previous work, where researchers found many discussions about vaccine hesitancy but ultimately found most tweets to have a positive sentiment [45]. This study also revealed trust as the dominant emotion found in tweets. This supports the results of a prior study that also found trust to be the dominant emotion expressed in tweets during an earlier period [46]. A comparison of these results shows that the vaccine conversation on Twitter remained relatively consistent over time.

Comparing COVID-19 vaccine perceptions on Twitter with attitudes in traditional public health surveys offers several benefits. Twitter serves as a platform for immediate and widespread dissemination of information. Analyzing vaccine perceptions on Twitter can help identify emerging issues or concerns related to COVID-19 vaccines at an early stage. This early detection allows public health authorities to address misconceptions, respond to emerging challenges, and promptly adapt their communication strategies. For example, the study results suggest that both models may be beneficial when deciding which cities to implement vaccine campaigns in, and the vaccine-compliant model can be used to better understand the role sentiments play in vaccination behaviors. This type of model can be used to craft effective social media messages related to COVID-19 vaccination.

Twitter provides a platform for a wide range of voices and opinions, including those of individuals with varying backgrounds, beliefs, and experiences. Comparing Twitter data with survey data allows for the exploration of diverse perspectives and can uncover viewpoints that may not be captured through traditional surveys alone. This broader range of perspectives enhances the understanding of the complexities surrounding vaccine perceptions. Twitter data also allow for the real-time monitoring of public sentiment and reactions toward COVID-19 vaccines. This timely information can provide valuable insights into evolving trends, emerging concerns, and the impact of specific events or interventions. By comparing Twitter discussions with survey responses, researchers can identify shifts in public opinion and monitor the effectiveness of public health communication strategies in real time.

Comparison with Twitter data can complement the findings of traditional surveys, providing a more comprehensive understanding of vaccine perceptions. Twitter data can provide contextual information, qualitative insights, and real-world examples that enrich the analysis of survey responses. The combination of both sources offers a more nuanced and holistic understanding of public attitudes toward COVID-19 vaccines.

This study provides further evidence for the benefits of using social media data for public health research. The overarching contribution of this work suggests the adoption of alternative data sources and NLP techniques to assist in public health decision-making.

Limitations and Future Work

Considering the limitations of this study may lead to future, related work. This study emphasizes the use of Twitter as a data source, but the lack of representation among Twitter users leads to bias in the sample and contributes to sampling errors. For example, Twitter users tend to be younger, be more educated, have higher incomes, and be more liberal [47]. The lack of representation among Twitter users suggests the limited generalizability of the results to the larger population. Adding to this lack of representation is the limited sample of tweets available to the public via the Twitter Streaming application programming interface, which makes available a random sample of 1% of all tweets made by Twitter users at any given time [48]. In addition, in studies assessing COVID-19 vaccine perceptions using social media data, individuals who do not have access to social media are systematically excluded from the analysis sample.

The lack of demographic information on Twitter users is also a limitation to using Twitter as a data source. The absence of demographic information, such as age, gender, income, and education makes it challenging to understand the characteristics of the Twitter users who generate the data. This lack of information may lead to biased or incomplete analyses and limit the generalizability of the findings. In addition, the absence of demographic data makes it difficult to compare Twitter data with data from other sources that do contain demographic information, such as survey data. Despite these limitations, Twitter data can still be useful in certain contexts.

It should also be acknowledged that the HPS data are also subject to sampling errors due to sample design, nonresponse, weighting adjustments, and measurement errors [49]. As a result, the true relationship between aggregate attitudes extracted from social media and vaccine attitudes collected via surveys may be different from what was revealed in this study.

Future studies should endeavor to use other NLP approaches, such as topic modeling, to compare public perceptions of the COVID-19 vaccine on social media with those found in surveys. The survey used in this study, the HPS, presented respondents with in-depth questions related to why they were vaccine hesitant, so applying topic models to tweets may reveal some of the same attitudes and themes as those expressed in the survey. Future studies may also involve pulling data from other social media platforms, such as Facebook, and comparing the overall perceptions reflected across all media.

Conclusions

The ongoing COVID-19 pandemic requires consistent monitoring and data-driven public health policies. To slow the spread of the virus, public health officials have stressed that vaccines are essential in the worldwide battle against COVID-19. However, vaccine hesitancy continues to be a barrier to effective and consistent vaccine rollout programs. Prior efforts have used surveys to gauge attitudes toward the COVID-19 vaccine, but this study suggests that these public perceptions may also be extracted from a readily available, low-cost data source, social media. In this study, we validated social media as a data source by evaluating the relationship between the attitudes expressed among Twitter users and attitudes expressed among respondents to the HPS as well as the ability of attitudes expressed among Twitter users to predict vaccine compliance and hesitancy among the HPS respondents. Leveraging Twitter data alongside traditional surveys can provide a more comprehensive and nuanced understanding of COVID-19 vaccine perceptions, facilitating evidence-based decision-making and tailored public health strategies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Model diagnostics.
[DOC File , 3561 KB - infodemiology_v3i1e43700_app1.doc]

References

1. Household pulse survey COVID-19 vaccination tracker. United States Census Bureau. 2021 Dec 22. URL: https://www.census.gov/library/visualizations/interactive/household-pulse-survey-covid-19-vaccination-tracker.html [accessed 2022-01-13]
2. Garett R, Young SD. 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]
3. Pershad Y, Hangge PT, Albadawi H, Oklu R. Social medicine: Twitter in healthcare. J Clin Med 2018 May 28;7(6):121 [FREE Full text] [doi: 10.3390/jcm7060121] [Medline: 29843360]
4. Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. Am J Trop Med Hyg 2012 Jan;86(1):39-45 [FREE Full text] [doi: 10.4269/ajtmh.2012.11-0597] [Medline: 22232449]
5. Hong L, Fu C, Torrens P, Frias-Martinez V. Understanding citizens' and local governments' digital communications during natural disasters: the case of snowstorms. In: Proceedings of the 2017 ACM on Web Science Conference. 2017 Jun Presented at: WebSci '17: ACM Web Science Conference; June 25-28, 2017; Troy, NY. [doi: 10.1145/3091478.3091502]
6. Ball-Rokeach SJ, DeFleur ML. A dependency model of mass-media effects. Commun Res 1976 Jan;3(1):3-21. [doi: 10.1177/009365027600300101]
7. Marston HR, Ko PC, Girishan Prabhu V, Freeman S, Ross C, Sharaievska I, et al. Digital practices by citizens during the COVID-19 pandemic: findings from an international multisite study. JMIR Ment Health 2023 Mar 06;10:e41304 [FREE Full text] [doi: 10.2196/41304] [Medline: 36877558]
8. Charquero-Ballester M, Walter JG, Nissen IA, Bechmann A. Different types of COVID-19 misinformation have different emotional valence on Twitter. Big Data Soc 2021 Sep 22;8(2):205395172110412 [FREE Full text] [doi: 10.1177/20539517211041279]
9. Bonnevie E, Gallegos-Jeffrey A, Goldbarg J, Byrd B, Smyser J. Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. J Commun Healthc 2020 Dec 15;14(1):12-19. [doi: 10.1080/17538068.2020.1858222]
10. 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]
11. Ward JK, Alleaume C, Peretti-Watel P, COCONEL Group. The French public's attitudes to a future COVID-19 vaccine: the politicization of a public health issue. Soc Sci Med 2020 Nov;265:113414 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113414] [Medline: 33038683]
12. Wang K, Wong EL, Ho KF, Cheung AW, Chan EY, Yeoh EK, et al. Intention of nurses to accept coronavirus disease 2019 vaccination and change of intention to accept seasonal influenza vaccination during the coronavirus disease 2019 pandemic: a cross-sectional survey. Vaccine 2020 Oct 21;38(45):7049-7056 [FREE Full text] [doi: 10.1016/j.vaccine.2020.09.021] [Medline: 32980199]
13. Ruiz JB, Bell RA. Predictors of intention to vaccinate against COVID-19: results of a nationwide survey. Vaccine 2021 Mar 12;39(7):1080-1086 [FREE Full text] [doi: 10.1016/j.vaccine.2021.01.010] [Medline: 33461833]
14. Arigo D, Pagoto S, Carter-Harris L, Lillie SE, Nebeker C. Using social media for health research: methodological and ethical considerations for recruitment and intervention delivery. Digit Health 2018 May 07;4:2055207618771757 [FREE Full text] [doi: 10.1177/2055207618771757] [Medline: 29942634]

15. Sigalo N, Awasthi N, Abrar SM, Frias-Martinez V. Using COVID-19 vaccine attitudes on Twitter to improve vaccine uptake forecast models in the United States: infodemiology study of Tweets. JMIR Infodemiology 2023 Aug 21;3:e43703 [FREE Full text] [doi: 10.2196/43703] [Medline: 37390402]
16. Wright KB. Researching internet-based populations: advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. J Comput Mediat Commun 2005 Apr 1;10(3):JCMC1034 [FREE Full text] [doi: 10.1111/j.1083-6101.2005.tb00259.x]
17. Enriquez D, Goldstein A. COVID-19's socioeconomic impact on low-income benefit recipients: early evidence from tracking surveys. Socius 2020 Nov 25;6:237802312097079 [FREE Full text] [doi: 10.1177/2378023120970794]
18. Morales DX, Morales SA, Beltran TF. Food insecurity in households with children amid the COVID-19 pandemic: evidence from the household pulse survey. Soc Curr 2021 Apr 23;8(4):314-325 [FREE Full text] [doi: 10.1177/23294965211011593]
19. Household pulse survey technical documentation. United States Census Bureau. URL: https://www.census.gov/programs-surveys/household-pulse-survey/technical-documentation.html [accessed 2023-05-09]
20. Daly M, Jones A, Robinson E. Public trust and willingness to vaccinate against COVID-19 in the US from October 14, 2020, to March 29, 2021. JAMA 2021 Jun 15;325(23):2397-2399 [FREE Full text] [doi: 10.1001/jama.2021.8246] [Medline: 34028495]
21. Silge J, Robinson D. tidytext: text mining and analysis using tidy data principles in R. J Open Source Softw 2016;1(3):37 [FREE Full text] [doi: 10.21105/joss.00037]
22. NRC word-emotion association lexicon. NRC Emotion Lexicon. URL: http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm [accessed 2022-01-13]
23. Kausar MA, Soosaimanickam A, Nasar M. Public sentiment analysis on Twitter data during COVID-19 outbreak. Int J Adv Comput Sci Appl 2021;12(2) [FREE Full text] [doi: 10.14569/IJACSA.2021.0120252]
24. Zeileis A, Cribari-Neto F, Gruen B, Kosmidis I, Simas AB, Rocha AV. Package 'betareg'. The Comprehensive R Archive Network. 2022 Oct 12. URL: https://cran.r-project.org/web/packages/betareg/betareg.pdf [accessed 2023-11-10]
25. Rizopoulos D. GLMMadaptive: generalized linear mixed models using adaptive gaussian quadrature. The Comprehensive R Archive Network. 2023. URL: https://cloud.r-project.org/web/packages/GLMMadaptive/index.html [accessed 2023-05-09]
26. Li B, Lingsma HF, Steyerberg EW, Lesaffre E. Logistic random effects regression models: a comparison of statistical packages for binary and ordinal outcomes. BMC Med Res Methodol 2011 May 23;11(1):77 [FREE Full text] [doi: 10.1186/1471-2288-11-77] [Medline: 21605357]
27. Harrison XA, Donaldson L, Correa-Cano ME, Evans J, Fisher DN, Goodwin CE, et al. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 2018 May 23;6:e4794 [FREE Full text] [doi: 10.7717/peerj.4794] [Medline: 29844961]
28. Chen K, Cheng Y, Berkout O, Lindhiem O. Analyzing proportion scores as outcomes for prevention trials: a statistical primer. Prev Sci 2017 Apr;18(3):312-321 [FREE Full text] [doi: 10.1007/s11121-016-0643-6] [Medline: 26960687]
29. IRB Definitions.: Division of Research - University of Maryland URL: https://research.umd.edu/resources/research-compliance/institutional-review-board-irb/irb-definitions [accessed 2023-11-27]
30. Bari A, Heymann M, Cohen RJ, Zhao R, Szabo L, Apas Vasandani S, et al. Exploring coronavirus disease 2019 vaccine hesitancy on Twitter using sentiment analysis and natural language processing algorithms. Clin Infect Dis 2022 May 15;74(Suppl_3):e4-e9. [doi: 10.1093/cid/ciac141] [Medline: 35568473]
31. Reichert M, Brüßler S, Reinhard I, Braun U, Giurgiu M, Hoell A, et al. The association of stress and physical activity: mind the ecological fallacy. Ger J Exerc Sport Res 2022 May 06;52(2):282-289 [FREE Full text] [doi: 10.1007/s12662-022-00823-0]
32. Huang J, Kumar S, Hu C. A literature review of online identity reconstruction. Front Psychol 2021 Aug 23;12:696552 [FREE Full text] [doi: 10.3389/fpsyg.2021.696552] [Medline: 34497560]
33. Kim HW, Zheng JR, Gupta S. Examining knowledge contribution from the perspective of an online identity in blogging communities. Comput Hum Behav 2011 Sep;27(5):1760-1770 [FREE Full text] [doi: 10.1016/j.chb.2011.03.003]
34. Hu C, Zhao L, Huang J. Achieving self-congruency? Examining why individuals reconstruct their virtual identity in communities of interest established within social network platforms. Comput Hum Behav 2015 Sep;50:465-475 [FREE Full text] [doi: 10.1016/j.chb.2015.04.027]
35. Bargh JA, McKenna KY, Fitzsimons GM. Can you see the real me? Activation and expression of the "true self" on the internet. J Soc Issues 2002 Dec 17;58(1):33-48. [doi: 10.1111/1540-4560.00247]
36. Donath JS. Identity and deception in the virtual community. In: Kollock P, Smith M, editors. Communities in Cyberspace. London, UK: Routledge; Dec 17, 1998.
37. Schau HJ, Gilly MC. We are what we post? Self-presentation in personal web space. J Consum Res 2003 Dec;30(3):385-404. [doi: 10.1086/378616]
38. Govindan V, Balakrishnan V. A machine learning approach in analysing the effect of hyperboles using negative sentiment tweets for sarcasm detection. J King Saud Univ Comput Inf Sci 2022 Sep;34(8):5110-5120 [FREE Full text] [doi: 10.1016/j.jksuci.2022.01.008]
39. Morland K, Filomena S. Disparities in the availability of fruits and vegetables between racially segregated urban neighbourhoods. Public Health Nutr 2007 Dec;10(12):1481-1489. [doi: 10.1017/S1368980007000079] [Medline: 17582241]

40. Lewis LB, Sloane DC, Nascimento LM, Diamant AL, Guinyard JJ, Yancey AK, et al. African Americans' access to healthy food options in South Los Angeles restaurants. Am J Public Health 2005 Apr;95(4):668-673. [doi: 10.2105/AJPH.2004.050260] [Medline: 15798128]
41. Sigalo N, St Jean B, Frias-Martinez V. Using social media to predict food deserts in the United States: infodemiology study of tweets. JMIR Public Health Surveill 2022 Jul 05;8(7):e34285 [FREE Full text] [doi: 10.2196/34285] [Medline: 35788108]
42. Sattar NS, Arifuzzaman S. COVID-19 vaccination awareness and aftermath: public sentiment analysis on Twitter data and vaccinated population prediction in the USA. Appl Sci 2021 Jun 30;11(13):6128 [FREE Full text] [doi: 10.3390/app11136128]
43. Liu S, Liu J. Public attitudes toward COVID-19 vaccines on English-language Twitter: a sentiment analysis. Vaccine 2021 Sep 15;39(39):5499-5505 [FREE Full text] [doi: 10.1016/j.vaccine.2021.08.058] [Medline: 34452774]
44. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al. Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: observational study. J Med Internet Res 2021 Apr 05;23(4):e26627 [FREE Full text] [doi: 10.2196/26627] [Medline: 33724919]
45. Yousefinaghani S, Dara R, Mubareka S, Papadopoulos A, Sharif S. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. Int J Infect Dis 2021 Jul;108:256-262 [FREE Full text] [doi: 10.1016/j.ijid.2021.05.059] [Medline: 34052407]
46. 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]
47. Freedman DA. Local food environments: they're all stocked differently. Am J Community Psychol 2009 Dec;44(3-4):382-393. [doi: 10.1007/s10464-009-9272-6] [Medline: 19838789]
48. Nguyen QC, Li D, Meng HW, Kath S, Nsoesie E, Li F, et al. Building a national neighborhood dataset from geotagged Twitter data for indicators of happiness, diet, and physical activity. JMIR Public Health Surveill 2016 Oct 17;2(2):e158 [FREE Full text] [doi: 10.2196/publichealth.5869] [Medline: 27751984]
49. Bradley VC, Kuriwaki S, Isakov M, Sejdinovic D, Meng XL, Flaxman S. Unrepresentative big surveys significantly overestimated US vaccine uptake. Nature 2021 Dec;600(7890):695-700 [FREE Full text] [doi: 10.1038/s41586-021-04198-4] [Medline: 34880504]

Abbreviations

GLM: generalized linear model
GLMM: generalized linear mixed model
HPS: Household Pulse Survey
NLP: natural language processing
NRC: Natural Language Understanding Research Consortium
RMSE: root mean square error

Edited by R Cuomo; submitted 20.10.22; peer-reviewed by F Kreute, D Steel, Y Mejova, H Mondal; comments to author 16.02.23; revised version received 18.05.23; accepted 14.10.23; published 30.11.23.

Please cite as:
Sigalo N, Frias-Martinez V
Using COVID-19 Vaccine Attitudes Found in Tweets to Predict Vaccine Perceptions in Traditional Surveys: Infodemiology Study
JMIR Infodemiology 2023;3:e43700
URL: https://infodemiology.jmir.org/2023/1/e43700
doi:10.2196/43700
PMID:37903294

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org