JMIR Infodemiology

Journal Impact Factor (JIF) (2023): 3.5
Volume 2 (2022), Issue 2 ISSN 2564-1891 Editor-in-Chief: Tim Ken Mackey, MAS, PhD

Contents

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

Review

Media Data and Vaccine Hesitancy: Scoping Review

Jason Dean-Chen Yin[1], BSc, MSc

School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Jason Dean-Chen Yin, BSc, MSc
School of Public Health
Li Ka Shing Faculty of Medicine
The University of Hong Kong
7 Sassoon Road
Pok Fu Lam
Hong Kong
China (Hong Kong)
Phone: 852 97907044
Email: jdyin@hku.hk

Abstract

Background: Media studies are important for vaccine hesitancy research, as they analyze how the media shapes risk perceptions and vaccine uptake. Despite the growth in studies in this field owing to advances in computing and language processing and an expanding social media landscape, no study has consolidated the methodological approaches used to study vaccine hesitancy. Synthesizing this information can better structure and set a precedent for this growing subfield of digital epidemiology.

Objective: This review aimed to identify and illustrate the media platforms and methods used to study vaccine hesitancy and how they build or contribute to the study of the media's influence on vaccine hesitancy and public health.

Methods: This study followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. A search was conducted on PubMed and Scopus for any studies that used media data (social media or traditional media), had an outcome related to vaccine sentiment (opinion, uptake, hesitancy, acceptance, or stance), were written in English, and were published after 2010. Studies were screened by only 1 reviewer and extracted for media platform, analysis method, the theoretical models used, and outcomes.

Results: In total, 125 studies were included, of which 71 (56.8%) used traditional research methods and 54 (43.2%) used computational methods. Of the traditional methods, most used content analysis (43/71, 61%) and sentiment analysis (21/71, 30%) to analyze the texts. The most common platforms were newspapers, print media, and web-based news. The computational methods mostly used sentiment analysis (31/54, 57%), topic modeling (18/54, 33%), and network analysis (17/54, 31%). Fewer studies used projections (2/54, 4%) and feature extraction (1/54, 2%). The most common platforms were Twitter and Facebook. Theoretically, most studies were weak. The following five major categories of studies arose: antivaccination themes centered on the distrust of institutions, civil liberties, misinformation, conspiracy theories, and vaccine-specific concerns; provaccination themes centered on ensuring vaccine safety using scientific literature; framing being important and health professionals and personal stories having the largest impact on shaping vaccine opinion; the coverage of vaccination-related data mostly identifying negative vaccine content and revealing deeply fractured vaccine communities and echo chambers; and the public reacting to and focusing on certain signals—in particular cases, deaths, and scandals—which suggests a more volatile period for the spread of information.

Conclusions: The heterogeneity in the use of media to study vaccines can be better consolidated through theoretical grounding. Areas of suggested research include understanding how trust in institutions is associated with vaccine uptake, how misinformation and information signaling influence vaccine uptake, and the evaluation of government communications on vaccine rollouts and vaccine-related events. The review ends with a statement that media data analyses, though groundbreaking in approach, should supplement—not supplant—current practices in public health research.

(JMIR Infodemiology 2022;2(2):e37300) doi:10.2196/37300

KEYWORDS

review; social media; traditional media; vaccine hesitancy; natural language processing; digital epidemiology

XSL•FO
RenderX

Introduction

Media and Public Health

The media are important for public health research. They are a source of information, a broadcasting station, an issue identifier, and a perception molder, among many things. Exposure to the media can thus shape health-related perceptions and, therefore, behaviors. This area of research has extended from the fields of psychology and social psychology and primarily looks at *effects of media* [1]. It primarily asks the following question: what are the consequences of media exposure at an individual, group, institutional, and social system level? This question highlights the different levels at which communication occurs.

At an individual (or micro) level, there are three interwoven theoretical areas: expectancy value, information processing, and message effect [1]. Expectancy value theories posit that health behaviors are motivated by beliefs and expectancies regarding an outcome and the values placed on it. Theories such as the health belief model (HBM) [2], theory of planned behavior [3], and theory of reasoned action [4] all account for how media exposure can affect the motivations, attitudes, and behaviors of individuals regarding a decision. Information processing focuses on how psychological processing occurs and leads to either changes or reinforcements in attitude. Examples include the elaboration likelihood model (ELM) [5], extended parallel processing model [6], and protective action decision model [7], which focus on how cues and the environment affect cognitive processes in decision-making, whether this induces a deliberate and thoughtful or passive and peripheral processing of information. These types of studies also focus on how messaging units and the different manifestations (eg, text and images) influence information processing. This alludes to the last theoretical area, message effects, which looks at how the construction of messaging influences information processing [8]. The most common approach in this area is the study of framing, which involves understanding how the media encodes messages through signs and symbols, thereby characterizing an issue and indirectly characterizing how entities should perceive it. These 3 areas, although presented separately, are tightly linked: message effects will affect processing and, thereby, expectations and values placed on outcomes.

At a societal (macro) level, much work has been done on the media's role in agenda setting. In agenda setting theory, the media can influence the importance of topics to the public and, thus, the topic's prioritization as a social problem [9]. This process unfurls in two simultaneous steps—framing and amplification. As stated earlier, the construction and characterization of messages shape public perception of the issue. This has a spillover effect of priming the audience to reconsider their evaluation of an outcome of or the value placed on a topic. When the media are broadcast on different channels, they inadvertently amplify those framed signals, highlighting the media's inherent nature as an amplification station. This concept was captured succinctly in the social amplification of risk framework (SARF) by Kasperson [10], focusing on how topics, events, or hazards interact with psychological, social, institutional, and cultural processes that result in amplification or attenuation of the perception of said topics, events, or hazards. In this process, the media is an institution that acts as an amplification station bringing attention to issues. Amplifying, coupled with framing, shapes public opinion.

Although the schema of micro and macro analyses is separated for presentation, emphasis should be placed on their interconnectedness, especially in a complicated media landscape. The agenda and framing of topics and their subsequent propagation through media channels may shape public and individual opinions. These upstream effects proceed to mold individual processing, expectations, and values around the topic. However, the media, presented as a monolithic concept thus far, can be deconstructed. The growth of alternative social media channels for communication has blurred who or what is considered media. Individual users can act as amplification stations and create content for access on large scales, upending the monopoly traditional media channels had on agenda setting, framing, and amplification. In short, everyone is a purveyor of information. This landscape shapes the mosaic of perceptions of an issue [11]. The next question is then what issue is important for public health?

Vaccine Hesitancy

The World Health Organization (WHO) listed vaccine hesitancy—a "delay in acceptance or refusal of vaccination despite availability of vaccination services" [12]—as one of the top 10 threats to global health in 2019 [13]. In a paper published by the WHO Strategic Advisory Group of Experts on Immunization, they proposed a matrix of determinants that identified three *categories* of influences—contextual, individual and group, and vaccine-specific—that shape the decision to accept, delay, or outright reject vaccines [12]. Several factors nested within these categories point to the media as potentially influencing vaccine uptake. For example, in contextual influences, "communication and media environment" explicitly highlights media as a contextual influence; the individual and group influence category contains "immunization as a social norm," which can be shaped by media portrayals; and vaccine-specific issues include the factors "introduction of a new vaccine," "the strength of recommendation," and "risk/benefit from scientific evidence," all of which are potentially shaped by media coverage and portrayal. Thus, the media and vaccine hesitancy are linked.

Although not a new phenomenon, vaccine hesitancy has been brought back into the limelight through 2 developments. The first development is the growth of social media as a platform for information consumption. The capacity of the individual to assume the role of media in information creation and propagation has complicated the information landscape. These complications include the credibility of the news source and the sheer increase in the size of information production. A resulting externality that may influence vaccine hesitancy is the exposure of the public to misinformation, both unintentional and deliberate. Another externality is exposure to the platforms' algorithms that perpetuate information to reinforce existing beliefs, encouraging polarization (echo chambers). The second development thrusting the vaccine debate to center page is the

XSL•FO
RenderX

COVID-19 pandemic. Although SARS-CoV-2 stagnated economies through 2020 and 2021, the vaccine was thought to be the exit strategy. However, this was not without marring public criticism regarding its development, efficacy, side effects, and necessity, among other concerns. Throughout the cycle of new variants and boosters after the initial introductions, vaccine-hesitant speech and behavior continued to propagate. Much of this was fueled on social media, which further amplified messaging.

Objectives

Alongside the public discussion was the proliferation of academic studies analyzing social media to better understand vaccine hesitancy. This proliferation is due in part to the growing number of media platforms but is also the result of paralleling advances in computing and analysis tools that process and handle big data. To date, there have been no studies that catalog the types of media platforms and analysis methods used to study vaccine hesitancy and if there are any consistent findings. To bridge this gap, the objectives of this study were to answer what platforms are studied and how the data contained are analyzed. The aim of this review was to understand how using these platforms and methods *builds* or *contributes* to the existing knowledge of the literature on the media's influence on vaccine hesitancy and, thereby, public health.

Methods

Overview

This review summarized studies on vaccine hesitancy using any form of *media data*—a catchall term for traditional and social media. Traditional media are loosely defined as any media before the advent of digital media. This review followed the guidelines proposed by Arskey and O'Malley [14] and the Joanna Briggs Institute [15]. All reporting of findings is in accordance with the guidelines specified by the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews; Multimedia Appendix 1) [16]. The protocol for the search is available from the corresponding author upon request and has not been registered.

Inclusion and Exclusion Criteria and Search Strategy

Several inclusion and exclusion criteria were specified to narrow the search. For inclusion, studies must have used any *media data* (see the definition in the *Overview* section) as their data source. The outcomes in the study must be related to vaccine sentiment, opinion, uptake, or hesitancy. Although the aim of this study was to look at vaccine hesitancy, this was often done in indirect ways of asking about sentiments regarding vaccines. Uptake can also be another proxy for vaccine acceptance. As social media became a phenomenon in the late 2000s, the search was limited to the year 2010, chosen arbitrarily but corresponding loosely to the year of the H1N1 influenza pandemic in which a vaccine was developed. Imposing a time restriction intentionally did two things: (1) it focused the search on social media platforms (although this is specified in the search terms) and (2) it weighted the search toward capturing more big data methods. Despite the imposed time cutoff and bias toward these methods, non–big data methods for analyzing texts were expected to appear in the search. Regarding exclusion, studies that used social media platforms for recruitment of participants for survey data collection were excluded. Studies using media platforms to conduct natural experiments (eg, introducing social media campaigns) were also excluded. Unpublished manuscripts, protocols, editorials, letters, case reports, commentaries, opinion pieces, narrative reviews, clinical guidelines, and books were also not analyzed.

The search strategy broadly consisted of 2 sets of terms. The first set captured the specified platform of interest to obtain the most popular messaging channels. The second set captured the concept of vaccine hesitancy using synonymic terms. These terms are expressions of the hesitancy concept in a different way. It is important to note that, although these terms are nuanced (eg, antivaccination connotes an absolute rejection of vaccines), they are still part of the overall vaccine hesitancy spectrum. Thus, they were included in their wildcard form. The same search was performed in two different databases: PubMed and Scopus. A summary of the exclusion and inclusion criteria can be found in Textbox 1, and the specific searches can be found in Multimedia Appendix 2.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Use of any media data (social media or traditional media) as data source
- Outcome must be related to vaccine sentiment (eg, opinion, uptake, hesitancy, acceptance, or stance)
- Written in English
- Published after 2010

Exclusion criteria

- No use of media data as data source
- Use of survey data (asking about social media use as a questionnaire item)
- Use of social media to recruit participants
- Use of social media platform as natural experiment
- Unpublished papers, protocols, editorials, letters, case reports, commentaries, opinion pieces, narratives, clinical guidelines, and books

XSL•FO
RenderX

Study Selection

Two-step screening was implemented after removing duplicates found in the three databases. Titles and abstracts were screened first as a quick filter for eligibility. Any study not meeting the inclusion criteria (or meeting the exclusion criteria) was removed. Subsequently, the remaining full texts were extracted. Studies that did not meet the eligibility criteria (Textbox 1) during extraction were further removed. All removed studies were classified on their reasons for exclusion. Only JDY screened the articles because of manpower limitations.

Data Extraction

To reiterate, this review summarized *what* platforms were studied, *how* the data contained were analyzed, and how the studies *built* or *contributed* to the existing work on the media's influence on vaccine hesitancy. This loosely corresponds to the "concept" portion in the Population-Concept-Context framework of the Joanna Briggs Institute [15]. Accordingly, the four main extracted elements were (1) media platform, (2) analysis method, (3) theories, and (4) findings. Other variables such as (5) the country of focus and (6) language were also included and can be thought to correspond to "context" given the foreseeable diversity in languages and regions of focus. All data were synthesized and charted in Covidence.

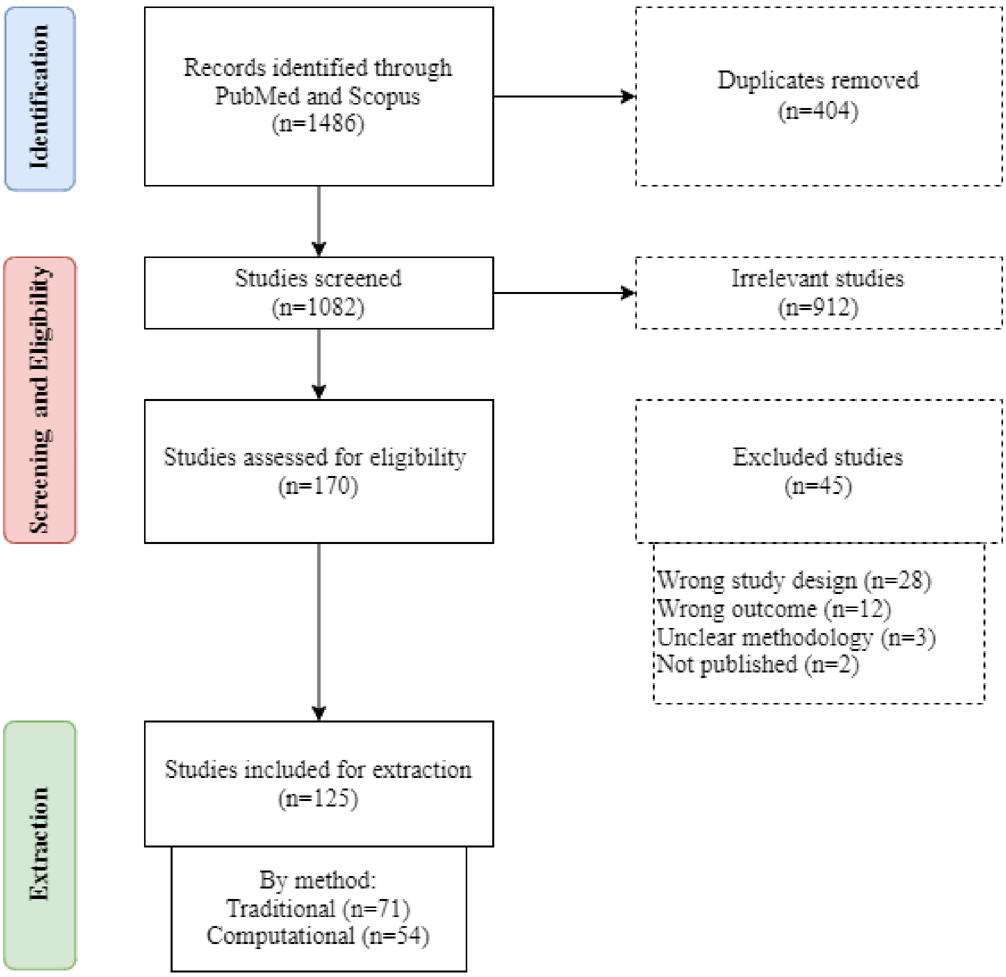
Presentation of Results

The results were separated according to what type of *media* data were used: traditional media or social media. Within each type of media data, a cross-tabulation of the platforms and data analysis methods was presented with accompanying descriptive statistics that illustrated notable trends. As studies can contain one or more platforms or methods, cells in the cross-tabulation are not mutually exclusive and present overlaps. Fully detailed extractions can be found in Multimedia Appendix 3 [17-80] and Multimedia Appendix 4 [81-134]. Trends in any theory were presented descriptively in the text in addition to the countries and languages represented. The *Discussion* section summarizes the major findings and gaps in the literature that uses media data for vaccine hesitancy research and proposes a method moving forward.

Results

The results of the screening and selection process are presented in the PRISMA-ScR chart (Figure 1). A total of 125 studies were included in this scoping review, of which 71 (56.8%) used traditional methods and 54 (43.2%) used computational methods.

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flow diagram for this scoping review.



XSL•FO
RenderX

Traditional Methods

Overview

Before the advent of computational big data approaches to analyze media data, several *traditional media data analysis methods* (hereon, *traditional methods* or *noncomputational methods*) were used to research vaccine-related topics. This term distinguishes studies that use media data in a manual way; that is, a way that requires the researcher to individually sort through each data point to extract data. These can be further decomposed into two types: tangential studies and directly related studies.

Tangential Studies

Studies with tangential relations include a discussion on vaccines or use of a vaccine-related variable but may not specifically focus on a vaccine outcome as the main variable of interest. There are three subtypes: a focus on a specific population, understanding the nature of information processing, and systematic reviews.

Regarding studies focusing on populations, Leader et al [135] tried to understand the role of "influencers" or "key opinion leaders" on spreading vaccine-related messages in groups of mothers through focus group interviews. They found that influencers posting on vaccine-related issues preferred using information from alternative sources and search engines as opposed to using mainstream information.

Another type of study focused on the nature of information processing in line with the aforementioned category of media studies. An example is the study by Domgaard and Park [136] analyzing how infographic versus SMS text messages may equip users with heightened ability to verify false news in relation to vaccines. Qian et al [137] look at how exposure to negative information may enforce preheld biases and how positive information exposure affects vaccine decision-making. These studies, by focusing more on the psychology of discernment and decision-making, found that the medium (infographics vs text) and connotation (positive or negative) of information transmission are associated with eventual vaccine uptake.

The last type of study was systematic reviews. A Cochrane review looked at the effectiveness of social media in public health interventions [138], with inconclusive findings on overall effectiveness but identifying that studies do not focus on the adverse effects of these interventions. Another systematic review focused on the different methods used for social media monitoring in relation to vaccines [139]. The last review looked specifically at digital interventions with the intention of increasing influenza vaccination among pregnant women [140]. The findings from these 3 studies are largely broad and inconclusive on any effect that public health interventions via social media have on either health outcomes or uptake of vaccination. This can be due to the lack of high-quality, comparable studies that have the same outcome. Notably, two of these systematic reviews consolidated information on experiments, and they were excluded from this review.

Although these studies can be argued to have vaccine-related outcomes as they include vaccine-related data, they are mentioned separately as the primary objectives do not focus on vaccine-related outcomes. Despite their exclusion, these studies highlight the potential of social media–type studies to broaden the scope of research at the public health level, specifically focusing on populations of users, processing of types of information, and public health outcomes from interventions. These factors—populations, processing, and interventions—are all tied closely to the 5 themes identified later.

Directly Related Studies

Most studies (65/125, 52%) focused on a direct vaccine outcome and encompassed a variety of countries and languages. The most represented countries were from Europe (France: 3/65, 5%; Italy: 6/65, 9%) as well as the United States (8/65, 12%). Fewer studies came from Asia (3/65, 5%), Africa (3/65, 5%), and the Middle East (1/65, 2%). This diversity in location was also represented in the different languages (where the country or language was not explicitly stated, an inference was made depending on the search terms or the national language of the country): Mandarin Chinese, Cantonese, French, Danish, Italian, Spanish, Hebrew, and English, with the most common being Italian (6/65, 9%) and English because of the multiple English-speaking nations (41/65, 63%). This language diversity will not be reflected in the computational study results, as will be seen.

The platforms and methods used in these studies are summarized in a cross-tabulation (Table 1). Most studies used manual content analysis (43/65, 66%), with a focus on any important themes, topics, frames, or discourse (column 2), and sentiment analysis (21/65, 32%), including any analysis of the tone of vaccine messages, stance on vaccination, polarity in comments, or sentiment classification (column 3) to analyze texts, with few touching on campaign evaluations (5/65, 8%). In the fourth column, some studies track search activity related to vaccines, vaccine coverage, and spread or reach of vaccine-related information (12/65, 18%), highlighting the importance of the SARF framework by Kasperson [10] in vaccine research. The studies included in the table were conducted over a wide assortment of platforms, from traditional media (print media, newspapers, web-based news, and talk shows) to social media (Facebook, Weibo, and Google).

XSL•FO
RenderX

Table 1. Traditional analysis methods and media platforms for studies with a direct vaccine-related outcome (N=65)[a,b].

Media platforms	Analysis methods			
	Content, theme, frame, and discourse analysis	Sentiment, stance, tone, and polarity coding	Activity on the web, media coverage, coverage of vaccines, and misinformation spread	Campaign evaluation
Weibo and Twitter	• Becker et al [17] • Bonnevie et al [18] • Criss et al [19] • Griffith et al [20] • Hou et al [21] • Keim-Malpass et al [22] • Marchetti et al [23] • Sundstrom et al [24]	• Becker et al [17] • Criss et al [19] • Keim-Malpass et al [22] • Marchetti et al [23] • Gori et al [25]	• Sundstrom et al [24] • Aquino et al [26]	• Sundstrom et al [24]
YouTube	• Marchetti et al [23] • Basch et al [27] • Fieselmann et al [28] • Lahouati et al [29]	• Marchetti et al [23] • Lahouati et al [29] • Covolo et al [30] • Donzelli et al [31]	• Basch et al [27] • Donzelli et al [31]	—[c]
Facebook	• Marchetti et al [23] • Sundstrom et al [24] • Fieselmann et al [28] • Bradshaw et al [32] • Jamison et al [33] • Kalichman et al [34] • Orr et al [35] • Tustin et al [36] • Wawrzuta et al [37] • Wiyeh et al [38]	• Marchetti et al [23] • Tustin et al [36] • Loft et al [39] • Luisi [40]	• Sundstrom et al [24] • Aquino et al [26] • Loft et al [39] • Luisi [40,41]	• Sundstrom et al [24] • Loft et al [39] • Pedersen et al [42]
Websites, mixed media, and blogs	• Marchetti et al [23] • Orr et al [35] • Aechtner [43] • Bruel et al [44] • Larson et al [45] • Moran et al [46] • Nugier et al [47] • Toth [48] • Ward and Budarick [49]	• Marchetti et al [23] • Larson et al [45] • Karapetiantz et al [50] • Panatto et al [51] • Shoup et al [52]	• Suppli et al [53]	—
Q&A[d] site	• Sharon et al [54]	• Sharon et al [54]	—	—
Google (search, results, and trends)	• Ruiz and Bell [55] • Sajjadi et al [56]	—	• Aquino et al [26] • Suppli et al [53] • Diaz et al [57]	—
Pinterest	• Guidry et al [58]	—	• Guidry et al [58] • Mahroum et al [59]	—
Print media, newspapers, and web-based news	• Ashwell and Murray [60] • Basch et al [61] • Casciotti et al [62] • Catalan-Matamoros and Elías [63] • Colón-López et al [64] • Court et al [65] • Das et al [66] • Kummervold et al [67] • Meyer et al [68] • Odone et al [69] • Olufowote [70] • Stephenson et al [71]	• Ward and Budarick [49] • Casciotti et al [62] • Catalan-Matamoros and Elías [63] • Das et al [66] • Catalan-Matamoros and Peñafiel-Saiz [72]	—	• Ward and Budarick [49]
Television talk show	• Toth [48]	—	—	—
Documentary	• Bradshaw et al [73]	—	—	—

XSL•FO
RenderX

Media platforms	Analysis methods			
	Content, theme, frame, and discourse analysis	Sentiment, stance, tone, and polarity coding	Activity on the web, media coverage, coverage of vaccines, and misinformation spread	Campaign evaluation
TikTok or Instagram	• Fieselmann et al [28] • Basch et al [74]	—	• Basch et al [74]	—

[a]This table does not include the tangential studies mentioned in the *Tangential Studies* section.

[b]The cells are not mutually exclusive. Studies may appear twice across cells.

[c]Not available. No studies exist using this media platform and analysis method.

[d]Q&A: question and answer.

A minority explicitly stated a theoretical framework that drives the analysis. Ward and Budarick [49] used a discursive legitimization strategy and ideological square theories to evaluate the use of anecdote and emotionality by The Daily Telegraph to push provaccine messaging in a campaign to increase vaccination. A study focusing on discourse used repertoire analysis to understand how parents' repertoires in distrust contribute to a delegitimization of systems propping up medical services, research, and government authorities [48]. Another study on repertoire echoes those using framing theories to understand how positive or negative framing could coerce behavior [60]. In total, 3% (2/65) of the studies looked at the influence of persuasion as a tactic in the delivery of text [46] and as a guide to framing certain cues to influence vaccine uptake behavior [43]. These studies used persuasion theory and the ELM of persuasion to guide discussion. Persuasion theory also connects to other influence theories such as social influence theory, in which individuals change their behaviors to meet the demands of a social environment. In total, 2% (1/65) of the studies analyzed how mothers changed their behaviors within Facebook networks around antivaccination advocates [32]. A total of 3% (2/65) of the studies, conducted by Luisi [40,41], directly used the SARF and the HBM to operationalize concepts within each framework, using human papillomavirus (HPV) vaccination discussions on Facebook as data.

Studies using content and discourse analysis have strong theoretical roots in the social sciences. However, few studies in which these methods were used to study vaccines explicitly mentioned a theory driving their study (12/65, 18%). If manual analyses, which are limited to the physical capacity of data processing, are already theoretically shaky, we expect an even weaker theoretical focus using computational methods.

Computational Methods

A total of 43.2% (54/125) of the studies used computational (big data) methods. There were obvious trends in language, region, and which vaccines were studied. Most of the studies (36/54, 67%) used English-language media data, with a small representation from other European languages (Italian: 5/54, 9%; Dutch: 1/54, 2%; Polish: 1/54, 2%; French: 2/54, 4%), which are often studied alongside English in the same study. Italian is an exception as it is studied independently of English compared with the other European languages. Several East Asian languages were represented as well, with simplified Mandarin Chinese (5/54, 9%), Korean (2/54, 4%), and Japanese (1/54, 2%). In total, 2% (1/54) of the studies used multiple languages from various contents to do a comparison by region as well [81].

In media data analysis, the geographical location or region of study (and, thereby, the population) is not often explicitly stated and, even when done so, it can be ambiguous. Most often, "geography" is determined by explicit mention of a region of interest or inference through pulling of data with a geographical focus (eg, pulling tweets from geotagged posts from the United States) or a language focus (eg, parsing data from a platform published mostly in Japanese). As a result, language often correlates with region, but this is not always the case, especially for a *lingua franca* such as English, which disallows mapping one-to-one because of the many countries that speak it. Despite this deductive approach, 26% (14/54) of the studies did not specify any location but contained English-language media data. Most studies were conducted with the United States as a geographic region of interest (17/54, 31%), followed by China (5/54, 9%) and Italy (5/54, 9%). In total, 9% (5/54) of the studies took a comparative approach and contained multiple jurisdictions of comparative interest, even including 20% (1/5) that adopted a global comparative approach [82]. Compared with studies using traditional methods, we observed limited representativeness of countries and languages. This was due in part to the necessity of parsing and understanding a large quantity of language and the limited language processing tools developed for smaller languages. For countries that are primarily English-speaking—or English-expressing, for capturing web-based information—but not represented here, there are likely to be more studies in these regions as language processing tools are popularized in public health.

Regarding the types of vaccines studied, it is important to note that time censoring of the review would bias the data set to more recent vaccine issues. Most studies (20/54, 37%) focused on the COVID-19 vaccine and were published within the last 2 years. The other popular category of vaccines was not any specific vaccine but, rather, vaccines in general (17/54, 31%), focusing on the overall sentiment and topics related to vaccination. A smaller minority focused on HPV (4/54, 7%); influenza (3/54, 6%); childhood vaccinations (3/54, 6%); maternal vaccinations (2/54, 4%); and the measles, mumps, and rubella vaccine (1/54, 2%).

All the studies included in this section (54/54, 100%) were published in or after 2016. Among them, a diverse selection of platforms and analysis methods were used. Table 2 cross-tabulates these 2 variables in a similar fashion to Table

XSL•FO
RenderX

1, revealing some trends. Overwhelmingly, Twitter was the most popular platform, with 57% (31/54) of studies using it. It is also more represented across the different analysis methods relative to other platforms. This is different when compared with the traditional methods table, where Twitter studies were uncommon. This trend was the opposite for print and news media and web-based news, with less representation as a platform when computational methods were used. The other platforms were novel in Table 2. For example, different search engines appeared: Baidu (China) and Naver (Korea). Parler, a microblogging platform, was also novel.

What types of analysis methods were used? The methods were categorized into the following eight broadly non–mutually exclusive groups: (1) sentiment analysis, (2) topic modeling, (3) semantic network analysis, (4) projections, (5) feature extraction, (6) image analysis, (7) descriptive studies, and (8) machine classification. Sentiment analysis studies (31/54, 57%) assessed various issues, such as stance [81,85,86], emotions [89,117], and polarity [91,123], and the following algorithms, which were used to determine the aforementioned issues, were diverse: Bidirectional Encoder Representations from Transformers, classification tree, K-nearest neighbors, multinomial naïve Bayes, random forest, robust optimized Bidirectional Encoder Representations from Transformers pretraining approach, support vector machine, and Valence Aware Dictionary and Sentiment Reasoner. Topic modeling (18/54, 33%) was a close second in popularity and focused on distilling latent topics within a corpus. The most common method for topic modeling was latent Dirichlet allocation coupled with other methods to look at topic clustering (related to semantic network analysis) or at inter- and intratopic distinctiveness [97]. The studies focused on sentiment analysis and topic modeling were, in part, a continued momentum of traditional research methods that focused on distilling these aspects from the text.

Semantic network analysis (17/54, 31%) focused on understanding the interaction and transfer of information and ideas within specific networks. Methods ranged from cluster analysis using Gelphi [103,106], latent space modeling [100], exponential random graph modeling [126], and the Louvain algorithm for community detection [102,105,106,126]. The remaining analysis types were represented in smaller numbers. A total of 2% (1/54) of the studies used a behavioral dynamics model—inspired by epidemiological models on susceptibility and infected and resistant states of being—to analyze opinion transmission models [116]. Another 4% (2/54) of projection studies used media data and regression models to predict vaccination rates and epidemic size [108,120]. Feature extraction was only found in 2% (1/54) of the studies, in which Lyu et al [109] extracted variables such as demographics, social capital, income, and political affiliation from a corpus of tweets and associated these features with vaccine stance using logistic regression. Image analysis, also known as *computer vision*, was represented in 2% (1/54) of the studies, in which Wang et al [132] used a multimodal network analysis to detect antivaccine messages on Instagram.

In total, 2 methods were included as separate groups despite their overlap with other methods. For example, all the studies likely contained a descriptive portion in their results. As such, descriptive studies were those that were only descriptive of their categories of interest, sentiment analysis aside. Examples include those describing group counts, changes over time, or other unique ways of data visualization [81,110,111,118,121,130]. Similarly, sentiment analysis studies sometimes include the development of a supervised machine learning model. Thus, the machine classifier method only contained 2% (1/54) of the studies that focused exclusively on machine classifying, which detailed the development of a classification model that identifies false HPV information [112].

Although diverse, computational studies also share a unifying theme with traditional method studies, which is the deficiency of the theoretical focus driving these studies. Even fewer studies using computational methods had a theoretical basis (6/54, 11%). Of the 6 studies that did, only 1 (17%) focused on a health behavior model [110], and the others used more generalized theories [81,97,108] and marketing [121].

XSL•FO
RenderX

Table 2. Computational analysis methods and media platforms (N=54).

Media platform	Analysis methods							
	Sentiment analysis	Topic modeling	Semantic network analysis	Projection	Feature extraction	Image analysis	Description[a]	Machine classifier[b]
Twitter	• Martin et al [81] • Liew and Lee [82] • Ajovalasit et al [83] • Blankenship et al [84] • Cotfas et al [85] • Deiner et al [86] • Du et al [87] • Gesualdo et al [88] • Hu et al [89] • Lyu et al [90] • Monselise et al [91] • Piedrahita-Valdés et al [92] • Shim et al [93] • Tavoschi et al [94] • Yan et al [95] • Yousefinaghani et al [96]	• Blankenship et al [84] • Cotfas et al [85] • Hu et al [89] • Lyu et al [90] • Monselise et al [91] • Shim et al [93] • Yan et al [95] • Argyris et al [97] • Dunn et al [98] • Guntuku et al [99] • Jiang et al [100]	• Martin et al [81] • Jiang et al [100] • Benis et al [101] • Boucher et al [102] • Featherston et al [103] • Germani et al [104] • Gunaratne et al [105] • Lutkenhaus et al [106] • Marcec and Likic [107]	• Pananos et al [108]	• Lyu et al [109]	—[c]	• Martin et al [81] • Guidry et al [110] • Kummervold et al [111]	• Ajovalasit et al [83] • Du et al [87] • Argyris et al [97] • Dunn et al [98] • Tomaszewski et al [112]
Weibo	• Chen et al [113] • Zhang et al [114]	• Hu et al [115]	—	• Yin et al [116]	—	—	—	—
Facebook	• Deiner et al [86] • Klimiuk et al [117] • Schmidt et al [118] • Zhang et al [119]	• Schmidt et al [118] • Zhang et al [119]	• Bar-Lev et al [120]	• Bar-Lev et al [120]	—	—	• Furini [121]	—
Parler	—	• Baines et al [122]	—	—	—	—	—	—

XSL•FO
RenderX

Media platform	Analysis methods							
	Sentiment analysis	Topic modeling	Semantic network analysis	Projection	Feature extraction	Image analysis	Description[a]	Machine classifier[b]
Forum, blog, or Reddit	• Martin et al [81] • Yan et al [95] • Melton et al [123]	• Yan et al [95]	• Martin et al [81]	—	—	—	• Martin et al [81]	—
Websites (assorted)	• Kang et al [124]	• Okuhara et al [125]	• Bar-Lev et al [120] • Kang et al [124] • Cafiero et al [126]	• Bar-Lev et al [120]	—	—	—	—
Web-based news or media cloud	• Getman et al [127]	—	• Getman et al [127]	—	—	—	—	—
Baidu, Google, or Naver	• Chen et al [113] • Lee et al [128] • Porreca et al [129]	—	• Porreca et al [129]	—	—	—	—	—
Multiple	• Powell et al [130]	• DeDominicis et al [131]	• DeDominicis et al [131]	—	—	—	• Powell et al [130]	—
Instagram	• Lee et al [128]	—	—	—	—	• Wang et al [132]	—	—
Q&A[d] site	• Luo et al [133]	• Luo et al [133]	—	—	—	—	—	—
YouTube	• Porreca et al [129]	—	• Porreca et al [129]	—	—	—	—	—

[a]Nearly all studies included descriptive statistics of their data set and outcome of interest. The descriptive classification included studies that were descriptive of categories other than sentiment (eg, categories of vaccine confidence or vaccine stance).

[b]Machine classifier studies only developed a classifier without further analysis. Often, but not always, studies using sentiment analysis or topic modeling also used machine classifiers; however, this table does not distinguish this.

[c]Not available. No studies exist using this social media platform and analysis method.

[d]Q&A: question and answer.

Themes

Although the studies were diverse in methods and their outcomes of focus, five themes were distilled that summarize this diversity: antivaccination themes, provaccination themes, framing, coverage and activity, and response of activity to certain events.

A set of studies (39/125, 31.2%) focused on what antivaccination topics arose. The most commonly recurring theme was a distrust of government institutions [18,20,21,36,37,43,70,75,76,81,102,124,131] or health institutions [32,35,36,49,67,93,100,125] or the idea of pharmaceutical companies profiteering off individuals [20,29,122,125]. This spilled into a related conversation about the infringement of civil liberties when individuals feel they are forced or mandated to receive a vaccine [37,46,47,64,73,131]. Often, the narratives can also be full of misinformation [27,73] or conspiracy theories [35,37,58,70,100,117], both featuring heavily when in antivaccination messages and accompanied by anger, fear, or frustration [96]. These sentiments were also paralleled by a

XSL•FO
RenderX

general concern about specific vaccines themselves, especially in relation to their overall perceived safety or efficacy (including side effects) [35,38,44,45,51,75,81,93,102,117,122,123,125,128], their constitution or ingredients [18], the adverse events around them [18,19,27,61,66], and their unnaturalness [32,47].

By contrast, other studies (8/125, 6.4%) focused on provaccination topics that emerged (although in fewer studies, understandably so as vaccine hesitancy was the focus). The most common theme was the use of scientific research and a constant reinforcement of vaccine safety and efficacy [19,62,77,100,125]. The second most common theme was how having an empathetic connection may lead to perception of vaccines in a more positive light. A total of 0.8% (1/125) of the studies found that knowing an afflicted person with HPV appeared in provaccination messages [38]. Another study (1/125, 0.8%) found that, for childhood vaccinations, vaccine advocates focused on the impact of vaccine hesitancy on children to encourage others to vaccinate their children [131].

The existence of antivaccination and provaccination topics alludes to the importance of who is delivering a message and how it is delivered [54]—captured in the third theme, framing. Regarding who delivers the message, general practitioners are used so that transmitted messages are more reliable [44,54,129] and engaged with [133], and sources from governments or professional associations are most used for credibility or transparency [56,63]. The opposite is true, where negative information is usually associated with less professional institutions [51,52]. This may be especially important in a landscape in which posts or content are likely generated by lay consumers or users [22,125,133]. The use of parents and mothers as messengers elicits a better generation of concern [49]. In addition, writers and journalists influenced by both provaccination and antivaccination camps are shown to continually reignite the debate on vaccination [106]. Regarding how the message is conveyed, personal stories, which are shown to be more engaging [39], are a tool used by both sides to enforce their viewpoints as correct [24] (such as the use of anecdotes on antivaccination websites [46] or the use of personal stories to encourage positive vaccination dialogue [39,42]). Another tactic used for framing, especially from the antivaccination side, is the use of shocking images or appeals to emotion through testimony to convince others of the antivaccination agenda [47,104]. Often, these antivaccination messages misuse scientific evidence [121,127] and loss-framed messaging [112] to transmit their ideas. These tactics may allude to a more generalized use of risk-amplifying messages to elicit reactions [41]. Framing also inadvertently occurs when using certain terms. In total, 0.8% (1/125) of the studies looked at how antivaccination characterizes vaccine-hesitant groups as ignorant, deviant, lacking access to vaccination (as opposed to being unwilling), pitied, and needing help [65]. In summary, who delivers the message, their background, and how they say it are all important in vaccine hesitancy research.

Closely related to framing is the relative amount of coverage, activity, or engagement on the web of provaccination and antivaccination communities. Most studies in this theme (12/125, 9.6%) found that any negative or antivaccination coverage or messages were generally more prevalent and engaged with (shared, viewed, retweeted, and liked) [21,25,29-31,33,50,55,68,71,78,84]. There were 1.6% (2/125) of studies in the opposite direction, finding that positive vaccination messages received more engagement [58,96] despite the existence of a higher quantity of antivaccination videos. Some studies (8/125, 6.4%) went further to establish an association between coverage—both the type and amount of coverage—and vaccine uptake. In total, 0.8% (1/125) of the studies found that a higher number of tweets, Facebook posts, and internet searches in an area were associated with lower measles, mumps, and rubella vaccine coverage [26]. This was corroborated by 1.6% (2/125) of the studies: an infodemic study that found an association between higher social media traffic and higher hesitancy [120] and a study that found that more exposure to HPV-related tweets explained variance in coverage [134]. Another study (1/125, 0.8%) found that more negative coverage meant less uptake of childhood vaccination [72]. This was corroborated by 1.6% (2/125) of the studies—a study looking at how adverse event reporting meant less vaccination [53] and a study that showed that discourse on HPV vaccines focusing on negative tones was correlated with more barriers to HPV vaccination [40]. However, the opposite was found in a Chinese study, which noted that increasing vaccine-related discussions correlated with an increasing number of vaccinated individuals [95]. Another study (1/125, 0.8%) found that more tailored messages to specific communities would lead to higher proactiveness in certain parts of the population to get vaccinated [101]. Another set of studies (9/125, 7.2%) looked at how vaccine-discussing communities engaged with each other. An example of this is the finding that antivaccination groups discussed vaccination issues much earlier [34]; are deeply fragmented in their beliefs, which spiral into radical communities [126]; and are part of a larger robust network of vaccine-hesitant individuals [97,98,127]. This robustness is also found in provaccination networks [124]. Overall, vaccines are a very polarizing topic, partly because of the ideological isolation and minimal interaction between provaccination and antivaccination groups [105], as well as other minority groups [99], and the existence of echo chambers that arise because of selective consumption of vaccine information [118].

The last theme captures how discussion of vaccines clusters around events, indicating a reactive public over time [64,87,89-91,94,107]. Overall, the conversation around vaccines usually follows certain occurrences or events in what is termed as crisis phases by Furini [121]. Diaz et al [57] found that there was increased search activity regarding vaccines and infertility following the US Center for Disease Control and Prevention emergency approval of COVID-19 vaccines. Interactions on Twitter increased in response to political events, suggesting disorientation [83,85]. Mahroum et al [59] found that, in an influenza vaccine scandal (the Fluad case), regions affected by the scandal had more related web search activity, suggesting a localized search behavior. Odone et al [69] corroborate this by highlighting that reports of deaths were the main signal that prompted more searches on the topic. A similar finding was also noted by Deiner et al [86], who showed that provaccination posts were correlated with a reporting of US cases (with antivaccination posts constantly happening in the background). There is also a focus on the associations of this increased

XSL•FO
RenderX

activity. Chen et al [113] looked at how the vaccine crisis of the Kangtai hepatitis B virus raised public attention and negative sentiments on the web in China. Dunn et al [134] found that exposure to more HPV-related tweets explained a variance in coverage of the HPV vaccine. Adverse event reporting also produces a more emotional response that leads to a decline in positive sentiments about vaccines [114]. Another set of studies (2/125, 1.6%) looked at the content of the messages, which overlaps with the aforementioned negative topic theme. In total, 0.8% (1/125) of the studies looked at how, during the peak season for influenza, more conspiracy theories about vaccination would occur [110]. Another study (1/125, 0.8%) looked at how the public had episodic expressions of distrust toward the Chinese government immediately after a vaccine-related scandal [115]. Although this theme discusses how the public reacts, there is also considerable overlap with the other themes in terms of what is being said as a reaction.

Discussion

Principal Findings

This review consolidated the current literature on the use of media data—both traditional and social media—to study vaccine hesitancy. This was done through three objectives: (1) summarizing media platforms; (2) summarizing analysis methods; and (3) understanding how the included studies build or contribute to the body of knowledge of the media's influence on vaccine hesitancy and, thereby, on public health. In doing so, this study aimed to bridge the fields of health behavior, computer science, and public health. A total of 125 studies were included, of which 71 (56.8%) used traditional research methods and 54 (43.2%) used big data (computational methods). The studies focused on the following five themes: identifying antivaccination topics; identifying provaccination topics; framing (who says what and how); the coverage, activity, and engagement in provaccination and antivaccination communities; and how the public reacts to events.

Overall, there is plurality in the analytical methods used. Several methods prevailed. For the traditional methods, most studies (43/71, 61%) focused on using content analysis, thematic analysis, or framing analysis, with other methods such as sentiment, stance, tone, or polarity coding also being popular. This preference was extended, perhaps naturally because of momentum in the field of vaccine hesitancy to focus on sentiment and topics, to studies using computational methods. Studies using network analysis and feature extraction were present but fewer (16/54, 30%). This could be due to a time lag in the arrival of big data analysis tools for academic research in this direction. Interestingly, all studies using computational methods (54/125, 43.2%) were published in or after 2016, indicating a relatively recent interest in this area. In the coming years, there may be growth in the computational field, especially regarding more advanced network analyses and feature extraction. This growth offers new insights to researchers, enabling them to reach new conclusions and challenge existing theories, thereby revolutionizing the way vaccine hesitancy studies are conducted.

However, this revolution is not only due to advances in computing. In parallel, the creation of new platforms will also shape the ways in which users engage with information. The different platforms used in the included studies span blogging sites, microblogging sites, newspapers, image-based social media platforms (Instagram), video-based social media platforms (YouTube), search engines, and question-answering sites. The growth of live streaming on platforms such as Instagram reels, TikTok, Bilibili (Chinese video streaming platform), and Twitch is likely to pivot analysis methods in the direction of *computer vision*, and preferences for more advanced methods may follow suit. In this review, this shift was observed. The studies using manual methods (71/125, 56.8%) focused more on traditional media, whereas those using computational methods targeted social media and microblogging platforms. Thus, the diversification of platforms parallels the advances in methods. Together, their parallel growth synergistically shapes the epistemological paradigms of media use in vaccine hesitancy research.

Despite fervor on the growth of this field, a glaring shortcoming misroutes it—a lack of theoretical foundation. Missing a theoretical focus portends the use of methods only for the sake of novelty and not necessarily informativeness. A corroborating finding speaking to this point is some studies' justification of publication on the grounds of a novel approach to data analysis when the analysis only applied methods to a different data source or platform. Another corroboration is the inadvertent lack of computational methods used to analyze traditional media, possibly because of the attractiveness of big data methods (ie, preferences to analyze social media because of novelty). Although this contributes to an overall body of knowledge in vaccine hesitancy research, it disorganizes the trajectory of the field as findings are not built on the cornerstones already set by theories in health behavior, vaccine hesitancy, and public health. Thus, it makes it difficult to draw any conclusive findings on the media's real influence on vaccine hesitancy as measured variables and outcomes differ. Using a theory-driven approach can counter this trend, making the consolidation of findings more cogent. By anchoring these studies on health behavior or information proliferation theories, the parallel development of media data and public health research can be bridged while simultaneously addressing the blind spot of theoretical weakness.

Few studies in this review (19/125, 15.2%) exhibited a theory-driven approach. Bradshaw et al [73] used social influence theory to guide the discussion on how antivaccination advocates on Facebook inadvertently used informational and normative influence processes to shape first-time mothers' vaccination sentiments. The discussion extended to how the Facebook network, being geographically unrestricted, may promote vaccine refusal in line with digital identity formation, expanding the realm of influence on vaccination. Aechtner [43] focused on persuasion cues derived from the ELM for persuasion to label and guide discussion of an Australian countervaccine lobby group. In total, 1.6% (2/125) of the studies, conducted by Luisi [40,41], also used two of the most prominent theories in health behavior and psychology to guide coding. Of these 2 studies, 1 (50%) looked at the amplification potential of

XSL•FO
RenderX

messages by measuring the concepts of the SARF on Facebook posts [41]. The other study used a similar method but used the HBM to guide the labeling of the concepts present in the messages on Facebook [40]. Pananos et al [108] took an entirely different approach, not using a health behavior model but rather one from mathematics. In their study, they used the theory of critical transitions and Twitter data to predict how critical periods in the vaccination course (around a “tipping point”) may affect the course of epidemics. If a study did not explicitly invoke a theory, it could arrive at one or more conclusions that were captured in one or more theoretical concepts. One example of this is the study finding that there are emotion-based risk expressions in antivaccination groups (risk as an emotion concept) [96]. These studies are only a sample of what can be done with theoretical guidance.

There are 2 additional implications of a theory-based approach. The first is that these novel methods in media analysis are unlikely to replace existing methods in vaccine hesitancy research; rather, they are an extension and complement to them. Survey methods have been validated in public health for the past 50 years in its research, and guided questions have been drafted to draw conclusions on the complex relationship among factors that drive behaviors. As such, there are some conclusions drawn from survey data that are difficult to obtain using media data. An example is the causative analysis of vaccine perceptions and uptake. Media data are just 1 factor in a complex information network, and there are confounding issues (demographics, preconceived beliefs, and heuristics, to name a few) in drawing causative conclusions on individual or population vaccine hesitancy because of exposure to information. From this review, it is apparent that there is a paucity of studies exploring associative links between a specific media channel and vaccine uptake. Thus, media data analyses will likely only complement the existing public health research paradigm until more advances are made. The second implication is that refocusing on theory (a defocusing on methods) allows for a better identification of gaps in the literature. Researchers are better able to identify which platforms, concepts, or relationships need stronger testing and empirical support if structured by a framework. These 2 implications delineate the scope of what media studies in vaccine research accomplish in terms of pushing forward the vaccine hesitancy research agenda.

Future Directions

On the basis of these themes, there are several open research areas for further exploration. The first is to understand how trust and distrust toward institutions (government and health care) may influence vaccination. A common theme of antivaccination worldwide appears to be rooted in distrust and suspicion—which translates to fear or disobedience—on the part of the public. This may translate to conspiracy theories and misinformation within antivaccination communities. Although media data can aid in the identification and classification of topics and understanding how they spread in networks, there is pending work on understanding the association between trust and adherence to public health measures. Second, and closely related, is research on understanding how misinformation spreads. This field of work will likely involve health psychologists, computer scientists, public health experts, and media researchers as it involves understanding how information signals are generated, spread, and processed; what signals are important in shaping risk perception; and how the timing of this matters. This field of misinformation and understanding how to combat it, with the implications for public health, will be a huge challenge in the era of big data and public health. The last area is the effective communication of governments and the pharmaceutical industry in addressing any vaccine concerns, from constitution and side effects to any other vaccine-related events. Evaluations of governments on vaccine communication should be performed and benchmarked against WHO-prescribed standards such as those laid out in the COVID-19 Vaccine Safety Manual [141] or the Managing Vaccine-Related Events guide [142], with the aim of identifying successful case studies on vaccine communications. These are several areas of suggested research on vaccine hesitancy moving forward.

Limitations

There are several limitations to this study. First, data were extracted by only 1 reviewer. This affects the inclusion criteria and extraction process through a combination of selection bias and manpower limitations. A predefined standardized extraction form partially diminished any biases in data extraction. In addition, as the review only consolidates and describes platforms, methods, and contributions to the field of study without concluding about results or effect size, the introduced bias has a marginal influence on the findings.

The second limitation is the left censorship of year in the search criteria. By including studies only conducted after 2010, there is a stronger representation of studies that used social media platforms and computational methods. As vaccine hesitancy and traditional media analysis are not new issues (ie, they were present before social media), there are relevant studies that have not been included. However, this is intentional. An objective was to have a closer look at the diversity in platforms and methods in recent years. Imposing a time restriction homed in on reaching this objective. Regarding concerns about the representativeness of the included studies, there already was an emerging trend in preference for platform and analysis method without the necessity to include every study (ie, a saturation in data findings). This saturation also diminished the biases of only having 1 reviewer.

However, this saturation in data does not preclude that rapid changes in the field will produce new uses of platforms and analyses methods, especially as new developments happen in the fields of computer science and natural language processing. The third limitation extending from the second is the inclusion of studies only in English. There is evidence in the review that analyses in other widely spoken languages such as Chinese and French are emerging. The language of publication is important as the foundation of media studies and natural *language* processing tries to parse meaning from language, with different languages analyzed through a different set of linguistic tools. These different tools, coupled with an inherent difference in language structure, may reveal alternative approaches to distill meaning and connotations from words. Furthermore, studies using non-English languages to analyze vaccine hesitancy could also have implications for global health as many of these

XSL•FO
RenderX

non-English languages are spoken by a large portion of the global population (eg, Chinese and Spanish). For these reasons, excluding non-English papers biases the comprehensiveness of the methods and platforms presented.

The last major limitation is the exclusion of studies that used surveys or cross-sectional data. This was explicitly included in the search terms to exclude studies that used surveys to ask about the use of media or the effects of media and to focus the body of studies on those that only used media studies as the main source of data. Although successful, this search excluded studies that used both survey and media data to study vaccine hesitancy. Thus, this major limitation restricts the comprehensiveness of the included studies, and a separate scoping review assessing the dual use of traditional tools and media data is required.

Conclusions

Our findings illustrate a variegation of media platforms and analysis methods for vaccine hesitancy research as well as 5 themes of focus. The first was the focus of antivaccination themes on the distrust of institutions, violations of civil liberties, the spread of misinformation and conspiracy theories, and concerns about specific vaccines. The second was the focus of provaccination themes on the use of scientific literature to support vaccine safety. The third was the importance of who delivers the message and how the way it is framed shapes the reception of vaccine opinion. The fourth was that coverage mostly centers on negative content and also circulates within echo chambers in both vaccination camps, indicating deeply fractured communities. The last theme was that the public responds to focusing events, suggesting volatile periods in which misinformation and conspiracy information can circulate. Despite the diversity in study types and platforms, these findings are consistent across both traditional and computational methods.

This burgeoning field—known as *digital epidemiology* or *infodemiology*—will continue diversifying as new media platforms arise and more tools from computer science trickle and become commonplace in public health research. This heterogeneity, although inspiring for new avenues of research, should also be met with cautious excitement. Researchers inclined to join this field should fully understand that media data analysis methods are meant to supplement—not supplant—current practices in public health research. A way to ensure this understanding is to establish a theoretical focus of the research before method or platform selection. In doing so, the mentality of adopting trending methods is avoided, there is a systematic consolidation in the synthesis of findings, and a coherent paradigm in the subfield of media data research on vaccine hesitancy can be established.

Acknowledgments

The author acknowledges Professor Gabriel Leung and Dr Thorin Jake Duffin for their proofreading.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[PDF File (Adobe PDF File), 101 KB - infodemiology_v2i2e37300_app1.pdf]

Multimedia Appendix 2
Search terms and dates.
[PDF File (Adobe PDF File), 16 KB - infodemiology_v2i2e37300_app2.pdf]

Multimedia Appendix 3
Detailed data extraction for traditional method studies.
[PDF File (Adobe PDF File), 199 KB - infodemiology_v2i2e37300_app3.pdf]

Multimedia Appendix 4
Detailed data extraction for computational method studies.
[PDF File (Adobe PDF File), 188 KB - infodemiology_v2i2e37300_app4.pdf]

References

1. Finnegan J, Viswanath K. Communication theory and health behavior change: the media studies framework. In: Health Behavior and Health Education Theory, Research, and Practice. 4th edition. Hoboken, New Jersey, United States: Wiley; 2008.
2. Champion V, Skinner C. The health belief model. In: Health Behavior and Health Education: Theory, Research, and Practice. Hoboken, New Jersey: Jossey-Bass; 2008.

XSL•FO
RenderX

3. Godin G, Kok G. The theory of planned behavior: a review of its applications to health-related behaviors. Am J Health Promot 1996 Aug 25;11(2):87-98. [doi: 10.4278/0890-1171-11.2.87] [Medline: 10163601]
4. Ajzen I, Fishbein M. Understanding Attitudes and Predicting Social Behavior. Hoboken, New Jersey, United States: Prentice-Hall; 1980.
5. Petty R, Barden J, Wheeler S. The elaboration likelihood model of persuasion: developing health promotions for sustained behavioral change. In: Emerging Theories in Health Promotion Practice and Research (2nd Edition). San Francisco: Jossey-Bass; 2002.
6. Popova L. The extended parallel process model: illuminating the gaps in research. Health Educ Behav 2012 Aug 14;39(4):455-473. [doi: 10.1177/1090198111418108] [Medline: 22002250]
7. Lindell MK, Perry RW. The protective action decision model: theoretical modifications and additional evidence. Risk Anal 2012 Apr;32(4):616-632. [doi: 10.1111/j.1539-6924.2011.01647.x] [Medline: 21689129]
8. Cappella J. Integrating message effects and behavior change theories: organizing comments and unanswered questions. J Commun 2006;56(suppl_1):S265-S279. [doi: 10.1111/j.1460-2466.2006.00293.x]
9. McCombs M, Shaw D. The agenda-setting function of mass media. Public Opin Q 1972;36(2):176-187 [FREE Full text] [doi: 10.1086/267990]
10. Kasperson RE, Renn O, Slovic P, Brown HS, Emel J, Goble R, et al. The social amplification of risk: a conceptual framework. Risk Anal 1988 Jun;8(2):177-187. [doi: 10.1111/j.1539-6924.1988.tb01168.x]
11. Reese S. The framing project: a bridging model for media research revisited. J Commun 2007;57(1):148-154. [doi: 10.1111/j.1460-2466.2006.00334.x]
12. 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]
13. 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 2022-02-10]
14. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Social Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
15. JBI Manual for Evidence Synthesis. Adelaide, SA, Australia: JBI Global; 2020.
16. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
17. Becker BF, Larson HJ, Bonhoeffer J, van Mulligen EM, Kors JA, Sturkenboom MC. Evaluation of a multinational, multilingual vaccine debate on Twitter. Vaccine 2016 Dec 07;34(50):6166-6171. [doi: 10.1016/j.vaccine.2016.11.007] [Medline: 27840012]
18. 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]
19. Criss S, Nguyen TT, Norton S, Virani I, Titherington E, Tillmanns EL, et al. Advocacy, hesitancy, and equity: exploring U.S. Race-related discussions of the COVID-19 vaccine on Twitter. Int J Environ Res Public Health 2021 May 26;18(11):5693 [FREE Full text] [doi: 10.3390/ijerph18115693] [Medline: 34073291]
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. 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]
22. Keim-Malpass J, Mitchell EM, Sun E, Kennedy C. Using Twitter to understand public perceptions regarding the #HPV vaccine: opportunities for public health nurses to engage in social marketing. Public Health Nurs 2017 Jul 06;34(4):316-323. [doi: 10.1111/phn.12318] [Medline: 28261846]
23. Marchetti F, Verazza S, Brambilla M, Restivo V. Rotavirus and the web: analysis of online conversations in Italy during 2020. Hum Vaccin Immunother 2022 Dec 31;18(1):2002087 [FREE Full text] [doi: 10.1080/21645515.2021.2002087] [Medline: 34856884]
24. Sundstrom B, Cartmell KB, White AA, Well H, Pierce JY, Brandt HM. Correcting HPV vaccination misinformation online: evaluating the social media campaign. Vaccines (Basel) 2021 Apr 06;9(4):352 [FREE Full text] [doi: 10.3390/vaccines9040352] [Medline: 33917512]
25. Gori D, Durazzi F, Montalti M, Di Valerio Z, Reno C, Fantini MP, et al. Mis-tweeting communication: a Vaccine Hesitancy analysis among twitter users in Italy. Acta Biomed 2021 Oct 05;92(S6):e2021416 [FREE Full text] [doi: 10.23750/abm.v92iS6.12251] [Medline: 34739459]
26. Aquino F, Donzelli G, De Franco E, Privitera G, Lopalco PL, Carducci A. The web and public confidence in MMR vaccination in Italy. Vaccine 2017 Aug 16;35(35 Pt B):4494-4498. [doi: 10.1016/j.vaccine.2017.07.029] [Medline: 28736200]
27. Basch CH, Zybert P, Reeves R, Basch CE. What do popular YouTube videos say about vaccines? Child Care Health Dev 2017 Jul 19;43(4):499-503. [doi: 10.1111/cch.12442] [Medline: 28105642]

XSL•FO
RenderX

28. 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 [FREE Full text] [doi: 10.1186/s12889-022-13265-y] [Medline: 35484619]
29. Lahouati M, De Coucy A, Sarlangue J, Cazanave C. Spread of vaccine hesitancy in France: what about YouTube™? Vaccine 2020 Aug 10;38(36):5779-5782. [doi: 10.1016/j.vaccine.2020.07.002] [Medline: 32682617]
30. Covolo L, Ceretti E, Passeri C, Boletti M, Gelatti U. What arguments on vaccinations run through YouTube videos in Italy? A content analysis. Hum Vaccin Immunother 2017 Jul 03;13(7):1693-1699 [FREE Full text] [doi: 10.1080/21645515.2017.1306159] [Medline: 28362544]
31. Donzelli G, Palomba G, Federigi I, Aquino F, Cioni L, Verani M, et al. Misinformation on vaccination: a quantitative analysis of YouTube videos. Hum Vaccin Immunother 2018 Jul 03;14(7):1654-1659 [FREE Full text] [doi: 10.1080/21645515.2018.1454572] [Medline: 29553872]
32. Bradshaw AS, Shelton SS, Wollney E, Treise D, Auguste K. Pro-Vaxxers get out: anti-vaccination advocates influence undecided first-time, pregnant, and new mothers on Facebook. Health Commun 2021 May 10;36(6):693-702. [doi: 10.1080/10410236.2020.1712037] [Medline: 31920115]
33. 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]
34. Kalichman S, Eaton L, Earnshaw V, Brousseau N. Faster than warp speed: early attention to COVD-19 by anti-vaccine groups on Facebook. J Public Health (Oxf) 2022 Mar 07;44(1):e96-105 [FREE Full text] [doi: 10.1093/pubmed/fdab093] [Medline: 33837428]
35. Orr D, Baram-Tsabari A, Landsman K. Social media as a platform for health-related public debates and discussions: the Polio vaccine on Facebook. Isr J Health Policy Res 2016 Nov 10;5(1):34 [FREE Full text] [doi: 10.1186/s13584-016-0093-4] [Medline: 27843544]
36. Tustin JL, Crowcroft NS, Gesink D, Johnson I, Keelan J, Lachapelle B. User-driven comments on a Facebook advertisement recruiting Canadian parents in a study on immunization: content analysis. JMIR Public Health Surveill 2018 Sep 20;4(3):e10090 [FREE Full text] [doi: 10.2196/10090] [Medline: 30249585]
37. Wawrzuta D, Jaworski M, Gotlib J, Panczyk M. What arguments against COVID-19 vaccines run on Facebook in Poland: content analysis of comments. Vaccines (Basel) 2021 May 10;9(5):481 [FREE Full text] [doi: 10.3390/vaccines9050481] [Medline: 34068500]
38. Wiyeh AB, Cooper S, Jaca A, Mavundza E, Ndwandwe D, Wiysonge CS. Social media and HPV vaccination: unsolicited public comments on a Facebook post by the Western Cape Department of Health provide insights into determinants of vaccine hesitancy in South Africa. Vaccine 2019 Oct 08;37(43):6317-6323. [doi: 10.1016/j.vaccine.2019.09.019] [Medline: 31521412]
39. Loft LH, Pedersen EA, Jacobsen SU, Søborg B, Bigaard J. Using Facebook to increase coverage of HPV vaccination among Danish girls: an assessment of a Danish social media campaign. Vaccine 2020 Jun 26;38(31):4901-4908. [doi: 10.1016/j.vaccine.2020.04.032] [Medline: 32362529]
40. Luisi ML. From bad to worse: the representation of the HPV vaccine Facebook. Vaccine 2020 Jun 15;38(29):4564-4573. [doi: 10.1016/j.vaccine.2020.05.016] [Medline: 32417141]
41. Luisi ML. From bad to worse II: risk amplification of the HPV vaccine on Facebook. Vaccine 2021 Jan 08;39(2):303-308. [doi: 10.1016/j.vaccine.2020.11.065] [Medline: 33303232]
42. Pedersen EA, Loft LH, Jacobsen SU, Søborg B, Bigaard J. Strategic health communication on social media: insights from a Danish social media campaign to address HPV vaccination hesitancy. Vaccine 2020 Jun 26;38(31):4909-4915. [doi: 10.1016/j.vaccine.2020.05.061] [Medline: 32482460]
43. Aechtner T. Distrust, danger, and confidence: a content analysis of the Australian Vaccination-Risks Network Blog. Public Underst Sci 2021 Jan 19;30(1):16-35. [doi: 10.1177/0963662520963258] [Medline: 33073708]
44. Bruel S, Peyrard-Chevrier K, Ginzarly M, Frappé P, Savall A. Human papillomavirus (HPV) vaccination: what can be found on the Web? Qualitative analysis of the Doctissimo.fr forum data. Rev Epidemiol Sante Publique 2020 Feb;68(1):25-31. [doi: 10.1016/j.respe.2019.09.005] [Medline: 31718831]
45. Larson HJ, Smith DM, Paterson P, Cumming M, Eckersberger E, Freifeld CC, et al. Measuring vaccine confidence: analysis of data obtained by a media surveillance system used to analyse public concerns about vaccines. Lancet Infect Dis 2013 Jul;13(7):606-613. [doi: 10.1016/s1473-3099(13)70108-7]
46. Moran MB, Lucas M, Everhart K, Morgan A, Prickett E. What makes anti-vaccine websites persuasive? A content analysis of techniques used by anti-vaccine websites to engender anti-vaccine sentiment. J Commun Healthc 2016 Oct 03;9(3):151-163. [doi: 10.1080/17538068.2016.1235531]
47. Nugier A, Limousi F, Lydié N. Vaccine criticism: presence and arguments on French-speaking websites. Med Mal Infect 2018 Feb;48(1):37-43. [doi: 10.1016/j.medmal.2017.09.003] [Medline: 29031651]
48. Toth C. Repertoires of vaccine refusal in Romania. Vaccines (Basel) 2020 Dec 13;8(4):757 [FREE Full text] [doi: 10.3390/vaccines8040757] [Medline: 33322096]

XSL•FO
RenderX

49. Ward S, Budarick J. Adapting to the postmodern medical paradigm: how The Daily Telegraph used emotion and anecdote to legitimise vaccination. Media Int Aust 2020 Nov 29;181(1):103-119. [doi: 10.1177/1329878x20975786]
50. Karapetiantz P, Audeh B, Lillo-Le Louët A, Bousquet C. Discrepancy between personal experience and negative opinion with human papillomavirus vaccine in web forums. Stud Health Technol Inform 2020 Jun 26;272:417-420. [doi: 10.3233/SHTI200584] [Medline: 32604691]
51. Panatto D, Amicizia D, Arata L, Lai PL, Gasparini R. A comprehensive analysis of Italian web pages mentioning squalene-based influenza vaccine adjuvants reveals a high prevalence of misinformation. Hum Vaccin Immunother 2018 Apr 03;14(4):969-977 [FREE Full text] [doi: 10.1080/21645515.2017.1407483] [Medline: 29172967]
52. Shoup JA, Narwaney KJ, Wagner NM, Kraus CR, Gleason KS, Albright K, et al. Social media vaccine websites: a comparative analysis of public and moderated websites. Health Educ Behav 2019 Jun;46(3):454-462 [FREE Full text] [doi: 10.1177/1090198118818253] [Medline: 30596265]
53. 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]
54. Sharon AJ, Yom-Tov E, Baram-Tsabari A. Vaccine 2020 Mar 10;38(12):2691-2699. [doi: 10.1016/j.vaccine.2020.02.010] [Medline: 32061384]
55. Ruiz JB, Bell RA. Understanding vaccination resistance: vaccine search term selection bias and the valence of retrieved information. Vaccine 2014 Oct 07;32(44):5776-5780. [doi: 10.1016/j.vaccine.2014.08.042] [Medline: 25176640]
56. Sajjadi NB, Shepard S, Ottwell R, Murray K, Chronister J, Hartwell M, et al. Examining the public's most frequently asked questions regarding COVID-19 vaccines using search engine analytics in the United States: observational study. JMIR Infodemiology 2021 Aug 4;1(1):e28740 [FREE Full text] [doi: 10.2196/28740] [Medline: 34458683]
57. Diaz P, Reddy P, Ramasahayam R, Kuchakulla M, Ramasamy R. COVID-19 vaccine hesitancy linked to increased internet search queries for side effects on fertility potential in the initial rollout phase following Emergency Use Authorization. Andrologia 2021 Oct 28;53(9):e14156 [FREE Full text] [doi: 10.1111/and.14156] [Medline: 34181273]
58. Guidry JP, Carlyle K, Messner M, Jin Y. On pins and needles: how vaccines are portrayed on Pinterest. Vaccine 2015 Sep 22;33(39):5051-5056. [doi: 10.1016/j.vaccine.2015.08.064] [Medline: 26319742]
59. Mahroum N, Watad A, Rosselli R, Brigo F, Chiesa V, Siri A, et al. An infodemiological investigation of the so-called "Fluad effect" during the 2014/2015 influenza vaccination campaign in Italy: ethical and historical implications. Hum Vaccin Immunother 2018 Mar 04;14(3):712-718 [FREE Full text] [doi: 10.1080/21645515.2017.1420448] [Medline: 29293392]
60. Ashwell D, Murray N. When being positive might be negative: an analysis of Australian and New Zealand newspaper framing of vaccination post Australia's No Jab No Pay legislation. Vaccine 2020 Jul 31;38(35):5627-5633 [FREE Full text] [doi: 10.1016/j.vaccine.2020.06.070] [Medline: 32653274]
61. Basch CH, Kecojevic A, Wagner VH. Reporting of recombinant adenovirus-based COVID-19 vaccine adverse events in online versions of three highly circulated US newspapers. Hum Vaccin Immunother 2021 Dec 02;17(12):5114-5119. [doi: 10.1080/21645515.2021.1979847] [Medline: 34714719]
62. Casciotti DM, Smith KC, Klassen AC. Topics associated with conflict in print news coverage of the HPV vaccine during 2005 to 2009. Hum Vaccin Immunother 2014 Feb 10;10(12):3466-3474 [FREE Full text] [doi: 10.4161/21645515.2014.979622] [Medline: 25668659]
63. Catalan-Matamoros D, Elías C. Vaccine hesitancy in the age of coronavirus and fake news: analysis of journalistic sources in the Spanish quality press. Int J Environ Res Public Health 2020 Nov 04;17(21):8136 [FREE Full text] [doi: 10.3390/ijerph17218136] [Medline: 33158059]
64. Colón-López V, Rivera-Figueroa V, Arroyo-Morales GO, Medina-Laabes DT, Soto-Abreu R, Rivera-Encarnación M, et al. Content analysis of digital media coverage of the human papillomavirus vaccine school-entry requirement policy in Puerto Rico. BMC Public Health 2021 Jul 01;21(1):1286 [FREE Full text] [doi: 10.1186/s12889-021-11311-9] [Medline: 34210308]
65. Court J, Carter SM, Attwell K, Leask J, Wiley K. Labels matter: use and non-use of 'anti-vax' framing in Australian media discourse 2008-2018. Soc Sci Med 2021 Dec;291:114502. [doi: 10.1016/j.socscimed.2021.114502] [Medline: 34715625]
66. Das MK, Singh D, Sharma S. Media news on vaccines and vaccination: the content profile, sentiment and trend of the online mass media during 2015–2020 in India. Clin Epidemiol Global Health 2021 Apr;10:100691. [doi: 10.1016/j.cegh.2020.100691]
67. Kummervold PE, Schulz WS, Smout E, Fernandez-Luque L, Larson HJ. Controversial Ebola vaccine trials in Ghana: a thematic analysis of critiques and rebuttals in digital news. BMC Public Health 2017 Aug 07;17(1):642 [FREE Full text] [doi: 10.1186/s12889-017-4618-8] [Medline: 28784109]
68. Meyer SB, Violette R, Aggarwal R, Simeoni M, MacDougall H, Waite N. Vaccine hesitancy and Web 2.0: exploring how attitudes and beliefs about influenza vaccination are exchanged in online threaded user comments. Vaccine 2019 Mar 22;37(13):1769-1774. [doi: 10.1016/j.vaccine.2019.02.028] [Medline: 30826142]

XSL•FO
RenderX

69. Odone A, Tramutola V, Morgado M, Signorelli C. Immunization and media coverage in Italy: an eleven-year analysis (2007-17). Hum Vaccin Immunother 2018 Jul 11;14(10):2533-2536 [FREE Full text] [doi: 10.1080/21645515.2018.1486156] [Medline: 29993316]
70. Olufowote JO. Local resistance to the global eradication of polio: newspaper coverage of the 2003-2004 vaccination stoppage in northern Nigeria. Health Commun 2011 Dec;26(8):743-753. [doi: 10.1080/10410236.2011.566830] [Medline: 21667367]
71. Stephenson N, Chaukra S, Katz I, Heywood A. Newspaper coverage of childhood immunisation in Australia: a lens into conflicts within public health. Critical Public Health 2018 Mar 07;28(4):472-483. [doi: 10.1080/09581596.2018.1446510]
72. Catalan-Matamoros D, Peñafiel-Saiz C. Exploring the relationship between newspaper coverage of vaccines and childhood vaccination rates in Spain. Hum Vaccin Immunother 2020 May 03;16(5):1055-1061 [FREE Full text] [doi: 10.1080/21645515.2019.1708163] [Medline: 32017659]
73. Bradshaw AS, Treise D, Shelton SS, Cretul M, Raisa A, Bajalia A, et al. Propagandizing anti-vaccination: analysis of vaccines revealed documentary series. Vaccine 2020 Feb 18;38(8):2058-2069. [doi: 10.1016/j.vaccine.2019.12.027] [Medline: 31980194]
74. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. Hum Vaccin Immunother 2021 Aug 03;17(8):2373-2377 [FREE Full text] [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
75. Ward JK, Peretti-Watel P, Larson HJ, Raude J, Verger P. Vaccine-criticism on the internet: new insights based on French-speaking websites. Vaccine 2015 Feb 18;33(8):1063-1070. [doi: 10.1016/j.vaccine.2014.12.064] [Medline: 25583384]
76. Yeung MW, Yau AH. 'This year's vaccine is only 10% effective': a study of public discourse on vaccine effectiveness in Hong Kong. J Commun Healthc 2020 Aug 20;14(1):78-89. [doi: 10.1080/17538068.2020.1809316]
77. Filice E, Dubé E, Graham JE, MacDonald NE, Bettinger JA, Greyson D, et al. Vaccination discourses among chiropractors, naturopaths and homeopaths: a qualitative content analysis of academic literature and Canadian organizational webpages. PLoS One 2020 Aug 12;15(8):e0236691 [FREE Full text] [doi: 10.1371/journal.pone.0236691] [Medline: 32785248]
78. Yiannakoulias N, Slavik CE, Chase M. Expressions of pro- and anti-vaccine sentiment on YouTube. Vaccine 2019 Apr 03;37(15):2057-2064. [doi: 10.1016/j.vaccine.2019.03.001] [Medline: 30862365]
79. Wong LP, Lin Y, Alias H, Bakar SA, Zhao Q, Hu Z. COVID-19 anti-vaccine sentiments: analyses of comments from social media. Healthcare (Basel) 2021 Nov 09;9(11):1530 [FREE Full text] [doi: 10.3390/healthcare9111530] [Medline: 34828576]
80. Zhou F, Zhang W, Cai H, Cao Y. Portrayals of 2v, 4v and 9vHPV vaccines on Chinese social media: a content analysis of hot posts on Sina Weibo. Hum Vaccin Immunother 2021 Nov 02;17(11):4433-4441. [doi: 10.1080/21645515.2021.1971016] [Medline: 34543155]
81. Martin S, Kilich E, Dada S, Kummervold PE, Denny C, Paterson P, et al. "Vaccines for pregnant women…?! Absurd" - mapping maternal vaccination discourse and stance on social media over six months. Vaccine 2020 Sep 29;38(42):6627-6637. [doi: 10.1016/j.vaccine.2020.07.072] [Medline: 32788136]
82. Liew TM, Lee CS. Examining the utility of social media in COVID-19 vaccination: unsupervised learning of 672,133 twitter posts. JMIR Public Health Surveill 2021 Nov 03;7(11):e29789 [FREE Full text] [doi: 10.2196/29789] [Medline: 34583316]
83. Ajovalasit S, Dorgali VM, Mazza A, d'Onofrio A, Manfredi P. Evidence of disorientation towards immunization on online social media after contrasting political communication on vaccines. Results from an analysis of Twitter data in Italy. PLoS One 2021 Jul 9;16(7):e0253569 [FREE Full text] [doi: 10.1371/journal.pone.0253569] [Medline: 34242253]
84. Blankenship E, Goff ME, Yin J, Tse ZT, Fu K, Liang H, et al. Sentiment, contents, and retweets: a study of two vaccine-related twitter datasets. Perm J 2018;22:17-138 [FREE Full text] [doi: 10.7812/TPP/17-138] [Medline: 29911966]
85. Cotfas L, Delcea C, Gherai R. COVID-19 vaccine hesitancy in the month following the start of the vaccination process. Int J Environ Res Public Health 2021 Oct 04;18(19):10438 [FREE Full text] [doi: 10.3390/ijerph181910438] [Medline: 34639738]
86. 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 17;25(3):1116-1132 [FREE Full text] [doi: 10.1177/1460458217740723] [Medline: 29148313]
87. Du J, Xu J, Song H, Tao C. Leveraging machine learning-based approaches to assess human papillomavirus vaccination sentiment trends with Twitter data. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):69 [FREE Full text] [doi: 10.1186/s12911-017-0469-6] [Medline: 28699569]
88. Gesualdo F, D'Ambrosio A, Agricola E, Russo L, Campagna I, Ferretti B, et al. How do Twitter users react to TV broadcasts dedicated to vaccines in Italy? Eur J Public Health 2020 Jun 01;30(3):510-515 [FREE Full text] [doi: 10.1093/eurpub/ckaa022] [Medline: 32073598]
89. Hu T, Wang S, Luo W, Zhang M, Huang X, Yan Y, et al. Revealing public opinion towards COVID-19 vaccines with twitter data in the United states: spatiotemporal perspective. J Med Internet Res 2021 Sep 10;23(9):e30854 [FREE Full text] [doi: 10.2196/30854] [Medline: 34346888]
90. Lyu JC, Han EL, Luli GK. COVID-19 vaccine-related discussion on Twitter: topic modeling and sentiment analysis. J Med Internet Res 2021 Jun 29;23(6):e24435 [FREE Full text] [doi: 10.2196/24435] [Medline: 34115608]

XSL•FO
RenderX

91. Monselise M, Chang C, 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]
92. Piedrahita-Valdés H, Piedrahita-Castillo D, Bermejo-Higuera J, Guillem-Saiz P, Bermejo-Higuera JR, Guillem-Saiz J, et al. Vaccine hesitancy on social media: sentiment analysis from June 2011 to April 2019. Vaccines (Basel) 2021 Jan 07;9(1):28 [FREE Full text] [doi: 10.3390/vaccines9010028] [Medline: 33430428]
93. Shim J, Ryu K, Lee SH, Cho E, Lee YJ, Ahn JH. Text mining approaches to analyze public sentiment changes regarding COVID-19 vaccines on social media in Korea. Int J Environ Res Public Health 2021 Jun 18;18(12):6549 [FREE Full text] [doi: 10.3390/ijerph18126549] [Medline: 34207016]
94. Tavoschi L, Quattrone F, D'Andrea E, Ducange P, Vabanesi M, Marcelloni F, et al. Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. Hum Vaccin Immunother 2020 May 03;16(5):1062-1069 [FREE Full text] [doi: 10.1080/21645515.2020.1714311] [Medline: 32118519]
95. Yan C, Law M, Nguyen S, Cheung J, Kong J. Comparing public sentiment toward COVID-19 vaccines across Canadian cities: analysis of comments on reddit. J Med Internet Res 2021 Sep 24;23(9):e32685 [FREE Full text] [doi: 10.2196/32685] [Medline: 34519654]
96. 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]
97. Argyris YA, Monu K, Tan P, Aarts C, Jiang F, Wiseley KA. Using machine learning to compare provaccine and antivaccine discourse among the public on social media: algorithm development study. JMIR Public Health Surveill 2021 Jun 24;7(6):e23105 [FREE Full text] [doi: 10.2196/23105] [Medline: 34185004]
98. 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]
99. Guntuku SC, Buttenheim AM, Sherman G, Merchant RM. Twitter discourse reveals geographical and temporal variation in concerns about COVID-19 vaccines in the United States. Vaccine 2021 Jul 05;39(30):4034-4038 [FREE Full text] [doi: 10.1016/j.vaccine.2021.06.014] [Medline: 34140171]
100. Jiang X, Su M, Hwang J, Lian R, Brauer M, Kim S, et al. Polarization over vaccination: ideological differences in Twitter expression about COVID-19 vaccine favorability and specific hesitancy concerns. Social Media Soc 2021 Sep 30;7(3):205630512110484. [doi: 10.1177/20563051211048413]
101. Benis A, Chatsubi A, Levner E, Ashkenazi S. Change in threads on twitter regarding influenza, vaccines, and vaccination during the COVID-19 pandemic: artificial intelligence-based infodemiology study. JMIR Infodemiology 2021 Oct 14;1(1):e31983 [FREE Full text] [doi: 10.2196/31983] [Medline: 34693212]
102. Boucher J, Cornelson K, Benham JL, Fullerton MM, Tang T, Constantinescu C, et al. Analyzing social media to explore the attitudes and behaviors following the announcement of successful COVID-19 vaccine trials: infodemiology study. JMIR Infodemiology 2021;1(1):e28800 [FREE Full text] [doi: 10.2196/28800] [Medline: 34447924]
103. Featherstone JD, Ruiz JB, Barnett GA, Millam BJ. Exploring childhood vaccination themes and public opinions on Twitter: a semantic network analysis. Telemat Inform 2020 Nov;54:101474. [doi: 10.1016/j.tele.2020.101474]
104. 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]
105. Gunaratne K, Coomes EA, Haghbayan H. Temporal trends in anti-vaccine discourse on Twitter. Vaccine 2019 Aug 14;37(35):4867-4871. [doi: 10.1016/j.vaccine.2019.06.086] [Medline: 31300292]
106. 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]
107. Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J 2022 Jul 09;98(1161):544-550 [FREE Full text] [doi: 10.1136/postgradmedj-2021-140685] [Medline: 34373343]
108. Pananos AD, Bury TM, Wang C, Schonfeld J, Mohanty SP, Nyhan B, et al. Critical dynamics in population vaccinating behavior. Proc Natl Acad Sci U S A 2017 Dec 26;114(52):13762-13767 [FREE Full text] [doi: 10.1073/pnas.1704093114] [Medline: 29229821]
109. Lyu H, Wang J, Wu W, Duong V, Zhang X, Dye TD, et al. Social media study of public opinions on potential COVID-19 vaccines: informing dissent, disparities, and dissemination. Intell Med 2022 Feb;2(1):1-12 [FREE Full text] [doi: 10.1016/j.imed.2021.08.001] [Medline: 34457371]
110. Guidry JP, Austin LL, O'Donnell NH, Coman IA, Lovari A, Messner M. Tweeting the #flushot: beliefs, barriers, and threats during different periods of the 2018 to 2019 flu season. J Prim Care Community Health 2020 Oct 22;11:2150132720932722 [FREE Full text] [doi: 10.1177/2150132720932722] [Medline: 33089738]
111. Kummervold PE, Martin S, Dada S, Kilich E, Denny C, Paterson P, et al. Categorizing vaccine confidence with a transformer-based machine learning model: analysis of nuances of vaccine sentiment in twitter discourse. JMIR Med Inform 2021 Oct 08;9(10):e29584 [FREE Full text] [doi: 10.2196/29584] [Medline: 34623312]

XSL•FO
RenderX

112. Tomaszewski T, Morales A, Lourentzou I, Caskey R, Liu B, Schwartz A, et al. Identifying false human papillomavirus (HPV) vaccine information and corresponding risk perceptions from Twitter: advanced predictive models. J Med Internet Res 2021 Sep 09;23(9):e30451 [FREE Full text] [doi: 10.2196/30451] [Medline: 34499043]
113. Chen B, Zhang JM, Jiang Z, Shao J, Jiang T, Wang Z, et al. Media and public reactions toward vaccination during the 'hepatitis B vaccine crisis' in China. Vaccine 2015 Apr 08;33(15):1780-1785. [doi: 10.1016/j.vaccine.2015.02.046] [Medline: 25731787]
114. Zhang Z, Feng G, Xu J, Zhang Y, Li J, Huang J, et al. The impact of public health events on COVID-19 vaccine hesitancy on Chinese social media: national infoveillance study. JMIR Public Health Surveill 2021 Nov 09;7(11):e32936 [FREE Full text] [doi: 10.2196/32936] [Medline: 34591782]
115. Hu D, Martin C, Dredze M, Broniatowski DA. Chinese social media suggest decreased vaccine acceptance in China: an observational study on Weibo following the 2018 Changchun Changsheng vaccine incident. Vaccine 2020 Mar 17;38(13):2764-2770 [FREE Full text] [doi: 10.1016/j.vaccine.2020.02.027] [Medline: 32093982]
116. Yin F, Wu Z, Xia X, Ji M, Wang Y, Hu Z. Unfolding the determinants of COVID-19 vaccine acceptance in China. J Med Internet Res 2021 Jan 15;23(1):e26089 [FREE Full text] [doi: 10.2196/26089] [Medline: 33400682]
117. Klimiuk K, Czoska A, Biernacka K, Balwicki Ł. Vaccine misinformation on social media - topic-based content and sentiment analysis of Polish vaccine-deniers' comments on Facebook. Hum Vaccin Immunother 2021 Jul 03;17(7):2026-2035 [FREE Full text] [doi: 10.1080/21645515.2020.1850072] [Medline: 33517844]
118. 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]
119. Zhang J, Xue H, Calabrese C, Chen H, Dang JH. Understanding human papillomavirus vaccine promotions and hesitancy in Northern California through examining public Facebook pages and groups. Front Digit Health 2021 Jun 17;3:683090 [FREE Full text] [doi: 10.3389/fdgth.2021.683090] [Medline: 34713153]
120. Bar-Lev S, Reichman S, Barnett-Itzhaki Z. Prediction of vaccine hesitancy based on social media traffic among Israeli parents using machine learning strategies. Isr J Health Policy Res 2021 Aug 23;10(1):49 [FREE Full text] [doi: 10.1186/s13584-021-00486-6] [Medline: 34425894]
121. Furini M. Identifying the features of ProVax and NoVax groups from social media conversations. Comput Human Behav 2021 Jul;120:106751. [doi: 10.1016/j.chb.2021.106751]
122. Baines A, Ittefaq M, Abwao M. #Scamdemic, #Plandemic, or #Scaredemic: what Parler social media platform tells us about COVID-19 vaccine. Vaccines (Basel) 2021 Apr 22;9(5):421 [FREE Full text] [doi: 10.3390/vaccines9050421] [Medline: 33922343]
123. Melton CA, Olusanya OA, Ammar N, Shaban-Nejad A. Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: a call to action for strengthening vaccine confidence. J Infect Public Health 2021 Oct;14(10):1505-1512 [FREE Full text] [doi: 10.1016/j.jiph.2021.08.010] [Medline: 34426095]
124. Kang GJ, Ewing-Nelson SR, Mackey L, Schlitt JT, Marathe A, Abbas KM, et al. Semantic network analysis of vaccine sentiment in online social media. Vaccine 2017 Jun 22;35(29):3621-3638 [FREE Full text] [doi: 10.1016/j.vaccine.2017.05.052] [Medline: 28554500]
125. Okuhara T, Ishikawa H, Okada M, Kato M, Kiuchi T. Japanese anti- versus pro-influenza vaccination websites: a text-mining analysis. Health Promot Int 2019 Jun 01;34(3):552-566. [doi: 10.1093/heapro/day015] [Medline: 29584863]
126. Cafiero F, Guille-Escuret P, Ward JK. “I’m not an antivaxxer, but…”: spurious and authentic diversity among vaccine critical activists. Soc Networks 2021 May;65:63-70. [doi: 10.1016/j.socnet.2020.11.004]
127. 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 Aug;45(4):599-606. [doi: 10.1177/1090198117739673] [Medline: 29267129]
128. Lee H, Noh EB, Park SJ, Nam HK, Lee TH, Lee GR, et al. COVID-19 vaccine perception in south Korea: web crawling approach. JMIR Public Health Surveill 2021 Sep 08;7(9):e31409 [FREE Full text] [doi: 10.2196/31409] [Medline: 34348890]
129. Porreca A, Scozzari F, Di Nicola M. Using text mining and sentiment analysis to analyse YouTube Italian videos concerning vaccination. BMC Public Health 2020 Feb 19;20(1):259 [FREE Full text] [doi: 10.1186/s12889-020-8342-4] [Medline: 32075631]
130. Powell GA, Zinszer K, Verma A, Bahk C, Madoff L, Brownstein J, et al. Media content about vaccines in the United States and Canada, 2012-2014: an analysis using data from the Vaccine Sentimeter. Vaccine 2016 Dec 07;34(50):6229-6235. [doi: 10.1016/j.vaccine.2016.10.067] [Medline: 27817958]
131. DeDominicis K, Buttenheim AM, Howa AC, Delamater PL, Salmon D, Omer SB, et al. Shouting at each other into the void: a linguistic network analysis of vaccine hesitance and support in online discourse regarding California law SB277. Soc Sci Med 2020 Dec;266:113216 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113216] [Medline: 33126093]
132. Wang Z, Yin Z, Argyris YA. Detecting medical misinformation on social media using multimodal deep learning. IEEE J Biomed Health Inform 2021 Jun;25(6):2193-2203. [doi: 10.1109/jbhi.2020.3037027]
133. Luo C, Ji K, Tang Y, Du Z. Exploring the expression differences between professionals and laypeople toward the COVID-19 vaccine: text mining approach. J Med Internet Res 2021 Aug 27;23(8):e30715 [FREE Full text] [doi: 10.2196/30715] [Medline: 34346885]

XSL•FO
RenderX

134. Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. Vaccine 2017 May 25;35(23):3033-3040 [FREE Full text] [doi: 10.1016/j.vaccine.2017.04.060] [Medline: 28461067]
135. Leader AE, Burke-Garcia A, Massey PM, Roark JB. Understanding the messages and motivation of vaccine hesitant or refusing social media influencers. Vaccine 2021 Jan 08;39(2):350-356 [FREE Full text] [doi: 10.1016/j.vaccine.2020.11.058] [Medline: 33280856]
136. Domgaard S, Park M. Combating misinformation: the effects of infographics in verifying false vaccine news. Health Educ J 2021 Aug 14;80(8):974-986. [doi: 10.1177/00178969211038750]
137. Qian D, Li O. The relationship between risk event involvement and risk perception during the COVID-19 outbreak in China. Appl Psychol Health Well Being 2020 Dec 23;12(4):983-999 [FREE Full text] [doi: 10.1111/aphw.12219] [Medline: 32829535]
138. Petkovic J, Duench S, Trawin J, Dewidar O, Pardo Pardo J, Simeon R, et al. Behavioural interventions delivered through interactive social media for health behaviour change, health outcomes, and health equity in the adult population. Cochrane Database Syst Rev 2021 May 31;5:CD012932 [FREE Full text] [doi: 10.1002/14651858.CD012932.pub2] [Medline: 34057201]
139. 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]
140. Parsons J, Griffiths S, Thomas N, Atherton H. How effective are digital interventions in increasing flu vaccination among pregnant women? A systematic review and meta-analysis. J Public Health (Oxf) 2021 Jun 23:2021-2014. [doi: 10.1093/pubmed/fdab220] [Medline: 34164686]
141. COVID-19 Vaccines: Safety Surveillance ManualCOVID-19 Vaccine Safety Communication. Geneva: World Health Organization; 2020.
142. Vaccine Safety Events: managing the communications response. World Health Organization. 2013. URL: https://www.euro.who.int/__data/assets/pdf_file/0007/187171/Vaccine-Safety-Events-managing-the-communications-response.pdf [accessed 2022-01-24]

Abbreviations

ELM: elaboration likelihood model
HBM: health belief model
HPV: human papillomavirus
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
SARF: social amplification of risk framework
WHO: World Health Organization

Edited by T Purnat; submitted 15.02.22; peer-reviewed by S Wei, SS Amritphale, A Naser, Y Wang, M Monselise; comments to author 24.05.22; revised version received 16.06.22; accepted 14.07.22; published 10.08.22.

Please cite as:
Yin JDC
Media Data and Vaccine Hesitancy: Scoping Review
JMIR Infodemiology 2022;2(2):e37300
URL: https://infodemiology.jmir.org/2022/2/e37300
doi:10.2196/37300
PMID:37113443

©Jason Dean-Chen Yin. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 10.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Social Listening to Enhance Access to Appropriate Pandemic Information Among Culturally Diverse Populations: Case Study From Finland

Anna-Leena Lohiniva[1*], MA, MSc; Katja Sibenberg[1*], MSS; Sara Austero[1], BA; Natalia Skogberg[1], PhD

Finnish Institute for Health and Welfare, Helsinki, Finland

*these authors contributed equally

Corresponding Author:
Anna-Leena Lohiniva, MA, MSc
Finnish Institute for Health and Welfare
166 Mannerheimintie
Helsinki, 00300
Finland
Phone: 358 295247191
Email: anna-leena.lohiniva@thl.fi

Abstract

Background: Social listening, the process of monitoring and analyzing conversations to inform communication activities, is an essential component of infodemic management. It helps inform context-specific communication strategies that are culturally acceptable and appropriate for various subpopulations. Social listening is based on the notion that target audiences themselves can best define their own information needs and messages.

Objective: This study aimed to describe the development of systematic social listening training for crisis communication and community outreach during the COVID-19 pandemic through a series of web-based workshops and to report the experiences of the workshop participants implementing the projects.

Methods: A multidisciplinary team of experts developed a series of web-based training sessions for individuals responsible for community outreach or communication among linguistically diverse populations. The participants had no previous training in systematic data collection or monitoring. This training aimed to provide participants with sufficient knowledge and skills to develop a social listening system based on their specific needs and available resources. The workshop design took into consideration the pandemic context and focused on qualitative data collection. Information on the experiences of the participants in the training was gathered based on participant feedback and their assignments and through in-depth interviews with each team.

Results: A series of 6 web-based workshops was conducted between May and September 2021. The workshops followed a systematic approach to social listening and included listening to web-based and offline sources; rapid qualitative analysis and synthesis; and developing communication recommendations, messages, and products. Follow-up meetings were organized between the workshops during which participants could share their achievements and challenges. Approximately 67% (4/6) of the participating teams established social listening systems by the end of the training. The teams tailored the knowledge provided during the training to their specific needs. As a result, the social systems developed by the teams had slightly different structures, target audiences, and aims. All resulting social listening systems followed the taught key principles of systematic social listening to collect and analyze data and used these new insights for further development of communication strategies.

Conclusions: This paper describes an infodemic management system and workflow based on qualitative inquiry and adapted to local priorities and resources. The implementation of these projects resulted in content development for targeted risk communication, addressing linguistically diverse populations. These systems can be adapted for future epidemics and pandemics.

(JMIR Infodemiology 2022;2(2):e38343) doi:10.2196/38343

KEYWORDS

infodemic; social listening; pandemic preparedness; cultural diversity; vulnerable populations

XSL•FO
RenderX

Introduction

Background

An infodemic is defined as an overabundance of information, some accurate and some not, which occurs during an epidemic [1]. During the COVID-19 pandemic, the infodemic has been rapidly expanding and evolving, particularly via social media channels. It has been estimated that rumors are 3 times more likely to be spread via social media than accurate information [2]. An infodemic poses a challenge for public health authorities who must continuously produce trustworthy and relevant information to inform the public about risks, influence behavioral change, and encourage compliance with emergency measures. Successful infodemic management saves lives and ultimately plays a major role in pandemic mitigation efforts. In contrast, failures in infodemic management during a pandemic can lead to misinterpreted messages, failed warnings, false rumors, and inconsistent information, which can negatively influence adherence to preventive behaviors of the public, which can be life-threatening and can negatively affect the trajectory of the pandemic [1,3]. Numerous infodemic examples have been documented during the pandemic, some of which have permeated geographic, cultural, and socioeconomic boundaries and required health care resources that, in most parts of the world, were already limited because of the pandemic. For example, a rumor that consuming highly concentrated alcohol could disinfect the body and kill the coronavirus [4] led to hospitalizations and fatalities after people ingested methanol in several countries, including Iran, Turkey, India, South Korea, and Qatar [5-8]. In addition, conspiracy theories circulated widely across the globe, often intentionally spreading disinformation [9].

Similarly, Finland witnessed a widespread infodemic during the COVID-19 pandemic. For example, anecdotal data point out that at the beginning of the COVID-19 vaccination campaigns, there were rumors that the vaccines were offered to harm people of particular ethnic backgrounds. Other widespread rumors claimed that the pandemic was created by pharmaceutical companies to make money by selling vaccines or that certain countries were responsible for the pandemic [10].

The World Health Organization (WHO) infodemic management framework advances equity as it highlights the importance of social listening and the need to identify context-specific information to tailor culturally appropriate infodemic responses [11]. Context specificity in infodemic management is of utmost importance as the pandemic has disproportionately affected ethnic minorities and a broad range of other populations that were already at a social disadvantage before the epidemic [12]. Members of minority groups may also be more resistant to following the guidance of authorities as they are often economically and socially more vulnerable than others [13]. Moreover, effective communication with various population subgroups tends to require tailored approaches [14].

Finland has become increasingly culturally and linguistically diverse in the past few decades. At the end of 2020, people speaking languages other than the official languages of Finnish, Swedish, or Sami constituted approximately 8% of the total Finnish population [12]. The increasingly linguistically and culturally diverse environment adds to the complexity of risk communication disseminated by health authorities. As in many other high-income countries, individuals of migrant origin in Finland were reported to have a higher incidence of COVID-19 infections and lower vaccine uptake than the general population [12,15]. This raised the need to better inform and engage people of migrant origins in risk communication planning and dissemination. Occasionally, individuals of migrant origin have also received negative attention and have been stigmatized as careless and unwilling to follow the guidance of health authorities, which has also raised awareness about the need to gain a better understanding of how various subgroups think. During the pandemic, health authorities have frequently highlighted the need for equity in health information and access to acceptable and appropriate information for everyone.

Social listening, a continuous process of collecting web-based and offline data using standard tools, has increased during the pandemic. Social listening projects have taken many forms worldwide. For example, in Vietnam, a social listening project was set up to explore public attention toward the pandemic, whereas another recent study concentrated on English-language tweets to identify the main pandemic topics globally [16,17]. Some projects use big data and dashboards to present the findings, such as the Red Cross COVID-19 dashboards piloted in some countries and the WHO Early Artificial Intelligence–Supported Response with Social Listening that monitors COVID-19–related web-based discussions in 30 countries [18-21]. Other projects have focused on smaller data sets based on manual internet browser searches and qualitative methods [3,22]. In addition, some projects have collected field-based data on rumors such as a real-time rumor-tracking pilot in Côte d'Ivoire, which leverages existing structures, including hotlines and community health workers, to submit rumors to a central database for rapid coding and visualization of the findings on dashboards [23].

In February 2020, the Finnish Institute for Health and Welfare (*Terveyden ja hyvinvoinnin laitos*; THL) initiated a social listening project to monitor pandemic perceptions of the public to provide recommendations for public authorities and risk communicators. The process comprised qualitative data collection and analysis in real time based on social media posts and information inquiries emailed by the public to THL. The results were shared and discussed with a group of public health and risk communication experts to determine appropriate infodemic responses every 2 to 4 weeks [3]. In May 2021, the project expanded to include training for regional health authorities and nongovernmental organizations that communicate COVID-19–related information to culturally and linguistically diverse populations. The extended social listening training project was initiated in response to public health experts' frequent concerns that pandemic-related information may not be reaching all population groups equally in Finland. Although community outreach workers have been listening to their target audiences even before joining the extended social listening project, there has not been a systematic and structured way of collecting, analyzing, and using the data. THL's multilingual and multichannel communication task force

XSL•FO
RenderX

implemented the expanded social listening project with the goal of supporting organizations in designing targeted communication for people from various ethnic backgrounds during a crisis.

Objectives

This study describes the development of systematic social listening training during the COVID-19 pandemic through a series of web-based workshops. It also reports the experiences of the workshop participants in implementing the projects.

Methods

Training Workshops

The overall concept behind the social listening workshops was that generic risk communication messages are not effective in reaching various subpopulations or changing the behaviors of target groups. Accordingly, having a systematic process that helps health authorities better understand the needs and motivations of various subpopulations will lead to more effective communication that is likely to lead to more sustainable behavior changes. Context-specific messaging also ensures less misunderstanding between health authorities and the public, which may help build increased trust between the two [24]. The training workshops were based on the conceptual framework of the WHO in infodemic management, which includes social listening, translating knowledge to practice, and quantifying impact [4] to expel misinformation and support targeted communication during crises.

The workshops were designed by a multidisciplinary team of experts from the THL. An anthropologist with a background in behavioral sciences and experience in social listening was mainly responsible for the content of the training workshops. An expert in pedagogy and risk communication was responsible for the contents of risk communication and for the overall structure of the workshops, including the timing and methods applied in group exercises. In addition, an expert in migration and cultural diversity was the main coordinator who also critically reviewed the social listening projects, ensuring that they were culturally appropriate.

The workshop design took into account the pandemic context. For example, training had to be short and intensive for participants to have time to participate. The aim of the training was to provide participants with sufficient knowledge and skills to develop their own social listening project based on their priorities, systems, and available resources. The workshops were based on a careful mix of tools that promoted the participation and internalization of knowledge and its application in a real-life project. The workshop structure was based on the principles of active learning [25]; for example, the rather rapid pace of alternating between activities aimed at maintaining active learning among the participants and motivating them to continue the training. The workshop methodology also used team-based learning pedagogy by introducing a systematic approach to building social listening projects in teams [26]. The workshop used Microsoft Teams and additional digital platforms, such as Howspace, for group exercises and for compilation of all workshop materials and suggested materials for further learning that participants could access after the training workshops as well.

The design also included homework that was meant to allow participants to practice what they had learned during the workshop and thus advance in developing their own social listening project design. The expected outcome of the workshops was a draft project plan by each team, including project flow, goals and objectives, data collection, and an analysis plan. The social listening methodology was based on qualitative data collection and analysis, with the notion that in-depth qualitative data provide a rich base for risk communication content development [27]. However, qualitative methods can be time consuming and complex [28], which requires adapting rapid qualitative data collection, recording, and analysis methods. Accordingly, the workshop encouraged the participants to adapt and test the taught methods of social listening to identify the best possible type of data collection and analysis for their specific needs. During the workshops, a substantial amount of time was allocated to teaching strategies on how to simplify the qualitative data collection and analysis processes.

The social listening training was designed for individuals responsible for community outreach work and for communication disseminating information to culturally diverse populations. The THL's multilingual and multichannel communications task force invited their collaborators from various cities and the Finnish Red Cross (FRC), which coordinated multilingual and multichannel projects among 20 local nongovernmental organizations. The invitation included a request to form a team that included those who could collect and analyze data and those who could develop communication messages and products. Each team was requested to have outreach workers and at least one communication expert.

Experiences in Implementing Social Listening Projects

Data on experiences in social listening project implementation were collected from the final presentations that each team shared at the end of the training, followed by 30-minute, one-on-one telephone interviews with each participant conducted by the corresponding author (ALL) in November 2021. Teams that did not continue implementing social listening followed by the workshops were requested to explain the reasons for this. The author also analyzed the interviews thematically using NVivo (QSR International) and shared the findings with the teams for verification [29].

Ethics Approval

Social listening was implemented using publicly available data that did not contain personal or sensitive data. Each team conducting social listening complied with ethical considerations based on the guidance from their own institutions (City of Helsinki, City of Espoo, City of Vantaa, and FRC). All teams ensured confidentiality through the anonymization of their data. No personal identifiers were collected. Confidentiality was maintained throughout the entire project cycle from data collection to reporting.

XSL•FO
RenderX

Results

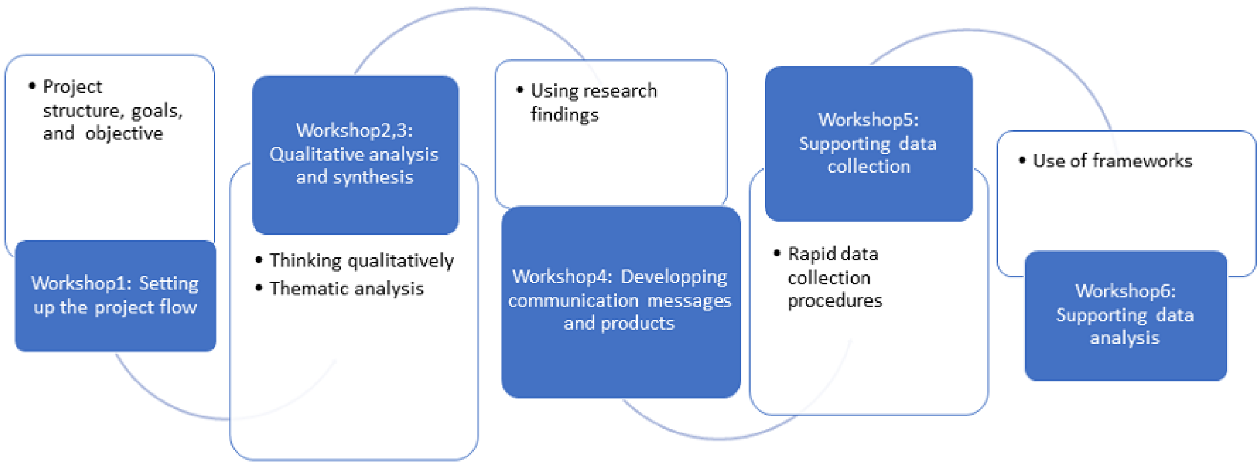
Workshops

A series of 6 web-based workshops was delivered between May and September 2021, by a multidisciplinary team of 3 experts from THL responsible for the methodological design of the workshops. The time between the first 4 workshops ranged from 1 to 2 weeks. Workshops 5 and 6 were conducted after the summer holidays, resulting in a nearly 2-month break from the earlier workshops. Each workshop lasted for a maximum of 2 hours and included short lectures, discussions, and exercises. Between the workshops, participants were assigned homework that focused on the implementation of the techniques learned during the workshops. The workshops were developed based on the following structure: (1) setting up a social listening project with roles and responsibilities and defining goals and objectives; (2) learning about the use of qualitative methodology and how to think qualitatively; (3) qualitative analysis and synthesis; (4) developing communication recommendations, messages, and products; (5) focusing on how to facilitate qualitative data collection procedures; and (6) learning to facilitate qualitative data collection procedures. A total of 2 consultations were organized by the trainers in between the workshops, during which participants shared their achievements and challenges. All team members were invited to the consultations (Figure 1 provides the structure of the training workshops).

All 6 workshops were conducted on the web with a carefully planned set of learning objectives, themes, and structures to ensure that they kept participants interested and occupied during the training. Accordingly, the workshop structure was based on short activities that started with a theory session followed by interactive group work. During the group work, teams were required to apply the theory through reflections and discussions to ensure that the theory and its application were truly internalized. Group work was always conducted within the participants' own social listening teams, with the exception of qualitative analysis in which each individual practiced data coding themselves. The participants also completed 3 homework assignments in between the workshops. Each team presented their homework to the others during the following workshop. A detailed plan of the workshops is presented in Table S1 in Multimedia Appendix 1.

The workshop participants also received a list of resources that they could use to deepen their understanding of the methods, techniques, and conceptual frameworks that were introduced during the workshop, as shown in Multimedia Appendix 2.

Figure 1. Structure of the training workshops.



Social Listening Projects

Overview

A total of 6 social listening teams joined the workshops from different geographic locations in Finland. A total of 4 groups implemented the project. Of the 2 groups that dropped out, one did not have sufficient human resources to conduct the project, whereas the other group did not know how to reach their target audience. The remaining teams comprised outreach workers and communication professionals. The participants had little or no prior experience in applying the research methods in their work. The following section describes how the teams formulated their social listening processes following the training. A summary of these projects is provided in Table 1. Details of the social listening team composition and the resources invested are provided in Table 2.

XSL•FO
RenderX

Table 1. Summary of social listening projects.

Project	Objectives	Data sources	Methodology
City of Helsinki	• To identify key concerns among the public to prepare and disseminate appropriate information • To detect and correct misinformation	A number of social media channels and face-to-face encounters with clients	• The weekly manual process included reviewing posts and extracting relevant posts to a spreadsheet followed by team discussions about the type of communication actions needed
City of Vantaa	• To track the main discussions and themes related to the pandemic and COVID-19 vaccines	Face-to-face encounters with clients	• Data collection through discussions and field-based observations by outreach workers • Project team meeting once a month to share, analyze, and brainstorm how to use the observations, followed by the development of messages and materials that are distributed through known channels to target audiences
Espoo	• To create vaccine demand	Face-to-face with client encounters and social media	• Informal discussions conducted during routine meetings are used to develop communication responses
Finnish Red Cross	• To develop targeted communication materials	Face-to-face discussions with partner organizations working with different language groups	• Partner organizations shared their experiences of encounters with different language groups with Finnish Red Cross project management who develop communication materials based on those encounters

Table 2. Average time and human resources spent on social listening.

Participants	City of Espoo	City of Helsinki	City of Vantaa	Finnish Red Cross
The number of people who participated in the training	6	6	9	3
Number of people involved in data collection	6	5	12-20	2
Number of people involved in data analysis	4	6	12	2
Number of communication experts	2	0	4	1
Time per week spent on the social listening project per person, mean (SD)	+1 or –1	N/A[a]	+1 or −1	4

[a]N/A: not available.

City of Helsinki

The social listening projects of the city of Helsinki aimed to listen to the information needs of people from linguistically diverse populations. The team comprised social media experts, whose data sources included one social media site run by the city of Helsinki and several open sites where linguistically diverse populations communicate. In addition, data were collected during face-to-face encounters with their clients, such as at the information desk of the Helsinki main library. Each data source had a focal point who recorded data independently on their work laptop and summarized the data into a joint Microsoft Excel sheet located on a secured server that could be accessed only by the team members. No identifiers were collected. The content of the sheet was discussed on a weekly basis to guide the planning of information provision for linguistically diverse populations. During the piloting period, social listening identified critical information voids that were addressed in open webinars and pop-up consultations that the city was organizing to boost the COVID-19 vaccine uptake. This project has opened new social media channels that can be included in future social listening activities.

City of Vantaa

The social listening project of the city of Vantaa focused on listening to the COVID-19 vaccine and pandemic-related discussions among linguistically diverse populations. The social listening project team comprised a communication expert specializing in multilingual communication and approximately 20 outreach workers with diverse cultural backgrounds across the city who work with linguistically diverse populations, such as groups of migrant origin, or with projects targeting audiences of all major language groups under one focused theme such as employment creation. Data collection was based on face-to-face encounters with target audiences and manual notetaking during these encounters. Each outreach worker was responsible for keeping the notes in a secure location at the workplace. The notes included no names or any other identifiable information. Social listening was introduced as a continuous process, not based on any specific schedule or weekly time limit but on ad hoc opportunities to chat with the target audience. The social

XSL•FO
RenderX

listening group met monthly to share, analyze, and reflect on field observations on an agreed-upon topic through an open discussion that culminated in jointly agreed-upon communication messages and actions. During the piloting period, social listening pinpointed a number of factors that influenced COVID-19 uptake among Russian-speaking and Somali-speaking clients, which were used to develop targeted messages for discussion events organized by the city. At the beginning of the process, the team had a joint platform to record data that were later omitted from the process as it was too time consuming. The project staff highlighted that the project was beneficial for the team members as it opened up opportunities for outreach workers to influence communication and for the communication experts to develop context-specific messaging. Following the pilot, the group plans to continue the project by improving the working modalities and developing checklists that can better focus on observations and analysis in the future.

City of Espoo

The city of Espoo introduced the social listening model to a number of different working groups and projects that communicated pandemic-related information to linguistically diverse populations; however, it did not formalize the system. Instead, social listening is considered a tool that can be used periodically when needed. Outreach team members verbally discussed the outcomes of encounters with their clients regarding the pandemic during a routine biweekly meeting to improve the messages that they communicated. Social listening was based on recall, and no notes were taken. The team members highlighted that the COVID-19 vaccination program has benefited from the social listening system by using the findings to create content for their COVID-19 and vaccination webinars. The process is still in the testing phase; however, future plans include formalizing social listening and creating a more formal structure to help monitor the process.

The FRC Organization

Social listening was part of a multilingual communication project of the FRC. The project aimed to develop pandemic-related multilingual materials based on data gathered by the FRC district offices and >10 partner organizations across the country. Partner organizations included 2 large umbrella organizations that covered a number of smaller organizations, all of which were in direct contact with people from culturally and linguistically diverse backgrounds. The focal points of the organizations were brought together with the FRC project coordinator to share their observations. The system was built upon the partners' requests for no formal structures, written data collection, or documentation, and it used existing meetings of the partners, which ensured minimum use of resources. Data collectors were not requested to take notes or use a specific amount of time for observations, and they shared insights based on what they memorized at the time of the meeting. All discussions were confidential. The observations were discussed, and materials were developed and shared with the focal points for their feedback before finalization. The focal points could also be consulted on a one-on-one basis to seek their opinions about certain messages and materials. The system has regularly fed into communication content and FRC, which develops materials, and their partners have realized the benefit of discussing topics of interest before they are implemented as communication messages and materials. The new working modalities are expected to expand to include other topics and collaborations.

Discussion

Principal Findings

This study provides valuable insights into a series of rapid social listening workshops designed to provide training participants with the knowledge and skills to develop their own social listening systems based on qualitative data collection and analysis. To our knowledge, this is the first paper describing the structure and contents of social listening training that focused on qualitative data analysis and synthesis and targeted individuals without previous research experience, who were responsible for conducting community outreach work. This paper further demonstrates how the training workshop participants adapted the knowledge and skills gained in the workshops in different contexts for implementing social listening programs among culturally and linguistically diverse populations during a crisis. The projects had different aims, target audiences, data sources, and working modalities; however, they were all able to produce meaningful insights that were further used to develop acceptable and appropriate communication messages for people belonging to different cultural and linguistic groups.

Approximately 67% (4/6) of teams that continued with designing their own social listening projects completed all the offered training workshops, designed their own social listening plans, and then successfully implemented these plans. A web-based methodology with short but focused sessions made participation logistically easier. Distant learning may have been even a prerequisite for participation for some of the participants in the midst of their hectic pandemic work schedule. However, at the same time, the training participants did not have the opportunity to learn about one another. Thus, informal learning between the teams that often occurs during coffee break discussions in face-to-face training was lacking. It is likely that in future crisis situations, similar training should also be conducted on the web. However, more mixing of the teams during group work exercises and reflection sessions could be embedded in the workshops to facilitate peer reflection and learning across different teams. The trainers did not mandate the workshop participants to keep their cameras on during the training. However, in the future, training workshops could be mandatory to foster communality.

The social listening methodology introduced in this project was based on qualitative inquiry, which is often perceived as difficult to implement [28]. Accordingly, during the workshops, a substantial amount of time was spent learning about the importance of qualitative data and various modalities that can be used to simplify the processes. The findings showed that projects had adopted rapid but systematic data collection and analysis processes, including the use of recording data in a joint Microsoft Excel sheet, handwritten notes, or memorizing data and verifying data weekly or monthly in a joint meeting. Interestingly, digital platforms such as Microsoft Teams, Google Docs, or other technological tools such as voice messaging were

XSL•FO
RenderX

not widely used when implementing the projects. In contrast, some projects found them more time consuming than traditional paper and pen notetaking. Two projects used a joint platform where individuals organized their data for discussion. The use of joint platforms to display data has been commonly used in other social listening projects, such as in Kenya, which had a messaging matrix available for all who communicate. In Sierra Leone, a cloud-based data collection resulted in a real-time message dashboard [2,30].

However, the projects did not use any particular behavioral frameworks or checklists that were introduced during the training, which were meant to facilitate data collection and analysis processes. It would be important to investigate the reasons for this as they have been found to be helpful in social listening projects elsewhere. For example, this was the case in Côte d'Ivoire, where phone hotline–based data managers coded rumors nearly in real time according to behavioral theory frameworks [23]. In addition, the projects did not develop procedures that would show how the data were interpreted or synthesized. Knowing that rapid and rigorous data analysis is a particularly daunting task, more tools could be introduced in future training. Such tools may be, for example, a rigorous and accelerated data reduction technique that converts raw textual data into a more manageable and user-friendly format, which involves systematic analysis during each step of the process [31]. As all training materials were provided to the trainees, the use of frameworks and improved recording of data analysis are issues that the teams can also develop later on once they want to start improving their methodology and systems.

Knowledge co-creation was a key feature of each project, highlighting the understanding of the essence of qualitative approaches that appreciate reflection [32]. All projects invested time to discuss the findings and to co-create messages and products, which is likely to ensure that the findings and resulting recommendations were seen by the target audience as salient, legitimate, and credible [33]. Knowledge co-creation also promoted shared learning, which is likely to result in more impact-driven risk communication [34,35]. Knowledge co-creation further emphasized the importance of a multiprofessional composition [36] of the teams, including field-based data collectors who have direct contact and access to target populations and communication experts with the ability to produce quality messages and materials. One of the teams mentioned that the collaboration between field teams and communication experts was an entirely new experience that was beneficial for both parties.

Examples of training participants' use of social listening data in communication with people from linguistically diverse backgrounds indicate that they internalized the very essence of cultural relativism, namely, valued the ideas of the target audience instead of judging them against expert opinions [37,38]. They used the thinking of their target audience to create communication that facilitated 2-way dialog. The realization of a lay perspective is also likely to decrease potential misunderstandings that are common when scientific or expert information is communicated to the general public [39]. Generic messages are rarely effective in changing attitudes and behaviors, unlike focused messages, which are based on an understanding of the target audience's needs [40,41].

All projects avoided highly structured systems in favor of informal and flexible approaches to make the data collection and recording process less time consuming and complex. Flexible structures are more adaptable to changing topics and target audiences, which is highly beneficial for social listening projects that aim to provide real-time information about relevant topics. Formal structures are likely to be seen as commitments that the organizations are not willing or able to make without a dedicated budget that none of them had for social listening purposes. The more flexible social listening projects were merged within their structures and ongoing activities, the more cost-effective the projects were. A formal structure would allow the institutionalization of social listening as a part of routine risk communication during future crises [42].

All teams demonstrated having developed targeted communication materials based on social listening after attending the workshops. These included content for webinars, pop-up consultations, face-to-face meetings, and printed materials. Pilots from other parts of the world with similar social listening projects, which have triangulated insights from digital and nondigital sources, have also developed meaningful communication; however, impact evaluations have not been conducted [43,44]. As 2 out of 6 teams dropped out without piloting social listening, it would be important in the future to better define the selection criteria to participate in the workshops and follow up on the selection of the workshop participants. Future plans could involve the development and testing of a joint platform across organizations that can share real-time data for communication purposes. Thus far, there has been no monitoring system to track changes in the attitudes or behaviors of target audiences. In the future, it would be important to integrate rigorous monitoring and evaluation components into projects to understand how targeted messaging influences the audience. It is also important to continue testing and learning from different project modalities. More efforts should be made to increase the use of multiple data sources to establish an integrated analysis that can further strengthen the quality of the data analysis and the recommendations [1].

Conclusions

A series of training workshops was designed to implement social listening based on qualitative data collection and analysis for individuals responsible for community outreach and for communication specialists who had little or no prior experience in research methods. Over the course of the training, the participants adapted the frameworks and techniques introduced during the training to design their own adapted social listening systems. These social listening systems were based on their specific priorities and resources. The implementation of these systems resulted in content development for targeted communication messages addressing linguistically diverse populations. They can be adapted for use in future epidemics and crises. Future studies should aim for more long-term follow-up of the implementation and impact assessment of the projects.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank the members of the city of Espoo social listening team, including Sari Hammar, Ulla-Kaisa Pihlaja, Mari Ala-Mikkula, Arja Filpus, Kati Ojeda, and Maria Tiilikkala, and the city of Helsinki team, including Marja Säikkä, Aino Hautsalo, Hanna-Kristina Einmann, Shangzhi Sun, Susanna Kaajaluoma, and Sari Kettunen for their outstanding efforts in developing and piloting a social listening project. The authors would also like to extend their gratitude to the city of Vantaa social listening team, including Larissa Mäkinen, Kristamaria Hellström, Salli Allen, Paula Hännikäinen, Tiina Kentta, Ilona Korhonen, Päivi Oikarinen, Sanna Heikkinen, Katariina Kuronen, Arttu Rajala, Katja Bister, and Kaisa Sunimento, and the Finnish Red Cross team, including Erika-Evely Ee Eisen, Marian Ismail, and Petri Öhman from a partner Central Organization of Russian speakers in Finland.

SA and NS were funded by the "Coping of disabled persons and persons of migrant origin in exceptional and crisis situations—Building the future based on experiences during the COVID-19 (Building the Future)" project supported by the European Social Fund. The funding body was not involved in the review or approval of this manuscript.

Authors' Contributions

ALL, KS, and SA conceptualized the social listening workshops to be implemented in the Building the Future project led by NS. ALL conceptualized and wrote the first draft of the manuscript. All authors provided critical comments and participated in the revision of the manuscript. All the authors approved the final version of the manuscript sent to the journal.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The curriculum of a series of workshops preparing participants to develop social listening projects.
[DOCX File , 21 KB - infodemiology_v2i2e38343_app1.docx]

Multimedia Appendix 2
Resources for trainees.
[DOCX File , 16 KB - infodemiology_v2i2e38343_app2.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. Sommariva S, Vamos C, Mantzarlis A, Đào LU, Martinez Tyson D. Spreading the (Fake) news: exploring health messages on social media and the implications for health professionals using a case study. Am J Health Educ 2018 Jun 07;49(4):246-255. [doi: 10.1080/19325037.2018.1473178]
3. Lohiniva AL, 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]
4. An ad hoc WHO technical consultation managing the COVID-19 infodemic: call for action, 7-8 April 2020. World Health Organization. Geneva, Switzerland: World Health Organization; 2020 Sep 15. URL: https://www.who.int/publications/i/item/9789240010314 [accessed 2022-03-09]
5. Iran: Over 700 dead after drinking alcohol to cure coronavirus. Aljazeera. 2020 Apr 27. URL: https://www.aljazeera.com/news/2020/04/iran-700-dead-drinking-alcohol-cure-coronavirus-200427163529629.html [accessed 2022-03-09]
6. Delirrad M, Mohammadi AB. New methanol poisoning outbreaks in Iran following COVID-19 pandemic. Alcohol Alcohol 2020 Jun 25;55(4):347-348 [FREE Full text] [doi: 10.1093/alcalc/agaa036] [Medline: 32400874]
7. Nine More Poeple Died from Pre Alcohol. CNN Türk. 2020 Mar 25. URL: https://www.cnnturk.com/video/turkiye/9-kisi-daha-saf-alkolden-oldu [accessed 2022-03-09]
8. Ahmed Siddiqui MY, Mushtaq K, Mohamed MF, Al Soub H, Hussein Mohamedali MG, Yousaf Z. "Social media misinformation"-an epidemic within the COVID-19 pandemic. Am J Trop Med Hyg 2020 Aug;103(2):920-921 [FREE Full text] [doi: 10.4269/ajtmh.20-0592] [Medline: 32534600]
9. Chen X, Zhang SX, Jahanshahi AA, Alvarez-Risco A, Dai H, Li J, et al. Belief in a COVID-19 conspiracy theory as a predictor of mental health and well-being of health care workers in Ecuador: cross-sectional survey study. JMIR Public Health Surveill 2020 Jul 21;6(3):e20737 [FREE Full text] [doi: 10.2196/20737] [Medline: 32658859]

XSL•FO
RenderX

10. HS: "Dangerous" disinformation on Covid-19 spreading in Finland. Helsinki Times. 2020 Nov 26. URL: https://www.helsinkitimes.fi/finland/finland-news/domestic/18361-hs-dangerous-disinformation-on-covid-19-spreading-in-finland.html [accessed 2022-03-09]
11. 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://tinyurl.com/5fxuvbhk [accessed 2022-03-09]
12. Holmberg V, Salmi H, Kattainen S, Ollgren J, Kantele A, Pynnönen J, et al. Association between first language and SARS-CoV-2 infection rates, hospitalization, intensive care admissions and death in Finland: a population-based observational cohort study. Clin Microbiol Infect 2022 Jan;28(1):107-113 [FREE Full text] [doi: 10.1016/j.cmi.2021.08.022] [Medline: 34949510]
13. Shadmi E, Chen Y, Dourado I, Faran-Perach I, Furler J, Hangoma P, et al. Health equity and COVID-19: global perspectives. Int J Equity Health 2020 Jun 26;19(1):104 [FREE Full text] [doi: 10.1186/s12939-020-01218-z] [Medline: 32586388]
14. Ashwell DJ. The challenges of science journalism: the perspectives of scientists, science communication advisors and journalists from New Zealand. Public Underst Sci 2016 Apr;25(3):379-393. [doi: 10.1177/0963662514556144] [Medline: 25387869]
15. Kontio M, Skogberg N, Sivelä J, Nohynek H, Salo H, Mäntymaa P, et al. Koronarokotuskattavuuteen vaikuttavat tekijät. Finnish Institute for Health and Welfare. 2021. URL: https://www.julkari.fi/bitstream/handle/10024/143238/URN_ISBN_978-952-343-753-1.pdf?sequence=1&isAllowed=y [accessed 2022-03-09]
16. Tran HT, Lu SH, Tran HT, Van Nguyen B. Social media insights during the COVID-19 pandemic: infodemiology study using big data. JMIR Med Inform 2021 Jul 16;9(7):e27116 [FREE Full text] [doi: 10.2196/27116] [Medline: 34152994]
17. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. J Med Internet Res 2020 Apr 21;22(4):e19016 [FREE Full text] [doi: 10.2196/19016] [Medline: 32287039]
18. Ahmad AR, Murad HR. The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. J Med Internet Res 2020 May 19;22(5):e19556 [FREE Full text] [doi: 10.2196/19556] [Medline: 32369026]
19. Steffens MS, Dunn AG, Leask J, Wiley KE. Using social media for vaccination promotion: practices and challenges. Digit Health 2020 Nov 3;6:2055207620970785 [FREE Full text] [doi: 10.1177/2055207620970785] [Medline: 35173976]
20. El Hamouch M. Rumor Tracking: A Digital Transformation Approach 2020. 510 Blog. 2021 Apr 15. URL: https://www.510.global/rumor-tracking-a-digital-transformation-approach/[accessed [accessed 2022-03-09]
21. Early AI-supported Response with Social Listening. World Health Organization. 2020. URL: https://www.who-ears.com/ [accessed 2022-03-09]
22. Liu J, Lu S, Lu C. Exploring and monitoring the reasons for hesitation with COVID-19 vaccine based on social-platform text and classification algorithms. Healthcare (Basel) 2021 Oct 12;9(10):1353 [FREE Full text] [doi: 10.3390/healthcare9101353] [Medline: 34683033]
23. Tibbels N, Dosso A, Allen-Valley A, Benie W, Fordham C, Brou JA, et al. Real-time tracking of COVID-19 rumors using community-based methods in Côte d'Ivoire. Glob Health Sci Pract 2021 Jun 30;9(2):355-364 [FREE Full text] [doi: 10.9745/GHSP-D-21-00031] [Medline: 34038385]
24. Buzoianu C, Bîră M. Using social media listening in crisis communication and management: new methods and practices for looking into crises. Sustainability 2021 Nov 25;13(23):13015. [doi: 10.3390/su132313015]
25. Harris BN, McCarthy PC, Wright AM, Schutz H, Boersma KS, Shepherd SL, et al. From panic to pedagogy: using online active learning to promote inclusive instruction in ecology and evolutionary biology courses and beyond. Ecol Evol 2020 Nov;10(22):12581-12612 [FREE Full text] [doi: 10.1002/ece3.6915] [Medline: 33250996]
26. Sannathimmappa MB, Nambiar V, Aravindakshan R, Kumar A. Are Online Synchronous Team-Based-Learning (TBL) pedagogy effective?: perspectives from a study on medical students in Oman. J Adv Med Educ Prof 2022 Jan;10(1):12-21 [FREE Full text] [doi: 10.30476/JAMP.2021.92361.1481] [Medline: 34981001]
27. Hawking MK, Timmis A, Wilkins F, Potter JL, Robson J. Improving cardiovascular disease risk communication in NHS Health Checks: a qualitative study. BMJ Open 2019 Sep 03;9(8):e026058 [FREE Full text] [doi: 10.1136/bmjopen-2018-026058] [Medline: 31481364]
28. Taylor B, Henshall C, Kenyon S, Litchfield I, Greenfield S. Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? A mixed methods study comparing rapid and thematic analysis. BMJ Open 2018 Oct 08;8(10):e019993 [FREE Full text] [doi: 10.1136/bmjopen-2017-019993] [Medline: 30297341]
29. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014 Oct 16;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
30. Jalloh MF, Sengeh P, James N, Bah S, Jalloh MB, Owen K, et al. Integrated digital system for community engagement and community-based surveillance during the 2014-2016 Ebola outbreak in Sierra Leone: lessons for future health emergencies. BMJ Glob Health 2020 Dec;5(12):e003936 [FREE Full text] [doi: 10.1136/bmjgh-2020-003936] [Medline: 33355270]
31. Watkins DC. Rapid and rigorous qualitative data analysis: the “RADaR” technique for applied research. Int J Qual Methods 2017 Jun 08;16(1):1-9. [doi: 10.1177/1609406917712131]

XSL•FO
RenderX

32. Leask CF, Sandlund M, Skelton DA, Altenburg TM, Cardon G, Chinapaw MJ, GrandStand, Safe Step and Teenage Girls on the Move Research Groups. Framework, principles and recommendations for utilising participatory methodologies in the co-creation and evaluation of public health interventions. Res Involv Engagem 2019 Jan 9;5:2 [FREE Full text] [doi: 10.1186/s40900-018-0136-9] [Medline: 30652027]
33. Cash DW, Belloy PG. Salience, credibility and legitimacy in a rapidly shifting world of knowledge and action. Sustainability 2020 Sep 08;12(18):7376 [FREE Full text] [doi: 10.3390/su12187376]
34. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: literature review and case study. Milbank Q 2016 Jun;94(2):392-429 [FREE Full text] [doi: 10.1111/1468-0009.12197] [Medline: 27265562]
35. Burman A. The political ontology of climate change: moral meteorology, climate justice, and the coloniality of reality in the Bolivian Andes. J Polit Ecol 2017;24(1):921-930. [doi: 10.2458/v24i1.20974]
36. Dilling L, Lemos MC. Creating usable science: opportunities and constraints for climate knowledge use and their implications for science policy. Glob Environ Change 2011 May;21(2):680-689. [doi: 10.1016/j.gloenvcha.2010.11.006]
37. Koizumi H, Yamashita H. Deficit lay or deficit expert: how do “experts” in environmental projects perceive lay people and lay knowledge? SAGE Open 2021 Jul 07;11(3):215824402110231. [doi: 10.1177/21582440211023155]
38. Grundmann R. The problem of expertise in knowledge societies. Minerva 2017;55(1):25-48 [FREE Full text] [doi: 10.1007/s11024-016-9308-7] [Medline: 28239194]
39. Eriksson M. Lessons for crisis communication on social media: a systematic review of what research tells the practice. Int J Strateg Commun 2018 Aug 30;12(5):526-551. [doi: 10.1080/1553118x.2018.1510405]
40. Gallagher KM, Updegraff JA. Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. Ann Behav Med 2012 Feb;43(1):101-116. [doi: 10.1007/s12160-011-9308-7] [Medline: 21993844]
41. Schmid KL, Rivers SE, Latimer AE, Salovey P. Targeting or tailoring? Mark Health Serv 2008;28(1):32-37 [FREE Full text] [Medline: 18389854]
42. Oxman AD, Lavis JN, Lewin S, Fretheim A. SUPPORT tools for evidence-informed health policymaking (STP) 1: what is evidence-informed policymaking? Health Res Policy Syst 2009 Dec 16;7 Suppl 1:S1 [FREE Full text] [doi: 10.1186/1478-4505-7-S1-S1] [Medline: 20018099]
43. Gesser-Edelsburg A. How to make health and risk communication on social media more "social" during COVID-19. Risk Manag Healthc Policy 2021 Aug 25;14:3523-3540 [FREE Full text] [doi: 10.2147/RMHP.S317517] [Medline: 34471393]
44. Bukar UA, Jabar MA, Sidi F, Nor RB, Abdullah S, Ishak I. How social media crisis response and social interaction is helping people recover from Covid-19: an empirical investigation. J Comput Soc Sci 2022;5(1):781-809 [FREE Full text] [doi: 10.1007/s42001-021-00151-7] [Medline: 34746496]

Abbreviations

FRC: Finnish Red Cross
THL: Terveyden ja hyvinvoinnin laitos (Finnish Institute for Health and Welfare)
WHO: World Health Organization

Edited by T Purnat; submitted 29.03.22; peer-reviewed by E Diaz, A Parrilla, S Boender; comments to author 22.04.22; revised version received 27.05.22; accepted 24.06.22; published 08.07.22.

Please cite as:
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;2(2):e38343
URL: https://infodemiology.jmir.org/2022/2/e38343
doi:10.2196/38343
PMID:37113448

©Anna-Leena Lohiniva, Katja Sibenberg, Sara Austero, Natalia Skogberg. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 08.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Users' Modifications to Electronic Nicotine Delivery Systems: Content Analysis of YouTube Video Comments

Yachao Li[1,2], PhD; David L Ashley[3], PhD; Lucy Popova[3], PhD

[1]Department of Communication Studies, The College of New Jersey, Ewing, NJ, United States
[2]Department of Public Health, The College of New Jersey, Ewing, NJ, United States
[3]School of Public Health, Georgia State University, Atlanta, GA, United States

Corresponding Author:
Lucy Popova, PhD
School of Public Health
Georgia State University
140 Decatur Street
Atlanta, GA, 30302
United States
Phone: 1 404 413 9338
Email: lpopova1@gsu.edu

Abstract

Background: User modifications can alter the toxicity and addictiveness of electronic nicotine delivery systems (ENDSs). YouTube has been a major platform where ENDS users obtain and share information about ENDS modifications. Past research has examined the content and characteristics of ENDS modification videos.

Objective: This study aims to analyze the video comments to understand the viewers' reactions to these videos.

Methods: We identified 168 YouTube videos depicting ENDS modifications. Each video's top 20 most liked comments were retrieved. The final sample included 2859 comments. A content analysis identified major themes of the comment content.

Results: Most comments were directed to creators and interacted with others: 952/2859 (33.30%) expressed appreciation, 135/2859 (4.72%) requested more videos, 462/2859 (16.16%) asked for clarification, and 67/2859 (2.34%) inquired about product purchases. In addition, comments mentioned viewers' experiences of ENDS modifications (430/2859, 15.04%) and tobacco use (167/2859, 5.84%); about 198/2859 (6.93%) also indicated intentions to modify ENDSs and 34/2859 (1.19%) mentioned that they were "newbies." Moreover, comments included modification knowledge: 346/2859 (12.10%) provided additional information, 227/2859 (7.94%) mentioned newly learned knowledge, and 162/2859 (5.67%) criticized the videos. Furthermore, few comments mentioned the dangers of ENDS modifications (136/2859, 4.76%) and tobacco use (7/2859, 0.24%). Lastly, among the 15 comments explicitly mentioning regulations, 13/2859 (0.45%) were against and 2/2859 (0.07%) were supportive of regulations.

Conclusions: The results indicated acceptance and popularity of ENDS modifications and suggested that the videos might motivate current and new users to alter their devices. Few comments mentioned the risks and regulations. Regulatory research and agencies should be aware of online ENDS modification information and understand its impacts on users.

(JMIR Infodemiology 2022;2(2):e38268) doi:10.2196/38268

KEYWORDS

ENDS modifications; YouTube; comments; vaping; content analysis

Introduction

Electronic nicotine delivery systems (ENDSs), also known as vapes, e-cigarettes, and e-hookahs, have become increasingly popular in the United States [1]. Some smokers may utilize ENDSs to quit smoking and some randomized control trials suggest that, under certain conditions, ENDSs may improve smoking cessation compared with nicotine replacement therapy [2]. However, current evidence from population studies indicates no significant association between ENDS use and increased smoking cessation among cigarette smokers [3]. Emerging research also suggests that ENDS use has both short- and long-term health risks [4-6], such as burn injuries [7], lung inflammation, and pulmonary fibrosis [8], and low birth weight associated with parental ENDS use [9]. Moreover, ENDS products often include highly modifiable features that allow

XSL•FO
RenderX

users to alter device, liquid, and aerosol characteristics, which may cause even more harmful consequences [10]. Indeed, The Centers for Disease Control and Prevention (CDC) urges ENDS users not to add substances or modify the products not intended by the manufacturer [11].

ENDS modifications include product misuse and tampering unintended by the manufacturers, as well as alteration, customization, adjustment, and user choice of e-liquid or accessories made within manufacturer parameters [12,13]. For instance, some users may alter the liquid materials to be aerosolized, such as making their own e-juice, adding substances such as cannabis, or substituting manufactured liquids with materials of unknown composition and origin [12,14]. Other common practices include modifying heating coils and changing battery voltage to increase levels of nicotine delivery, produce larger clouds, and experience different throat hit [12,14]. While users consider the ability to customize nicotine levels and flavors an attractive feature of ENDS products [10], ENDS modifications can expose users to higher levels of harmful substances in the aerosol when they increase power to the coil [15,16]. Other harms related to ENDS modification include overheating and explosion-related injuries [17,18], use of illicit substances [19], and clinical nicotine toxicity [20]. Moreover, the availability of certain flavors encourages youth use [1]. Thus, ENDS modifications could change the toxicity and addictiveness of the products, inhibit cessation, and increase initiation of ENDS use.

Given the popularity of ENDS and the health risks related to modifications, more research is needed to understand users' attitudes toward modifications and behaviors when modifying ENDS. As one of the primary video-sharing sites across the world, YouTube has been identified as a major platform where ENDS users share and obtain information about ENDS products and modifications [12,21,22]. To date, 2 studies [13,23] have examined the characteristics and content of YouTube videos depicting ENDS modifications. One study found that videos depicting unorthodox use (unintended by the manufacturer) were 3 times more prevalent than videos depicting orthodox use (intended by the manufacturer) [23]. Another study highlighted several concerning trends in ENDS modification videos, including lack of warnings, adding marijuana derivatives to e-liquids, and positive portrayal of ENDS devices and modifications [13]. While both studies provide valuable insights into how users modify ENDSs and identify features of the modification videos, we know little about viewers' reactions to those YouTube videos, which may be mined for understanding the potential impacts of online information on people's ENDS modification attitudes and behaviors.

One way to understand viewers' reactions to YouTube videos is to study the comments left on the videos. The comment function allows users to directly respond to the video content and to express their opinions [24]. The data collection is also unobtrusive, providing relatively accurate valuable insights into how viewers naturally think about the videos and the issues portrayed in the videos. Thus, we conducted a content analysis of user comments on ENDS modification videos. The research objectives are to identify common themes of those comments and to explore YouTube users' attitudes toward ENDS use and modifications. The results would complement prior research on ENDS modifications and studies of YouTube modification videos, as well as provide a better understanding of the potential effects of the videos on viewers.

Methods

Data Collection

We searched YouTube on March 15, 2019, to identify videos depicting ENDS modifications. A new account was created on the incognito (private) mode to minimize the impacts of browsing history on search results. A total of 28 search phrases, derived from interviews with ENDS users [12] and the literature [25], were utilized, such as "vape DIY (ie, do it yourself)," "vape e-juice custom build," "vape dripping DIY," and "vaping modif* custom build." The full search terms and video identification procedures are reported elsewhere [13]. The top 10 most viewed videos for each search phrase were identified (n=280 videos). Trained coders then reviewed the videos and removed duplicates, non-English videos, and those not presenting ENDS modifications, resulting in 168 videos. The oldest video was posted on May 1, 2013, and the most recent video was posted on March 14, 2019. A video featured one or multiple types of ENDS modifications. Specifically, modifications to the coil were the most frequently portrayed in the videos (70%), followed by modification of e-liquids (26%), battery modifications (8%), and refilling nonreusable pods with e-liquids (5%) [13].

Sorted by the number of likes to the comments, each video's top 20 comments were retrieved. Notably, only initial comments directed to the videos were retrieved. Replies to comments were rare in our data collection and excluded. If a video had fewer than 20 comments, then all comments were included. A total of 3103 comments were eventually identified. After removing comments that were non-English, the final data set included 2859 comments on ENDS modification videos. The oldest comment was posted 8 years ago, and the most recent comment was posted on March 14, 2019. Results on the content analyses of videos have been previously reported [13].

Coding Procedures

The first author and a research assistant served as coders for this study. First, the 2 coders reviewed the comments multiple times to become familiar with the data. Next, open and axial coding [26] was conducted to identify prevalent types of comment content and create categories. As an interpretive process, open coding involves describing, naming, and classifying the observed data. Axial coding is an inductive process aimed at identifying higher-level concepts that organize subordinate types into broader categories [26]. In the open coding process, each coder independently generated a list of topics. Then, they compared the degree of overlap between their lists. During the axial coding phase, common topics were combined into overarching categories. In total, we identified 13 subordinate types of comment content, which were then grouped into 4 major categories. See Table 1 for examples and frequencies.

XSL•FO
RenderX

Table 1. Examples and frequencies of types of content in the video comments (n=2859).

Category/type[a]	Example	Value, n (%)[b]
1. Interactions with creators and others		1527 (53.41)
1.1. Appreciation and compliment	*Thank you for showing how to change cotton.*	952 (33.30)
1.2. Request for more videos	*Hey man! Love your tutorials! Was wondering if you would make a video for the sx33 chip?*	135 (4.72)
1.3. Clarification and advice seeking	*Do you think there is a way to allow the resistance meter to be used longer than 3s?*	462 (16.16)
1.4. Purchase inquiry	*Where can you purchase replacement parts?*	67 (2.34)
2. Modification and tobacco use behaviors		782 (27.35)
2.1. Experiences of ENDS[c] modifications	*Just made my own e-juice today. Working on prototyping adjustable airflow.*	430 (15.04)
2.2. Experiences of tobacco use	*I started smoking at 14 but switched to vaping in college to quit smoking cigarettes.*	167 (5.84)
2.3. Modification intentions	*The coil I ordered just arrived today. Now I'll try it the way you suggested.*	198 (6.93)
2.4. New to ENDS use and modifications	*I'm brand new to vaping. This video helped me out a lot. Thanks dude.*	34 (1.19)
3. Modification knowledge		727 (25.43)
3.1. Providing additional info	*Cool! Also, if you choose to buy flavoring for E-Liquid use, they must be Di-acetyl Free.*	346 (12.10)
3.2. Gained new knowledge or skills	*Wow! I finally ran into the right video and know how to build now.*	227 (7.94)
3.3. Criticism and "better" alternatives	*This is wrong. For Scottish roll, you should remove hard parts. Better flavor and wicking.*	162 (5.67)
4. Risks and safety		143 (5)
4.1. Dangers of ENDS modifications	*This causes way too much dry burning. It is ridiculously dangerous and stupid.*	136 (4.76)
4.2. Health risks of tobacco use	*Now that's how you get popcorn lung.*	7 (0.24)
5. Regulation attitudes		15 (0.52)
5.1. Antiregulation	*US government is absolutely ridiculous, pretty soon they're going to regulate our toilet paper.*	13 (0.45)
5.2. Neutral	N/A[d]	0 (0)
5.3. Proregulation	*Just forbid them (combustible products) as they're a major health concern and can cause several types of cancer.*	2 (0.07)

[a]Categories and types are not mutually exclusive. A comment may include multiple types of content (eg, having both Types 1.1 and 3.1).

[b]The value represents the number (percentage) of comments that included at least one content type, which may be smaller than adding up the numbers (percentages) of all comment types.

[c]ENDS: electronic nicotine delivery system.

[d]N/A: not applicable.

A codebook was developed based on the open and axial coding processes. We were also interested in whether and how viewers mentioned tobacco and ENDS modification regulations in their comments. Thus, a fifth category, "regulation attitudes," was added to the codebook, including 3 types of content: antiregulations, neutral toward regulations, and proregulations (Table 1). The unit of analysis was each comment. We coded for the presence (1=presence, 0=absence) of each type of content. Types of content were not mutually exclusive. For instance, the comment, "I'm not new to vaping but I am new to the RDA (rebuildable dripping atomizer) and this video was very helpful and informative. Thank you for doing the video. I learned A LOT from you," was coded as "appreciation and compliment" and "gained new knowledge or skills." The 2 coders both coded a random 30% of the sample (n=900). Intercoder reliability was high, with Krippendorff α values ranging from .89 to .92. Disagreements were resolved through discussion. The remaining 1959 comments were divided evenly and randomly assigned to each coder.

Data Analysis

Data were analyzed using SPSS version 27 (IBM, Inc.). We performed descriptive statistics to assess the frequency of each comment type and category.

Ethical Considerations

Ethics approval is not needed because the study only analyzed publicly available data and the results do not contain any identifiable information.

XSL•FO
RenderX

Results

Comments on ENDS modification YouTube videos included 5 categories: interactions with creators and others, modification and tobacco use behaviors, modification knowledge, risks and safety, and regulation attitudes. Each category comprised several subordinate types of comment content. On average, each comment included 17.35 words (SD 25.81), with a range from range 1 to 569.

Interactions With Creators and Others

More than half (1527/2859, 53.41%) of the comments included the content directed to video creators or interactions with other audience members, in which the viewers complimented and appreciated the creators, requested more videos from the producers, asked for clarification, and inquired about product purchases. Specifically, about 1 in 3 comments (952/2859, 33.30%) thanked the video producers for creating the videos or praised the video content and attributes: "Thank you for showing how to change cotton," "I really appreciate this well-detailed tutorial," and "You put a lot of work in your video. Thanks." A few (135/2859, 4.72% of comments) requested the creators to make more videos in general or about a particular product or ENDS modification. For instance, one comment read, "We'd like to see a Final Boss Vapes Review!" Others asked, "Can you do a review on the goblin mini-RTA please?" and "Can you do a video about how to change the coil and cotton?"

In addition, viewers asked questions about ENDS modifications in their comments. Some (462/2859, 16.16%) sought advice on or clarification for ENDS modifications. For instance, "(If I) wire the computer supply to 5v, do you think it will be safe?" "So if I have two 35a batteries, will they not work with the fuses?" and "So you are saying that the old coil is not recyclable/re-usable??" Others (67/2859, 2.34%) inquired what and where to purchase products for ENDS use and modifications. Examples included, "Can you give me a list of things to salvage them from or where to buy them online?" "How can I buy it?" and "Smok alien vs. vaporesso revenger, which one should I buy?" Notably, given that the comments are often visible to everyone, while the comments might initially be directed to the video creators, those questions could be reviewed and answered by both the creators and other viewers.

Modification and Tobacco Use Behaviors

More than 1 in 4 (782/2859, 27.35%) comments mentioned viewers' own experiences and intentions of ENDS modifications and tobacco use. This category represents how users were involved with and planned to engage in ENDS modifications. Specifically, 15.04% (430/2859) of comments described viewers' past and current experiences of modifying ENDSs. For example, one comment stated, "I've just made a 4x32awg rods build then I twisted both ends to make it more of a clapped cable look that ohm'd out at .51." Another said, "I made one last week to see how I was vaping in 213. And I can surely say that I'm much happier with my builds now, with a lot of surface area and organic cotton." Moreover, 5.84% (167/2859) of comments mentioned viewers' tobacco use experiences without explicitly referring to ENDS modifications. Some described their current ENDS use (eg, "I vape at 25 watts on my baby alien"), whereas others mentioned switching from smoking to vaping: "I used to smoke when I was 15, then switched to vapes," "I switched from smoking 6 a day to vaping and I'm a thousand times healthier for it," and "Vape is mainly for quitting cigarettes but honestly it could be enjoyed by anyone."

Notably, 6.93% (198/2859) of comments indicated viewers' intentions to modify ENDS devices or to change how they modified their ENDS products, especially after watching the modification videos. One comment said, "I made that coil in minutes with 24 gauge. You are right (that) it burns hot but easy to build. It was too much for my RDA. Thanks for the video! I will try it again with a better bigger RDA." Other examples included, "Downloaded your pdf and printed it, I will try a build soon!" "I will try lowering the % of flavorings and see if that does the trick," and "Thanks for the tutorial, lady, I am going to have fun with those." In addition, 1.19% (34/2859) of comments explicitly indicated that the viewers were new to ENDS use or ENDS modifications: "I just got my DOVPO 5., a beginner actually starting vaping a week ago," "I am a newbie at the vape game so thank you!" and "I am a newbie to cloud chasing." The comments suggest that ENDS modification videos may motivate ENDS use and modifications.

Modification Knowledge

In addition to modification behaviors, a quarter (727/2859, 25.43%) of comments addressed users' knowledge of ENDS modifications. In these comments, viewers provided additional information to complement the video content, indicated that they gained new knowledge or skills, and criticized the video and offered "better" ways to modify ENDS. First, 12.10% (346/2859) of the comments added more tips or recommendations for ENDS modifications. One comment read, "This is probably the best way to mix especially when dealing with small batches like when developing a new flavor. The smaller the batch the more important it is to keep every measurement as accurate as possible...All new DIYers should take the time to get those gravity numbers and mix by mass instead of volume." Another comment added to the video, "You are right, there is some basic stuff to know when vaping, and quality of CBD is important too!...I also buy on plantandhemp.com, you know them? Good brands and they only work with quality verified brands, I use them a lot!"

Viewers stated in their comments (227/2859, 7.94%) that they have learned specific knowledge and skills related to ENDS and ENDS modification after watching the videos. Examples included, "In the first 3 minutes you gave me a better understanding of voltage/p.d. than I ever had in physics class," "Thanks so much for this. This helped explain ohms law in a very easy to understand way and helped me get started on everything I need to know," and "This is exactly what I was looking for to fully understanding the principles of vaping." However, 5.67% (162/2859) of comments disagreed with the videos and mentioned "better" alternatives to ENDS modifications. One comment read, "You are soooo incredibly wrong on PG (Propylene Glycol). Do a little more research prior to posting a video on such a vast platform." Another comment stated, "Awesome. But I wouldn't use the syringe to stir, because

XSL•FO
RenderX

it'll clog and gunk it up quicker and you'll have to replace the syringes more often."

Risks and Safety

Only 5% (143/2859) of comments mentioned the health risks and safety concerns of ENDS modifications and tobacco use. Among those comments, the majority (4.76%, 136/2859, of all comments) focused on the risks of ENDS modifications. Some viewers raised their concerns: "Copper? Isn't it toxic...Will cause cancer," "Very cool for learning purposes but I would not recommend using that. Very dangerous elements you are using such as copper and zinc," and "This is how you burn your [expletive] hands." Others mentioned the actual adverse consequences of ENDS modifications: "My friend is in hospital because of this," and "Did this step-by-step and caught on fire." In addition, 7/2859 comments (0.24%) explicitly mentioned the negative health effects of tobacco products: "Just forbid them (combustible products) as they're a major health concern and can cause several types of cancer," "Vaping is too dangerous," and "It (vaping) is as bad as smoking. It has toxic chemicals too."

Regulation Attitudes

We added this last category to the codebook to explore whether and how viewers mentioned tobacco and ENDS modification regulations in their comments. Only 15/2859 (0.52%) comments explicitly mentioned regulations, among which 13/2859 (0.45%) comments were against tobacco and ENDS modification regulations. One comment stated, "If the FDA bans flavors, then the US economy will sink simply because it's keeping a LOT of people working in vape shops, e-liquid makers, marketing people who make labels, etc. It will be the worst decision the FDA has ever made." Another comment also expressed concerns about ENDS regulations, "Vaping is becoming so popular that the FDA now doesn't think it's a good enough option to quit smoking and wants to ban flavored liquid other than tobacco :(." No comments were neutral toward regulation. Only 2/2859 (0.07%) comments were supportive of regulations: "Seriously, this is probably why the FDA is fighting to cripple the vaping industry. Vaping may be a safer alternative to smoking, but it won't stay that way for long if these mods keep getting more, and more powerful. If this keeps up, smoking may eventually become the safer alternative."

Discussion

Principal Findings

Given the emerging evidence that ENDS modifications may result in adverse health consequences [10,15,16] and the popularity of YouTube to share and obtain information about ENDS products and modifications [12,21,22], this study aimed to explore how viewers respond to modification videos and discuss ENDS use and modifications. Specifically, we analyzed the common topics of users' comments left on YouTube modification videos, most of which featured coil, e-liquid, and battery modifications. A content analysis identified 5 common categories and various subordinate types of comment content. The results suggest several concerning trends, including the positive reactions to ENDS modifications, potential motivating effects of modification videos, and lack of mentions of ENDS risks and regulations.

Our results showed that about 1 in 5 comments mentioned viewers' own experiences of ENDS modifications and use. In addition, nearly 1 in 3 comments thanked the creators for the videos. About 5% (135/2859, 4.72%) of comments also included requests for more videos. An interview study of ENDS enthusiasts showed that while the prevalence of ENDS modifications might have peaked a few years ago, some hobbyists continued to build their own coils and batteries, and many users continued to misuse e-liquids [12]. Likewise, our results also revealed the acceptance and popularity of coil modification, battery alternation, and e-liquid customization. Moreover, in 18.50% (529/2859) of the comments, viewers inquired about product purchases, asked for clarifications, and sought advice about ENDS modifications from the video creators and other viewers. This demonstrates the demand for and appreciation of modification information.

Another concerning finding is that by providing ENDS modification information in an educational form, YouTube modification videos may motivate viewers, especially young audiences, to use ENDSs and engage in ENDS modifications. A prior content analysis revealed that most YouTube videos had positive portrayals of ENDS modifications without safety warnings [13]. Our results showed that 7.94% (227/2859) of comments explicitly mentioned that the viewers had learned new skills and knowledge related to modifications of their devices and e-liquids. Nearly 7% (198/2859, 6.93%) of the comments also indicated that the viewers intended to modify their ENDS devices or changed their modification activities after watching the videos. Thus, exposure to modification videos may result in positive attitudes toward ENDS modifications, increased modification knowledge, and in turn greater intentions to use and modify ENDS devices.

Moreover, in the United States, about 77% of individuals aged 18-25 years use YouTube [27]. This demographic group often experiments with cigarettes and ENDSs [28] and is frequently targeted by tobacco companies [29,30]. Indeed, when sharing their experiences of tobacco use and ENDS modifications, some viewers indicated that they started to use tobacco products at an early age (eg, "I used to smoke when I was 15," and "At 16, my brother and I began to vape."). It is alarming that modification videos, which may encourage ENDS use and modifications, are accessible to young audiences across the world. Thus, more attention should be devoted to the impacts of online modification information, such as YouTube videos, on people's ENDS use and modifications.

In contrast to the positivity and support shown in most comments, only 5% (143/2859) of the comments directly mentioned the health risks and safety concerns of ENDS modifications and tobacco use. Moreover, among the 15 comments that explicitly mentioned regulations, 13 were against regulations of ENDS use and modifications. Only 2 comments clearly stated that certain ENDS products and modifications should be banned or regulated. The results are not surprising given that most modifications videos did not include a safety warning [13], and many viewers of the videos, especially those

XSL•FO
RenderX

who left a comment, may have already had modification experiences as well as hold positive attitudes toward ENDS use and modifications. However, the small percentage of comments mentioning risks and regulations are indeed alarming. Tobacco regulatory sciences and agencies should be aware of the YouTube ENDS videos and investigate the cognitive, emotional, and behavioral effects of those modification videos on YouTube viewers.

Limitations

There are several limitations of this study. First, we only collected comments directed to the videos and excluded replies to existing comments. While reply comments were infrequent in our data collection and our approach helped focus on how viewers respond to video content, we left the potential interactions and dynamics between viewers and creators for future research. Moreover, our decision to collect the top 20 comments of each video on one day may not capture the dynamics and full landscape of the comments. In addition, we focused on English comments only. Yet, many YouTube videos are accessible across the globe. We were unable to explore how non-English speaking viewers react to and think about ENDS use and modifications. Furthermore, we did not know the demographics and other characteristics of viewers who left comments. Moreover, although many comments mentioned that viewers had gained new knowledge and intended to modify their ENDS devices, no causal relationships can be established in a content analysis. Our results were descriptive in nature. Future experimental studies should explicitly investigate how modification videos affect viewers. Lastly, our sample was collected in early 2019, before the report of the first e-cigarette or vaping use-associated lung injury (EVALI) case and the COVID-19 pandemic. We do not know how those public health crises may affect ENDS modification activities.

Conclusion

Notwithstanding the limitations, to the best of our knowledge, this is the first study exploring the content of YouTube ENDS modification video comments. Our results indicated the acceptance and popularity of ENDS modifications among users and potential users. Some comments also suggested that the videos motivated current and new ENDS users to alter their ENDS devices. Few comments mentioned the health risks and safety concerns of ENDS modification, and very few mentioned ENDS product and modification regulations, among which only 2 comments clearly supported regulations. Tobacco regulatory researchers and agencies should be aware of online ENDS modification information. More research and attention should be devoted to understanding the impacts of online modification messages.

Acknowledgments

We thank Tiffany A Olszuk for her assistance in the coding process. Research reported in this publication was supported by the National Institute of Drug Abuse of the National Institutes of Health and Food and Drug Administration Center for Tobacco Products (R01DA047397). 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 Food and Drug Administration.

Conflicts of Interest

DLA has received funds for work done for the World Health Organization Tobacco Free Initiative, as a Special Government Employee of the US Food and Drug Administration, as a consultant for Pfizer, as an employee of Cherokee National Operational Systems, and as an independent contractor for McKing Consulting.

References

1. Walley SC, Wilson KM, Winickoff JP, Groner J. A Public Health Crisis: Electronic Cigarettes, Vape, and JUUL. Pediatrics 2019 Jun 23;143(6):e20182741. [doi: 10.1542/peds.2018-2741] [Medline: 31122947]
2. Hajek P, Phillips-Waller A, Przulj D, Pesola F, Myers Smith K, Bisal N, et al. A Randomized Trial of E-Cigarettes versus Nicotine-Replacement Therapy. N Engl J Med 2019 Feb 14;380(7):629-637. [doi: 10.1056/nejmoa1808779]
3. Patil S, Arakeri G, Patil S, Ali Baeshen H, Raj T, Sarode SC, et al. Are electronic nicotine delivery systems (ENDs) helping cigarette smokers quit?-Current evidence. J Oral Pathol Med 2020 Mar;49(3):181-189. [doi: 10.1111/jop.12966] [Medline: 31642553]
4. National Academies of Sciences, Engineering, and Medicine. Public health consequences of e-cigarettes. Washington, DC: The National Academies Press; 2018.
5. Goniewicz ML, Smith DM, Edwards KC, Blount BC, Caldwell KL, Feng J, et al. Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes. JAMA Netw Open 2018 Dec 07;1(8):e185937 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.5937] [Medline: 30646298]
6. McConnell R, Barrington-Trimis JL, Wang K, Urman R, Hong H, Unger J, et al. Electronic Cigarette Use and Respiratory Symptoms in Adolescents. Am J Respir Crit Care Med 2017 Apr 15;195(8):1043-1049 [FREE Full text] [doi: 10.1164/rccm.201604-0804OC] [Medline: 27806211]
7. Wang B, Liu ST, Rostron B, Hayslett C. Burn injuries related to E-cigarettes reported to poison control centers in the United States, 2010-2019. Inj Epidemiol 2020 Jul 20;7(1):36 [FREE Full text] [doi: 10.1186/s40621-020-00263-0] [Medline: 32684171]

XSL•FO
RenderX

8. Eid MS, Ewees MG, El-Mahdy MA, Zweier JL. Chronic exposure to electronic cigarettes induces lung oxidative stress, inflammation, fibrosis,impaired DNA repair. The FASEB Journal 2022 May 13;36(1):S1. [doi: 10.1096/fasebj.2022.36.S1.L8022]
9. Regan A, Bombard J, O'Hegarty M, Smith R, Tong V. Adverse Birth Outcomes Associated With Prepregnancy and Prenatal Electronic Cigarette Use. Obstet Gynecol 2021 Jul 01;138(1):85-94. [doi: 10.1097/AOG.0000000000004432] [Medline: 34259468]
10. Brown CJ, Cheng JM. Electronic cigarettes: product characterisation and design considerations. Tob Control 2014 May;23 Suppl 2:ii4-i10 [FREE Full text] [doi: 10.1136/tobaccocontrol-2013-051476] [Medline: 24732162]
11. Centers for Disease Control and Prevention (CDC). Outbreak of lung injury associated with the use of e-cigarette, or vaping, products. CDC. 2020. URL: https://www.cdc.gov/tobacco/basic_information/e-cigarettes/severe-lung-disease.html [accessed 2022-06-28]
12. Li Y, Fairman RT, Churchill V, Ashley DL, Popova L. Users' Modifications to Electronic Nicotine Delivery Systems (ENDS): Interviews with ENDS Enthusiasts. Int J Environ Res Public Health 2020 Feb 02;17(3):918 [FREE Full text] [doi: 10.3390/ijerph17030918] [Medline: 32024230]
13. 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]
14. Massey ZB, Fairman RT, Churchill V, Ashley DL, Popova L. "It's Cool, Modifying and All, but I Don't Want Anything Blowing Up on Me:" A Focus Group Study of Motivations to Modify Electronic Nicotine Delivery Systems (ENDS). Int J Environ Res Public Health 2021 Nov 09;18(22):11735 [FREE Full text] [doi: 10.3390/ijerph182211735] [Medline: 34831491]
15. Talih S, Balhas Z, Salman R, Karaoghlanian N, Shihadeh A. "Direct Dripping": A High-Temperature, High-Formaldehyde Emission Electronic Cigarette Use Method. Nicotine Tob Res 2016 Apr;18(4):453-459 [FREE Full text] [doi: 10.1093/ntr/ntv080] [Medline: 25863521]
16. Kosmider L, Sobczak A, Fik M, Knysak J, Zaciera M, Kurek J, et al. Carbonyl compounds in electronic cigarette vapors: effects of nicotine solvent and battery output voltage. Nicotine Tob Res 2014 Oct;16(10):1319-1326 [FREE Full text] [doi: 10.1093/ntr/ntu078] [Medline: 24832759]
17. Ahmed AR, Etchey B, Ahmed M. Explosions, Burn Injuries and Adverse Health Effects of Electronic Nicotine Delivery Systems: A Review of Current Regulations and Future Perspectives. J Pharm Pharm Sci 2021 Sep 05;24:462-474 [FREE Full text] [doi: 10.18433/jpps32242] [Medline: 34499601]
18. Flores C, Chestovich P, Saquib S, Carroll J, Al-Hamad M, Foster K, et al. Electronic Cigarette-Related Injuries Presenting to Five Large Burn Centers, 2015-2019. J Burn Care Res 2021 Nov 24;42(6):1254-1260. [doi: 10.1093/jbcr/irab114] [Medline: 34143185]
19. Almazrouei ES, Bintamim AA, Khalil SEA, Alremeithi R, Gewily S. The identification of drugs of abuse in E-cigarette samples seized in Dubai between 2016 and 2020. Forensic Sci Int 2022 Apr;333:111233. [doi: 10.1016/j.forsciint.2022.111233] [Medline: 35228141]
20. Bendel G, Hiller H, Ralston A. Nicotine Toxicity Secondary to Aftermarket Modifications to a Vaping Device. Mil Med 2022 Jul 01;187(7-8):e1007-e1010. [doi: 10.1093/milmed/usab223] [Medline: 34114039]
21. Collins L, Glasser AM, Abudayyeh H, Pearson JL, Villanti AC. E-Cigarette Marketing and Communication: How E-Cigarette Companies Market E-Cigarettes and the Public Engages with E-cigarette Information. Nicotine Tob Res 2019 Jan 01;21(1):14-24 [FREE Full text] [doi: 10.1093/ntr/ntx284] [Medline: 29315420]
22. McCausland K, Maycock B, Leaver T, Jancey J. The Messages Presented in Electronic Cigarette-Related Social Media Promotions and Discussion: Scoping Review. J Med Internet Res 2019 Feb 05;21(2):e11953 [FREE Full text] [doi: 10.2196/11953] [Medline: 30720440]
23. 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]
24. Madden A, Ruthven I, McMenemy D. A classification scheme for content analyses of YouTube video comments. Journal of Documentation 2013 Sep 02;69(5):693-714. [doi: 10.1108/JD-06-2012-0078]
25. 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]
26. Strauss A, Corbin J. Basics of qualitative research: Techniques and procedures for developing grounded theory. 2nd ed. Newbury Park, CA: Sage; 1990.
27. Ceci L. Percentage of U.S. internet users who use YouTube as of 3rd quarter 2020, by age group. Statista. 2022. URL: https://www.statista.com/statistics/296227/us-youtube-reach-age-gender/ [accessed 2022-02-15]
28. Centers for Disease Control and Prevention (CDC). E-cigarette use triples among middle and high school students in just one year. CDC. 2022. URL: https://www.cdc.gov/media/releases/2015/p0416-E-cigarette-use.html [accessed 2022-02-23]

XSL•FO
RenderX

29. Henriksen L, Schleicher NC, Dauphinee AL, Fortmann SP. Targeted advertising, promotion, and price for menthol cigarettes in California high school neighborhoods. Nicotine Tob Res 2012 Jan;14(1):116-121 [FREE Full text] [doi: 10.1093/ntr/ntr122] [Medline: 21705460]
30. Pierce JP, Messer K, James LE, White MM, Kealey S, Vallone DM, et al. Camel No. 9 cigarette-marketing campaign targeted young teenage girls. Pediatrics 2010 Apr;125(4):619-626. [doi: 10.1542/peds.2009-0607] [Medline: 20231181]

Abbreviations

CDC: The Centers for Disease Control and Prevention
EVALI: e-cigarette or vaping use-associated lung injury

Edited by T Mackey; submitted 25.03.22; peer-reviewed by K Luo, L Laestadius; comments to author 09.06.22; revised version received 28.06.22; accepted 26.07.22; published 12.08.22.

Please cite as:
Li Y, Ashley DL, Popova L
Users' Modifications to Electronic Nicotine Delivery Systems: Content Analysis of YouTube Video Comments
JMIR Infodemiology 2022;2(2):e38268
URL: https://infodemiology.jmir.org/2022/2/e38268
doi:10.2196/38268
PMID:35992739

©Yachao Li, David L Ashley, Lucy Popova. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 12.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Confounding Effect of Undergraduate Semester–Driven “Academic" Internet Searches on the Ability to Detect True Disease Seasonality in Google Trends Data: Fourier Filter Method Development and Demonstration

Timber Gillis[1], BSc; Scott Garrison[1], BASc, MD, PhD

Department of Family Medicine, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Scott Garrison, BASc, MD, PhD
Department of Family Medicine
University of Alberta
6-60 University Terrace 8303 112 Street NW
Edmonton, AB, T6G 2T4
Canada
Phone: 1 780 248 1853
Email: scott.garrison@ualberta.ca

Abstract

Background: Internet search volume for medical information, as tracked by Google Trends, has been used to demonstrate unexpected seasonality in the symptom burden of a variety of medical conditions. However, when more technical medical language is used (eg, diagnoses), we believe that this technique is confounded by the cyclic, school year–driven internet search patterns of health care students.

Objective: This study aimed to (1) demonstrate that artificial “academic cycling” of Google Trends’ search volume is present in many health care terms, (2) demonstrate how signal processing techniques can be used to filter academic cycling out of Google Trends data, and (3) apply this filtering technique to some clinically relevant examples.

Methods: We obtained the Google Trends search volume data for a variety of academic terms demonstrating strong academic cycling and used a Fourier analysis technique to (1) identify the frequency domain fingerprint of this modulating pattern in one particularly strong example, and (2) filter that pattern out of the original data. After this illustrative example, we then applied the same filtering technique to internet searches for information on 3 medical conditions believed to have true seasonal modulation (myocardial infarction, hypertension, and depression), and all bacterial genus terms within a common medical microbiology textbook.

Results: Academic cycling explains much of the seasonal variation in internet search volume for many technically oriented search terms, including the bacterial genus term [“Staphylococcus”], for which academic cycling explained 73.8% of the variability in search volume (using the squared Spearman rank correlation coefficient, $P<.001$). Of the 56 bacterial genus terms examined, 6 displayed sufficiently strong seasonality to warrant further examination post filtering. This included (1) [“Aeromonas” + “Plesiomonas”] (nosocomial infections that were searched for more frequently during the summer), (2) [“Ehrlichia”] (a tick-borne pathogen that was searched for more frequently during late spring), (3) [“Moraxella”] and [“Haemophilus”] (respiratory infections that were searched for more frequently during late winter), (4) [“Legionella”] (searched for more frequently during midsummer), and (5) [“Vibrio”] (which spiked for 2 months during midsummer). The terms [“myocardial infarction”] and [“hypertension”] lacked any obvious seasonal cycling after filtering, whereas [“depression”] maintained an annual cycling pattern.

Conclusions: Although it is reasonable to search for seasonal modulation of medical conditions using Google Trends’ internet search volume and lay-appropriate search terms, the variation in more technical search terms may be driven by health care students whose search frequency varies with the academic school year. When this is the case, using Fourier analysis to filter out academic cycling is a potential means to establish whether additional seasonality is present.

(JMIR Infodemiology 2022;2(2):e34464) doi:10.2196/34464

XSL•FO
RenderX

KEYWORDS

Google Trends; seasonality; Fast Fourier transform; FFT; pathogenic bacteria; depression; Google search; Google; health information; health information seeking; internet search

Introduction

Google Trends and Disease Seasonality

Google Trends is an open access portal that allows researchers to explore how the public's quest for information on specific topics varies with time. The data made available by Google Trends is the "volume" (number) of searches for a specific search term entered by the public into the Google search engine per unit time (eg, per week), provided as a percentage of the highest search volume for that term over the period of interest (eg, last 5 years). The data are anonymous and collated geographically, and, given the public use of Google to search for health information [1,2], has been used to establish unexpected seasonality in the symptom burden or incidence of a variety of chronic conditions [3-5]. To describe such population-level investigations of disease processes using web-based data sources, Eysenbach [6] has coined the term "infodemiology."

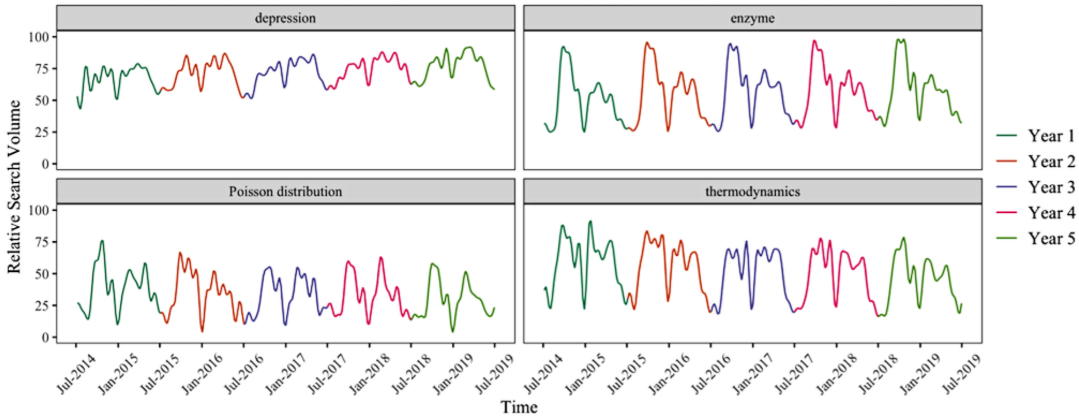
"Seasonality" in symptom burden refers to an annual periodicity, or modulation, in some measurable aspect of those symptoms. Much of this modulation may result from seasonal variation in environmental factors that convey the risk of disease. Respiratory viral illnesses are one of the best examples of this [7]. Humidity, temperature, and wind speed are all seasonally modulated, and each factor influences the spread of air-borne pathogens [8]. Mammals additionally have some seasonal modulation of their physiology (eg, body weight, fur thickness, and estrus). While this is not commonly thought of for humans, some studies suggest that even our physiology has some seasonality. Examples of this include higher long-bone growth in children during summer, retention of extracellular water starting in spring, continuing into the summer for patients on dialysis and increased immune system reactivity during winter [9-11]. Outside of infectious diseases, seasonality has also been observed in depression, cardiovascular disease, and overall mortality [12-14]. Recognizing and trying to understand the driving forces behind disease seasonality helps deliver insights that might lead to more effective prevention or treatment of seasonally modulated conditions.

Google Trends has become a popular tool for investigation of disease seasonality. An early use in this area was rapid real-time surveillance of influenza-like illness [15], something that continues to be worked on to augment conventional public health surveillance measures [16]. Others have sought to uncover unexpected seasonality in common conditions such as nocturnal leg cramps, ankle swelling, dental carries, and various mental health disorders [4,5,17-19]. However, a variety of things can confound the use of big data sources such as Google Trends search volume for health information as a proxy for symptom burden [20]. Search terms, for instance, might have dual meanings. Shingles is a disease, but they are also roofing tiles, whose use and related searches might be seasonal in Northern (snow experiencing) climates. Medical conditions can also be more or less newsworthy (eg, when celebrities are involved), and news coverage can sometimes drive search volume more than personal experience with the condition [21]. Influenza surveillance, for instance, has been inconsistent in its predictive ability when compared to hospital-based viral detection [22].

In our use of Google Trends to explore disease seasonality, we have come across an important potential confounder, which has yet to be described. This confounder is the searches for health information carried out by students who are taking courses at the undergraduate level. Such searches can be expected to be low in volume during the summer and winter break (in most countries) and high in volume during the final examination season. We have repeatedly observed such a biphasic seasonal pattern, which we will refer to as "academic cycling," in many academic-oriented search terms (ie, fairly technical terms that are less commonly used in lay conversation such as proper diagnoses). Such academic cycling spans all fields of study. Some examples from health care, mathematics, and physics are shown in Figure 1. This same academic cycling pattern is clearly present in some of the infodemiology literature, but, even when it appears to be the main driver of the variation in search volume, it is either not acknowledged as such or not accounted for when its presence is recognized [18,23,24]. In this study, we (1) used the fast Fourier transform (FFT) on Google Trends search volume data with strong academic cycling, (2) identified the frequency domain pattern of that academic cycling, (3) searched for and removed that pattern from the frequency domain of search terms where seasonality is of clinical interest, and (4) recreated the time series data for the terms of clinical interest, with the academic cycling component removed. In so doing, we seek to empower researchers with strategies to investigate whether the seasonal trend they see in their Google Trends data is true, disease-related seasonality, or merely a confounding search pattern introduced by academic, school year–driven search volume.

XSL•FO
RenderX

Figure 1. Google Trends search volume for terms with strong academic cycling in the 5 years prior to onset of the COVID-19 pandemic. Searches are limited to the United States, and each color represents a period of 1 year. A high-frequency filter has been applied to remove fluctuations with a period less than 5 weeks (this smooths the curve and eliminates current event driven search volume spikes that last less than 2.5 weeks).



Fourier Analysis and Filtering

One of the pillars of signal processing is the recognition that time-series data can be represented as the sum of many different sinusoidal waves, each with its own amplitude and phase difference. FFT is a software tool that does just that, representing a given time series (such as our 5-year Google Trends search volume) in the "frequency domain," by showing what sinusoidal waves would need to be added together to produce the same curve [25]. FFT also lets users go backwards from the frequency domain representation and recreate the time series again (the "inverse FFT"). The advantage of this frequency domain representation is that we can think of our data as having a variety of driving forces and, if the frequencies of those driving forces are unique and can be identified, we can potentially remove them in the frequency domain and put the time series back again without the contribution of the unwanted component. A simple example of this, if one is listening to the radio, would be removing high- or low-frequency "noise" from the radio waves to hear voices more clearly. A more complex use of the same technique might be adding or removing an antipiracy frequency domain watermark from a piece of music or an image [26]. FFT has previously been applied to Google Trends data in order to identify the dominant frequency in time series data describing urinary tract infections and chronic lifestyle diseases [27,28].

Methods

Overview

We first demonstrated our filtering process in detail using the term ["thermodynamics"], which was chosen because of its strong academic cycling and helped each step to be visualized. The initial step involved preprocessing of the Google Trends data before FFT could be applied and involved shifting the time-series data down by subtracting the mean value. The resulting transformed data had the same shape as the original time series, but the data were now represented by positive and negative numbers that had a mean value of 0. Although not strictly necessary, we also chose to filter out high-frequency "noise" with FFT to make patterns more visible to the naked eye. These 2 preprocessing steps were applied to both the term of interest and to the control terms that represent the academic cycling that we wish to remove. We then identified how much of the academic cycling component was present in the term of interest by using a least squares regression analysis, subtracted that component in the frequency domain, and recreated the time series with inverse FFT. Following this demonstration, we applied the same technique to a selection of clinically relevant examples.

Google Trends Data Collection

Google Trends time series data are freely downloadable and presented as the relative search volume (RSV) for the specified search terms per unit time (month, week, day, or hour). An RSV of 0 indicates little to no search volume, and an RSV of 100 indicates the highest volume for that term in the period of interest. We used weekly data for the 5-year period from July 3, 2016, to June 30, 2021. We restricted our analysis to the United States since it was the country with the largest internet search volume and since a single geographic region was needed for most residents to have a shared experience of the changing of the seasons and school year. Our 5-year window was selected to capture 5 full academic years. Although Google Trends provides the option of having search terms represent "topics" (in which case Google Trends aggregates a variety of searches they feel capture the same topic area), this option is not available for all search terms. Hence, for consistency, unless otherwise indicated, we did not use the "topic" search feature. Our search term nomenclature is in accordance with previous literature [29].

"Fingerprint" Frequency Filtering

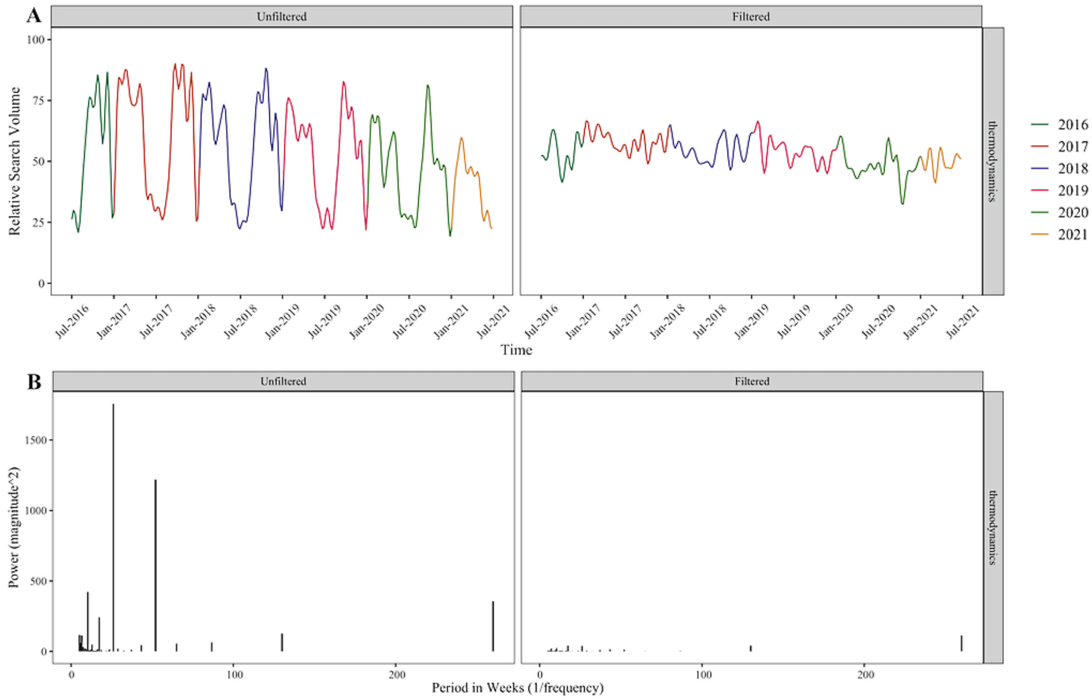
Overview

Our frequency filtering program was built using R (version 4.0.2) within the RStudio interactive development environment (version 1.4.1106). The process for filtering out academic cycling, every time it was applied, used the following steps. We will illustrate each step using the example term ["thermodynamics"], which displays strong academic cycling. When we refer to the time domain, we mean how the data look as a time series (ie, the way Google Trends initially presents the data in their web browser). When we refer to the frequency domain, we mean the way the data are visualized using the FFT, which is as a series of spikes showing how much of each

XSL•FO
RenderX

frequency is present in the data for all of the sinusoids that would need to be combined to create it (Figure 2). Our filtering process involved the following 7 steps.

Figure 2. (A) Time series representation of ["thermodynamics"] Google Trends data both before and after removal of academic cycling; color indicates a calendar year. (B) Frequency domain representations of the same time series. Each frequency domain spike is the amplitude of the sinusoids that would need to be combined to produce the time series shown.



Transformation

We first shifted and scaled the data such that it moved up and down around a mean value of 0 using the following formula:

Transformed RSV = [RSV – mean(RSV)]/mean(RSV)

Once filtering was complete, we applied this transformation in reverse to return to the original scaling.

High-Frequency Filtering

Assuming that most high-frequency fluctuations in search volume (ie, sudden changes) are not biologically driven [21], we removed frequencies with a period less than 5 weeks. This effectively removed spikes in search volume, which rose and fell in less than 2.5 weeks, a period we felt would cover most search volume surges triggered by sporadic events or media reports. The smoothing effect of this high-frequency filtering on the term ["thermodynamics"] is shown in Multimedia Appendix 1.

Converting to the Frequency Domain

After high-frequency filtering, we applied FFT as natively encoded in R [30] to produce the frequency domain representation shown in Figure 2 (which exhibits major frequency components at 52 weeks and 26 weeks). This representation, however, is a simplification that only shows the amplitude of the frequency component. It is also necessary to know the phase of the sinusoids with those frequencies. Numerically, FFT represents each frequency component with 2 numbers (a "real" and "imaginary" component) that define amplitude and the phase of each sinusoid in the same way that x and y coordinates could be used to define the length and position of the tip of a second hand on a clock's face.

Selection of Control Terms (Terms With Strong Academic Cycling)

The academic cycling pattern that we want to filter out could look different for different disciplines considering the school year and that examination schedules could differ. As such, we chose different control terms for our "thermodynamics" example than we did for our medically relevant examples (choosing ["binomial" + "integral" + "derivative"] as control terms for ["thermodynamics"] and ["gram stain" + "gram positive" + "gram negative"] as control terms for biomedical searches). In the Google Trends browser, using a "+" sign means "or"; that is, ["cat" + "dog"] would count any Google search in which the words "cat" or "dog" were included in the search phrase entered by the user.

Identification of the Frequency Domain "Academic Fingerprint" to Be Removed

Similar to our search terms of interest ("thermodynamics" in this example), the search volume for the control terms (ie, ["binomial" + "integral" + "derivative"]) also underwent the first 3 aforementioned steps. The frequency domain pattern of spikes for the control term is the "fingerprint" we intend to filter out of the data for our terms of interest.

Determining How Much of the Academic Fingerprint Frequencies to Remove

To best estimate how much of the academic fingerprint was present in a signal, we used a sum of squares minimization

XSL•FO
RenderX

approach using the optimize algorithm in R. That is, we took the frequency domain representations of both the term of interest and the control term, and scaled the control term components by an amount k, such that the sum of the squared differences in frequency components between term of interest and control was minimized (note that as shown in Textbox 1, this used the sum of the squared differences of each real and imaginary component and not just the amplitudes). For terms that do not display academic cycling, k was close to 0. For terms with a high degree of academic cycling, k was closer to 1. For our example term ["thermodynamics"], the scaling coefficient (k) for the control term ["binomial" + "integral" + "derivative"] was 0.8663. To remove the academic cycling component, we simply subtracted the scaled frequency components of the control from the same components in the term of interest.

Textbox 1. Scaling approach: minimize algorithm in R minimizes the sum of squared differences (SS2) for the scaling coefficient as represented by "k."

$SS^2 = \Sigma(\text{Real Test} - \text{Real Control}*k)^2 + (\text{Imaginary Test} - \text{Imaginary Control}*k)^2)$

Recreating the Time Series Without the Academic Cycling Component

The resultant filtered Fourier coefficients were back-transformed to the time domain using the inverse FFT algorithm, which is part of the same native R function. This allowed us to visualize the time series without the academic cycling, which appears to be eliminated in the "thermodynamics" example (Figure 2).

Selection of Clinically Relevant Terms to Explore

Pathogenic Bacteria

The genus names of pathogenic organisms could be searched for by both patients and providers, who encounter the organism in the usual course of care, and by students learning about such organisms during their training. It is also possible that the abundance of these organisms, their vectors, or the environments in which they are most easily transmitted undergo seasonal modulation. As such, we identified and analyzed 58 pathogenic bacterial genus terms discussed in a common medical microbiology textbook [31]. The genus term ["Bacillus"] was not used as it has a separate meaning in terms of bacterial morphology more generally. After data processing, we also chose to combine the terms ["Aeromonas" + "Plesiomonas"], recognizing both as water-borne pathogens that shared a common taxonomic identity in the past [32]. Recognizing that the search volume for many of these terms would be low, and hence the time series could appear too "noisy" to visibly observe larger trends, we also averaged the Google Trends data for each genus term together to average out random fluctuations and demonstrate whether academic cycling was indeed present in these terms.

Conditions Believed to Have Some Seasonality

We also applied our filtering technique to 3 conditions that appeared to have academic cycling and for which previous observational evidence suggests some seasonal modulation; these include depression, myocardial infarction, and hypertension [12,13,33].

Statistical and Graphical Analyses

Post filtering, for bacterial genus terms, we selected the 6 terms (top 10%) with the strongest annual cycling component (ie, genus names with the highest amplitude frequency domain peaks at 52 weeks) and displayed them graphically. To do this, since these terms generally had a low search volume, and hence a relatively high amount of noise (ie, more seemingly random fluctuations), we graphed the average monthly volume to help average out random fluctuations and make any annual patterns more visible. In order to demonstrate how much the academic searches were driving the search volume for bacterial genus terms, we also calculated the squared Spearman rank correlation coefficient between the time series for each bacterial term and the time series for our control term (ie, ["gram stain" + "gram positive" + "gram negative"]). The squared Spearman rank correlation coefficient was used to estimate the amount of variation in the test data set, which was explained by the variation in the control.

Results

Overview

Our filtering technique successfully removed academic cycling from a wide variety of Google Trends data where it is evident. Although the terms ["depression"], ["hypertension"], and ["myocardial infarction"] all had annual cycling prefiltering, this was only evident in searches for ["depression"] once academic cycling was removed. Of 56 pathogenic bacterial genus names, largely because of low search volumes, only 5 displayed substantial annual cycling prefiltering (["Clostridium"], ["Escherichia"], ["Mycobacterium"], ["Staphylococcus"], and ["Streptococcus"]), and none of these 5 genus names displayed seasonality after academic cycling was removed. After filtering all genus terms, 10% of them with the strongest seasonality (ie, strongest 1-year periodicity in the frequency domain) were ["Aeromonas" + "Plesiomonas"], ["Moraxella"], ["Haemophilus"], ["Ehrlichia"], ["Legionella"], and ["Vibrio"], each of which had search volume peaks consistent with what the clinical literature would predict.

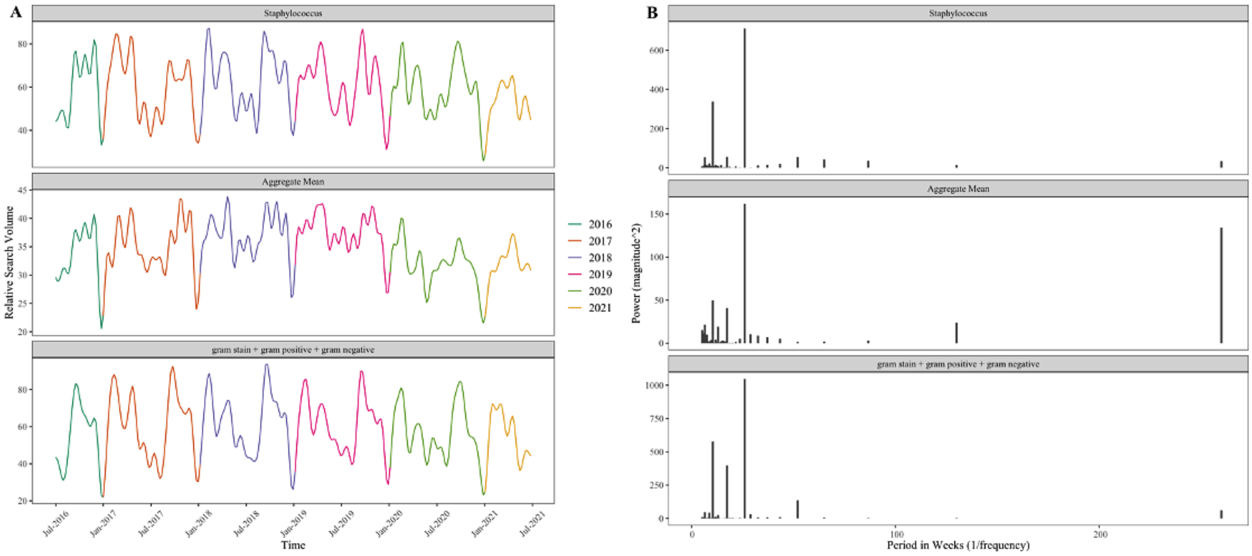
Pathogenic Bacteria

Owing to the relatively low search volume, few of our 56 bacterial genus terms displayed obvious academic cycling, with only 5 having a squared Spearman rank correlation coefficient of ≥0.5 with their corresponding control term. Academic cycling was clearly present, however, when the bacterial genus terms were averaged together and in the term ["Staphylococcus"] (Figure 3). Academic cycling explained three-quarters of the variation in the search volume for "Staphylococcus" (ie, R^2=0.74), and half of the variation in our aggregate of 56 other bacterial terms (R^2=0.55).

XSL•FO
RenderX

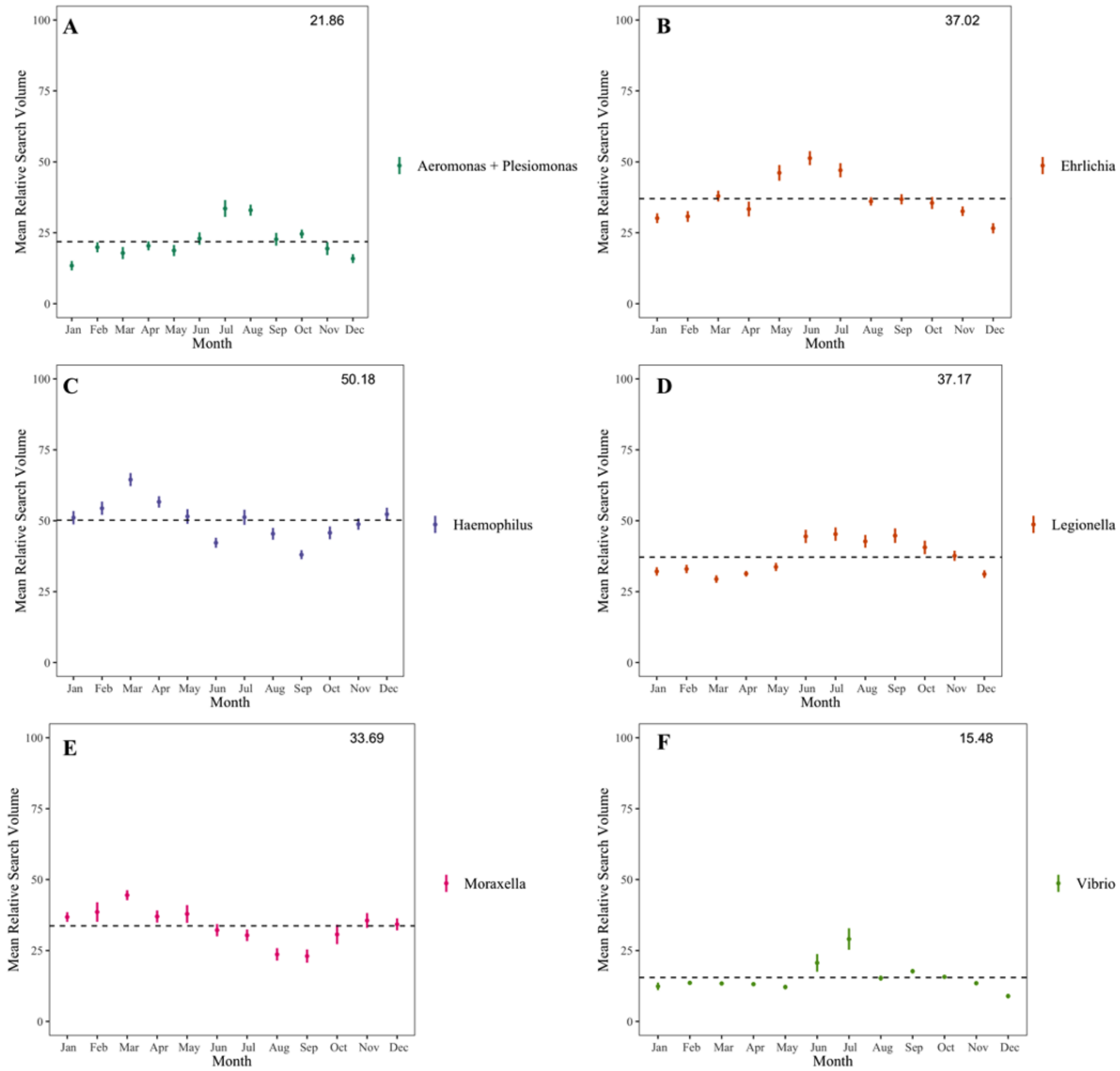
The top 10% of genus terms with the most annual cycling (ie, highest 52-week frequency domain peaks) after filtering out academic cycling are shown in Figure 4. ["Aeromonas" + "Plesiomonas"] searches increased during midsummer, ["Moraxella"] and ["Haemophilus"] searches increased during late winter, and "Ehrlichia" search volume spiked in late spring. ["Legionella"] searches had a slow, sustained peak throughout the summer months and during early fall, and ["Vibrio"] searches had a sharp spike during midsummer. All of these had no visible academic cycling and were essentially unaffected by our filter as demonstrated in Multimedia Appendix 1. Only 5 bacterial genus terms had obvious academic cycling, as demonstrated by a squared Spearman rank correlation coefficient of ≥0.5 for comparison with our control terms. These terms were ["Clostridium"], ["Escherichia"], ["Mycobacterium"], ["Staphylococcus"], and ["Streptococcus"], none of which displayed seasonality after filtering (Figure 5).

Figure 3. (A) High-frequency filtered Google Trends Internet search volumes for ["Staphylococcus"], the aggregate mean of 56 pathogenic bacterial genus term data (excluding ["Staphylococcus"]), and the ["gram stain" + "gram positive" + "gram negative"] control term used to identify academic cycling in such terms; color indicates a calendar year. (B) The frequency domain representation of the same time series, showing the amplitude of each sinusoid that would need to be summed to obtain the original signal.



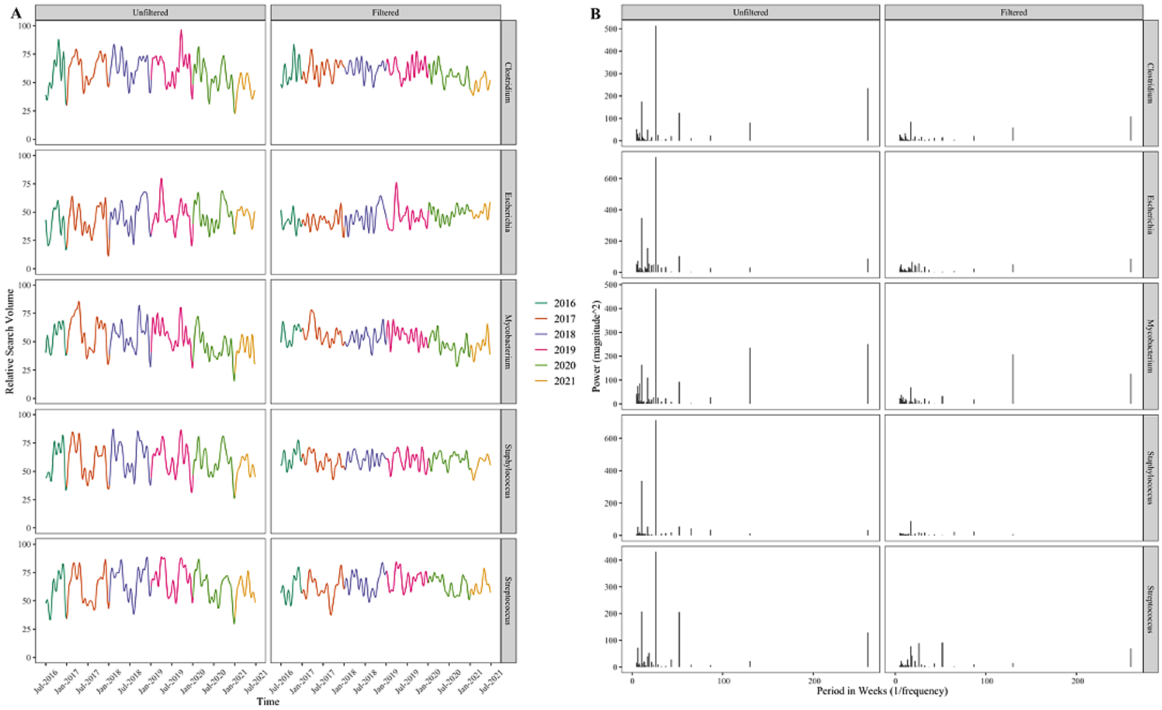
XSL•FO
RenderX

Figure 4. Google Trends internet relative search volume for various pathogenic bacteria, filtered to remove academic cycling, and averaged for each month over a 5-year span from July 3, 2016, to June 30, 2021. (A) ["Aeromonas" + "Plesiomonas"] (combined out of convenience owing to similar reservoirs, similar modes of infection, and historically common taxonomy). (B) ["Ehrlichia"]. (C) ["Haemophilus"]. (D) ["Legionella"]. (E) ["Moraxella"]. (F) ["Vibrio"]. The dotted line is the mean search volume across all 261 data points that are available for averaging. Numbers being averaged are the weekly search volume, obtained as a percentage value of the maximum weekly search volume for that term over the 5-year period.



XSL•FO
RenderX

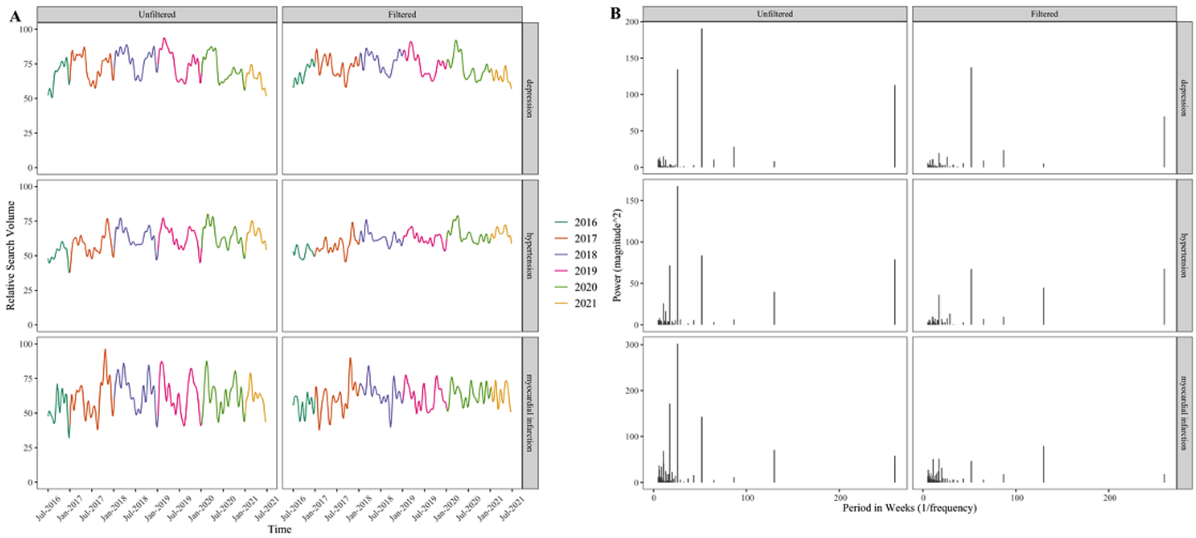
Figure 5. Google Trends internet relative search volume before and after filtering out academic cycling for the terms ["Clostridium"], ["Escherichia"], ["Mycobacterium"], ["Staphylococcus"], and ["Streptococcus"]. (A) These terms in the time domain. (B) The same terms in the frequency domain after applying the fast Fourier transform tool.



Depression, Hypertension, and Myocardial Infarction

Academic cycling is evident in searches for information on all 3 of these common conditions (Figure 6). However, after filtering, only ["depression"] displays what appears to be a strong seasonal pattern in the time domain (corresponding to a dominant 52-week peak in the frequency domain), with searches peaking during winter. The terms ["hypertension"] and ["myocardial infarction"] have small peaks at 52 weeks. This could represent a lesser degree of seasonality or perhaps some residual academic cycling that we failed to remove.

Figure 6. Google Trends internet relative search volume before and after filtering out academic cycling for the terms ["depression"], ["hypertension"], and ["myocardial infarction"]. (A) These terms in the time domain. (B) The same terms in the frequency domain after applying the fast Fourier transform tool.



Discussion

Biphasic academic cycling is commonly seen in Google Trends data when technical search terms are used. When this is the case, it can potentially be filtered out using FFT and an appropriate control. Although initially confounded by academic cycling, true seasonality in the public's searches for information on depression seems to be present. It is less obvious that seasonality is present in searches for information on myocardial infarction and hypertension. Seasonality is also present in searches for information on a variety of pathogenic bacteria.

Biphasic academic cycling patterns are clearly present in some published Google Trends data, but to date, those patterns have

XSL•FO
RenderX

been overlooked or given other interpretations. This includes an exploration of the influence of public health campaigns on searches for information on marijuana use, colorectal cancer, and HIV [23]. The search volume for cannabis use in Canada followed a clear biphasic pattern, peaking in the winter and fall, followed by a summer trough. The same is true for an exploration of worldwide searches for information on osteoporosis, where recognizing the academic search pattern, and interpreting it in the light of when school terms start and end in different countries, might have provided an alternate explanation for the observed seasonality of searches and the observed differences between countries [24]. Academic cycling is also clearly present in an exploration of searches for information about mental health conditions [18]. The authors acknowledged the potential for academic searches to confound their findings but considered that its effect would have been negligible.

The months during which we observed higher interest in internet searches on specific bacterial pathogens are consistent with those reported in the microbiology literature. In Hungary, cases of *Plesiomonas* and *Aeromonas* (water-borne pathogens) have been shown to peak between May and September [34]. In the United States, human Ehrlichiosis due to *Ehrlichia* (a tick-borne pathogen) peak in June and July [35]. In the United States and Belgium, *Legionella* respiratory infections rise in summer and autumn [36,37]. In Japan, *Moraxella* respiratory infections are more common in winter [38]. In patients with cystic fibrosis, *Haemophilus* respiratory infections peak in February and March [39]. Furthermore, in the United States, noncholera *Vibrio* gastroenteritis peaks in the summer [40]. Google Trends data have been used to identify the seasonality of searches for antibiotics and probiotics in general (both of which peak in winter), [41] and for tracking and real-time surveillance for viral infections such as influenza [15,16], but we are unaware of it having been used to track the bacterial pathogens we report here.

We chose ["depression"], ["hypertension"], and ["myocardial infarction"] as terms to explore because each has both academic cycling in Google Trends data and epidemiologic evidence of seasonal modulation. Depression and myocardial infarctions have been shown to be more common in winter [12,13], and blood pressure is higher at the same time [33]. This includes Google Trends data showing more searches for ["depression"] in winter in the northern hemisphere [19]. Although our analyses only found obvious seasonality in searches for depression, Google Trends has limitations when it comes to detecting seasonality. If the bulk of internet searches for information on myocardial infarction are not driven by clinical events, seasonality may not be evident. We also assumed that hypertension diagnoses would also be more common in winter because blood pressure in normotensive individuals is higher in winter. This may not be the case. Conceivably, individuals with hypertension might display less seasonal variation in blood pressure readings than do normotensive individuals. It is also possible that the small 52-week peaks that we observed could have resulted from averaging Google Trends data by month, as we did for the low search volume bacterial terms. This has been done by other investigators exploring ["hypertension"] in Google Trends data from Poland, Australia, and the United States [42]. In each case, a winter peak was demonstrated, with a dip in December that was attributed to people possibly being less concerned about their health during Christmas, a dip we would instead attribute to academic cycling.

Our filtering technique is limited by our ability to use an appropriate control. If the shape of the academic cycling in our control term does not match that of the term of interest, its removal would be imperfect or would introduce other seemingly seasonal components. We chose to use the control term ["gram stain" + "gram positive" + "gram negative"] for all our clinical examples because we believed microbiology-related searches would track with health care searching in general. While future researchers could choose to use this same control term to identify and filter out academic cycling, they may alternatively wish to build control terms that display strong academic cycling, which are more specific to the relevant specialty area. We can also only remove academic cycling when it is obviously present. For lower search volume terms, where there is vast higher-frequency "noise," our filtering method essentially left the waveform intact. As such, our method of averaging together the search volume on a monthly basis to remove some of the noise, and reinforce the seasonal component, would have also reinforced any academic cycling component that was present.

Google Trends internet search volume is a useful tool for detecting disease seasonality when symptoms, or diagnoses, can be expressed in lay terms that have no alternate meaning. Care should be taken, however, to ensure that any emerging cyclic patterns do not have the biphasic pattern that is highly characteristic of searches driven by the academic school year. This is particularly relevant when researchers use more technical terms, such as proper diagnoses. When this is the case, consideration could be given to using the filtering technique we present here, the R script for which is available in Multimedia Appendix 2. With such an approach, we are able to lessen the confounding influence of academic cycling in Google Trends time-series data and increase the likelihood that any residual cycling might have clinical relevance, perhaps being driven by previously unrecognized seasonality that is inherent in human physiology, in the virulence, abundance or reservoirs of pathogenic organisms, or other socioeconomic or behavioral factors that convey risk of illness. Uncovering such seasonality could open up new understanding of human physiology and disease etiology and new opportunities for disease prevention and treatment.

Acknowledgments

The authors would like to thank the Faculty of Medicine and Dentistry at the University of Alberta for providing TG with The David and Beatrice Reidford Research Scholarship to pursue this work.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures.
[DOCX File , 708 KB - infodemiology_v2i2e34464_app1.docx]

Multimedia Appendix 2
Open-source R code for our Fourier analysis filtering methodology adapted for reader use.
[ZIP File (Zip Archive), 3 KB - infodemiology_v2i2e34464_app2.zip]

References

1. Beck F, Richard J, Nguyen-Thanh V, Montagni I, Parizot I, Renahy E. Use of the internet as a health information resource among French young adults: results from a nationally representative survey. J Med Internet Res 2014 May 13;16(5):e128 [FREE Full text] [doi: 10.2196/jmir.2934] [Medline: 24824164]
2. Bernardo TM, Rajic A, Young I, Robiadek K, Pham MT, Funk JA. Scoping review on search queries and social media for disease surveillance: a chronology of innovation. J Med Internet Res 2013 Jul 18;15(7):e147 [FREE Full text] [doi: 10.2196/jmir.2740] [Medline: 23896182]
3. Patel JC, Khurana P, Sharma YK, Kumar B, Ragumani S. Chronic lifestyle diseases display seasonal sensitive comorbid trend in human population evidence from Google Trends. PLoS One 2018;13(12):e0207359 [FREE Full text] [doi: 10.1371/journal.pone.0207359] [Medline: 30540756]
4. Garrison SR, Dormuth CR, Morrow RL, Carney GA, Khan KM. Seasonal effects on the occurrence of nocturnal leg cramps: a prospective cohort study. CMAJ 2015 Mar 03;187(4):248-253 [FREE Full text] [doi: 10.1503/cmaj.140497] [Medline: 25623650]
5. Liu F, Allan GM, Korownyk C, Kolber M, Flook N, Sternberg H, et al. Seasonality of Ankle Swelling: Population Symptom Reporting Using Google Trends. Ann Fam Med 2016 Jul;14(4):356-358 [FREE Full text] [doi: 10.1370/afm.1953] [Medline: 27401424]
6. 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]
7. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of Respiratory Viral Infections. Annu Rev Virol 2020 Sep 29;7(1):83-101. [doi: 10.1146/annurev-virology-012420-022445] [Medline: 32196426]
8. Dbouk T, Drikakis D. Fluid dynamics and epidemiology: Seasonality and transmission dynamics. Phys Fluids (1994) 2021 Feb 01;33(2):021901 [FREE Full text] [doi: 10.1063/5.0037640] [Medline: 33746486]
9. Gelander L, Karlberg J, Albertsson-Wikland K. Seasonality in lower leg length velocity in prepubertal children. Acta Paediatr 1994 Dec;83(12):1249-1254. [doi: 10.1111/j.1651-2227.1994.tb13006.x] [Medline: 7734863]
10. Broers NJH, Usvyat LA, Marcelli D, Bayh I, Scatizzi L, Canaud B, on behalf the MONitoring Dialysis Outcomes (MONDO) Consortium. Season affects body composition and estimation of fluid overload in haemodialysis patients: variations in body composition; a survey from the European MONDO database. Nephrol Dial Transplant 2015 Apr;30(4):676-681. [doi: 10.1093/ndt/gfu367] [Medline: 25492895]
11. Dopico XC, Evangelou M, Ferreira RC, Guo H, Pekalski ML, Smyth DJ, et al. Widespread seasonal gene expression reveals annual differences in human immunity and physiology. Nat Commun 2015 May 12;6:7000 [FREE Full text] [doi: 10.1038/ncomms8000] [Medline: 25965853]
12. Patten SB, Williams JVA, Lavorato DH, Bulloch AGM, Fiest KM, Wang JL, et al. Seasonal variation in major depressive episode prevalence in Canada. Epidemiol Psychiatr Sci 2017 Apr;26(2):169-176 [FREE Full text] [doi: 10.1017/S2045796015001183] [Medline: 26751782]
13. Bhatia S, Bhatia S, Mears J, Dibu G, Deshmukh A. Seasonal Periodicity of Ischemic Heart Disease and Heart Failure. Heart Fail Clin 2017 Oct;13(4):681-689. [doi: 10.1016/j.hfc.2017.05.004] [Medline: 28865777]
14. Grant WB, Bhattoa HP, Boucher BJ. Seasonal variations of U.S. mortality rates: Roles of solar ultraviolet-B doses, vitamin D, gene exp ression, and infections. J Steroid Biochem Mol Biol 2017 Oct;173:5-12. [doi: 10.1016/j.jsbmb.2017.01.003] [Medline: 28088363]
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. Yang S, Santillana M, Kou SC. Accurate estimation of influenza epidemics using Google search data via ARGO. Proc Natl Acad Sci U S A 2015 Nov 24;112(47):14473-14478 [FREE Full text] [doi: 10.1073/pnas.1515373112] [Medline: 26553980]
17. Cruvinel T, Ayala Aguirre PE, Lotto M, Marchini Oliveira T, Rios D, Pereira Cruvinel AF. Digital behavior surveillance: Monitoring dental caries and toothache interests of Google users from developing countries. Oral Dis 2019 Jan;25(1):339-347. [doi: 10.1111/odi.12986] [Medline: 30270556]

XSL•FO
RenderX

18. Ayers JW, Althouse BM, Allem J, Rosenquist JN, Ford DE. Seasonality in seeking mental health information on Google. Am J Prev Med 2013 May;44(5):520-525. [doi: 10.1016/j.amepre.2013.01.012] [Medline: 23597817]
19. Soreni N, Cameron DH, Streiner DL, Rowa K, McCabe RE. Seasonality Patterns of Internet Searches on Mental Health: Exploratory Infodemiology Study. JMIR Ment Health 2019 Apr 24;6(4):e12974 [FREE Full text] [doi: 10.2196/12974] [Medline: 31017582]
20. Fan J, Han F, Liu H. Challenges of Big Data Analysis. Natl Sci Rev 2014 Jun;1(2):293-314 [FREE Full text] [doi: 10.1093/nsr/nwt032] [Medline: 25419469]
21. Cervellin G, Comelli I, Lippi G. Is Google Trends a reliable tool for digital epidemiology? Insights from different clinical settings. J Epidemiol Glob Health 2017 Sep;7(3):185-189 [FREE Full text] [doi: 10.1016/j.jegh.2017.06.001] [Medline: 28756828]
22. Jabour AM, Varghese J, Damad AH, Ghailan KY, Mehmood AM. Examining the Correlation of Google Influenza Trend with Hospital Data: Retrospective Study. J Multidiscip Healthc 2021;14:3073-3081 [FREE Full text] [doi: 10.2147/JMDH.S322185] [Medline: 34754195]
23. Ling R, Lee J. Disease Monitoring and Health Campaign Evaluation Using Google Search Activities for HIV and AIDS, Stroke, Colorectal Cancer, and Marijuana Use in Canada: A Retrospective Observational Study. JMIR Public Health Surveill 2016 Oct 12;2(2):e156 [FREE Full text] [doi: 10.2196/publichealth.6504] [Medline: 27733330]
24. Wang P, Xu Q, Cao R, Deng F, Lei S. Global Public Interests and Dynamic Trends in Osteoporosis From 2004 to 2019: Infodemiology Study. J Med Internet Res 2021 Jul 5;23(7):e25422. [doi: 10.2196/25422]
25. Rockmore D. The FFT: an algorithm the whole family can use. Comput Sci Eng 2000;2(1):60-64. [doi: 10.1109/5992.814659]
26. Rao KR, Hwang JJ, Kim DN. Fast Fourier Transform - Algorithms and Applications. New York: Springer; 2010.
27. Rossignol L, Pelat C, Lambert B, Flahault A, Chartier-Kastler E, Hanslik T. A method to assess seasonality of urinary tract infections based on medication sales and google trends. PLoS One 2013;8(10):e76020 [FREE Full text] [doi: 10.1371/journal.pone.0076020] [Medline: 24204587]
28. Patel JC, Khurana P, Sharma YK, Kumar B, Ragumani S. Chronic lifestyle diseases display seasonal sensitive comorbid trend in human population evidence from Google Trends. PLoS One 2018;13(12):e0207359 [FREE Full text] [doi: 10.1371/journal.pone.0207359] [Medline: 30540756]
29. Nuti SV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of google trends in health care research: a systematic review. PLoS One 2014;9(10):e109583 [FREE Full text] [doi: 10.1371/journal.pone.0109583] [Medline: 25337815]
30. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. Math Comp 1965 Jan 01;19(90):297-301. [doi: 10.1090/s0025-5718-1965-0178586-1]
31. Ryan K, Weissman S, Ahmad N, Alspaugh J, Drew W, Lagunof M. Sherris Medical Microbiology (7th edition). New York: McGraw-Hill Education; 2018.
32. Janda JM, Abbott SL, McIver CJ. Plesiomonas shigelloides Revisited. Clin Microbiol Rev 2016 Apr;29(2):349-374 [FREE Full text] [doi: 10.1128/CMR.00103-15] [Medline: 26960939]
33. Sega R, Cesana G, Bombelli M, Grassi G, Stella ML, Zanchetti A, et al. Seasonal variations in home and ambulatory blood pressure in the PAMELA population. Pressione Arteriose Monitorate E Loro Associazioni. J Hypertens 1998 Nov;16(11):1585-1592. [doi: 10.1097/00004872-199816110-00004] [Medline: 9856358]
34. Gajdács M. Resistance trends and epidemiology of and infections (RETEPAPI): a 10-year retrospective survey. Infect Dis (Lond) 2019 Sep;51(9):710-713. [doi: 10.1080/23744235.2019.1640389] [Medline: 31298069]
35. Nichols Heitman K, Dahlgren FS, Drexler NA, Massung RF, Behravesh CB. Increasing Incidence of Ehrlichiosis in the United States: A Summary of National Surveillance of Ehrlichia chaffeensis and Ehrlichia ewingii Infections in the United States, 2008-2012. Am J Trop Med Hyg 2016 Jan;94(1):52-60 [FREE Full text] [doi: 10.4269/ajtmh.15-0540] [Medline: 26621561]
36. Fisman DN, Lim S, Wellenius GA, Johnson C, Britz P, Gaskins M, et al. It's not the heat, it's the humidity: wet weather increases legionellosis risk in the greater Philadelphia metropolitan area. J Infect Dis 2005 Dec 15;192(12):2066-2073. [doi: 10.1086/498248] [Medline: 16288369]
37. Braeye T, Echahidi F, Meghraoui A, Laisnez V, Hens N. Short-term associations between Legionnaires' disease incidence and meteorological variables in Belgium, 2011-2019. Epidemiol Infect 2020 Apr 29;148:e150 [FREE Full text] [doi: 10.1017/S0950268820000886] [Medline: 32345387]
38. Hirai J, Kinjo T, Koga T, Haranaga S, Motonaga E, Fujita J. Clinical characteristics of community-acquired pneumonia due to Moraxella catarrhalis in adults: a retrospective single-centre study. BMC Infect Dis 2020 Nov 10;20(1):821 [FREE Full text] [doi: 10.1186/s12879-020-05564-9] [Medline: 33172398]
39. Psoter KJ, De Roos AJ, Wakefield J, Mayer JD, Rosenfeld M. Seasonality of acquisition of respiratory bacterial pathogens in young children with cystic fibrosis. BMC Infect Dis 2017 Jun 09;17(1):411 [FREE Full text] [doi: 10.1186/s12879-017-2511-9] [Medline: 28599639]
40. Logar-Henderson C, Ling R, Tuite AR, Fisman DN. Effects of large-scale oceanic phenomena on non-cholera vibriosis incidence in the United States: implications for climate change. Epidemiol Infect 2019 Jan;147:e243 [FREE Full text] [doi: 10.1017/S0950268819001316] [Medline: 31364581]

XSL•FO
RenderX

41. Kamiński M, Łoniewski I, Marlicz W. Global Internet Data on the Interest in Antibiotics and Probiotics Generated by Google Trends. Antibiotics 2019 Oct 22;8(4):147 [FREE Full text] [doi: 10.3390/antibiotics8040190] [Medline: 31652664]
42. Płatek AE, Sierdziński J, Krzowski B, Szymański FM. Seasonal trends in hypertension in Poland: evidence from Google search engine query data. Kardiol Pol 2018;76(3):637-641 [FREE Full text] [doi: 10.5603/KP.a2017.0264] [Medline: 29297185]

Abbreviations

FFT: Fast Fourier transform
RSV: relative search volume
SS2: sum of squared differences

Edited by T Mackey; submitted 11.11.21; peer-reviewed by M Lotto, M K.; comments to author 31.03.22; revised version received 09.06.22; accepted 24.06.22; published 19.07.22.

Please cite as:
Gillis T, Garrison S
Confounding Effect of Undergraduate Semester–Driven "Academic" Internet Searches on the Ability to Detect True Disease Seasonality in Google Trends Data: Fourier Filter Method Development and Demonstration
JMIR Infodemiology 2022;2(2):e34464
URL: https://infodemiology.jmir.org/2022/2/e34464
doi:10.2196/34464
PMID:37113451

©Timber Gillis, Scott Garrison. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 19.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Web-Based Perspectives of Deemed Consent Organ Donation Legislation in Nova Scotia: Thematic Analysis of Commentary in Facebook Groups

Alessandro R Marcon[1], MA; Darren N Wagner[1], PhD; Carly Giles[1], BSc; Cynthia Isenor[2], MSCN

[1]Health Law Institute, Faculty of Law, University of Alberta, Edmonton, AB, Canada

[2]Nova Scotia Health, Halifax, NS, Canada

Corresponding Author:
Alessandro R Marcon, MA
Health Law Institute
Faculty of Law
University of Alberta
Office 470
Edmonton, AB, T6G 2H5
Canada
Phone: 1 613 620 8970
Email: marcon@ualberta.ca

Abstract

Background: The Canadian province of Nova Scotia recently became the first jurisdiction in North America to implement deemed consent organ donation legislation. Changing the consent models constituted one aspect of a larger provincial program to increase organ and tissue donation and transplantation rates. Deemed consent legislation can be controversial among the public, and public participation is integral to the successful implementation of the program.

Objective: Social media constitutes key spaces where people express opinions and discuss topics, and social media discourse can influence public perceptions. This project aimed to examine how the public in Nova Scotia responded to legislative changes in Facebook groups.

Methods: Using Facebook's search engine, we searched for posts in public Facebook groups using the terms "deemed consent," "presumed consent," "opt out," or "organ donation" and "Nova Scotia," appearing from January 1, 2020, to May 1, 2021. The finalized data set included 2337 comments on 26 relevant posts in 12 different public Nova Scotia–based Facebook groups. We conducted thematic and content analyses of the comments to determine how the public responded to the legislative changes and how the participants interacted with one another in the discussions.

Results: Our thematic analysis revealed principal themes that supported and critiqued the legislation, raised specific issues, and reflected on the topic from a neutral perspective. Subthemes showed individuals presenting perspectives through a variety of themes, including compassion, anger, frustration, mistrust, and a range of argumentative tactics. The comments included personal narratives, beliefs about the government, altruism, autonomy, misinformation, and reflections on religion and death. Content analysis revealed that Facebook users reacted to popular comments with "likes" more than other reactions. Comments with the most reactions included both negative and positive perspectives about the legislation. Personal donation and transplantation success stories, as well as attempts to correct misinformation, were some of the most "liked" positive comments.

Conclusions: The findings provide key insights into perspectives of individuals from Nova Scotia on deemed consent legislation, as well as organ donation and transplantation broadly. The insights derived from this analysis can contribute to public understanding, policy creation, and public outreach efforts that might occur in other jurisdictions considering the enactment of similar legislation.

(JMIR Infodemiology 2022;2(2):e38242) doi:10.2196/38242

KEYWORDS

organ donation; organ transplantation; deemed consent; presumed consent; social media; Facebook; public perceptions; public policy; thematic analysis

XSL•FO
RenderX

Introduction

Background

In 2019, the Canadian province of Nova Scotia became the first jurisdiction in North America to pass legislation for organ donation instituting deemed consent, otherwise commonly known as presumed consent or opt out [1]. Within Canada, Nova Scotia has relatively high rates of organ donation [2]; however, both the province and Canada as a whole have rates lower than many other comparable regions and nations [3]. Several jurisdictions within Canada and abroad have sought to establish deemed consent donation laws to remedy organ and tissue donation shortfalls but have faced considerable criticism about the effectiveness and public reception of such proposed laws [4,5].

Studies have shown that deemed consent legislation alone does not necessarily rectify organ donation shortages [2,6,7]. Canadian Blood Services have clearly articulated the key elements for successful deceased organ donation systems within 6 significant foundational concepts, and legislation is only 1 of these 6 concepts [8]. Crucial factors for improving donation rates include properly functioning donation registries, ethical organ allocation systems, and context-sensitive donation laws [4,7,9]. In the past, some nations that instituted deemed consent laws, including Singapore, Brazil, and Chile, did not successfully increase donation and transplantation rates following legislation changes [10]. Others, such as Wales and the Netherlands, are observing an increase in donation and transplantation rates [11]. Importantly, in the context of Wales's success, the United Kingdom has a well-established donation system infrastructure to support legislative changes [11]. The efficiency of any donation consent model depends on ancillary factors such as instilling trust in health care systems, accommodating next of kin, and creating effective public outreach [10].

Objectives

In response to the new legislation in Nova Scotia, Health Canada has funded a program of research, Legislative Evaluation: Assessment of Deceased Donation Reform, to evaluate "the implementation process and full impact of the deceased organ donation legislation and the health system transformation" and to "inform future legislative or administrative changes to donation and transplantation in other jurisdictions" [12]. Our research contributes to the Legislative Evaluation: Assessment of Deceased Donation Reform program by examining web-based public discussions on the legislative changes in Nova Scotia. Understanding web-based public perspectives is valuable as social media can influence how the public learns about, thinks about, and acts on health topics [13-15].

We observed a substantial number of discussions on Nova Scotia's deemed consent organ policy on Facebook. Facebook is a key platform for sharing views, exchanging information, and seeking advice about personal health actions and decisions, including at the intersection of political decisions with health ramifications [16-18]. Facebook, similar to many other social media platforms, involves community formation and group connections [19]. Numerous studies have shown how belonging to health-related Facebook groups can provide emotional support and increase social connectivity for participants [14,20]. Unlike more anonymous platforms, such as Reddit or newspaper comments sections, Facebook users commonly operate through personal profiles, which means that their activities are often seen by family and friends [21]. Research shows how Facebook users typically only follow, and participate in, a few pages in their Facebook activities [22], which demonstrates the sociological understanding of "homophily"—where people interact more with others similar to themselves [23]. Recent research has used "homophily" ideas to interpret social media interactions, showing how similar web-based interactions can strengthen ties between individuals [24]. Further research in health contexts, particularly during the COVID-19 pandemic, has shown how group formation around political or ideological lines can play an influential role in shaping perspectives and informing decisions [25], whereas other projects have demonstrated that scientific literacy and cognitive sophistication are also key drivers [26].

Facebook use is high among Canadians [27], and research shows that many Canadians use Facebook to access news stories [28]. Although research on the Canadian public demonstrates that Canadians do not commonly trust the information they come across on social media [29], it also shows that Canadians have high levels of trust in friends, family [30], and those in their local communities [31] and are willing to be persuaded by convincing arguments from individuals they trust [28]. Although organ donation is a relatively niche topic, certain Nova Scotian Facebook groups had lively discussions offering many public perspectives about the legislated changes to organ donation. However, Facebook can be a breeding ground for misinformation [32], which has raised concerns about the kinds of information with which users engage [18,33].

Observing and analyzing the deemed consent discourse in Nova Scotian Facebook groups allowed us to observe public perspectives, including how others responded to sentiments and opinions, including misinformation. Our research incorporated the user responses to Facebook posts, namely comments, replies, shares, and emoji reactions (eg, "Likes") [34]. Research shows how emoji reactions play an important communicative role on Facebook, helping forge connectedness and social intimacy among users [35,36], as well as how stories get promoted by Facebook's algorithm [37]. Future public information campaigns on deemed consent for organ donation will need to better understand web-based public discourse and be better prepared to effectively disseminate accurate information while countering and correcting misinformation. Our study elucidated these precise issues for Facebook discussions as the new organ donation legislation and policy rolled out in Nova Scotia.

Methods

Overview

Our project examined web-based commentary around the deemed consent legislation changes produced on public (as opposed to private) Nova Scotia–based Facebook groups. To the best of our knowledge, no other study has examined a social media platform for the public's web-based response to this new

XSL•FO
RenderX

legislation. We chose to investigate Facebook as the platform has a significant social and demographically diverse influence [18,20,25,38], and intensive observation revealed Facebook to be the primary social media platform where most relevant discussions concerning Nova Scotia occurred. It is well known that Facebook groups represent a popular way for individuals to congregate, discuss, and share information [14,20]. A growing body of research shows that Facebook can be a source and propagator of misinformation [32], and several studies have demonstrated how web-based discourse, including Facebook comments, provides valuable insights into public perceptions and decision-influencing practices [13,14,39]. Therefore, we used comments and responses to posts in public Nova Scotia–based Facebook groups to analyze public perspectives on legislative change.

Data Collection

We generated a sample of comments and replies for this study using the Facebook search function. We searched for posts in any public groups using the following inquires: ([“deemed consent” or “presumed consent” or “opt out” or “organ donation”] and “Nova Scotia,”) appearing between January 1, 2020, and May 1, 2021, extending from before the legislation came into force to the date the searches were performed.

All posts that appeared in public Nova Scotia–based Facebook groups were included in our data set. We did not include posts belonging to nationwide groups (such as the national Canadian Broadcasting Corporation) or groups affiliated with other provinces. As the *Results* section shows, our selected time frame encompassed the period of relevant public discussions, which occurred from late June to early July 2020 and again in mid-to-late January 2021, corresponding to when the legislative changes were implemented on January 18, 2021.

On May 25, 2021, we opened all the comments and replies on the respective Facebook post pages and took screenshots of all commentaries in discussions, saving this data in a Google Docs folder. All data were held in Google Docs folders accessible to all coders, and the analysis was conducted in shared Google Sheets.

The screenshots provided a fixed data set that could be subject to iterative analysis involving 3 coders over several months. Although usernames appeared in the screenshots, neither the usernames nor attributable accounts of individuals were included in the analysis to protect user privacy. For each post, we recorded the total counts of shares and the number of emoji reactions by type, which are *Like*, *Love*, *Care*, *Wow*, *Haha*, *Anger*, and *Sad*. Multimedia Appendix 1 provides visual images of these emojis.

Coding and Analysis

We performed 2 analytic procedures on the data set to answer our two central research questions: (1) what perspectives did Facebook users express in comments about the new deemed consent organ donation legislation in Nova Scotia, and (2) how did Facebook users respond to the commentary of others? First, we used thematic analysis [40-42] as a means of capturing the wide range of public perceptions evident in the discussions. Second, we performed content analysis [43] on the 3 comments that garnered the most reactions in each discussion to provide insights into the kinds of comments resonating most strongly among the users.

Thematic analysis is a flexible qualitative approach that provides a highly detailed and complex summary of rhetorical data without sacrificing a plurality in meanings [40-42]. It has been used in other health contexts to analyze web-based commentary, including on Facebook [44,45]. This analytical method was well suited for the analysis of web-based comments derived from numerous socially diverse Facebook groups. This enabled us to obtain a detailed overview of the diverse themes, defined as “central organizing concepts” [41] as they appeared across disparate groups and at different periods.

Performing thematic analysis requires choosing between an “inductive approach,” where the data dictate the themes that emerge through analysis, and a “deductive approach,” where ideas, concepts, and themes are brought to the data before analysis [40]. We blended the 2 approaches as 2 coders knowledgeable on the topic brought some concepts, topics, and expectations to the study before engaging the data. However, the coders were not limited to these previously obtained perspectives as they anticipated, and were willing, to observe new rhetoric, topics, and language indicating emergent themes. The 3 coders followed the 6 phases of thematic analysis described in detail by Braun and Clarke [40] and examined the data for “trustworthiness,” as outlined by Nowell et al [42]. Careful attention was paid to constructing themes that were “specific enough to be discrete” but sufficiently broad to capture “ideas contained in numerous text segments” [42].

For the content analysis, we first determined the 3 comments that elicited the most emoji reactions in each discussion and then conducted the content analysis [43] on these comments. We applied the previously conducted thematic analysis categorizing each comment as pro (promoting or supporting the legislation or donation more broadly), critique (critiquing the legislation or donation more broadly), or neutral (reflections that neither clearly promote nor critique the legislation). We looked for any particular trends in the themes to provide greater insight into the comments that generated the most reactions from Facebook users. The content analysis was first performed by one coder, and a second coder checked all coding. There were only 3 disagreements between the coders, resulting in an intercoder reliability of Cohen κ=0.92, which demonstrates “almost perfect” levels of agreement [46]. The 3 discrepancies were resolved in a consensus session [47].

Ethical Considerations

Ethics approval was not required for this research as the study involved analysis of publicly available data. The results do not contain any identifying information of commenters (eg, usernames), and the text examples have been paraphrased to further protect individuals’ privacy.

XSL•FO
RenderX

Results

Thematic Analysis

Overview

Our final data set included 26 posts with 2337 comments and replies from 12 different Facebook groups. Most comments appeared on Facebook groups belonging to either media companies or the Government of Nova Scotia. Some province-based community groups were also represented. The number of comments for each post ranged considerably (8-442; Table 1). The thematic analysis resulted in 4 principal themes and a total of 8 subthemes (Textbox 1). Textboxes 2-5 present each subtheme and illustrative excerpts from the Facebook comments.

Table 1. Complete data of Facebook groups and discussions[a].

Facebook group name	Number of discussions	Number of comments	Date of discussions by post date
Nova Scotia Department of Health and Wellness	6	333	June 30, 2020; December 22, 2020; January 12, 15, 18, and 18, 2021
Nova Scotia Government	2	579	June 30, 2020; January 18, 2021
Nova Scotia Health	1	8	July 2, 2020
CBC Nova Scotia	3	462	July 1, 2021; January 18 and 28, 2021
Q97.7	1	104	January 19, 2021
Halifax Muslims	4	168	August 13, 2020; January 15, 15, and 17, 2021
Black NS News	1	13	December 18, 2020
Global Halifax	2	118	June 30, 2020; January 15, 2021
CTV Atlantic News	1	172	January 18, 2021
The Chronicle Herald	2	207	January 19 and 19, 2021
Halifax Today	2	152	July 15, 2020; January 19, 2021
Cape Breton Daily News	1	21	January 18, 2021

[a]Total: 26 discussions, 2337 comments; 2020—8 discussions; 2021—18 discussions.

Textbox 1. Principal themes and respective subthemes.

Supporting and promoting donation and transplantation and the new donation legislation (theme 1)

- Caring about donation is caring about others
- The legislation isn’t a problem, and here’s why you naysayers are ignorant, stupid, selfish, and wrong

Raising issues with donation and transplantation broadly and critiquing the new donation legislation (theme 2)

- The legislation conflicts with my personal principles and world views
- They’re out to get us! They’re not to be trusted!
- Why fix what isn’t broken?! The changes aren’t needed or justified

Discussing particulars and pointing out issues (theme 3)

- Religious beliefs about donation and transplantation
- Is donation from gay men acceptable now?
- Family power is a benefit and a concern

Metacommentary, softer reflections, jokes, and questions (theme 4)

- Not applicable

XSL•FO
RenderX

Textbox 2. Paraphrased examples for theme 1 (supporting and promoting donation and transplantation and the new donation legislation).

Caring about donation is caring about others

- Losing a child who was waiting for a transplant was horrible. Nova Scotia's new initiative will be beneficial and "it's about time" something like this was done.
- 100% support for the legislative change; many more organs will be available, and lives will be saved.
- A friend died in a tragic accident and their donated organs helped five different people. The donation was enormous gift, and the donation brought comfort to his grieving parents.
- Donation doesn't just improve the quality of life for recipients but offers a means for grieving families to find comfort; donation provides hope to all.
- Waiting for organs is a serious struggle, and the new legislation is a splendid idea.

The legislation isn't a problem, and here's why you naysayers are ignorant, stupid, selfish, and wrong

- Those opting out should be ineligible to receive. "Selfish" people who don't want to help shouldn't get the chance to be helped.
- Having to check a simple box to opt out is not something to be upset about. It's ridiculous and silly to think your rights are being "taken away." Shut your whiny mouth
- Italy has done this for a long time, and many other jurisdictions should be like them.
- Those wanting to opt out are being "selfish Neanderthals".
- Giving consent to donate is ok but having to consent to not donate is a "big deal"? It makes no sense to not help a dying child and just have the organs "rot instead."
- It's inexplicable why people are upset about this. Your organs are going "to rot in a hole" and there's nothing science can do about it, so you may as well save someone else's life.
- It's a selfish position to not want to help save someone's life, and it's nonsensical why some see the new legislation as an issue.

XSL•FO
RenderX

Textbox 3. Paraphrased examples for theme 2 (raising issues with donation and transplantation broadly and critiquing the new donation legislation).

The legislation conflicts with my personal principles and world views

- I agree with donation, and I am a donor, but I believe it is a decision for each person to make. It's not right for others to "take" an organ unless you say no, and the new law acts as a "dangerous slippery slope."
- The legislation is wrong, and the Nova Scotia government is not the owners of others' bodies.
- Many other cultures, religions, and minorities care about how donation is done, and there is no clarity around how these processes will take place, especially as time is very sensitive in these contexts.
- It's a serious cause of frustration as I support donation but disagree with the government taking ownership of a body after, say, a brain injury. The issue is that opting out is the only way to protect a right to choose.
- The government is treating us like "fucking lab rats," robbing our graves, and assuming ownership over dead bodies!
- The new legislation is 100% WRONG! It's a dangerous situation to have a law that "removes sovereignty" and seizing the ownership rights of a body that has not yet died.

They're out to get us! They're not to be trusted!

- People need to wake up and realize it's about harvesting organs and selling them.
- They are trying to "trick people," hoping people won't know what's happening.
- Come on over, Russia, take our body parts after the government takes away our firearms.
- It was "not cool" how MacNeil put this through secretly.
- The doctors will determine who is worthy to live and will let some die to save others' lives. Say bye-bye to sick old people.
- It's big business to sell body parts but now they know where you live and don't even wait for you to die.
- Doctors don't know about all rare diseases, and for some people they can prevent more problems by choosing to opt out.
- A nurse advised me once to not sign a donor card, and I think it's because she saw a case of organs being taken before it was time for them to die. This legislation only works if people are lazy.
- This is "all about money." Doctors are crooked and harvesting is a way to make money.

Why fix what isn't broken?! The changes aren't needed or justified

- The option already in place was "just a check mark," and so if everyone agreed with being a donor why was the new law necessary? Why take away others' rights?
- There was no reason to change the old way. Assuming is not right. Some, like myself, will be very confused by the opt out system.
- There is nothing about consulting the family, just that not opting out means the organs get taken. This new change will be expensive for taxpayers, and there is nothing wrong that needs fixing.

XSL•FO
RenderX

Textbox 4. Paraphrased examples for theme 3 (discussing particulars and pointing out issues).

Religious beliefs about donation and transplantation

- Support for organ donation can be found in the majority of Muslim scholars.
- Loving God, who loves all children, means loving others and becoming organ donors.
- In my interpretation, if one's body is not for one to decide whether it lives or dies, why should a person decide they can give it to others? On this basis, organ donation is not permitted, and Muslim organs donated to non-Muslim bodies will no longer be "cleansed."
- Judaism frowns upon donation after death but not in live liver or kidney donations, but I understand that others choose to donate. For some cultures, giving away organs is like giving away part of the soul.
- Respect should be given for those who decided not to donate organs, whether that be for religious beliefs or other reasons. These people should not be attacked or called "monsters"; it's too much.
- There is no reason to opt out except for being selfish and having "faith."
- Some people can't donate because of their religious beliefs.

Is donation from gay men acceptable now?

- I guess that now they will start accepting blood donation from me as a gay man?
- Prohibited from donating blood as a gay person in Nova Scotia, I guess that since a gay person's blood isn't acceptable, neither are the organs.
- Being gay means my organs can't be donated.

Family power is a benefit and a concern

- The family veto issue is "pesky" as your next of kin's wishes should be seen as YOUR wishes!
- The new legislation means that a family will be consulted by a nurse about donation if a person hasn't registered a decision, and the family has the power to say no.
- I'm glad that this will be the default, but a family overriding the desires of a person is something they shouldn't be able to do.
- I had the most terrible experience with doctors pressuring us to "carve up" my brother, and for that reason I hope families can play a role in the donation decision.

Textbox 5. Paraphrased examples for theme 4.

Principal theme 4: metacommentary, softer reflections, jokes, and questions

- The number of donation arguments and opinions on the Internet is exhausting.
- It's not right to pass judgement on others, given how personal and emotional the donation decision is.
- People are incapable of having "calm" discussions now, and I've been watching this get worse over the years.
- I am still unsure about my decision on this topic.
- My abused liver offers nothing to nobody!
- Is any person too old for donation?
- While I might sign up to be a donor, what happens in the case of my children? Can I overrule the choice I might have made for them?

Principal Theme 1: Supporting and Promoting Donation and Transplantation and the New Donation Legislation

Caring About Donation Is Caring About Others

These comments demonstrated compassion and portrayed donation and transplantation as practices to be respected, promoted, and encouraged. Several personal and emotional anecdotes regarding donation and transplantation success were included. Users often implicitly and explicitly expressed altruistic sentiments, stating that organs from deceased individuals should be shared with others, which the new legislation would facilitate, thereby saving more lives. These comments showed a desire to help people and encouragement for others, including other provinces, to adopt a similar approach. Some comments expressed feelings of pride in Nova Scotia's initiative. This subtheme included reactions to critiques of donation and transplantation and typically expressed surprise, dismay, and disappointment at others' lack of altruism. Permeating this commentary was an implied trust in health care systems and workers, including physicians (examples in Textbox 2).

The Legislation Isn't a Problem, and Here's Why You Naysayers Are Ignorant, Idiotic, Selfish, and Wrong

Commentary in this subtheme was distinctly more aggressive and argumentative than in the first subtheme. These comments were often replies to other users expressing concerns or issues with the legislation. Users voiced arguments, frequently with

XSL•FO
RenderX

frustrated and angry tones, about why the legislation, and, broadly, donation, should be supported. It was commonly argued that the new legislation maintained choice and autonomy, that opting out would be easy, and that bodies with the potential to save lives could now be more readily used. For example, common references to dead bodies and organs "rotting in the ground" highlighted the perceived waste of a valuable resource in the absence of donation. A very common argument was that people who opt out should not be eligible to receive transplants.

Typical features in this subtheme included name-calling, labeling people who opt out as "selfish," and suggesting detractors are unintelligent or mentally ill. Some users emphasized that deemed consent is relatively common for other legal procedures, such as with wills and estates, and is in place for donation in other countries. In a few cases, comments included statistics to support arguments (eg, the fact that an individual is much more likely to need a transplanted organ than to be an organ donor). Many comments attempted to contest and correct misinformation presented by others (examples in Textbox 2).

Principal Theme 2: Raising Issues With Donation and Transplantation Broadly and Critiquing the New Donation Legislation

The Legislation Conflicts With My Personal Principles and World Views

Central to this subtheme was the idea that powerful entities (namely, governments) were usurping individuals' agency; acting against personal rights, autonomy, and freedom; overriding religious and spiritual beliefs and convictions; and diminishing people's ability to consent. In numerous instances, users presented the concept of consent in absolute terms, suggesting that consent can and should not be presumed or negotiated ("my body, my choice"). Common arguments included the idea of the government assuming "ownership" of individuals' bodies and the notion that this legislation was another example of the government increasingly encroaching on individual autonomy (eg, "slippery slope" and "what comes next?"). Tied into these sentiments was the idea that powerful entities would callously exploit bodies in undignified ways ("chopping up"), violating the perceived sanctity of the body and personal wishes upon death. Some users explicitly stated their desire to support organ donation while disapproving the new legislation. In a few cases, comments exhibited antialtruistic sentiments, such as explicitly stating their preference to not help others or to only help their family members (examples in Textbox 3).

They're Out to Get Us! They're Not to Be Trusted!

This subtheme centered on user comments, demonstrating a profound mistrust of the government and health care systems. These comments raised issues regarding the lack of transparency and consultation efforts of the government and health institutions. Common rhetoric included terms such as "tricky" and "secret" and phrases such as "hidden in legislation." These users often argued that the government was intentionally (and maliciously) trying to dupe the public. Some comments directly targeted Nova Scotia's then-governing Liberal Party and then-Premier Stephen McNeil (eg, labeling him a "dictator"). Some comments also disparaged the new legislation by comparing the changes with actions by foreign totalitarian governments.

As in the first subtheme, some users expressed concerns about the undesired exploitation and manipulation of bodies. However, such comments in the second subtheme underscored the nefarious objectives of public officials, including profit motives and sacrificing lives to save others ("harvesting"). Some comments suggested that increased organ procurement would cater to the needs of the rich (the poor would get worse service), take advantage of vulnerable populations (the homeless, youth, and those with mental health issues), involve transplanting infected and damaged organs unknowingly (eg, Lyme disease and HIV), result in fewer efforts to save lives to supply more donor organs, and cause data errors with serious consequences (eg, mishandling of individual health records). Many of these comments touched on conspiratorial ideas (examples in Textbox 3).

Why Fix What Isn't Broken?! The Changes Aren't Needed or Justified

This subtheme was characterized by an argumentative commentary about the need for a new model for donor consent. Users argued that if people wanted to donate, nothing in the old donation model would prevent them from doing so. Users also raised the parallel argument that the shortfall in donations was because people did not want to donate rather than merely forgetting or neglecting. In addition to questioning the legality of the new legislation (eg, "this won't hold up in the courts"), users criticized the new model's costliness. Comments typically argued that the old opt-in model was better—as it maintained personal choice and autonomy—and that the old model should instead be improved by, for instance, requiring the public to declare a donation preference when renewing a health card (examples in Textbox 3).

Principal Theme 3: Discussing Particulars and Pointing Out Issues

Religious Beliefs About Donation and Transplantation

Many of the discussions touched on religion; however, comments tended not to be specific to the new legislation. Users offered questions and observations about whether donation and transplantation align with the tenets of various religions, including concerns about donation conflicting with religious beliefs and the need to opt out for religious and spiritual reasons. It was uncommon for users to state their own religious convictions about donation. Rather, those commenting about religion typically generalized and assumed what others believed and felt. Such generalizations were often accompanied by the opinion that opting out for religious or spiritual reasons was an acceptable choice. Several users argued that specific religions, notably Islam and Christianity, allowed donation and transplantation and that refusing to donate might be contrary to principles of charity (examples in Textbox 4).

Is Donation From Gay Men Acceptable Now?

A few discussions raised the issue of whether donated organs from "gay" men would have specific restrictions, as with blood

XSL•FO
RenderX

donation. Users raising this concern expressed offense at such discriminatory policies, although these critiques were not specific to the new legislation (examples in Textbox 4).

Family Power Is a Benefit and a Concern

Comments in this subtheme related to the power granted to family members to make donation choices on behalf of an incapacitated person. Some users expressed comfort in such a safeguard, whereas others expressed concern about family members overriding an individual's decision (family veto). The importance of discussing donation decisions with one's family was raised often (examples in Textbox 4).

Principal Theme 4: Metacommentary, Softer Reflection, Jokes, and Questions

The core characteristic of this theme was a neutral stance on the new legislation, which included reflections on the discussions, requests for information, and attempts at humor. These reflections discussed donation and transplantation, as well as thoughts on Nova Scotia and the nature of the modern media. Comments about donation and transplantation, and specifically the new legislation, included requests for clarification on facts and common practices (eg, eligibility to donate) and requests for the opt-out link. Some users questioned their eligibility to donate, in some cases making self-deprecating remarks about personal health and the unsuitability of their organs for donation (examples in Textbox 5).

Content Analysis

Analysis of the top 3 comments with the most emoji reactions in each discussion (80/2337, 3.42%) demonstrated that positive emojis (*Like*, *Love*, or *Care*) were the most common, accounting for 95.45% (1112/1165) of all emoji reactions. Indeed, in the total sum of reactions (n=1165), negative reactions (*Anger* and *Sad*) only accounted for a small number (n=4, 0.34%). Multimedia Appendix 2 provides complete numbers. However, the types of comments that generated the nearly universal positive emoji reactions were a mix of responses to the new legislation or donation and transplantation broadly: positive (57/80, 71%), negative (13/80, 16%), and neutral (10/80, 13%; Multimedia Appendix 3). Thus, comments that were supportive, neutral, and critical toward the new legislation received positive emoji reactions from others.

The commentary that evoked the most positive reactions typically included both subthemes from theme 1, including one observable trend: 12% (7/57) offered a personal anecdote of donation and transplantation benefit, and 12% (7/57) exhibited an effort to correct misinformation. The oppositional commentary that provoked the most reactions was related to all 3 subthemes, including 2 with antialtruistic comments (not wanting to help others). The neutral commentary that garnered the most reactions was related to themes 3 and 4, including some discussions on religions and attempts at humor.

Discussion

Principal Findings

Nova Scotia is the first jurisdiction to pass deemed consent for organ donation in North America, and this study is one of the first studies to analyze web-based public discussions on the topic. The results of our analysis show that the new legislation generated controversy, with commentary displaying mixed reactions to the new legislation specifically and donation and transplantation broadly. A range of perspectives was expressed and fervently argued among Facebook group users. The principal themes that emerged from the analysis comprised being in favor and supportive, being opposed and critical, not being openly opposed or in favor but raising particular issues, and general commentary from a neutral perspective; some of these themes have been noted in the literature regarding deceased donation in general [48]. The subthemes constituting these principal themes, which touched on the topics of power, autonomy, government authority, religion and altruism, policy options, and argumentative strategy, provided key insights into how these diverse perspectives were supported and propagated, which is valuable for informing public outreach initiatives. These findings also demonstrate some key dynamics of user engagement with health policy news on social media.

Findings of Public Perception in Other Contexts

Our research findings need to be contextualized through comparisons with legislated changes to organ donation consent in other jurisdictions. Nova Scotia joins England, Scotland, Wales, and Northern Ireland, which also recently moved to deemed consent models. National jurisdictions that have implemented deemed consent legislation variously observed increases and decreases in donation rates. Notably, Brazil saw a sharp decrease, as did Chile and Singapore [10]. Conversely, the Netherlands and Hong Kong both experienced increases in donation rates [10]. However, the general consensus is that modifying the consent model is not the key action generating an increase in donation rates [6,10,49]. Nevertheless, changing consent models can affect cultural norms and social consciousness, shifting the default position toward universal donation. Importantly, trust in the health care system and regional government is crucial to the adoption or rejection of donation policies, which includes how changes are communicated and how data are managed, especially as different contexts show that there are diverse public perceptions around implementing deemed or presumed consent models [10,48,50,51].

Similar research on public perceptions of deemed consent was recently conducted in Scotland, Northern Ireland, and England [52]. These researchers performed thematic analyses on free-text responses by individuals stating whether they would opt in or out or they remained unsure of the newly legislated donation scheme. The themes observed in that study corroborate our findings. Users who supported the switch to deemed consent also stressed how the new legislation promoted altruism and gave arguments about eligible body parts saving lives rather than being "wasted." Our findings similarly revealed that these proponents included personal and emotional anecdotes about transplantation. Narrative messaging can have a powerful impact on how others perceive a range of issues [53,54]. As such, the sharing of positive personal anecdotes about donation and transplantation in public web-based spaces could be a valuable strategy for motivating others to consider donation [55]. Indeed, although not quantifiable, we observed that personal narratives

XSL•FO
RenderX

shifted the tone of the discussions. In addition, our content analysis showed that some of the most liked comments were positive personal stories of donation and transplantation.

Further corroborating our findings, the UK study [52] showed that supporters of deemed consent stressed the idea of "reciprocity," (ie, those willing to receive should be willing to donate) and commonly labeled people wanting to opt out as "selfish." We speculated that users voicing such opinions—that individuals opting out should not be eligible to receive transplants—might have been motivated by reading others' negative comments. We also considered name-calling and the labeling of detractors as "selfish" to be reactionary responses. However, contrary to our speculation, the UK study [51], as well as survey research in the United Kingdom [56], demonstrated that such sentiments constitute a core principle of equity for proponents of the policy. Therefore, any prospective change to opt-out consent models should acknowledge the potential for social tension to arise in the public discourse. Our study shows the tension between those who desire total public participation in donations and those who have reasons to opt out. The complexity of public perceptions and approaches has been previously reported in the literature [57]. Tension typically surfaces only with opt-out systems as opt-in models do not require people to actively or openly state their (perceived) opposition to donation. Rather, opt-in registries often do not require people to declare their donation preference, which seems to dilute this polemic.

The power of family members to veto an organ donor's wishes or to grant final authority for donation appeared as an important subtheme in both our analysis and in the UK research [52,56]. Respondents who wished to explicitly state their opt-in position within the deemed consent model often believed that such a declaration would aid family decisions about donation and protect their personal choice from family interference. Our findings showed that commenters were both relieved and alarmed by the rights afforded to family members as ultimate decision-makers.

Policies addressing the involvement of family members are important for organ donation consent models. Trust in the health care system and in the organ donation process is paramount to any consent model being effective in increasing overall donation rates [10,48,49]. Prohibiting all forms of decision-making by family members would likely be perceived as inflexible and autocratic. Having frontline health care workers enforce a donor's wishes against a family's contestation is highly impractical and ethically problematic. Speaking to this issue in the Canadian context, an expert study asked whether it is "unrealistic to assume the next-of kin refusal rate would decrease under opt-out legislation" [49]. What factors and circumstances would foster a culture where family interference with donations would decrease? Careful monitoring and evaluation of the deemed consent program implementation in Nova Scotia will help answer these questions.

The salient themes we observed among users expressing concerns and grievances about deemed consent are also corroborated by the findings of the thematic analysis conducted in the United Kingdom [52]. Opponents in Nova Scotia emphasized a mistrust of the health care system and criticized the government for infringing on individual freedom and autonomy. Indeed, mistrust of the health care system is known to be a significant barrier to organ donation [10,49], especially given that opt-out policies can be perceived as deceitful, manipulative, and restrictive [58]. In both the UK study [52] and our study, opposition to deemed consent included personal beliefs about government power, philosophical views about consent, and practical concerns about organ donation procedures. For example, both studies found that opponents expressed worries about the unequal provision of health care services, the contested "ownership" of bodies or body parts, and the perceived uncertainty around declaring brain death for donors. However, despite many shared themes, only our study found critiques of the government for profiteering from excised organs. In addition, ideas around health care incompetence (eg, mistakenly transplanting diseased organs) were also seemingly unique to the Nova Scotian context.

Acts of, for, and Against the Body

A pervasive feature of the Facebook discussions we analyzed was diverse and sometimes contrary perspectives associated with how the body is manipulated during organ donation and transplantation. The UK study [52] also observed similar issues. Both studies found that people addressed the body as something to be "used" or "recycled"—a valuable resource not to be "wasted." In our research, the phrase "rot in the ground" was commonly used to both support the commenter's prodonation position and criticize others for opting out ("wasting"). Interestingly, some of the most visceral, harsh, and argumentative language in our sample invoked the "rot" rhetoric. Some telling examples included comments about selfish people's "useful organs" rotting and getting "eaten by bugs" and organs rotting away in "holes" instead of saving lives. This rhetoric is open to several interpretations—as an argumentative tactic, commentary on differing perspectives of the afterlife, a means of virtue signaling, or even a shocking reminder of the inevitability of death. Although this kind of rhetoric is forceful and abrupt, it is unlikely to constructively change the discussion or the perspectives of opponents of the organ donation legislation. If anything, it serves to exacerbate tensions.

Our interpretation of criticisms of organ procurement that referred to the body differed slightly from the UK team's analysis [52]. For instance, we understood the "chopping up" rhetoric to suggest the undesirable and callous handling of bodies. The verb "chop" emphasizes the violent physicality of abuse occurring in organ removal and recovery. We also interpreted the often-repeated "harvesting" phrases to exemplify the impersonal corporeality described with terms such as "biopower," which refers to the state exploiting bodies for governmental objectives [59]. Certainly, the concept of "biopower" could be applied to many of the objections raised by those opposing the legislation on the grounds of ownership, ethical consent, and state abuse among others. Similar to our study, the UK study [52] identified the subtheme of a "violation to bodily integrity," which included participants who stressed a desire to have their bodies remain intact during and after death.

XSL•FO
RenderX

Unsurprisingly, much of the rhetoric about the body's sanctity intertwined with religious topics. We observed discussions on how donation and transplantation aligned or conflicted with various theologies, including those grounded in Islam, Judaism, Protestantism, and Catholicism. Facebook users offered religious doctrines and observances as arguments both for and against organ donation. For example, users discussing Islamic beliefs debated the conflicting priorities of maintaining the "sanctity" of the intact body and offering organs to others as a charitable duty [60,61]. Some argued that donation is "frowned upon" in Judaism, whereas others stated that "loving God's children" is well served through organ donation.

Research shows that religious beliefs can affect people's perspectives on organ donation [62]; however, caution should be exercised when generalizing the influence of religion, especially with regard to opposing donation [61,63]. Indeed, many commenters in the Facebook groups assumed that some people would opt out because of religious beliefs. Similarly, survey research in the United Kingdom found that most believed organ donation to conflict with religious beliefs [56]. These perceptions about religious opposition to donation are in contrast to the fact that no major world religion has a total prohibition on organ donation; rather, organ donation is often connected to concepts of altruism and the ability to save lives [61]. The Facebook discussions we analyzed typically seemed respectful and supportive of people choosing to opt out because of religious beliefs, although it remains unknown how widespread the opt-out position is among religious communities. Effective public outreach should certainly account for the role that religion plays in promoting opinions about donation by engaging communities respectfully and proactively while fostering transparency in health care systems [61].

Social Media and Misinformation

Analyzing web-based discourse provides insights into how these perspectives are expressed and propagated. Analyzing reactions via reaction tools (eg, *Likes*, *Loves*, and *Wows*) and reply chain discussions helps us understand how news posts and comments are received, debated, and refuted. In other words, how information is taken up, altered, and disseminated through these web-based spaces. Our research has some important media-related findings that are relevant to deemed consent donation laws and broadly to social media.

Most user comments appeared on Nova Scotia government Facebook groups, including the Nova Scotia Department of Health and Wellness (333/2337, 14.25% of all comments) and Nova Scotia Government (579/2337, 24.78% of all comments). These numbers demonstrate the value of government institutions using social media for public outreach, although other discussions might occur on private Facebook groups. There are benefits and challenges that arise with government entities hosting discussions. Hosting enables moderators from the organizations to analyze commentary; facilitate access and analysis from others; moderate discussions (deleting comments or blocking users if necessary); and respond to questions or comments, especially regarding misinformation.

In our study, moderators responded to questions about the donation and transplantation process broadly (eg, age limits for donation and transplant procedures), as well as specific aspects of the new legislation (eg, opt-out process and opt-out choices). Moderators, as well as other commenters, shared text and links to accurate websites, including the official government pages. Such information sharing is helpful to disseminate policy facts, for example, that people could opt out of donating some specific organs. Although not quantifiable in our research, we observed that moderator participation typically had a positive impact on discussions, especially when correcting misinformation or providing clarity around policies. However, as documented in other research, moderating poses numerous challenges, including how and when to engage commenters and on what grounds comments should be blocked or removed [64]. Ongoing research is to determine effective strategies that health practitioners and institutes can use in different web-based contexts [65].

The spread of misinformation through social media has been studied in a wide range of contexts. Information scholars have distinguished between misinformation and disinformation, where disinformation refers to an intent to spread inaccuracies and make facts appear ambiguous [66]. In this study, we observed numerous inaccurate comments. For example, some users suggested that the legislation was passed as a means of generating profits for the government and physicians. One such comment claimed that Canada is the leading exporter of heart valves and bone marrow, and the new legislation is focused on greed rather than helping fellow Canadians. Other comments stated that opting out was not possible after the legislation enactment date (January 18, 2021) and that the new legislation would not take into account the wishes of one's family. Others stated that physicians would intentionally let people die to obtain organs for others. A significant number of these inaccurate comments were frequently repeated by the same few users. Although research projects have likened this activity of users repeating inaccuracies to disinformation [39,67], it is not possible to draw that conclusion in this context as the potential intent to deceive remains uncertain.

We observed numerous efforts to correct or debunk misinformation, especially in terms of the opt-out date and family involvement. Users variously countered misinformation with links to official government websites; personal expertise; and, when someone suggested that physicians will kill potential donors, by detailing the Hippocratic oath—known in Canada as the Code of Ethics and Professionalism [68]. Interestingly, we did not see much countering of the messages around profiteering from the selling of transplant organs. Ideally, there would have been some forthright messages clarifying these issues, especially from moderators or experts in the field.

When correcting misinformation, some scholars have raised concerns around the "back fire effect" [69], which argues that correcting misinformation leads to increased adherence to the misinformation. However, such an effect and concern do not appear substantiated in research [70]. There is certainly value in attempting to correct misinformation and promoting accuracy [71,72], especially in this specific context, for more uncontestable facts (eg, a deadline to opt out). Indeed, our content analysis on the most reacted-to comments showed that some of the most liked comments included debunking efforts. Users liking the debunking comments indicated that many

XSL•FO
RenderX

participants valued their contributions to the discussions. However, what remains unclear is whether those comments had any influence on those who disagreed with the sentiments expressed. Although we did not perform an analysis on the different users participating in discussions, it would seem highly valuable for experts to weigh in on public discussions and provide clarity and accurate information whenever possible. Indeed, research specifically on Facebook has also shown how comments from experts receiving a relatively high number of likes are perceived as the most credible health messages [73].

Those few especially vocal commenters who spread inaccurate or conspiracy-tinged comments occasionally received backlash from other users in the group. In addition to correcting observed inaccuracies, some commenters also made concerted efforts to add accurate statistics to the discussions. For example, a commenter shared the statistic that a person is 6 times more likely to need an organ than to be an eligible donor. There were numerous instances where such a statistic might have usefully grounded abstract or polemic discussions. Indeed, new research has also argued that rather than countering misinformation, more effective public engagement would work toward improving the acceptance of reliable information [74]. This underscores the need for ongoing research to better understand how accuracy on the web can best be promoted, including a social media design that promotes critical reflection, and how public health agencies can productively engage the public on the web when dealing with polemic issues [66,72,75].

The Canadian Donation and Transplantation Research Program recently held a web-based workshop about public engagement on the web. Individuals from 3 health science organizations discussed their web-based public engagement strategies [76]. The workshop featured presentations from representatives of the United Kingdom's National Health Services, Nova Scotia Health, and a new Canadian-based initiative—Science Up First—which aims to educate the public and debunk misinformation on a range of health and science topics. The presenters from the National Health Services and Nova Scotia Health spoke on the topic of deemed consent organ donation legislation, whereas the presenter from Science Up First spoke on their public engagement efforts during the COVID-19 pandemic. Key messages from all presenters highlighted the need for careful monitoring of web-based conversations and carefully planned strategies to tackle the presence of misinformation. Importantly, they noted that ongoing monitoring of conversations allowed moderators and communicators to track the kinds of problematic sentiments shared, as well as accounts that, in some cases, needed to be blocked or banned from further participation. In the context of deemed consent in Nova Scotia specifically, web-based moderators encouraged dialog among participants and only engaged to provide clarity and accuracy (often by providing links to government web pages) or to remove comments and, in a few selective cases, ban users.

Vital to these efforts was the creation of a detailed frequently asked questions document used to train moderators, equipping them with the tools to answer questions in a timely manner. A carefully constructed terms of use policy document, which participants might not read but which moderators can provide as evidence to an offending individual who had comments removed or who was banned from a platform for breaking its rules (eg, using abusive language and repeated offenses), is also essential. Importantly, moderators chose to provide accuracy and clarity when encountering misinformation—as opposed to deleting comments—to make people feel that their voices and concerns were being heard. Indeed, moderators and provincial health officials often positively interpreted web-based debates as public engagement on the topic and as engagement that ultimately reached bigger audiences and raised more awareness about the legislative change. The presenter from Nova Scotia Health also stressed the need for moderators to engage in a neutral, emotionless tone, which would help maintain the focus on relaying accurate information. All presenters noted that countering misinformation was likely less effective as a means of changing the perspective of those sharing inaccuracies but very valuable to help stop the viral spread and influence of misinformation on the wider audience. All organizations stressed the need to engage diverse communities and build wide networks that collaboratively work toward transferring accurate information and heightening science and policy literacy.

Limitations

Our study analyzed comments in Nova Scotia Facebook groups during a relevant period using an approach consistent with research on similar objectives [52,77]. However, there are some limitations to consider when assessing this research. This project only analyzed comments in publicly accessible Facebook groups, and the demographics of those contributing comments are unknown. Therefore, the findings are not generalizable to the public. Analyzing user accounts can provide additional insights but raises ethical issues related to privacy and personal information. In addition to our analysis, conversations about this new legislation on private Facebook groups or even other social media sites might support or contradict our findings.

Although our analysis is particular to Nova Scotia, there are many similarities with research conducted in the United Kingdom [52]. As research on this topic continues to grow, we can better anticipate how the public in different jurisdictions might respond to similar legislation, thus enhancing the ability of policy makers and communication strategists to craft effective public outreach and engagement in policy and legislative reform.

Conclusions

Facebook users in public groups expressed diverse and passionate perspectives about the new deemed consent organ donation legislation in Nova Scotia. These perspectives touched on the topics of health care systems, communities, government authority, religions, the body, death, and the afterlife. Critical perspectives need to be corroborated with other research on public perspectives and actual opt-out rates. Notably, since the implementation of presumed consent, Nova Scotia has experienced an increase in tissue donation and organ referrals while data show that only 5.7% of residents have opted out [78]. The degree to which the increases can be attributed to the organ donation legislative changes requires further and ongoing examination.

XSL•FO
RenderX

Even if concerns around deemed consent are held by a small minority, the issues should not be ignored. Trust is an integral component of health care systems and needs to be maintained and strengthened wherever possible. In the Canadian context, it is well known that mistrust circulates among racialized communities, notably indigenous communities, which have been subjected to colonization and systemic racism. Listening to the concerns of these voices and addressing concerns with actions can only help improve trust in the health care system. Some of these efforts include engaging individuals and communities in dialog offline and on the web. A proactive approach involves listening to issues, clarifying doubts where possible, providing transparency regarding policies, and correcting misinformation. Social media is a hotbed of misinformation; however, it also has the potential to effectively inform the public through creative and accurate messaging.

Acknowledgments

The authors would like to thank and acknowledge Health Canada, Genome Canada, Genome Alberta, and the Canadian Institutes for Health Research for their generous support of Legislative Strategies to Improve Deceased Organ Donation in Canada: A Special Focus on Evaluating the Impact of Opt-Out Legislation in Canada Legislative Evaluation: Assessment of Deceased Donation Reform (LEADDR) and Precision Medicine CanPREVENT AMR: Applying Precision Medicine Technologies in Canada to Prevent Antibody Mediated Rejection and Premature Kidney Transplant Loss. The authors extend thanks to Timothy Caulfield, Robyn Hyde-Lay, Jade Dirk, Matthew Weiss, and Stephen Beed for their support with this project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A total of 7 Facebook emoji reactions.
[DOCX File , 71 KB - infodemiology_v2i2e38242_app1.docx]

Multimedia Appendix 2
Complete number of top 3 comments in each discussion based on the number of emoji reactions.
[DOCX File , 22 KB - infodemiology_v2i2e38242_app2.docx]

Multimedia Appendix 3
Sentiment expressed in 3 most reacted-to comments (n=80) in each discussion (n=26).
[PNG File , 113 KB - infodemiology_v2i2e38242_app3.png]

References

1. Human Organ and Tissue Donation Act Bill 133. Nova Scotia Legislature. 2019. URL: https://nslegislature.ca/legc/bills/63rd_2nd/1st_read/b133.html [accessed 2022-02-15]
2. Norris S. Organ donation and transplantation in Canada. Library of Parliament, Ottawa, Canada. 2018 Feb 14. URL: https://publications.gc.ca/collections/collection_2018/bdp-lop/bp/YM32-2-2018-13-eng.pdf [accessed 2022-09-06]
3. Donation activity charts. IRODaT. 2021. URL: https://www.irodat.org/?p=database#data [accessed 2022-02-15]
4. Arshad A, Anderson B, Sharif A. Comparison of organ donation and transplantation rates between opt-out and opt-in systems. Kidney Int 2019 Jun;95(6):1453-1460. [doi: 10.1016/j.kint.2019.01.036] [Medline: 31010718]
5. Matesanz R, Domínguez-Gil B. Opt-out legislations: the mysterious viability of the false. Kidney Int 2019 Jun;95(6):1301-1303. [doi: 10.1016/j.kint.2019.02.028] [Medline: 31122708]
6. Rithalia A, McDaid C, Suekarran S, Myers L, Sowden A. Impact of presumed consent for organ donation on donation rates: a systematic review. BMJ 2009 Jan 14;338(jan14 2):a3162 [FREE Full text] [doi: 10.1136/bmj.a3162] [Medline: 19147479]
7. Mone T. With renewed calls for presumed consent, let s look at the donation and transplant rates of opt-in vs opt-out countries. Transplantation 2017;101:S11. [doi: 10.1097/01.tp.0000524984.62432.48]
8. Organ and tissue donation and transplantation system progress report 2018. Canadian Blood Services. 2018. URL: https://tinyurl.com/mcwav9mt [accessed 2022-02-15]
9. Saab S, Saggi SS, Akbar M, Choi G. Presumed consent: a potential tool for countries experiencing an organ donation crisis. Dig Dis Sci 2019 May 5;64(5):1346-1355. [doi: 10.1007/s10620-018-5388-6] [Medline: 30519849]
10. Etheredge HR. Assessing global organ donation policies: opt-in vs opt-out. Risk Manag Healthc Policy 2021;14:1985-1998 [FREE Full text] [doi: 10.2147/RMHP.S270234] [Medline: 34012308]
11. Madden S, Collett D, Walton P, Empson K, Forsythe J, Ingham A, et al. The effect on consent rates for deceased organ donation in Wales after the introduction of an opt-out system. Anaesthesia 2020 Sep 05;75(9):1146-1152 [FREE Full text] [doi: 10.1111/anae.15055] [Medline: 32372409]

XSL•FO
RenderX

12. Weiss MJ, Krmpotic K, Cyr T, Dhanani S, Dieudé M, Dirk JS, et al. A program of research to evaluate the impact of deceased organ donation legislative reform in nova scotia: the LEADDR program. Transplant Direct 2021 Jan;7(1):e641 [FREE Full text] [doi: 10.1097/TXD.0000000000001093] [Medline: 33335980]
13. Chadwick A, Kaiser J, Vaccari C, Freeman D, Lambe S, Loe BS, et al. Online social endorsement and COVID-19 vaccine hesitancy in the United Kingdom. Soc Media Society 2021 Apr 05;7(2):205630512110088. [doi: 10.1177/20563051211008817]
14. Zhang Y, He D, Sang Y. Facebook as a platform for health information and communication: a case study of a diabetes group. J Med Syst 2013 Jun;37(3):9942. [doi: 10.1007/s10916-013-9942-7] [Medline: 23588823]
15. Bradshaw AS, Shelton SS, Wollney E, Treise D, Auguste K. Pro-Vaxxers get out: anti-vaccination advocates influence undecided first-time, pregnant, and new mothers on Facebook. Health Commun 2021 May;36(6):693-702. [doi: 10.1080/10410236.2020.1712037] [Medline: 31920115]
16. Sturm Wilkerson H, Riedl MJ, Whipple KN. Affective affordances: exploring Facebook reactions as emotional responses to hyperpartisan political news. Digital Journalism 2021 Apr 14;9(8):1040-1061. [doi: 10.1080/21670811.2021.1899011]
17. Kite J, Foley BC, Grunseit AC, Freeman B. Please like me: Facebook and public health communication. PLoS One 2016;11(9):e0162765 [FREE Full text] [doi: 10.1371/journal.pone.0162765] [Medline: 27632172]
18. Tustin JL, Crowcroft NS, Gesink D, Johnson I, Keelan J, Lachapelle B. User-driven comments on a Facebook advertisement recruiting Canadian parents in a study on immunization: content analysis. JMIR Public Health Surveill 2018 Sep 20;4(3):e10090 [FREE Full text] [doi: 10.2196/10090] [Medline: 30249585]
19. Voorveld HA, van Noort G, Muntinga DG, Bronner F. Engagement with social media and social media advertising: the differentiating role of platform type. J Advert 2018 Feb 13;47(1):38-54. [doi: 10.1080/00913367.2017.1405754]
20. Al Mamun M, Ibrahim HM, Turin TC. Social media in communicating health information: an analysis of Facebook groups related to hypertension. Prev Chronic Dis 2015 Jan 29;12:E11 [FREE Full text] [doi: 10.5888/pcd12.140265] [Medline: 25633486]
21. Kim J, Lewis SC, Watson BR. The imagined audience for and perceived quality of news comments: exploring the perceptions of commenters on news sites and on Facebook. Social Media Soc 2018 Mar 29;4(1):205630511876574. [doi: 10.1177/2056305118765741]
22. 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]
23. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. Annu Rev Sociol 2001 Aug;27(1):415-444. [doi: 10.1146/annurev.soc.27.1.415]
24. Figeac J, Favre G. How behavioral homophily on social media influences the perception of tie-strengthening within young adults' personal networks. New Media Soc 2021 Jun 25:146144482110206. [doi: 10.1177/14614448211020691]
25. Wang D, Marmo-Roman S, Krase K, Phanord L. Compliance with preventative measures during the COVID-19 pandemic in the USA and Canada: results from an online survey. Soc Work Health Care 2021;60(3):240-255. [doi: 10.1080/00981389.2020.1871157] [Medline: 33407057]
26. 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]
27. Facebook user in Canada. Napolean Cat. 2021 Feb. URL: https://napoleoncat.com/stats/facebook-users-in-canada/2021/02/ [accessed 2022-09-06]
28. Blevis M, Coletto D. Matters of opinion 2017: 8 things we learned about politics, the news, and the internet. Abacus Data. 2017 Feb 7. URL: https://abacusdata.ca/matters-of-opinion-2017-8-things-we-learned-about-politics-the-news-and-the-internet/ [accessed 2022-02-15]
29. Edelman Trust Barometer 2021. Edelman. 2021. URL: https://www.edelman.ca/sites/g/files/aatuss376/files/trust-barometer/2021%20Canadian%20Edelman%20Trust%20Barometer_0.pdf [accessed 2022-02-15]
30. Droman M. Does Serious Journalism Have a Future in Canada?. Ottawa, ON: Canada's Public Policy Forum; 2016.
31. Edelman Trust Barometer 2020. Edelman. 2020. URL: https://www.edelman.ca/sites/g/files/aatuss376/files/2020-02/2020%20Edelman%20Trust%20Barometer%20Canada%20-%20FINAL.pdf [accessed 2022-02-15]
32. Quattrociocchi ZF. Misinformation spreading on Facebook. In: Complex Spreading Phenomena in Social Systems. Computational Social Sciences. Cham: Springer; 2018.
33. Mitchell A, Jurkowitz M, Oliphant J, Shearer E. Americans who mainly get their news on social media are less engaged, less knowledgeable. Pew Research Center. 2020 Jul 30. URL: https://www.pewresearch.org/journalism/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/ [accessed 2022-02-15]
34. Kim C, Yang S. Like, comment, and share on Facebook: how each behavior differs from the other. Public Relations Rev 2017 Jun;43(2):441-449. [doi: 10.1016/j.pubrev.2017.02.006]
35. Tian Y, Galery T, Dulcinati G, Molimpakis E, Sun C. Facebook sentiment: reactions and emojis. In: Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media. 2017 Presented at: Fifth International Workshop on Natural Language Processing for Social Media; Apr, 2017; Valencia, Spain. [doi: 10.18653/v1/W17-1102]

XSL•FO
RenderX

36. Gesselman AN, Ta VP, Garcia JR. Worth a thousand interpersonal words: emoji as affective signals for relationship-oriented digital communication. PLoS One 2019 Aug 15;14(8):e0221297 [FREE Full text] [doi: 10.1371/journal.pone.0221297] [Medline: 31415640]
37. Merrill JB, Oremus W. Five points for anger, one for a 'like': how Facebook's formula fostered rage and misinformation. The Washington Post. 2021 Oct 26. URL: https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/ [accessed 2022-02-15]
38. Social media fact sheet. Pew Research Center. 2021 Apr 7. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/?menuItem=2fc5fff9-9899-4317-b786-9e0b60934bcf [accessed 2022-02-15]
39. Marcon AR, Caulfield T. The Hydroxychloroquine Twitter war: a case study examining polarization in science communication. First Monday 2021 Oct 04 [FREE Full text] [doi: 10.5210/fm.v26i10.11707]
40. Braun V, Clarke V. Thematic analysis. In: APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, D.C., United States: American Psychological Association; 2012.
41. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exercise Health 2019 Jun 13;11(4):589-597. [doi: 10.1080/2159676x.2019.1628806]
42. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. Int J Qual Methods 2017 Oct 02;16(1):160940691773384. [doi: 10.1177/1609406917733847]
43. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
44. Thomas-Meyer M, Mytton O, Adams J. Public responses to proposals for a tax on sugar-sweetened beverages: a thematic analysis of online reader comments posted on major UK news websites. PLoS One 2017;12(11):e0186750 [FREE Full text] [doi: 10.1371/journal.pone.0186750] [Medline: 29166399]
45. Ahmed S, Haines-Saah RJ, Afzal AR, Tam-Tham H, Al Mamun M, Hemmelgarn BR, et al. User-driven conversations about dialysis through Facebook: a qualitative thematic analysis. Nephrology (Carlton) 2017 Apr;22(4):301-307. [doi: 10.1111/nep.12780] [Medline: 27003709]
46. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
47. Cascio MA, Lee E, Vaudrin N, Freedman DA. A team-based approach to open coding: considerations for creating intercoder consensus. Field Methods 2019 Mar 27;31(2):116-130. [doi: 10.1177/1525822x19838237]
48. Ralph A, Chapman JR, Gillis J, Craig JC, Butow P, Howard K, et al. Family perspectives on deceased organ donation: thematic synthesis of qualitative studies. Am J Transplant 2014 Apr;14(4):923-935 [FREE Full text] [doi: 10.1111/ajt.12660] [Medline: 24612855]
49. Tennankore KK, Klarenbach S, Goldberg A. Perspectives on opt-out versus opt-in legislation for deceased organ donation: an opinion piece. Can J Kidney Health Dis 2021;8:20543581211022151 [FREE Full text] [doi: 10.1177/20543581211022151] [Medline: 34188947]
50. Runarsdottir K, Olafsson K, Arnarsson A. [Public attitudes towards presumed consent in organ donation in Iceland]. Laeknabladid 2014 Oct;100(10):521-525 [FREE Full text] [doi: 10.17992/lbl.2014.10.562] [Medline: 25310041]
51. Siminoff LA, Saunders Sturm CM. African-American reluctance to donate: beliefs and attitudes about organ donation and implications for policy. Kennedy Inst Ethics J 2000 Mar;10(1):59-74. [Medline: 11658155]
52. Miller J, Currie S, O'Carroll RE. 'If I donate my organs it's a gift, if you take them it's theft': a qualitative study of planned donor decisions under opt-out legislation. BMC Public Health 2019 Nov 06;19(1):1463 [FREE Full text] [doi: 10.1186/s12889-019-7774-1] [Medline: 31694604]
53. Caulfield T, Marcon A, Murdoch B, Brown J, Perrault S, Jarry J, et al. Health misinformation and the power of narrative messaging in the public sphere. Can J Bioeth 2019 Mar 20;2(2):52-60 [FREE Full text] [doi: 10.7202/1060911ar]
54. Tanner C, Munsie M, Sipp D, Turner L, Wheatland C. The politics of evidence in online illness narratives: an analysis of crowdfunding for purported stem cell treatments. Health (London) 2019 Jul 19;23(4):436-457. [doi: 10.1177/1363459319829194] [Medline: 30782021]
55. Dahlstrom MF, Scheufele DA. (Escaping) the paradox of scientific storytelling. PLoS Biol 2018 Oct;16(10):e2006720 [FREE Full text] [doi: 10.1371/journal.pbio.2006720] [Medline: 30300354]
56. Miller J, Currie S, O'Carroll RE. 'What if I'm not dead?' - myth-busting and organ donation. Br J Health Psychol 2019 Feb;24(1):141-158 [FREE Full text] [doi: 10.1111/bjhp.12344] [Medline: 30345605]
57. The potential impact of an opt out system for organ donation in the UK: an independent report from the Organ Donation Taskforce. Organ Donation Taskforce. 2008. URL: https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/4250/thepotentialimpactofanoptoutsystemfororgandonationintheuk.pdf [accessed 2022-02-15]
58. Yan H, Yates J. Improving acceptability of nudges: learning from attitudes towards opt-in and opt-out policies. Judgment Decision Making 2019;14(1):26-39 [FREE Full text]
59. Rabinow P, Rose N. Biopower today. BioSocieties 2006 Jun 19;1(2):195-217 [FREE Full text] [doi: 10.1017/s1745855206040014]

XSL•FO
RenderX

60. Sharif A. Organ donation and Islam-challenges and opportunities. Transplantation 2012 Sep 15;94(5):442-446. [doi: 10.1097/TP.0b013e31825f4474] [Medline: 22854424]
61. Oliver M, Woywodt A, Ahmed A, Saif I. Organ donation, transplantation and religion. Nephrol Dial Transplant 2011 Feb;26(2):437-444. [doi: 10.1093/ndt/gfq628] [Medline: 20961891]
62. Irving MJ, Tong A, Jan S, Cass A, Chadban S, Allen RD, et al. Community attitudes to deceased organ donation: a focus group study. Transplantation 2012 May 27;93(10):1064-1069. [doi: 10.1097/TP.0b013e31824db997] [Medline: 22510575]
63. Li MT, Hillyer GC, Husain SA, Mohan S. Cultural barriers to organ donation among Chinese and Korean individuals in the United States: a systematic review. Transpl Int 2019 Oct;32(10):1001-1018 [FREE Full text] [doi: 10.1111/tri.13439] [Medline: 30968472]
64. Myers West S. Censored, suspended, shadowbanned: user interpretations of content moderation on social media platforms. New Media Soc 2018 May 08;20(11):4366-4383. [doi: 10.1177/1461444818773059]
65. 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;111(3):389-391 [FREE Full text] [doi: 10.17269/s41997-020-00343-0] [Medline: 32519085]
66. 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]
67. Zannettou S, Caulfield T, De Cristofaro E, Sirivianos M, Stringhini G, Blackburn J. Disinformation warfare: understanding state-sponsored trolls on Twitter and their influence on the web. In: Companion Proceedings of The 2019 World Wide Web Conference. 2019 Presented at: WWW '19: The Web Conference; May 13 - 17, 2019; San Francisco USA URL: https://doi.org/10.48550/arXiv.1801.09288 [doi: 10.1145/3308560.3316495]
68. CMA Code of ethics and professionalism. Canadian Medical Association. 2018. URL: https://policybase.cma.ca/documents/PolicyPDF/PD19-03.pdf [accessed 2022-02-15]
69. 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]
70. Wood T, Porter E. The elusive backfire effect: mass attitudes' steadfast factual adherence. Polit Behav 2018 Jan 16;41(1):135-163 [FREE Full text] [doi: 10.1007/s11109-018-9443-y]
71. Ecker UK, O'Reilly Z, Reid JS, Chang EP. The effectiveness of short-format refutational fact-checks. Br J Psychol 2020 Feb;111(1):36-54 [FREE Full text] [doi: 10.1111/bjop.12383] [Medline: 30825195]
72. 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]
73. 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]
74. Acerbi A, Altay S, Mercier H. Research note: fighting misinformation or fighting for information? Harvard Kennedy School Misinformation Review. 2022 Jan 12. URL: https://misinforeview.hks.harvard.edu/article/research-note-fighting-misinformation-or-fighting-for-information/ [accessed 2022-02-15]
75. Epstein Z, Berinsly AJ, Cole R, Gully A, Pennycook G, Rand D. Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online. Harvard Kennedy School Misinformation Review. 2021 May 18. URL: https://misinforeview.hks.harvard.edu/article/developing-an-accuracy-prompt-toolkit-to-reduce-covid-19-misinformation-online/ [accessed 2022-02-15]
76. YouTube. 2022 May 18. URL: https://www.youtube.com/watch?v=YLrbioCsiCI [accessed 2022-08-10]
77. Cabrera LY, Brandt M, McKenzie R, Bluhm R. Online comments about psychiatric neurosurgery and psychopharmacological interventions: public perceptions and concerns. Soc Sci Med 2019 Jan;220:184-192. [doi: 10.1016/j.socscimed.2018.11.021] [Medline: 30453110]
78. The Canadian Press. Big jump in tissue donations, organ referrals after Nova Scotia presumed consent laws. Global News. 2022 Jan 20. URL: https://globalnews.ca/news/8526608/nova-scotia-organ-donation-presumed-consent/ [accessed 2022-02-15]

Edited by T Mackey; submitted 25.03.22; peer-reviewed by R Bluhm, K Uludag; comments to author 19.07.22; revised version received 10.08.22; accepted 15.08.22; published 14.09.22.

Please cite as:
Marcon AR, Wagner DN, Giles C, Isenor C
Web-Based Perspectives of Deemed Consent Organ Donation Legislation in Nova Scotia: Thematic Analysis of Commentary in Facebook Groups
JMIR Infodemiology 2022;2(2):e38242
URL: https://infodemiology.jmir.org/2022/2/e38242
doi:10.2196/38242
PMID:37113450

XSL•FO
RenderX

©Alessandro R Marcon, Darren N Wagner, Carly Giles, Cynthia Isenor. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 14.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring Factors That Predict Marketing of e-Cigarette Products on Twitter: Infodemiology Approach Using Time Series

Nnamdi C Ezike[1], BSc, MSc, PhD; Allison Ames Boykin[1], BSc, MSc, PhD; Page D Dobbs[1], BSc, MSc, PhD; Huy Mai[2], BSc; Brian A Primack[3], BA, EdM, MD, MSc, PhD

[1]College of Education and Health Professions, University of Arkansas, Fayetteville, AR, United States

[2]College of Engineering, University of Arkansas, Fayetteville, AR, United States

[3]College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States

Corresponding Author:
Nnamdi C Ezike, BSc, MSc, PhD
College of Education and Health Professions
University of Arkansas
751 W Maple Street
Fayetteville, AR, 72701
United States
Phone: 1 479 575 3586
Email: ncezike@uark.edu

Abstract

Background: Electronic nicotine delivery systems (known as electronic cigarettes or e-cigarettes) increase risk for adverse health outcomes among naïve tobacco users, particularly youth and young adults. This vulnerable population is also at risk for exposed brand marketing and advertisement of e-cigarettes on social media. Understanding predictors of how e-cigarette manufacturers conduct social media advertising and marketing could benefit public health approaches to addressing e-cigarette use.

Objective: This study documents factors that predict changes in daily frequency of commercial tweets about e-cigarettes using time series modeling techniques.

Methods: We analyzed data on the daily frequency of commercial tweets about e-cigarettes collected between January 1, 2017, and December 31, 2020. We fit the data to an autoregressive integrated moving average (ARIMA) model and unobserved components model (UCM). Four measures assessed model prediction accuracy. Predictors in the UCM include days with events related to the US Food and Drug Administration (FDA), non-FDA-related events with significant importance such as academic or news announcements, weekday versus weekend, and the period when JUUL maintained an active Twitter account (ie, actively tweeting from their corporate Twitter account) versus when JUUL stopped tweeting.

Results: When the 2 statistical models were fit to the data, the results indicate that the UCM was the best modeling technique for our data. All 4 predictors included in the UCM were significant predictors of the daily frequency of commercial tweets about e-cigarettes. On average, brand advertisement and marketing of e-cigarettes on Twitter was higher by more than 150 advertisements on days with FDA-related events compared to days without FDA events. Similarly, more than 40 commercial tweets about e-cigarettes were, on average, recorded on days with important non-FDA events compared to days without such events. We also found that there were more commercial tweets about e-cigarettes on weekdays than on weekends and more commercial tweets when JUUL maintained an active Twitter account.

Conclusions: e-Cigarette companies promote their products on Twitter. Commercial tweets were significantly more likely to be posted on days with important FDA announcements, which may alter the narrative about information shared by the FDA. There remains a need for regulation of digital marketing of e-cigarette products in the United States.

(JMIR Infodemiology 2022;2(2):e37412) doi:10.2196/37412

KEYWORDS

tobacco; electronic cigarettes; social media; marketing; time series; youth; young adults; infodemiology; infoveillance; digital marketing; advertising; Twitter; promote; e-cigarette

XSL•FO
RenderX

Introduction

Use of electronic nicotine delivery systems (known as electronic cigarettes, vapes, or e-cigarettes) has increased substantially over the past decade, particularly among young populations (youth, those aged under 18 years, and young adults, those aged 18-24 years) [1,2]. E-cigarettes use among these young populations is particularly concerning due to the risks of cardiovascular and respiratory illnesses that these devices can have for those who would not otherwise use tobacco products [3-5]. Further, the addiction potential of these novel tobacco products, especially newer models that contain excessive levels of nicotine, has caused many in the public health community to question if this new technology could create a new generation of smokers, reversing declines in smoking rates and hard-fought public health milestones [6,7].

Recent data suggest that e-cigarette use is most common among those aged 18 to 44 years [2]. People in these age groups are the most active users of Twitter, one of the most popular social media platforms [8]. As of April 2021, 76% of Twitter's 300 million active users were aged 18 to 49 years. With a maximum of 280-character length, messages containing personal information or views about products such as e-cigarettes can be shared by users. Users' posts on Twitter are referred to as tweets.

Emery and colleagues [9] suggest that, when compared to non–e-cigarette users, users of e-cigarette products were more likely to be exposed to information about e-cigarettes via social media platforms, such as Twitter and Facebook, and other mediums like television content, email, and the internet. e-Cigarette content to which social media users are exposed includes tobacco marketing and promotional material [10-12]. This type of advertising on social media helps tobacco companies target users based on their demographic information [13,14]. However, although there has been significant work around the content analysis of commercial tweets about e-cigarettes on social media [15-17], little is known about the factors that drive how often manufacturers of e-cigarettes promote their products on social media.

In 1971, the US Congress outlawed tobacco advertisements on radio and television. Since that time, manufacturers of tobacco products have sought alternative ways to market their products, including marketing campaigns on the internet and social media. Digital marketing, currently unregulated in the United States, offers tobacco (and e-cigarette) companies the opportunity to reach a wide audience [10,18]. This includes social media platforms such as Facebook, Twitter, YouTube, and TikTok [11,19,20]. For example, Huang and colleagues [20] examined the marketing of e-cigarettes on Twitter and found 89.6% of e-cigarette tweets to be commercial tweets. Similarly, Kim and colleagues [11] identified 1.7 million tweets about e-cigarettes spanning over 5 years and found that 93.4% of these tweets advertised e-cigarettes. Social media, therefore, provides a largely unguarded platform for marketing e-cigarette products that has important public health implications.

Social media marketing of e-cigarette products may come from individual accounts, paid corporate advertisements, and paid corporate "influencers" [21]. For example, Jackler and colleagues [22] noted that JUUL, a major e-cigarette company, paid influencers (private social media users with large numbers of followers) to "increase brand awareness and inspire sales." This type of marketing has been associated with the use of e-cigarettes, especially among adolescent audiences [23]. Social media platforms such as Facebook, Instagram, and Twitter prohibit advertisement of tobacco products [24,25]. This restriction only applies to paid advertising. This means that tobacco companies may still market their products on social media via posts and tweets but cannot use paid advertising, which can be specifically used to target users of certain demographic groups.

Although e-cigarette advertisements are currently not regulated, the US Food and Drug Administration (FDA) has the authority to regulate tobacco products in the United States, including manufacture, distribution, and marketing. On March 17, 2021, the FDA requested that 4 e-cigarette companies disclose information about their marketing practices [26]. Part of the request included information on social media advertising and marketing plans, as well as plans to target specific audiences. Given the FDA's limitations on exploring each e-cigarette company's social media marketing, research is needed to understand the factors that predict how tobacco companies conduct brand marketing of their products on social media. Kim and colleagues [11] described the features of commercial tweets about e-cigarettes, including the type of products contained in the advertisement, the number of active accounts, and the type of advertising (promotion, coupon, percent off, and discount). Although these features capture the characteristics of the commercial tweets, they contain little information about the factors that trigger these commercial companies to aggressively promote their products. Thus, the purpose of this study was to determine the best approach for modeling commercial Twitter data on marketed e-cigarette products. This study also sought to explore factors associated with commercial Twitter marketing of e-cigarette products.

Methods

Data Collection and Annotation

The data analyzed in this study are tweets about e-cigarettes between January 1, 2017, and December 31, 2020. The tweets were collected daily using the real-time infoveillance of Twitter health messages (RITHM) open-source software [27]. Using the Twitter streaming application programming interface, the RITHM software gathers key information about each tweet, including the number of duplicate tweets based on the tweet ID, where the software automatically saves duplicate tweets as 1 single tweet record. This was crucial to our analysis as it prevented the factor of tweets or retweets with the same text from influencing our findings. We used search terms that capture Twitter chatter related to e-cigarettes, similar to past research [28-30], including words such as vape, vapes, vaper, vapers, vaping, JUUL, JUULs, JUULing, and tobacco. A total of 1% (n=2401) of the tweets posted between August 23, 2019, and September 25, 2019, were selected for annotation by 2 independent researchers. The date range was selected based on

XSL•FO
RenderX

a particularly high volume of tweets posted for the given dates. Further, selected tweets were stratified by day to account for volume changes in the number of tweets and to accurately represent Twitter discussions over time. Previous work [27,31,32] established that this sample size and selection method provided adequate representation of tweets made within the selected time frame.

The procedures developed by Crabtree and Miller [33] for public health qualitative research served as a guide for developing the codebook used for human annotation. The first step involved an inductive procedure [34]. Using in vivo coding, 3 researchers explored 200 tweets searching for nuanced information related to e-cigarette–related tweets. Next, the team refined the codebook by adding, splitting, expanding, or deleting codes, an inductive procedure used during qualitative data analyses [34,35]. Relevant tweets were coded as dichotomous indicators, denoting whether the tweet referred to vaping in the context of e-cigarettes. For example, the following tweet was classified as a relevant tweet: "Omg!!!!! Mine is getting interrupted by a vaping special. Coming on at 11pm here. _emoj_weary_ _emoj_weary_ _emoj_weary_ I am tired." If the tweet did not mention e-cigarettes or referred to vapor in an unrelated context, it was removed from further analysis. Subsequently, we identified promotional posts about tobacco products that appeared to be advertisements or marketing for vaping products. These posts were classified as commercial tweets. For example, the following tweet was classified as a commercial tweet: "COCO THC CBD Oil # Vape System New pod Style THC # CBD Oil System 4 empty tanks that are easy to fill and a 220ohm slim battery. Share !"

Two coders were provided with online versions of the 2401 tweets for annotation using a qualitative content analysis approach. Coders were also provided with retweets, which are tweets that are in response to other users' tweets. Coding 500 tweets each week, annotators classified tweets as commercial if the tweets were commercial promotion of e-cigarettes and noncommercial if otherwise. Cohen kappa [36] measure of interrater agreement reveal a high coder agreement (κ>.80) on classification of relevant and commercial tweets, indicating over 80% agreement between coders after accounting for chance agreement.

Classification of Tweets

Tweets annotated by human coders were used to train a model to classify the remaining tweets. In this study, classification was performed using a classifier that was pretrained and fine-tuned on BERTweet, a variation of Google AI Language's bidirectional encoder representations from transformers (BERT). Pretrained on English tweets, BERTweet improves on other transformer models used for natural language processing tasks by enhancing the transformer's capability of recognizing important words in a given text sequence [37]. This is accomplished by the masking and next sentence prediction objectives performed in the pretraining layers of BERTweet [38], along with the pretraining optimizations of the "robustly optimized BERT pretraining approach" to address the significant undertraining of BERT [39]. As the model uses the encoder representation of a transformer, BERTweet can be fine-tuned for classification tasks.

Ethics Approval

This study did not use human participants. Data were collected from publicly available platforms and require no ethics approval.

Modeling Techniques

One of the goals of this study was to find the best approach for modeling time series data to predict commercial Twitter activities about e-cigarettes and vaping. Time series models can provide tools to predict or forecast future events based on past trends. Time series modeling has been extensively used in public health research to predict coronavirus disease spread, study Zika epidemic case counts, and understand changes in public health opinions due to coronavirus restrictions [40-42]. This study compared the performance of the autoregressive integrated moving average (ARIMA) [43] model and unobserved components model (UCM) [44] in predicting commercial Twitter activities about e-cigarettes and vaping.

ARIMA Approach

The ARIMA model can be expressed as

$$yt = \varphi0 + \varphi1yt{-}1 + ... + \varphi pyt{-}p + \theta1\varepsilon t{-}1 + ... + \theta q\varepsilon t{-}q + \varepsilon t \quad (1)$$

where t is the time point, y_t is the forecast variable which is the frequency of commercial tweets at time t, φ_i is the coefficient for the autoregressive term p, θ_j is the coefficient for the moving average term q, and ε_t is the random error at time t. The ARIMA modeling technique consists of 3 steps: model identification, parameter estimation, and model diagnostic checking. These steps were performed to optimize the ARIMA model for assessing the frequency of commercial tweets. First, the amount of differencing and the lag size were determined at the model identification stage. ARIMA models are based on the assumption of stationarity of the differenced series [45]. Second, we verified that the stationarity and homoscedasticity assumptions were satisfied after model estimation. Third, diagnostic plots such as autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were examined to assess if the fitted models were appropriate. The ACF plot provides the correlation between observations at time t and at time t–k (where k is the number of lags). It is preferred to have autocorrelations near zero for all lags. The PACF plot provides the correlation between observations at time t and the residuals at previous lags. Essentially, PACF removes the components that have been explained by previous lags. The PACF plot is a useful tool for determining the order of the autoregressive term. Finally, we selected the appropriate autoregressive (AR) parameter (p) and moving average (MA) parameter (q) based on the ACF and PACF plots.

UCM Approach

One of the main advantages of the UCM approach over the ARIMA approach is that researchers can identify and introduce additional explanatory variables. The explanatory variables could be intervention variables that are useful in explaining patterns in the series [44]. In addition, UCM is efficient in

XSL•FO
RenderX

handling missing observations [45]. In the UCM modeling framework, the series is decomposed into trend, seasonal, cyclical, and autoregressive components. In addition, the UCM models regression effects due to the predictor series. The UCM can be expressed as

where t is the time point, y_t is the forecast variable which is the frequency of commercial tweets at time t, μ_t is the trend component, γ_t is the seasonal component, and ψ_t is the cyclical component. The term is used to model the autoregressive regression component based on past observations of the series. The term captures explanatory regression predictors where x_{jt} is the observed value of predictor x_j at time t and β_j is the regression slope for predictor x_j. Finally, ε_t is a white noise error term.

We included 4 explanatory variables in the UCM used in this study: (1) FDA-related events, (2) other (non-FDA) events, (3) day of the week, and (4) JUUL.

FDA Variable

Drug Watch International and Consumer Advocates for Smoke-Free Alternatives Association (CASAA) maintain a timeline of events of vaping and e-cigarettes. We reviewed the timeline to identify days with FDA-related events such as announcements about vaping/e-cigarettes, campaigns, and court rulings. The FDA variable was dummy coded. Days in which there were FDA-related events were coded as 1 and 0 if otherwise.

Other Variable

The Drug Watch and CASAA timeline of events was also used to create a variable for other events. These events were events of high importance that were non–FDA-related. For example, other events included state legislative actions controlling the use of e-cigarette products and significant scientific research studies reported in national news. The variable on other events about e-cigarettes was also dummy coded. A value of 1 was used to indicate a day with such an event.

JUUL Variable

We also included a variable referred to as JUUL in the model. We included this variable in order to understand the impact of JUUL's tweet activities on the frequency of commercial tweets about e-cigarettes. JUUL is the most popular e-cigarette brand accounting for 76% of e-cigarette retail sales [46]. JUUL has a corporate Twitter page. Of note, JUUL stopped tweeting from its corporate Twitter account on August 29, 2019. We included a dummy coded variable by assigning a value of 1 to indicate periods that JUUL was tweeting and 0 for the period when they stopped tweeting (ie, after August 29, 2019). We will refer to the periods when JUUL was tweeting as "active" and the periods of prolonged inactivity as "inactive."

Day Variable

Finally, a dummy-coded day variable was included in the model to indicate whether the commercial tweet was promoted on a weekend (value of 1) or weekday (value of 0).

Data Analysis

All analyses were performed in SAS (version 9.4, SAS Institute Inc). There were 1401 out of 1460 days with complete data. A RITHM software outage resulted in failure to collect 59 days of data. Missing observations may bias the forecasting ability of time series models. Jalles [45] noted that it is difficult to the use ARIMA model in the presence of missing data. However, the UCM procedure handles missing values efficiently and can be extended to ARIMA models [47,48]. Both the ARIMA model and UCM were fitted using the UCM procedure in SAS [48].

We took an iterative modeling approach to determine the best fitting UCM. First, we specified a UCM with trend and irregular components. Next, we examined the parameter estimates of the components to determine whether to treat them as stochastic or deterministic. Nonsignificant (deterministic) components were removed from the model. Finally, the 4 explanatory variables used in this study were included in the model (ie, day, FDA event, non-FDA event, and JUUL). At each step, the ACF and PACF plots served as diagnostic tools for assessing the fitted models.

Model Evaluation

The performance of our models was evaluated using root mean square error (RMSE), mean absolute percentage error (MAPE), mean absolute deviation (MAD), and coefficient of determination (R^2).

Root Mean Square Error

RMSE gives the overall measure of accuracy of how well the model predicts the frequency of daily commercial tweets. The RMSE for each model was computed using

where y_t is the frequency of commercial tweets at time t, is the predicted frequency of commercial tweets at time t based on the fitted model, and n is the number of observations.

Mean Absolute Percentage Error

MAPE measures the accuracy of the model in terms of percentage error. The MAPE for each model was computed using

where y_t is the frequency of commercial tweets at time t, is the predicted frequency of commercial tweets at time t based on the fitted model, and n is the number of observations. Smaller values of the MAPE indicate fewer prediction errors, hence the best fitting model will have a smaller MAPE.

XSL•FO
RenderX

Mean Absolute Deviation

MAD is the average of the absolute value of the deviation between the observed frequency of commercial tweets and the predicted frequency of commercial tweets based on the fitted model. Essentially, MAD provides the amount of prediction errors in the same units as the observed counts. The MAD for each model was computed using

where y_t is the frequency of commercial tweets at time t, is the predicted frequency of commercial tweets at time t based on the fitted model, and n is the number of observations. Smaller values of the MAD are preferred.

Coefficient of Determination

The R^2 (coefficient of determination) statistic measures the proportion of variance in the frequency of commercial tweets which is accounted for by the predictors. The R^2 statistic is computed as

where y_t is the frequency of commercial tweets at time t, is the average frequency of commercial tweets, is the predicted frequency of commercial tweets at time t based on the fitted model, and n is the number of observations. A larger R^2 statistic is preferred.

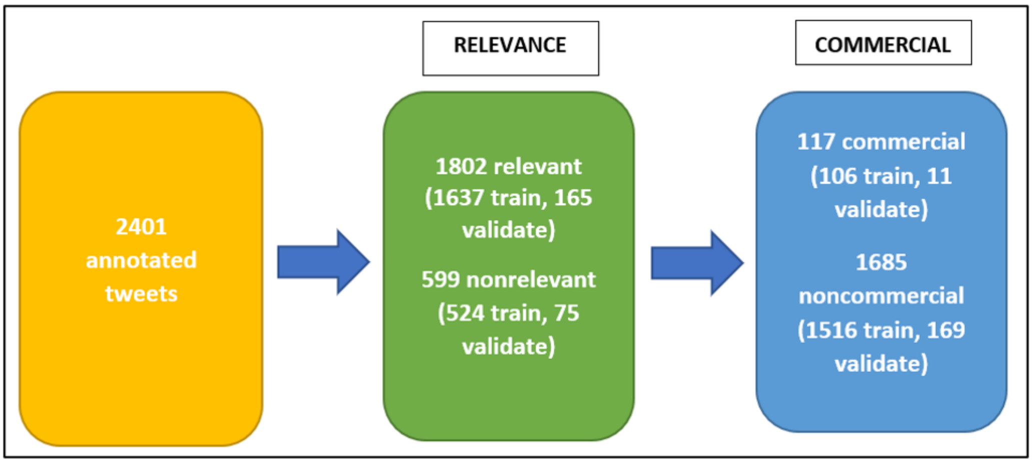
Results

Tweet Classification Results

Classifier Settings

Two BERTweet classifiers were trained using the set of annotated tweets: one for relevance and another for commercial. The number of tweets used to train and validate each classifier is provided in Figure 1. The sets of tweets for relevance and commercial were each split randomly to where 90% of the tweets were used to train and fine-tune the model while the remaining 10% was used to validate the model. For the hyperparameters, each BERTweet classifier was trained for 20 epochs with a batch size of 32 and learning rate of 5×10^{-5}. For comparison, we used the long short-term memory (LSTM) model proposed by Visweswaran et al [28], which was trained for 5 epochs and a batch size of 64 under the same splits on the annotated data set as the BERTweet classifiers. As part of a previous study analyzing the trend in the commercial nature of tweets related to vaping, this LSTM model was found to have the highest classification accuracy when tested against other deep learning classifiers such as convolutional neural network (CNN), LSTM-CNN, and bidirectional LSTM [28].

Figure 1. Filtering process of the 2401 tweets used to train and validate the BERTweet classifiers.



Classifier Results

We measured the performance of the classifiers using F_1, which is a function of precision and recall, and area under the receiver operating characteristic (AUROC), which measures the discrimination of the classifiers. For the task of classifying a tweet as relevant or nonrelevant, the BERTweet classifier obtained an F_1-score of 0.976 and an AUROC score of 0.945 while the LSTM model had an F_1-score of 0.924 and an AUROC score of 0.924. In classifying tweets as commercial or noncommercial, the BERTweet classifier produced an F_1-score of 0.990 and an AUROC score of 0.993. In comparison, the LSTM classifier achieved an F_1-score of 0.727 and an AUROC score of 0.903.

Descriptive Statistics

A total of 1,821,603 commercial e-cigarette tweets were recorded from January 1, 2017, to December 31, 2020. Figure 2 presents the daily frequency of commercial tweets. On average, there were 1300 commercial tweets per day, and the frequency of tweets was highly variable with a standard deviation of 718. Figure 3 presents a visual comparison of the daily frequency of relevant (ie, tweets that referred to vaping in the context of e-cigarettes) and commercial tweets about e-cigarettes. On average, 26% (SD 9.3%) of the relevant tweets were brand marketing of e-cigarette products. Brand marketing of e-cigarettes on Twitter declined over the 4-year period. In 2017, the average percentage of commercial tweets was 35% (SD 3.5%). This dropped to an average of 30% (SD 8.9%) in 2018 and an average of 20% (SD 7.3%) in 2019. Finally,

XSL•FO
RenderX

following Twitter's ban on paid advertising, only 19% (SD 3.2%) of the relevant tweets in 2020 were classified as commercial tweets.

Table 1 presents the descriptive statistics of the explanatory variables investigated. On average, the mean frequency of daily commercial tweets on days with FDA-related events was 1447.60 (SD 659.08) compared to 1295.10 (SD 719.61) on days without FDA events. Similarly, on average, there were more commercial tweets on days with other non-FDA events (mean 1336.21, SD 604.61) and on weekdays (mean 1390.20, SD 585.85). The average number of daily commercial tweets when JUUL maintained an active account was over 1000 tweets higher than when JUUL stopped tweeting from its corporate account.

Figure 2. Daily frequency of commercial tweets from January 1, 2017, to December 31, 2020.

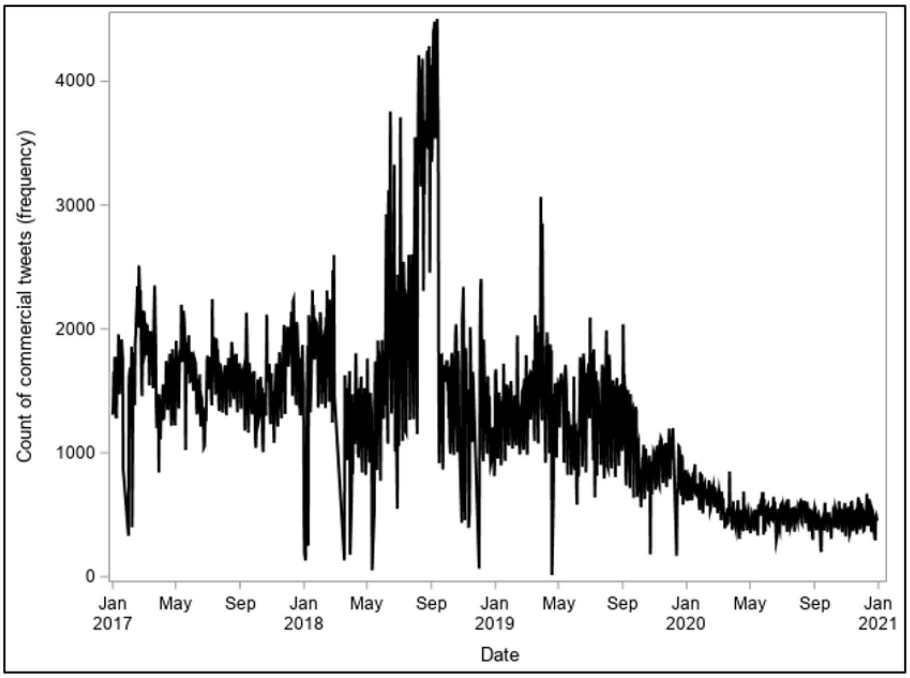
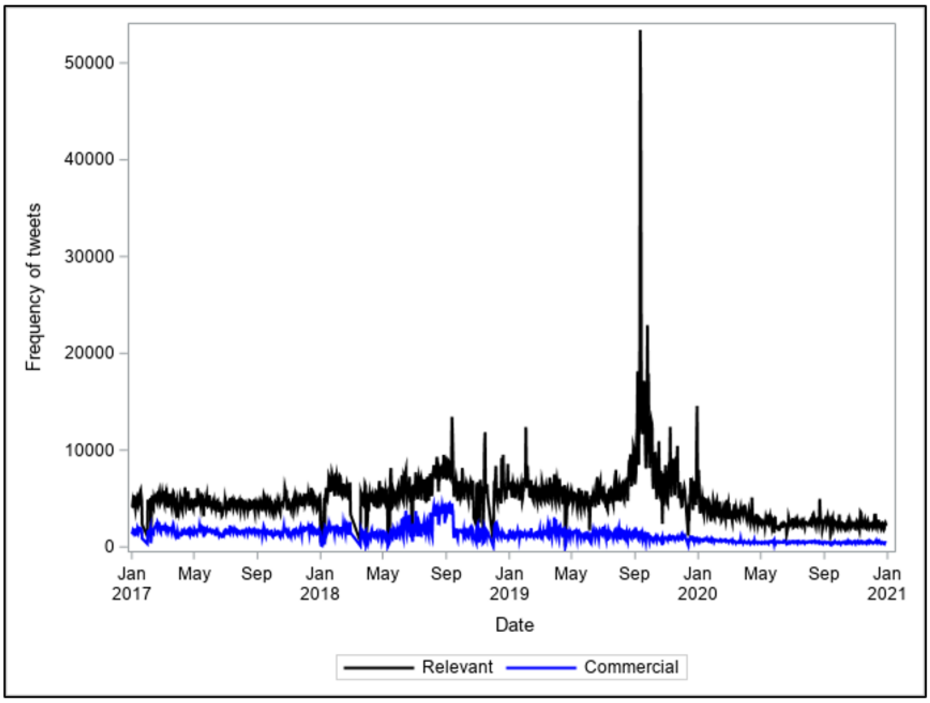


Figure 3. Daily frequency of relevant and commercial tweets from January 1, 2017, to December 31, 2020.



XSL•FO
RenderX

Table 1. Summary of daily frequency of commercial tweets for each predictor.

Predictor and description of level	Values, mean (SD)
FDA[a]	
FDA event (n=47)	1447.60 (659.08)
No FDA event (n=1354)	1295.10 (719.61)
Other	
Other event (n=137)	1336.21 (604.61)
No other event (n=1264)	1296.31 (729.30)
JUUL	
Active[b] account (n=920)	1648.76 (630.19)
Inactive account (n=481)	633.56 (254.82)
Day	
Weekend (n=395)	1071.04 (744.84)
Weekday (n=1006)	1390.20 (585.85)

[a]FDA: US Food and Drug Administration.

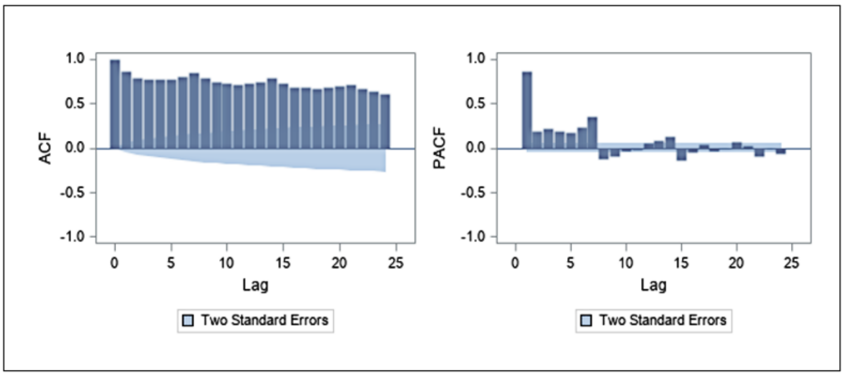
[b]Active is defined as periods when JUUL was tweeting from its corporate Twitter account.

Model Estimation Summary

ARIMA Approach

The frequency of daily commercial tweets shown in Figure 1 does not appear to suggest the presence of seasonal or cyclical trends in the data. The identification stage of the data showed that the series is nonstationary, as depicted in the ACF and PACF plots in Figure 4. The ACF plot of a stationary series will decay to zero relatively quickly, which is not the case in Figure 4. We performed a first-order differencing of the series in order to establish stationarity (see Figure 5). The differenced series suggests that AR(7) and MA(1) were appropriate for the data. This suggests that the model uses commercial tweets about e-cigarette for the past 7 days to forecast the frequency of commercial tweets for the next day. The ACF and PACF plots of the final higher order ARIMA model with p=7 and q=1 are presented in Figure 6. These plots suggest that the fitted model yields a better fit to the data.

Figure 4. Autocorrelation function (left panel) and partial autocorrelation function (right panel) plots for daily commercial tweets about e-cigarettes before differencing for the autoregressive integrated moving average model. ACF: autocorrelation function; PACF: partial autocorrelation function.



XSL•FO
RenderX

Figure 5. First-order differenced frequency of commercial tweets from January 1, 2017, to December 31, 2020.

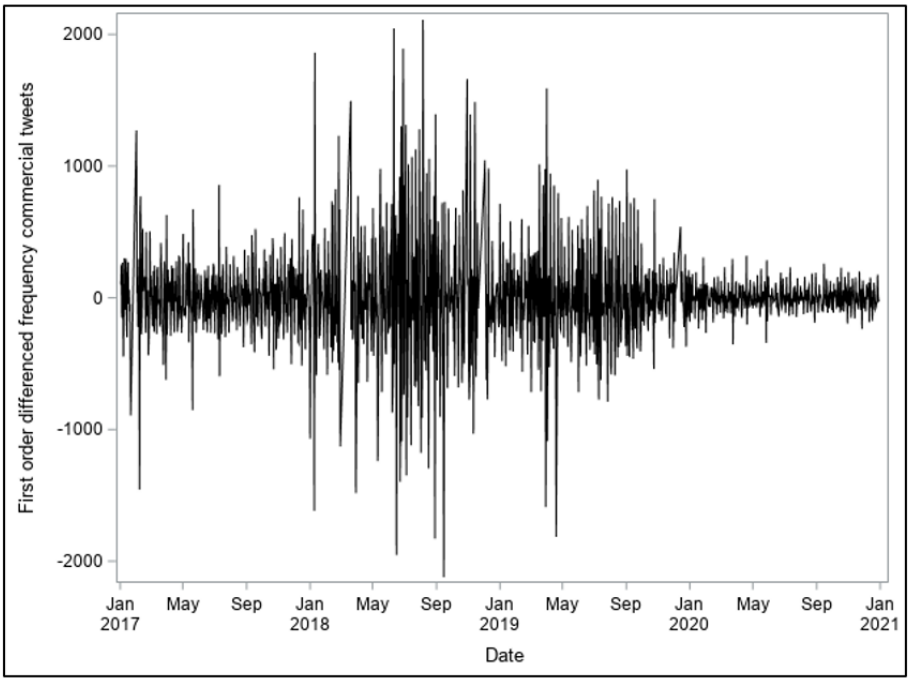
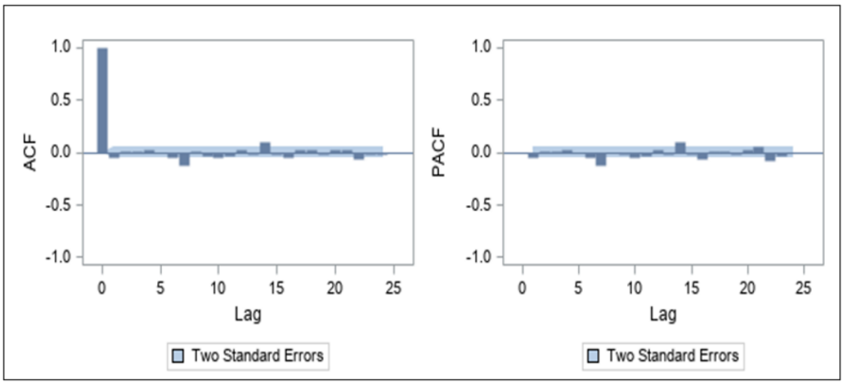


Figure 6. Autocorrelation function (left panel) and partial autocorrelation function (right panel) plots for daily commercial tweets about e-cigarettes after first-order differencing for the autoregressive integrated moving average model. ACF: autocorrelation function; PACF: partial autocorrelation function.



UCM Approach

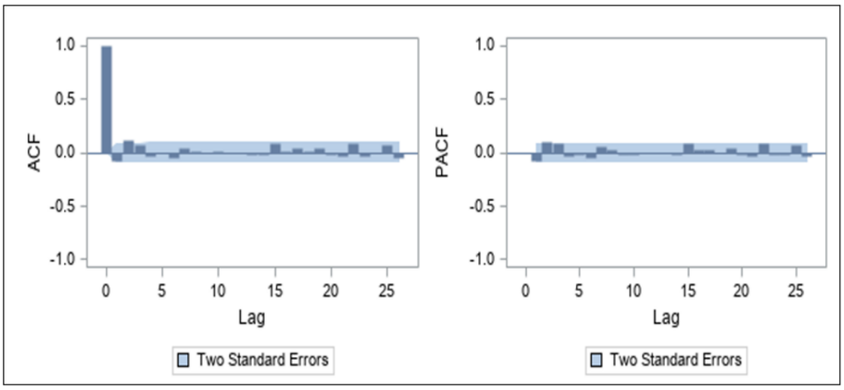
The first fitted UCM included only the trend and irregular components. The final estimates of the free parameters for the UCM with only irregular and trend components are presented in Table 2. This table shows the variances of the irregular, slope, and level components. The results suggest fixing the variance of the slope component to zero (=0.00, *P*=.99) while inferring stochastic irregular (=82530, *P*<.001) and stochastic level (=13043, *P*<.001) components. Subsequent specification of the UCM, by fixing to zero, suggests dropping the slope component from the model (χ^2_1=0.06, *P*=.81). The final specified UCM, after dropping the slope component, includes irregular and level components and all 4 predictors (ie, FDA events, other events, day, and JUUL). The ACF and PACF plots shown in Figure 7 suggest that the specified UCM with all 4 predictors was a good fit to the data.

Table 2. Final estimates of free parameters of the unobservable components model.

Component	Parameter	Estimate	SE	*t*-score	*P* value
Irregular		82530	4938.90	16.71	<.001
Level		13043	2533.70	5.15	<.001
Slope		0	0	0.01	.99

XSL•FO
RenderX

Figure 7. Autocorrelation function (left panel) and partial autocorrelation function (right panel) plots for daily commercial tweets about e-cigarettes for the unobservable components model. ACF: autocorrelation function; PACF: partial autocorrelation function.



Model Comparison

Four measures were used to evaluate the predictive performance of the ARIMA model and UCM. The prediction accuracy of the models is summarized in Table 3. The results show that the UCM outperformed the ARIMA model. From Table 3, the MAPE indicates that, on average, the predicted values of the UCM are only off by about 12% compared to 31% for the ARIMA model. Similarly, the UCM produced the smallest RMSE (102.47) estimates, indicating that the UCM is more appropriate for our data. The MAD suggests that the UCM resulted in the smallest MAD (65.08) between the predicted frequency of commercial tweets and the observed frequency of commercial tweets. Finally, the findings show that 84% of the variability in the commercial tweets is well-described components in the UCM compared to 79% when the data were fitted with ARIMA model.

Table 3. Fit indices based on residuals for various models.

Criterion	Model	
	ARIMA[a]	UCM[b]
RMSE[c]	314.62	102.47
MAPE[d] (%)	31.20	11.98
MAD[e]	190.60	65.08
R^{2}[f]	0.79	0.84

[a]ARIMA: autoregressive integrated moving average.
[b]UCM: unobservable component modeling.
[c]RMSE: root mean squared error.
[d]MAPE: mean absolute percentage error.
[e]MAD: mean absolute deviation.
[f]R^2: coefficient of determination.

Predictors of Commercial Tweets About e-Cigarettes

All 4 explanatory variables included in the UCM were significant predictors of the frequency of commercial tweets about e-cigarettes. The results of the predictors are presented in Table 4. The results indicate that, on average, commercial tweets about e-cigarette on the days with FDA events were significantly higher by around 20 tweets per day after accounting for other variables (β=19.32, *P*<.001). The coefficient associated with "other" event was 7.74. This implies that commercial tweets about e-cigarette on the days with other major events were significantly higher by around 8 tweets per day, after accounting for other variables, on average (β=7.74, *P*=.001). Compared to weekdays, the results show that there were significantly fewer commercial tweets about e-cigarettes on weekends by around 5 tweets after accounting for other variables (β=–4.73, *P*=.001). Furthermore, we found that, on average, commercial tweets about e-cigarettes when JUUL's Twitter account was active were significantly higher by around 171 tweets per day, after accounting for other variables (β=170.68, *P*<.001).

XSL•FO
RenderX

Table 4. Unobservable components model analysis summary for explanatory variables in the model.

Predictors	Estimate	SE	*t*-score	*P* value
FDA[a] event	19.32	3.65	5.29	<.001
Other event	7.74	2.35	3.30	.001
JUUL	170.68	38.29	4.46	<.001
Weekend	–4.73	1.48	–3.19	.001

[a]FDA: US Food and Drug Administration.

Discussion

Principal Findings

Brand marketing and promotion of e-cigarette products on social media are currently unregulated in the United States. The lack of social media surveillance means that youths are continually exposed to digital marketing of e-cigarette products. As one study reports, Twitter expanded the reach of information about e-cigarettes by 10-fold [49]. Our study contributes to knowledge about factors that drive how commercial companies engage in brand marketing and advertising of e-cigarette products. This analysis used the UCM to model the daily frequency of commercial tweets about e-cigarettes. Previous studies that explored brand marketing and advertising of e-cigarettes only used descriptive statistics to describe the frequency of tweets [11,20,50]. Thus, a strength of this study is the use of 4 explanatory variables to predict the daily frequency of commercial tweets about e-cigarettes. We used data on commercial tweets about e-cigarettes collected over a 4-year period to investigate this.

We found that the daily frequency of commercial tweets was, on average, higher on days with FDA-related events and other non-FDA important events. One possible explanation of this result is that manufacturers of e-cigarette products flood the Twitter space with digital marketing on days with major FDA announcements. For example, there were 3782 commercial tweets about e-cigarettes on September 11, 2018. This was the highest frequency of commercial tweets recorded during an FDA-related event within our data collection period (ie, from January 1, 2017, to December 31, 2020). Remarkably, there were 2 important FDA-related events on this day. First, the FDA issued a statement on “new steps to address epidemic of youth e-cigarette use” [51]. Second, the FDA issued warning letters to more than 1300 retailers and 5 major manufacturers for their roles in perpetuating youth access [52]. There was a noteworthy spike in the number of commercial tweets on the same day that the FDA issued these letters. Research has shown that manufacturers of e-cigarettes use paid social media influencers to promote their products. The spike recorded on September 11, 2018, may suggest that FDA-related events or other major events are a part of marketing plans of e-cigarette manufacturers. In a March 17, 2021, brief, the FDA requested marketing documents from 4 manufacturers of e-cigarette products to understand how these commercial companies engage their users on social media. This analysis provides evidence of trends in brand marketing and advertisement of e-cigarette products when there are important announcements.

In late 2019, some social media platforms restricted paid advertising of tobacco products on their platforms. Twitter’s policy states that “Twitter prohibits the promotion of tobacco products, accessories, and brands globally” [25]. We observed a decline in the frequency of commercial tweets after these social media platforms restricted paid advertising of tobacco products. Interestingly, JUUL stopped tweeting from its corporate account on August 29, 2019, coinciding with the period that some social media companies moved to ban paid advertising of tobacco products on their platforms. We observed that there were, on average, 1000 fewer commercial tweets about e-cigarettes in 2020 compared to the previous years in this study (ie, 2017 to 2019). This demonstrates that tobacco companies still get around these policies through nonpaid advertisements and use of paid social media influencers [21,22].

Adequately modeling our data was essential to provide policymakers with appropriate tools to forecast daily patterns in commercial tweets about e-cigarettes. To find the best-fitting model for our data, we compared the prediction accuracies of 2 statistical models: ARIMA and UCM. The prediction accuracies of the ARIMA model and UCM were judged using MAPE, MAD, RMSE, and R^2 statistics. The results demonstrate the utility of UCM in predicting daily commercial tweets about e-cigarettes. We showed that UCM was an improvement over ARIMA. Unfortunately, forecasting in ARIMA is limited to past behavior of the variable (ie, frequency of commercial tweets). This implies that the effects of other factors or confounding variables cannot be modeled in ARIMA. In addition, outliers are difficult to forecast in ARIMA [45]. The UCM compensates for ARIMA as it provided the luxury to capture different components in the series. In addition, we included 4 explanatory variables in the UCM. All 4 explanatory variables that we examined significantly predicted the daily frequency of commercial tweets about e-cigarettes.

Limitations

One limitation of this analysis is that commercial content was investigated using Twitter only. Future studies could explore other social media platforms commonly used among young audiences such as Facebook, Snapchat, and YouTube [53]. Another limitation of this study is the limited period of selected tweets for annotation. Tweets between August 23, 2019, and September 25, 2019, were selected for annotation and subsequent training of the classifier. Another limitation is that we did not develop any mechanisms for filtering out suspicious “bot” accounts, which may include newly opened accounts or accounts with zero followers. The public health community has called for increased surveillance of social bots, which are

XSL•FO
RenderX

automated accounts relying on sophisticated artificial intelligence to influence discussion, ideas, or products [54,55]. However, a previous study on e-cigarettes revealed that tweets posted by bot accounts were less than 5% since 2012 [56]. For this reason, we did not use bot detection but see this approach as an important step in future research. We acknowledge that the search terms we used to capture Twitter chatter related to e-cigarettes may not have been exhaustive. Some tweets related to e-cigarettes that did not include any of the search terms that we used may have been missed during data collection. Additional search terms from recent research and trending hashtags should be considered in future work.

Research has shown that manufacturers of e-cigarette products use the services of social media influencers to market e-cigarette products. Our study did not distinguish among type of commercial tweet (eg, whether the tweet was from a corporate marketing account or other accounts such as paid social media influencers). In addition, the classifier developed for this study did not include specific marketing themes of commercial tweets (eg, flavors or price promotions). These could serve as areas of consideration for future studies, especially with the FDA seeking to understand the social media advertising and marketing plans of manufacturers of e-cigarette products. Despite these limitations, the UCM is promising in modeling predictors of commercial tweets about e-cigarettes.

Conclusion

The aim of this study was to investigate factors that predict changes in daily frequency of commercial tweets about e-cigarettes using time series modeling techniques. Data collected were fitted using 2 time series models, ARIMA and UCM. The results of the UCM, which proved to be the best fitting model, showed that brand advertisement and marketing of e-cigarettes on Twitter was significantly higher on days with FDA-related events compared to days without FDA events after accounting for other variables. In addition, we found higher marketing of e-cigarette products on days with important national news like state legislative actions controlling the use of e-cigarette products and significant scientific research studies. We conclude that e-cigarette companies may increase brand marketing of their products on days with important FDA announcements related to e-cigarettes and days with important national news about e-cigarettes, possibly to alter the narrative about the information shared by the FDA or other important news reporting on e-cigarettes. Our results also reveal significantly higher marketing of e-cigarette products on weekdays compared to weekends. Previous work showed that the use of e-cigarette products decreased during weekends [57]. This leads us to believe that e-cigarette companies, more likely than not, target their audience the most during weekdays.

Acknowledgments

This work was supported through grant R01CA225773 from the National Cancer Institute.

Conflicts of Interest

None declared.

References

1. Bandi P, Cahn Z, Goding Sauer A, Douglas CE, Drope J, Jemal A, et al. Trends in e-cigarette use by age group and combustible cigarette smoking histories, U.S. adults, 2014-2018. Am J Prev Med 2021 Feb;60(2):151-158. [doi: 10.1016/j.amepre.2020.07.026] [Medline: 33032869]
2. Dai H, Leventhal AM. Prevalence of e-cigarette use among adults in the United States, 2014-2018. JAMA 2019 Nov 12;322(18):1824-1827 [FREE Full text] [doi: 10.1001/jama.2019.15331] [Medline: 31524940]
3. Alzahrani T, Pena I, Temesgen N, Glantz S. Association between electronic cigarette use and myocardial infarction. Am J Prev Med 2018 Oct;55(4):455-461 [FREE Full text] [doi: 10.1016/j.amepre.2018.05.004] [Medline: 30166079]
4. Wills TA, Pagano I, Williams RJ, Tam EK. E-cigarette use and respiratory disorder in an adult sample. Drug Alcohol Depend 2019 Jan 01;194:363-370 [FREE Full text] [doi: 10.1016/j.drugalcdep.2018.10.004] [Medline: 30472577]
5. Layden JE, Ghinai I, Pray I, Kimball A, Layer M, Tenforde M, et al. Pulmonary illness related to e-cigarette use in Illinois and Wisconsin—final report. N Engl J Med 2020 Mar 05;382(10):903-916. [doi: 10.1056/NEJMoa1911614] [Medline: 31491072]
6. Loukas A, Marti CN, Cooper M, Pasch KE, Perry CL. Exclusive e-cigarette use predicts cigarette initiation among college students. Addict Behav 2018 Jan;76:343-347. [doi: 10.1016/j.addbeh.2017.08.023] [Medline: 28892771]
7. Doran N, Brikmanis K, Petersen A, Delucchi K, Al-Delaimy WK, Luczak S, et al. Does e-cigarette use predict cigarette escalation? A longitudinal study of young adult non-daily smokers. Prev Med 2017 Jul;100:279-284 [FREE Full text] [doi: 10.1016/j.ypmed.2017.03.023] [Medline: 28583658]
8. Auxier B, Anderson M. Social Media Use in 2021. Washington: Pew Internet and American Life Project; 2021 Apr 07. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2021/04/PI_2021.04.07_Social-Media-Use_FINAL.pdf [accessed 2022-01-23]
9. Emery SL, Vera L, Huang J, Szczypka G. Wanna know about vaping? Patterns of message exposure, seeking and sharing information about e-cigarettes across media platforms. Tob Control 2014 Jul;23 Suppl 3:17-25 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051648] [Medline: 24935893]

XSL•FO
RenderX

10. Chu K, Unger JB, Cruz TB, Soto DW. Electronic cigarettes on twitter—spreading the appeal of flavors. Tob Regul Sci 2015 Apr;1(1):36-41 [FREE Full text] [doi: 10.18001/TRS.1.1.4] [Medline: 27853734]
11. Kim AE, Hopper T, Simpson S, Nonnemaker J, Lieberman AJ, Hansen H, et al. Using Twitter data to gain insights into e-cigarette marketing and locations of use: an infoveillance study. J Med Internet Res 2015 Nov;17(11):e251 [FREE Full text] [doi: 10.2196/jmir.4466] [Medline: 26545927]
12. Kim A, Miano T, Chew R, Eggers M, Nonnemaker J. Classification of Twitter users who tweet about e-cigarettes. JMIR Public Health Surveill 2017 Sep 26;3(3):e63 [FREE Full text] [doi: 10.2196/publichealth.8060] [Medline: 28951381]
13. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
14. Mangold WG, Faulds DJ. Social media: the new hybrid element of the promotion mix. Business Horizons 2009 Jul;52(4):357-365. [doi: 10.1016/j.bushor.2009.03.002]
15. McCausland K, Maycock B, Leaver T, Wolf K, Freeman B, Thomson K, et al. E-Cigarette promotion on Twitter in Australia: content analysis of Tweets. JMIR Public Health Surveill 2020 Nov 05;6(4):e15577 [FREE Full text] [doi: 10.2196/15577] [Medline: 33151159]
16. Sowles SJ, Krauss MJ, Connolly S, Cavazos-Rehg PA. A content analysis of vaping advertisements on Twitter, November 2014. Prev Chronic Dis 2016 Sep 29;13:E139 [FREE Full text] [doi: 10.5888/pcd13.160274] [Medline: 27685432]
17. Williams RS, Derrick J, Liebman AK, LaFleur K. Content analysis of e-cigarette products, promotions, prices and claims on Internet tobacco vendor websites, 2013-2014. Tob Control 2017 Nov 03;27:34-40. [doi: 10.1136/tobaccocontrol-2017-053762] [Medline: 29101294]
18. Richardson A, Ganz O, Stalgaitis C, Abrams D, Vallone D. Noncombustible tobacco product advertising: how companies are selling the new face of tobacco. Nicotine Tob Res 2014 May 30;16(5):606-614. [doi: 10.1093/ntr/ntt200] [Medline: 24379146]
19. Hébert ET, Case KR, Kelder SH, Delk J, Perry CL, Harrell MB. Exposure and engagement with tobacco- and e-cigarette-related social media. J Adolesc Health 2017 Sep;61(3):371-377 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.04.003] [Medline: 28669801]
20. Huang J, Kornfield R, Szczypka G, Emery SL. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control 2014 Jul;23 Suppl 3:26-30 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051551] [Medline: 24935894]
21. Vogel E, Ramo D, Rubinstein M, Delucchi KL, Darrow SM, Costello C, et al. Effects of social media on adolescents' willingness and intention to use e-cigarettes: an experimental investigation. Nicotine Tob Res 2021 Mar 19;23(4):694-701 [FREE Full text] [doi: 10.1093/ntr/ntaa003] [Medline: 31912147]
22. Jackler R, Chau C, Getachew B. JUUL advertising over its first three years on the market. 2019. URL: https://tobacco-img.stanford.edu/wp-content/uploads/2021/07/21231836/JUUL_Marketing_Stanford.pdf [accessed 2022-07-06]
23. Mantey DS, Cooper MR, Clendennen SL, Pasch KE, Perry CL. E-cigarette marketing exposure is associated with e-cigarette use among US youth. J Adolesc Health 2016 Jun;58(6):686-690 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.03.003] [Medline: 27080732]
24. Helping creators turn their passion into a living. 2021 Nov 14. URL: https://business.instagram.com/blog/helping-creators-turn-their-passion-into-a-living/ [accessed 2022-05-27]
25. Tobacco and tobacco accessories. URL: https://business.twitter.com/en/help/ads-policies/ads-content-policies/tobacco-and-tobacco-accessories.html [accessed 2022-05-27]
26. FDA In Brief: FDA Requires four e-cigarette brands to provide critical information on social media practices. US Food and Drug Administration. 2021 Mar 17. URL: https://www.fda.gov/news-events/fda-brief/fda-brief-fda-requires-four-e-cigarette-brands-provide-critical-information-social-media-practices [accessed 2021-12-24]
27. Colditz JB, Chu K, Emery SL, Larkin CR, James AE, Welling J, et al. Toward real-time infoveillance of Twitter health messages. Am J Public Health 2018 Aug;108(8):1009-1014. [doi: 10.2105/AJPH.2018.304497] [Medline: 29927648]
28. Visweswaran S, Colditz JB, O'Halloran P, Han N, Taneja SB, Welling J, et al. Machine learning classifiers for Twitter surveillance of vaping: comparative machine learning study. J Med Internet Res 2020 Aug 12;22(8):e17478 [FREE Full text] [doi: 10.2196/17478] [Medline: 32784184]
29. Dobbs PD, Schisler E, Colditz JB, Primack BA. Miscommunication about the US federal Tobacco 21 law: a content analysis of Twitter discussions. Tob Control 2022 Feb 16:057099. [doi: 10.1136/tobaccocontrol-2021-057099] [Medline: 35173067]
30. Sidani JE, Colditz JB, Barrett EL, Shensa A, Chu K, James AE, et al. I wake up and hit the JUUL: analyzing Twitter for JUUL nicotine effects and dependence. Drug Alcohol Depend 2019 Nov;204:107500. [doi: 10.1016/J.DRUGALCDEP.2019.06.005]
31. Colditz JB, Welling J, Smith NA, James AE, Primack BA. World vaping day: contextualizing vaping culture in online social media using a mixed methods approach. J Mixed Methods Res 2017 Apr 09;13(2):196-215. [doi: 10.1177/1558689817702753]
32. Sidani JE, Colditz JB, Barrett EL, Chu K, James AE, Primack BA. JUUL on Twitter: analyzing tweets about use of a new nicotine delivery system. J Sch Health 2020 Feb 11;90(2):135-142 [FREE Full text] [doi: 10.1111/josh.12858] [Medline: 31828791]

XSL•FO
RenderX

33. Moltu C, Stefansen J, Svisdahl M, Veseth M. Negotiating the coresearcher mandate: service users' experiences of doing collaborative research on mental health. Disabil Rehabil 2012;34(19):1608-1616. [doi: 10.3109/09638288.2012.656792] [Medline: 22489612]
34. Azungah T. Qualitative research: deductive and inductive approaches to data analysis. Qual Res J 2018 Oct 31;18(4):383-400. [doi: 10.1108/qrj-d-18-00035]
35. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
36. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 2016 Jul 02;20(1):37-46. [doi: 10.1177/001316446002000104]
37. Nguyen D, Vu T, Tuan NA. BERTweet: a pre-trained language model for English Tweets. Proc 2020 Conf Empir Methods Natural Lang Proc 2020:9-14 [FREE Full text] [doi: 10.18653/v1/2020.emnlp-demos.2]
38. Baker W. Using Large Pre-Trained Language Models to Track Emotions of Cancer Patients on Twitter [Thesis]. Fayetteville: University of Arkansas; 2021.
39. Liu Y, Ott M, Goyal N. RoBERTa: a robustly optimized BERT pretraining approach. ArXiv Preprint posted online on July 26, 2019 [FREE Full text] [doi: 10.48550/arXiv.1907.11692]
40. Singh RK, Rani M, Bhagavathula AS, Sah R, Rodriguez-Morales AJ, Kalita H, et al. Prediction of the COVID-19 pandemic for the top 15 affected countries: advanced autoregressive integrated moving average (ARIMA) model. JMIR Public Health Surveill 2020 May 13;6(2):e19115 [FREE Full text] [doi: 10.2196/19115] [Medline: 32391801]
41. Chen S, Xu Q, Buchenberger J, Bagavathi A, Fair G, Shaikh S, et al. Dynamics of health agency response and public engagement in public health emergency: a case study of CDC tweeting patterns during the 2016 Zika epidemic. JMIR Public Health Surveill 2018 Nov 22;4(4):e10827 [FREE Full text] [doi: 10.2196/10827] [Medline: 30467106]
42. Chum A, Nielsen A, Bellows Z, Farrell E, Durette P, Banda JM, et al. Changes in public response associated with various COVID-19 restrictions in Ontario, Canada: observational infoveillance study using social media time series data. J Med Internet Res 2021 Aug 25;23(8):e28716 [FREE Full text] [doi: 10.2196/28716] [Medline: 34227996]
43. Box G, Jenkins G. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day; 1970.
44. Harvey AC. Forecasting, Structural Time Series Models and the Kalman Filter. Oxford: Cambridge University Press; 1989.
45. Jalles JT. Structural Time Series Models and the Kalman Filter: A Concise Review. 2009 Jun 19. URL: https://run.unl.pt/bitstream/10362/11569/1/wp541.pdf [accessed 2021-12-24]
46. 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 31;28(2):146-151 [FREE Full text] [doi: 10.1136/tobaccocontrol-2018-054382] [Medline: 29853561]
47. Selukar R. Structural analysis of time series using the SAS/ETS UCM procedure. SAS Institute Inc. 2009. URL: https://support.sas.com/resources/papers/proceedings09/306-2009.pdf [accessed 2022-07-06]
48. The UCM Procedure: SAS/ETS 13.2 User's Guide. SAS Institute Inc. 2018. URL: https://support.sas.com/documentation/onlinedoc/ets/151/autoreg.pdf [accessed 2021-12-23]
49. Chu K, Unger JB, Allem J, Pattarroyo M, Soto D, Cruz TB, et al. Diffusion of messages from an electronic cigarette brand to potential users through twitter. PLoS One 2015 Dec 18;10(12):e0145387 [FREE Full text] [doi: 10.1371/journal.pone.0145387] [Medline: 26684746]
50. Jo CL, Kornfield R, Kim Y, Emery S, Ribisl KM. Price-related promotions for tobacco products on Twitter. Tob Control 2016 Jul;25(4):476-479. [doi: 10.1136/tobaccocontrol-2015-052260] [Medline: 26130449]
51. Statement from FDA Commissioner Scott Gottlieb, M.D., on new steps to address epidemic of youth e-cigarette use. 2018 Sep 11. URL: https://www.fda.gov/news-events/press-announcements/statement-fda-commissioner-scott-gottlieb-md-new-steps-address-epidemic-youth-e-cigarette-use [accessed 2021-12-24]
52. FDA takes new steps to address epidemic of youth e-cigarette use, including a historic action against more than 1,300 retailers and 5 major manufacturers for their roles perpetuating youth access. 2018 Sep 11. URL: https://tinyurl.com/y8fnn7xp [accessed 2021-12-24]
53. Anderson M, Jiang J. Teens, social media & technology. Washington: Pew Internet and American Life Project; 2018 May 31. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2022-01-23]
54. Allem J, Ferrara E. Could social bots pose a threat to public health? Am J Public Health 2018 Aug;108(8):1005-1006. [doi: 10.2105/AJPH.2018.304512] [Medline: 29995482]
55. Allem J, Ferrara E, Uppu SP, Cruz TB, Unger JB. e-Cigarette surveillance with social media data: social bots, emerging topics, and trends. JMIR Public Health Surveill 2017 Dec 20;3(4):e98 [FREE Full text] [doi: 10.2196/publichealth.8641] [Medline: 29263018]
56. McCausland K, Maycock B, Leaver T, Wolf K, Freeman B, Jancey J. e-Cigarette advocates on Twitter: content analysis of vaping-related tweets. JMIR Public Health Surveill 2020 Oct 14;6(4):e17543 [FREE Full text] [doi: 10.2196/17543] [Medline: 33052130]
57. Dautzenberg B, Bricard D. Real-time characterization of e-cigarettes use: the 1 Million Puffs Study. J Addict Res Ther 2015;06(02):229. [doi: 10.4172/2155-6105.1000229]

XSL•FO
RenderX

Abbreviations

ACF: autocorrelation function
AR: autoregressive parameter
ARIMA: autoregressive integrated moving average
AUROC: area under the receiver operating characteristic
BERT: bidirectional encoder representations from transformers
CASAA: Consumer Advocates for Smoke-Free Alternatives Association
CNN: convolutional neural network
e-Cigarette: electronic cigarette
FDA: US Food and Drug Administration
LSTM: long short-term memory
MA: moving average
MAD: mean absolute deviation
MAPE: mean absolute percentage error
PACF: partial autocorrelation function
R^2: coefficient of determination
RITHM: real-time infoveillance of Twitter health messages
RMSE: root mean square error
UCM: unobservable components model

Edited by A Majmundar; submitted 19.02.22; peer-reviewed by H Dashtian, D Wu; comments to author 14.04.22; revised version received 04.06.22; accepted 17.06.22; published 22.07.22.

Please cite as:
Ezike NC, Ames Boykin A, Dobbs PD, Mai H, Primack BA
Exploring Factors That Predict Marketing of e-Cigarette Products on Twitter: Infodemiology Approach Using Time Series
JMIR Infodemiology 2022;2(2):e37412
URL: https://infodemiology.jmir.org/2022/2/e37412
doi:10.2196/37412
PMID:37113447

©Nnamdi C Ezike, Allison Ames Boykin, Page D Dobbs, Huy Mai, Brian A Primack. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 22.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Twitter Trends for Celiac Disease and the Gluten-Free Diet: Cross-sectional Descriptive Analysis

Monique Germone[1,2*], PhD; Casey D Wright[3*], PhD; Royce Kimmons[4], PhD; Shayna Skelley Coburn[5,6], PhD

[1]Digestive Health Institute, Colorado Center for Celiac Disease, Children's Hospital Colorado, Aurora, CO, United States

[2]Department of Psychiatry, University of Colorado, Anschutz Medical Campus, Aurora, CO, United States

[3]Department of Developmental Sciences, School of Dentistry, Marquette University, Milwaukee, WI, United States

[4]David O McKay School of Education, Instructional Psychology & Technology, Brigham Young University, Provo, UT, United States

[5]Department of Psychology and Behavioral Health, Children's National Hospital, Washington, DC, United States

[6]Department of Psychiatry and Behavioral Sciences, The George Washington School of Medicine and Health Sciences, Washington, DC, United States

[*]these authors contributed equally

Corresponding Author:
Monique Germone, PhD
Digestive Health Institute
Colorado Center for Celiac Disease
Children's Hospital Colorado
13123 East 16th Avenue
Aurora, CO, 80045
United States
Phone: 1 7207773711
Fax: 1 7207777311
Email: monique.germone@childrenscolorado.org

Abstract

Background: Few studies have systematically analyzed information regarding chronic medical conditions and available treatments on social media. Celiac disease (CD) is an exemplar of the need to investigate web-based educational sources. CD is an autoimmune condition wherein the ingestion of gluten causes intestinal damage and, if left untreated by a strict gluten-free diet (GFD), can result in significant nutritional deficiencies leading to cancer, bone disease, and death. Adherence to the GFD can be difficult owing to cost and negative stigma, including misinformation about what gluten is and who should avoid it. Given the significant impact that negative stigma and common misunderstandings have on the treatment of CD, this condition was chosen to systematically investigate the scope and nature of sources and information distributed through social media.

Objective: To address concerns related to educational social media sources, this study explored trends on the social media platform Twitter about CD and the GFD to identify primary influencers and the type of information disseminated by these influencers.

Methods: This cross-sectional study used data mining to collect tweets and users who used the hashtags #celiac and #glutenfree from an 8-month time frame. Tweets were then analyzed to describe who is disseminating information via this platform and the content, source, and frequency of such information.

Results: More content was posted for #glutenfree (1501.8 tweets per day) than for #celiac (69 tweets per day). A substantial proportion of the content was produced by a small percentage of contributors (ie, "Superuser"), who could be categorized as self-promotors (eg, bloggers, writers, authors; 13.9% of #glutenfree tweets and 22.7% of #celiac tweets), self-identified female family members (eg, mother; 4.3% of #glutenfree tweets and 8% of #celiac tweets), or commercial entities (eg, restaurants and bakeries). On the other hand, relatively few self-identified scientific, nonprofit, and medical provider users made substantial contributions on Twitter related to the GFD or CD (1% of #glutenfree tweets and 3.1% of #celiac tweets, respectively).

Conclusions: Most material on Twitter was provided by self-promoters, commercial entities, or self-identified female family members, which may not have been supported by current medical and scientific practices. Researchers and medical providers could potentially benefit from contributing more to this space to enhance the web-based resources for patients and families.

(JMIR Infodemiology 2022;2(2):e37924) doi:10.2196/37924

XSL•FO
RenderX

KEYWORDS

celiac disease; social media; Twitter; gluten-free; social networking site; diet; infodemiology; education; online; content; accuracy; credibility

Introduction

Chronic disease diagnoses often are coupled with a significant period of adjustment as patients learn how to manage and live with the condition. Having access to relevant and reliable information is important for educating and aiding new patients in disease management [1-3]. Over the past 16 years, many individuals with a chronic disease have been turning to Internet sources, such as social media, for education about their condition and treatment [4,5] despite a hesitancy from physicians and medical providers to use this resource for patient education [6]. Social media use among Americans has increased dramatically across adults of all genders, race, income, education level, and communities since the early 2000s [4]. The social media platform Twitter provides a unique data source whereby important questions can be asked and analyzed regarding how various participants are searching and sharing information, such as information related to patient education and disease management.

Having the technological ability to collect (ie, "mine") publicly available data on social media platforms such as Twitter provides an opportunity to systematically quantify and categorize information on such platforms into trends and useful information for interested parties (eg, patients with chronic diseases). One component of using these emerging methodologies to analyze social media information is through the use of "affinity spaces." Affinity spaces represent either physical or web-based gathering places (rather than geographic or identity-based communities) where people come together in a "common endeavor" to develop and share various types of knowledge, including individual, internal, and in-depth information [7].

The systematic application of common data mining techniques on social media platforms facilitates the analysis of disease management–related trends and information available to patients [5]. This is of relevance to those with celiac disease (or in British English, "coeliac disease"; CD). CD is a condition that requires extensive education around a dietary treatment steeped in stigma and myth [8]. CD is a chronic autoimmune condition wherein the ingestion of gluten results in an immune-mediated injury to the small intestine [9]. Damage to the small intestine leads to malabsorption of nutrients and can result in short- and long-term complications ranging from gastroenterological distress to cancer and even death [9]. It is estimated that CD affects approximately 1% of individuals worldwide [10]. To date, the only treatment is adherence to a strict gluten-free diet (GFD) [11]. CD is associated with heavy biopsychosocial demands and challenges following a CD diagnosis [12,13].

Prior work on broader internet-based sources for CD education is emerging and denotes concern for the information, and misinformation, that is presented by these sources [14-18]. Overall, information disseminated by the top websites found in web-based searches conducted by researchers are not entirely accurate, transparent, or reliable for interested consumers such as patients or providers, including dietitians [15,17,19]. Moreover, despite its potential to reach millions of viewers, the top videos on YouTube related to CD in 2019 lacked adequate or helpful information [14].

Given the high prevalence of CD and the heavy burden associated with managing CD and the GFD, many resources are available; nonetheless, it is difficult to identify credible educational information about the treatment for CD (a GFD). New methodologies from the field of computer science have emerged that allow for further exploration of patient education through not only the internet but also, more specifically, the social media space. The purpose of this study was to combine the fields of computer science and behavioral science to explore trends on Twitter as an educational source for patients with CD. This study conducted a preliminary evaluation of the scope and nature of information available on Twitter by (1) determining who the primary contributors are who lead the conversations about CD and GFD-related topics on Twitter, as well as (2) identifying what type of information (ie, content, source, and frequency) is being disseminated by these contributors.

Methods

Selecting an Internet Information Source

The social media platform Twitter allows for broader access to data than other social media platforms. Additionally, the nature of "tweets" (posts from Twitter users) and user profile descriptors is text-based versus image-based (such as content found on Instagram), which allows for more ready analysis of the data. Despite not being the most widely used platform, as is YouTube (81%) or Facebook (69%), Twitter is used by approximately a quarter (23%) of American adults and relatively equally among self-identified men and women and racial groups [5]. A 2021 survey of US adults demonstrated that young adults (18 to 29 years) are the predominant users of social media [5]. However, use by older adults (>65 years) has increased in recent years to 45% of older adults in 2021, which indicates that they use at least 1 social media site [5]. Given the ready availability of the data and wide use of users including individuals with CD, Twitter was chosen as the social media source for this study.

Defining Affinity Spaces

An increasingly common research practice has been to examine affinity spaces found on the popular social media platform Twitter through the use of hashtags (an author's use of the hash symbol followed by the subject of a message) as a way to categorize and group messages; eg, #celiac and #glutenfree) [20,21]. These hashtags are conceptualized as a type of affinity space to explicate how these organic web-based spaces are used by communities to communicate, share, and find information [20,21]. As an open platform with very few barriers to participation and 330 million monthly active users [22], Twitter encourages the organic development of affinity spaces around

XSL•FO
RenderX

topics and events via hashtagged keywords (eg, #celiac and #glutenfree).

The 2 topics most central to this study are "celiac" and "gluten-free." Information available on Twitter regarding these topics might exhibit different norms in terms of who participates in these affinity spaces and how (eg, someone might want information on a gluten-free diet for non–celiac-related reasons). Hence, the original tweets that were tagged by Twitter users with either the #celiac or #glutenfree hashtag were treated as 2 different affinity spaces rather than 1 collective affinity space. These affinity spaces were then analyzed individually and compared to each other. Moreover, recognizing that many other hashtags might be used synonymously with #glutenfree or #celiac, hashtags akin to either of these terms in their relative affinity spaces also were included (ie, #gluten-free, #glutenfreediet, and #gluten_free, with #glutenfree and including #celiacdisease, #celiacs, #celiacsdisease, #coeliac, and #coeliacdisease with #celiac).

Data Collection

This study collected Twitter user and tweet data using the public Twitter application programming interface. Researchers used custom PHP scripts to collect query results and store them to a MySQL relational database for cleaning and analysis. The data set used in this study consisted of tweets that included any of the hashtags listed above. Given the large number of such tweets, we limited the time frame of our study to 8 months (October 27, 2019, through June 8, 2020), allowing us to have sufficient data for analysis without being influenced too heavily by a single event (eg, the US Thanksgiving holiday season or the onset of the COVID-19 pandemic). Furthermore, because our main emphasis was to understand who was posting to these hashtags, "retweets" (a user who reposts a message created by a different user) were excluded to focus only on original posts and the users who generated them. This resulted in 334,907 and 15,602 original tweets containing #glutenfree and #celiac, respectively, including those from synonymous hashtags for analysis. Table 1 provides an overview of general user and Tweet metadata over the 8-month collection period.

Table 1. General user and tweet metadata over the 8-month data collection period.

	#glutenfree[a]			#celiac[b]		
Metric	Superuser	Contributor	Lurker	Superuser	Contributor	Lurker
User count, n	1718	16,947	145,246	44	394	3945
Overall tweets, %	25.5	25.2	49.3	28.7	35.2	36.1
Tweets per user, mean (SD)	49.8 (84.0)	5.0 (2.7)	1.1 (0.3)	101.7 (58.7)	13.9 (9.9)	1.4 (0.9)

[a]Tweets: n=334,907; tweets per user: mean 2.0, SD 10.0; users: n=163,911.

[b]Tweets: n=15,602; tweets per user: mean 3.6, SD 3.6; users: n=4383.

Data Analysis

As is standard in analyzing data gathered from Twitter to analyze affinity spaces [20,21], all tweet and author users' publicly available profile data (eg, Twitter handles and locations) were saved to a database. Descriptive statistics of tweet and author user objects were calculated to determine the method to use to classify users into user types for further analysis. Descriptive statistics revealed that users exhibited a highly positive skew in their posting activities. This behavior was expected given previous studies carried out on Twitter data [23]. Based on the positive skew, van Mierlo's [24] 90-9-1 Principle was selected to classify users in each affinity space into relative activity groups. Users were classified as follows: superusers (top 1% of users posting content), contributors (next 9% of users contributing content), or lurkers (the remaining 90% of users; see Table 1) [24]. Following the standard for affinity space analysis [20,21], basic language processing techniques were then used to (1) extract keywords from user biographies (eg, "doctor" or "blogger"), (2) identify co-occurring hashtags (eg, "#vegan" or "#recipe"), and (3) identify common domains that users linked to in their tweets (eg, celiac.com). A detailed description of these categories is provided below in the *Results* section.

Ethical Considerations

Ethics approval was obtained or determined to not be necessary by all author institutions owing to the public nature of the data.

Results

Aim 1: Examining the Primary Influencers on Twitter

User Activity Group: Superusers, Contributors, and Lurkers

Participation in each affinity space (ie, #glutenfree and #celiac) was evenly spread across the 3 groups, with superusers producing 25.5% of an overall 28.7% of posts containing #glutenfree and #celiac, contributors producing 25.2% of an overall 35.2% of posts, and lurkers producing 49.3% of an overall 36.1% posts. In other words, superusers (1% of users posting to the named affinity spaces) posted on average 10.0 times (#glutenfree) and 7.3 times (#celiac) more than contributors (the next 9% of users contributing), and contributors posted 4.5 times (#glutenfree) and 9.8 times (#celiac) more than lurkers (the other 90% of users posting to these spaces). Additionally, a comparison of raw tweet counts showed that Lurker behaviors were similar between the 2 hashtag groups but that #celiac superusers and contributors posted at least twice as often as their #glutenfree counterparts. #glutenfree represented more than 20 times the tweets as #celiac, but 40.3%

XSL•FO
RenderX

of tweets in #celiac were also cross-listed in the #glutenfree data set (Table 1).

Biographical Self-descriptors

To understand the professional backgrounds of Twitter users posting to these hashtags, each user's self-description was parsed out into a list of keywords [25] after removing stop words (eg, "a," "and," and "the"). Descriptions produced roughly 200,000 unique keywords (eg, "blogger" and "author"). The study team reviewed the most common 500 keywords for each hashtag and user activity group and then excluded those that did not suggest the author's expertise or were disassociated from the topic (eg, "director" and "vegan" were retained, while "music" and "www" were excluded). Descriptors related to family relationships were also retained (eg, "mother" was included), expecting that many family members of individuals with CD would participate in these affinity spaces to learn more about managing CD and the GFD. Specifically, tweets from users who self-identified with these keywords related to female family relationships (eg, mother or wife) represented 4.3% of tweets containing #glutenfree and 8% of those containing #celiac. Male family relationships (eg, father or husband) represented 1.5% of tweets containing #glutenfree and 1.2% of those containing #celiac.

Specific keywords that suggested an author's medical expertise (eg, "doctor," "physician," or "dietitian") or a terminal degree (eg, "MD" and "PhD") were also targeted [25]. Top results for each hashtag (#glutenfree and #celiac) and user category are provided in Tables 2 and 3; they indicated that "writer" (3.6% and 4.5%), "blogger" (1.4% and 2.4%), "author" (1.8% and 3.2%), and "advocate" (0.8% and 2.3%) were some of the most common self-descriptors. Targeted medical degrees and the term "doctor" were not widely used as self-descriptors by users and are provided in Tables 4 and 5, with "writers" and "bloggers" typically out-representing "PhDs" and "MDs" at a rate of 10-to-1 or more. The word stems "naturopath-" and "homeopath-" also accompanied many instances of "doctor" in both hashtags (5.5% and 14.0%, respectively). Overall, tweets from users who self-identified with keywords including "doctor," "dietitian," "physician," "PhD," or "MD" represented only 2.0% of tweets containing #glutenfree and 6.0% of those containing #celiac.

Recognizing that some users might identify terminal degrees and medical expertise in their name fields instead of their descriptions, a keyword search for variants of "Doctor," "Physician," "PhD," "MD," and "dietitian" on names was conducted. This showed that 0.4% of #glutenfree users and 2.1% of #celiac users self-identified with one of these terms in this way, but this calculation also included various distractors, such as multiple references to the television series "Doctor Who."

Table 2. Top 15 self-descriptive identifiers of user accounts posting to #glutenfree.

	Superuser (n=1718)		Contributor (n=16,947)		Lurker (n=145,246)	
Rank	Keyword	Posts, n	Keyword	Posts, n	Keyword	Posts, n
1	Blogger	109	Writer	718	Writer	4432
2	Vegan	103	Vegan	626	Lover	3883
3	Writer	65	Lover	543	Fan	3646
4	Author	64	Mom	511	Mom	3069
5	Mom	56	Blogger	458	Artist	2392
6	Lover	53	Author	434	Wife	2309
7	Creator	36	Fan	379	Author	2235
8	Chef	34	Artist	372	Enthusiast	2060
9	Foodie	32	Wife	365	Vegan	1744
10	Photographer	27	Enthusiast	238	Husband	1664
11	Fan	26	Chef	237	Student	1589
12	Wife	25	Mother	196	Blogger	1531
13	Owner	23	Photographer	188	Teacher	1465
14	Advocate	22	Owner	178	Father	1376
15	Coach	21	Coach	177	Mother	1358

XSL•FO
RenderX

Table 3. Top 15 self-descriptive identifiers of users posting to #celiac.

	Superuser (n=44)		Contributor (n=394)		Lurker (n=3945)	
Rank	Keyword	Posts, n	Keyword	Posts, n	Keyword	Posts, n
1	Blogger	5	Mom	27	Mom	156
2	Advocate	4	Blogger	23	Lover	150
3	Vegan	4	Advocate	21	Writer	149
4	Author	3	Writer	20	Wife	123
5	Mom	2	Wife	17	Author	101
6	Writer	2	Lover	17	Fan	95
7	Wife	2	Vegan	16	Mum	78
8	Mother	2	Author	16	Advocate	75
9	Chef	2	Dietitian	10	Blogger	69
10	Host	2	Mother	8	Husband	68
11	Dietitian	1	Editor	8	Vegan	67
12	Editor	1	Founder	8	Student	63
13	Mum	1	Physician	7	Mother	61
14	MD	1	Fan	6	Teacher	58
15	Teacher	1	Student	6	Dietitian	52

Table 4. Targeted medical degrees or terms that are self-descriptive identifiers of user accounts posting to #glutenfree.

Superuser (n=1718)		Contributor (n=16,947)		Lurker (n=145,246)	
Keyword	Posts, n	Keyword	Posts, n	Keyword	Posts, n
Dietitian	11	PhD	80	PhD	762
PhD	4	Dietitian	55	Doctor	347
MD	2	Doctor	45	Dietitian	187
Doctor	2	MD	22	MD	169
Physician	0	Physician	16	Physician	97

Table 5. Targeted medical degrees or terms that are self-descriptive identifiers of user accounts posting to #celiac.

Superuser (n=44)		Contributor (n=394)		Lurker (n=3945)	
Keyword	Posts, n	Keyword	Posts, n	Keyword	Posts, n
Dietitian	1	Dietitian	10	Dietitian	59
MD	1	Physician	7	PhD	46
Doctor	0	PhD	4	Doctor	29
PhD	0	MD	3	MD	14
Physician	0	Doctor	2	Physician	14

Aim 2: Examining the Type of Information Distributed on Twitter

Affinity Spaces: #glutenfree Versus #celiac

Comparing the 2 affinity spaces, #glutenfree was much more active, averaging 1501.8 (SD 223.2) tweets per day, while #celiac averaged 69.0 (SD 16.7) tweets per day. Users posting to #glutenfree represented 163,911 accounts, averaging 2.0 (SD 10.0) tweets per account for the time period, while users posting to #celiac represented 4383 accounts, averaging 3.6 (SD 12.4) tweets per account. At the user participation level, a noticeable overlap was found between affinity spaces, with 64.0% of #celiac posters also posting to #glutenfree in the time period (with 1.7% of #glutenfree users also posting to #celiac).

Co-occurring Hashtags

To better understand the nature of the tweets that were being posted in each affinity space, the use of co-occurring hashtags was analyzed for easy grouping. In other words, hashtags that

XSL•FO
RenderX

were used in tweets that did not have similar word stems to the targeted grouping hashtags (eg, #vegan was included in #glutenfree, while #gluten and #gf were ignored) were analyzed to identify groupings [26]. Percentages for each co-occurring hashtag were calculated by the likelihood that the hashtag would be used if any co-occurring hashtags existed at all (see Tables 6 and 7).

Tweets containing #celiac were highly represented in the #glutenfree data set, ranking at a similar level to mentions of paleo and keto diet hashtags, but overall results indicate that tweets containing #glutenfree focused heavily on a variety of other diets, including #vegan, #dairyfree, #plantbased, #keto, #paleo, #vegetarian, and #organic, suggesting that interest in GFDs was most commonly associated with a variety of weight loss and health regimens unrelated to CD (Tables 6 and 7). In the #celiac data set, gluten-related hashtags were dominant (with #glutenfree co-occurring in 50.5%-69% of tweets; see Tables 6 and 7), but other hashtags were more varied with some focusing on recipes (eg, #veganrecipes), others on symptoms (eg, #chronicpain), and other diseases (eg, #IBD and #IBS). These hashtags amounted to less than 1% of overall tweets.

Comparing the 2 affinity spaces, it appeared that #glutenfree was both more widely used but also more lifestyle based (eg, associated with other diet trends such as paleo or keto) than the #celiac space (see Tables 6 and 7).

Table 6. Top 15 co-occurring hashtags with #glutenfree.

	Superuser		Contributor		Lurker	
Rank	Hashtag	Posts, %	Hashtag	Posts, %	Hashtag	Posts, %
1	Vegan	20.9	Vegan	18.5	Vegan	13.2
2	Recipe	7.5	Dairyfree	6.4	Dairyfree	3.6
3	Dairyfree	7.4	Keto	2.6	Keto	1.8
4	Recipes	5.6	Celiac	2.5	Food	1.7
5	Food	5.3	Plantbased	2.5	Organic	1.7
6	Cooking	4.5	Recipe	2.3	Vegetarian	1.5
7	Keto	4.3	Paleo	2.3	Celiac	1.5
8	Lowcarb	4.0	Organic	2.2	Plantbased	1.5
9	Paleo	4.0	Vegetarian	2.1	Sugarfree	1.4
10	Celiac	3.5	Food	1.9	Baking	1.4
11	Delicious	3.3	Healthy	1.7	Healthy	1.3
12	Vegetarian	3.0	Lowcarb	1.6	Paleo	1.2
13	Cook	2.9	Coeliac	1.5	Recipe	1.1
14	Organic	2.6	Homemade	1.5	Homemade	1.0
15	Foodie	2.3	Sugarfree	1.5	Pizza	1.0

XSL•FO
RenderX

Table 7. Top 15 co-occurring hashtags with #celiac.

	Superuser		Contributor		Lurker	
Rank	Hashtag	Posts, %	Hashtag	Posts, %	Hashtag	Posts, %
1	GlutenFree	69.0	GlutenFree	66.3	GlutenFree	50.5
2	Gluten	14.5	Gluten	7.5	Gluten	9.0
3	Foodpics	12.7	Vegan	4.1	Vegan	4.4
4	Lovefood	12.3	Food	3.1	GF	2.5
5	Foodies	11.9	Foodie	3.0	Autoimmune	2.5
6	Vegan	10.4	GF	2.8	IBS	1.7
7	Freefrom	6.9	GlutenFreeLife	2.6	Covid19	1.7
8	Veganfood	6.9	Autoimmune	2.5	Disease	1.5
9	Health	6.8	Covid19	2.1	Dairyfree	1.5
10	Veganfriendly	6.8	Singluten	2.0	Health	1.5
11	Eggallergy	6.8	Dairyfree	1.9	Coronavirus	1.5
12	Veganrecipes	6.5	Colesbakeryandcafe	1.9	Foodallergy	1.4
13	Veganfoodlover	6.4	Freefrom	1.8	IBD	1.3
14	Eggfreefood	6.4	Foodallergies	1.8	Food	1.3
15	Chronicpain	6.3	Beer	1.7	Foodallergies	1.3

Shared Link Domains

To understand what resources users were sharing, the domains of unshortened links in tweets were analyzed. URL shorteners that were used as aliases rather than an actual direct link, and automated content providers were ignored (eg, bit.ly) [27]. Results for both affinity spaces revealed that links to social media and video sharing sites were common (eg, Instagram, Pinterest, and YouTube), and many blog, recipe, and other specialty sites were heavily linked to as well (see Table 8). Some of these domains were highly represented because many users were tweeting about them (eg, 1064 users tweeting YouTube videos in posts containing #glutenfree), but others were highly represented because a relatively small number of users were promoting a specific resource (eg, 1 user tweeting about foodgawker.com 136 times and promoting it to the #2 spot; Table 8).

Domains ending in ".com" (ie, commercial sites) were more prevalent (as opposed to nonprofit [.org] or government [.gov] domains). In fact, keyword searches for .com, .org, and .gov domains on the overall data set revealed that .com websites were linked to posts containing #glutenfree or #celiac 54.7 and 16.8 times more than .org sites and 1173.0 and 44.7 times more than .gov domains. This shows that the commercial influence seems to be much more apparent and disproportional to other influences in the #glutenfree space but that information in the #celiac space may also be heavily dominated by commercial interests.

XSL•FO
RenderX

Table 8. Most common linked domains.

#glutenfree			#celiac		
Domain	Tweets, n	Unique users, n	Domain	Tweets, n	Unique users, n
instagram.com	4446	2385	celiac.com	172	9
pinterest.com	2245	91	foodgawker.com	136	1
youtu.be	1924	1064	instagram.com	98	59
celiac.com	1454	18	wp.me	44	1
goo.gl	812	101	paper.li	34	10
simplygluten-free.com	632	42	youtu.be	30	24
untp.beer	534	474	gofundme.com	28	3
recipecialist.com	532	1	mygfguide.com	26	1
bloglovin.com	521	137	joshealthykitchen.com	22	1
amzn.to	475	118	theglutenfreeblogger.com	22	1
ntelikanis.com	465	1	glutenfreerespect.com	21	3
wp.me	437	24	ncbi.nlm.nih.gov	21	4
thisvivaciouslife.com	375	47	facebook.com	16	12
amazon.com	326	109	hamandeggerfiles.blogspot.com	16	2
sumo.ly	306	57	coeliac.org.uk	15	8
youtube.com	303	164	parenting.nytimes.com	12	11
facebook.com	255	192	drrobertpastore.com	12	1
mummytries.com	217	1	michellesglutenfreekitchen.wordpress.com	11	1
lifewaysvillage.com	206	3	medicalxpress.com	10	6
growingupgf.com	187	1	glutenfreepan.com	10	1

Discussion

Principal Findings

The purpose of this study was to combine methods from computer science and the behavioral sciences to begin to examine internet-based CD educational sources. As part of this initial investigation, this study describes information about CD and the GFD disseminated on the social media platform Twitter. With increasing use of social media as an educational resource and source of support for populations of individuals with chronic illness [28-30], it is crucial to understand the nature of information on platforms such as Twitter. Our findings emphasize the prominence of posts on both CD and the GFD, which appear to come from users focused on promotion of themselves (eg, identifying as vegan) or a business (eg, endorsing a restaurant) rather than from more traditional sources of information such as medical professionals or nonprofit organizations [19]. This supports previous findings regarding the hesitancy of medical providers to engage in social media as a form of medical education [6]. It also raises concerns about the quality of information individuals are receiving about CD and the GFD, as individuals with CD require the GFD for medical purposes [19]. This is likely not unique to CD as concerns have been raised in the field of food allergies [30]. We propose the need for a social media presence focused on providing high-quality, up-to-date, fact-checked information to users, particularly for those within the CD or other gluten-related diseases.

Clinical Implications

Based on our findings, there is an opportunity and arguably a demand for increased presence on social media and internet-based platforms among medical and nonprofit experts in CD to provide high-quality information to consumers. This has been executed among populations of individuals with other diseases, such as inflammatory bowel disease (IBD). For example, ImproveCareNow [31] is a community of clinicians, researchers, parents, and patients of children and youths with IBD. The main goal of this organization is to provide a platform to help this community learn about “more reliable, proactive IBD care” [31]. Their social media campaign involves accounts on various platforms, including a blog, Facebook, Twitter, and YouTube. The content posted on these platforms is monitored by the organization.

Guidelines have been developed by several organizations to help inform medical providers on social media best practices, including the Association for Healthcare Social Media [32]. The use of guidelines can best inform medical providers on the use of social media as a source of patient education. Other groups are working to develop competencies including advocacy and communication responsibilities that specialists in various

XSL•FO
RenderX

areas of health might develop in helping to educate certain patient populations [33].

Limitations and Future Directions

There are several additional considerations for this study in analyzing publicly available Twitter data. First, we collected our sample of data during a relatively narrow (8-month) time period, which may not account for natural variations across seasons and events (eg, holidays and major scientific or medical conferences). The activity and nature of posts may have changed as the COVID-19 pandemic has continued. Second, our analysis did not examine co-occurring words within individual user accounts. For instance, it is possible that one account may note being a "vegan," "blogger," and "mom." Future research could collect more detailed information about active members of social media to better understand "influencers" in this area. Furthermore, this study should be understood in light of the typical Twitter user. Twitter is used by about a quarter of American adults, both men and women of various racial groups, but we recognize that social media users may be younger and not necessarily representative of all ages and demographics [5]. Future work might examine the role of social media use in educating different subgroups of the population.

Additionally, we used established but relatively new methods of automated extraction and categorization of data rather than human coding, though we used human observation and judgment during the process of cleaning and synthesizing the data. This relied on algorithms based on anticipated data and did not allow for inductive reasoning by the human eye. Such an approach allowed the study team to rely on objective data rather than potential biases or a priori assumptions of individual experts [34]. Future studies may strengthen knowledge on this topic through expansion of data collection across a longer time span and further evaluation of the nature of users as well as the sentiments and accuracy of content within tweets.

Conclusions

To our knowledge, this was the first study evaluating Twitter data using the topics #celiac and #glutenfree. Given the popularity and broad use of social media, this is an important starting point for this research that generates several new hypotheses and research questions. Our findings emphasize the large volume of information communicated on social media. We suggest that platforms such as Twitter pose risks of spreading biased or inaccurate information to the public, particularly when the sources of information come from entities who may be influenced by commercial conflicts of interest.

Social media represents an immense opportunity to achieve open and clear dialogue between health care professionals and the public, which could be a major facilitator of future research and patient education about CD and the GFD.

Acknowledgments

We would like to thank all those who contributed tweets for our study, patients who have informed our research and practice, and the following funding support for coauthors of this study: the National Institute of Dental and Craniofacial Research (NIH-NIDCR F31 DE027859) for author CDW and the National Institute Of Diabetes And Digestive And Kidney Diseases of the National Institutes of Health (NIH-K23DK129826) for author SSC.

Conflicts of Interest

None declared.

References

1. Stenberg U, Haaland-Øverby M, Koricho AT, Trollvik A, Kristoffersen LR, Dybvig S, et al. How can we support children, adolescents and young adults in managing chronic health challenges? A scoping review on the effects of patient education interventions. Health Expect 2019 Oct;22(5):849-862 [FREE Full text] [doi: 10.1111/hex.12906] [Medline: 31131527]
2. Pinchera B, DelloIacono D, Lawless CA. Best practices for patient self-management: implications for nurse educators, patient educators, and program developers. J Contin Educ Nurs 2018 Sep 01;49(9):432-440. [doi: 10.3928/00220124-20180813-09] [Medline: 30148541]
3. Win KT, Hassan NM, Oinas-Kukkonen H, Probst Y. Online patient education for chronic disease management: consumer perspectives. J Med Syst 2016 Apr;40(4):88. [doi: 10.1007/s10916-016-0438-0] [Medline: 26846749]
4. Social Media Fact Sheet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2022-02-04]
5. Chirumamilla S, Gulati M. Patient education and engagement through social media. Curr Cardiol Rev 2021;17(2):137-143 [FREE Full text] [doi: 10.2174/1573403X15666191120115107] [Medline: 31752656]
6. Katz M, Nandi N. Social media and medical education in the context of the COVID-19 pandemic: scoping review. JMIR Med Educ 2021 Apr 12;7(2):e25892 [FREE Full text] [doi: 10.2196/25892] [Medline: 33755578]
7. Gee J. Semiotic social spaces and affinity spaces: From The Age of Mythology to today's schools. In: Barton D, Tusting K, editors. Beyond Communities of Practice: Language Power and Social Context (Learning in Doing: Social, Cognitive and Computational Perspectives). Cambridge: Cambridge University Press; 2005:214-232.
8. Pytell R. More than a diet: the struggles of celiac disease. Dialouges@RU. New Brunswick, NJ: Rutgers University; 2017. URL: https://dialogues.rutgers.edu/images/Journals_PDF/2017-18-dialogues-web_e6db3.pdf [accessed 2022-11-30]

XSL•FO
RenderX

9. Rubin JE, Crowe SE. Celiac disease. Ann Intern Med 2020 Jan 07;172(1):ITC1-ITC16 [FREE Full text] [doi: 10.7326/AITC202001070] [Medline: 31905394]
10. Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. Clin Gastroenterol Hepatol 2018 Jun;16(6):823-836.e2. [doi: 10.1016/j.cgh.2017.06.037] [Medline: 29551598]
11. Treatment for celiac disease. How do doctors treat celiac disease? National Institute of Diabetes and Digestive and Kidney Diseases. URL: https://www.niddk.nih.gov/health-information/digestive-diseases/celiac-disease/treatment [accessed 2022-01-07]
12. Berry N, Vaiphei K, Dhaka N, Sinha SK, Kochhar R. Quality of life in celiac disease and the effect of gluten-free diet. JGH Open 2018 Aug;2(4):124-128 [FREE Full text] [doi: 10.1002/jgh3.12056] [Medline: 30483576]
13. Theethira TG, Dennis M. Celiac disease and the gluten-free diet: consequences and recommendations for improvement. Dig Dis 2015;33(2):175-182. [doi: 10.1159/000369504] [Medline: 25925920]
14. Basch CH, Hillyer GC, Garcia P, Basch CE. Content of widely viewed YouTube videos about celiac disease. Public Health 2019 Feb;167:147-151. [doi: 10.1016/j.puhe.2018.11.004] [Medline: 30682698]
15. England CY, Nicholls AM. Advice available on the Internet for people with coeliac disease: an evaluation of the quality of websites. J Hum Nutr Diet 2004 Dec;17(6):547-559. [doi: 10.1111/j.1365-277X.2004.00561.x] [Medline: 15546433]
16. McNally SL, Donohue MC, Newton KP, Ogletree SP, Conner KK, Ingegneri SE, et al. Can consumers trust web-based information about celiac disease? Accuracy, comprehensiveness, transparency, and readability of information on the internet. Interact J Med Res 2012 Apr 04;1(1):e1 [FREE Full text] [doi: 10.2196/ijmr.2010] [Medline: 23611901]
17. Tangri V, Chande N. Quality of Internet-based information on gastrointestinal diseases. Can J Gastroenterol 2011 Feb;25(2):93-96 [FREE Full text] [doi: 10.1155/2011/345076] [Medline: 21321681]
18. van der Marel S, Duijvestein M, Hardwick JC, van den Brink GR, Veenendaal R, Hommes DW, et al. Quality of web-based information on inflammatory bowel diseases. Inflamm Bowel Dis 2009 Dec;15(12):1891-1896. [doi: 10.1002/ibd.20976] [Medline: 19462423]
19. Buseck A, Lebwohl B, Green PHR. Quality and content of online patient resources for celiac disease. Dig Dis Sci 2021 Jul;66(7):2209-2215. [doi: 10.1007/s10620-020-06537-3] [Medline: 32816213]
20. Greenhalgh SP, Koehler MJ. 28 Days later: Twitter hashtags as “Just in Time” teacher professional development. TechTrends 2016 Dec 19;61(3):273-281. [doi: 10.1007/s11528-016-0142-4]
21. Rosenberg JM, Greenhalgh SP, Koehler MJ, Hamilton ER, Akcaoglu M. An investigation of State Educational Twitter Hashtags (SETHs) as affinity spaces. E-Learning and Digital Media 2016 Oct 22;13(1-2):24-44. [doi: 10.1177/2042753016672351]
22. Lin Y. 10 Twitter statistics every marketer should know in 2022. Oberlo. 2022. URL: https://www.oberlo.com/blog/twitter-statistics [accessed 2022-11-21]
23. Veletsianos G, Kimmons R. Scholars in an increasingly open and digital world: how do education professors and students use Twitter? Internet High Educ 2016 Jul;30:1-10. [doi: 10.1016/j.iheduc.2016.02.002]
24. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014 Feb 04;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966] [Medline: 24496109]
25. Carpenter JP, Kimmons R, Short CR, Clements K, Staples ME. Teacher identity and crossing the professional-personal divide on twitter. Teach Teach Educ 2019 May;81:1-12. [doi: 10.1016/j.tate.2019.01.011]
26. Weng L, Menczer F. Topicality and impact in social media: diverse messages, focused messengers. PLoS One 2015;10(2):e0118410 [FREE Full text] [doi: 10.1371/journal.pone.0118410] [Medline: 25710685]
27. Kimmons R, Veletsianos G, Woodward S. Institutional uses of Twitter in U.S. higher education. Innov High Educ 2016 Aug 5;42(2):97-111. [doi: 10.1007/s10755-016-9375-6]
28. Hamshaw RJ, Barnett J, Lucas JS. Framing the debate and taking positions on food allergen legislation: the 100 chefs incident on social media. Health Risk Soc 2017 May 26;19(3-4):145-167. [doi: 10.1080/13698575.2017.1333088]
29. Hamshaw RJT, Barnett J, Lucas JS. Tweeting and eating: the effect of links and likes on food-hypersensitive consumers' perceptions of Tweets. Front Public Health 2018 Apr 23;6:118 [FREE Full text] [doi: 10.3389/fpubh.2018.00118] [Medline: 29740573]
30. Hamshaw RJT, Barnett J, Gavin J, Lucas JS. Perceptions of food hypersensitivity expertise on social media: qualitative study. Interact J Med Res 2019 Jun 28;8(2):e10812 [FREE Full text] [doi: 10.2196/10812] [Medline: 31254334]
31. ImproveCareNow. URL: https://www.improvecarenow.org/purpose-success [accessed 2022-01-07]
32. Association for Healthcare Social Media. URL: https://ahsm.org/ [accessed 2022-09-22]
33. Stellefson M, Paige SR, Chaney BH, Chaney JD. Evolving role of social media in health promotion: updated responsibilities for health education specialists. Int J Environ Res Public Health 2020 Feb 12;17(4):1153 [FREE Full text] [doi: 10.3390/ijerph17041153] [Medline: 32059561]
34. Kimmons R, Veletsianos G. Public internet data mining methods in instructional design, educational technology, and online learning research. TechTrends 2018 Jun 7;62(5):492-500. [doi: 10.1007/s11528-018-0307-4]

XSL•FO
RenderX

Abbreviations

CD: celiac disease
GFD: gluten-free diet
IBD: inflammatory bowel disease

Edited by T Mackey; submitted 11.03.22; peer-reviewed by R Bayoumi, R Hamshaw; comments to author 15.08.22; revised version received 04.11.22; accepted 15.11.22; published 05.12.22.

Please cite as:
Germone M, Wright CD, Kimmons R, Coburn SS
Twitter Trends for Celiac Disease and the Gluten-Free Diet: Cross-sectional Descriptive Analysis
JMIR Infodemiology 2022;2(2):e37924
URL: https://infodemiology.jmir.org/2022/2/e37924
doi:10.2196/37924
PMID:37113453

©Monique Germone, Casey D Wright, Royce Kimmons, Shayna Skelley Coburn. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 05.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implicit Incentives Among Reddit Users to Prioritize Attention Over Privacy and Reveal Their Faces When Discussing Direct-to-Consumer Genetic Test Results: Topic and Attention Analysis

Yongtai Liu[1], MS; Zhijun Yin[1,2], PhD; Zhiyu Wan[2], PhD; Chao Yan[2], PhD; Weiyi Xia[2], PhD; Congning Ni[1], ME; Ellen Wright Clayton[3,4,5], MD, JD; Yevgeniy Vorobeychik[6], PhD; Murat Kantarcioglu[7], PhD; Bradley A Malin[1,2,8], PhD

[1]Department of Computer Science, Vanderbilt University, Nashville, TN, United States

[2]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States

[3]School of Law, Vanderbilt University, Nashville, TN, United States

[4]Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN, United States

[5]Department of Health Policy, Vanderbilt University Medical Center, Nashville, TN, United States

[6]Department of Computer Science and Engineering, Washington University in St. Louis, St. Louis, MO, United States

[7]Department of Computer Science, University of Texas at Dallas, Richardson, TX, United States

[8]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, United States

Corresponding Author:
Bradley A Malin, PhD
Department of Biomedical Informatics
Vanderbilt University Medical Center
2525 West End Ave Room / Suite1030
Nashville, TN, 37203
United States
Phone: 1 615 343 9096
Email: b.malin@vumc.org

Abstract

Background: As direct-to-consumer genetic testing services have grown in popularity, the public has increasingly relied upon online forums to discuss and share their test results. Initially, users did so anonymously, but more recently, they have included face images when discussing their results. Various studies have shown that sharing images on social media tends to elicit more replies. However, users who do this forgo their privacy. When these images truthfully represent a user, they have the potential to disclose that user's identity.

Objective: This study investigates the face image sharing behavior of direct-to-consumer genetic testing users in an online environment to determine if there exists an association between face image sharing and the attention received from other users.

Methods: This study focused on r/23andme, a subreddit dedicated to discussing direct-to-consumer genetic testing results and their implications. We applied natural language processing to infer the themes associated with posts that included a face image. We applied a regression analysis to characterize the association between the attention that a post received, in terms of the number of comments, the karma score (defined as the number of upvotes minus the number of downvotes), and whether the post contained a face image.

Results: We collected over 15,000 posts from the r/23andme subreddit, published between 2012 and 2020. Face image posting began in late 2019 and grew rapidly, with over 800 individuals revealing their faces by early 2020. The topics in posts including a face were primarily about sharing, discussing ancestry composition, or sharing family reunion photos with relatives discovered via direct-to-consumer genetic testing. On average, posts including a face image received 60% (5/8) more comments and had karma scores 2.4 times higher than other posts.

Conclusions: Direct-to-consumer genetic testing consumers in the r/23andme subreddit are increasingly posting face images and testing reports on social platforms. The association between face image posting and a greater level of attention suggests that

XSL•FO
RenderX

people are forgoing their privacy in exchange for attention from others. To mitigate this risk, platform organizers and moderators could inform users about the risk of posting face images in a direct, explicit manner to make it clear that their privacy may be compromised if personal images are shared.

(JMIR Infodemiology 2022;2(2):e35702) doi:10.2196/35702

KEYWORDS

direct-to-consumer genetic testing; topic modeling; social media

Introduction

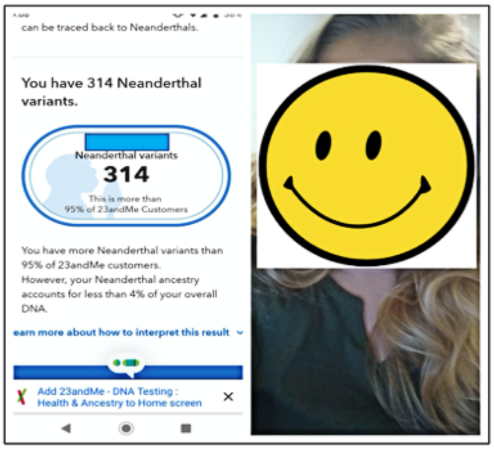
The cost of genome sequencing has steadily decreased over time [1], which, in turn, has enabled the emergence of direct-to-consumer genetic testing (DTC-GT) services available to the public [2]. DTC-GT allows consumers to learn about their genetic information without consulting with a health care provider [3]. The number of people who have participated in DTC-GT has increased dramatically, growing from 12 million in January 2018 to 26 million in January 2019 [4]. As of late 2021, the two largest DTC-GT companies, AncestryDNA and 23andme, had amassed over 20 million and 12 million clients, respectively [5]. Recent studies indicate that people pursue DTC-GT for various reasons, primarily to learn about their ancestry and to discover or confirm kinship [6,7].

As DTC-GT services have grown in popularity, consumers have increasingly relied upon online social platforms to discuss and share their test results (though not always the raw genome sequences) [8]. One particularly notable platform is Reddit, an online content rating and discussion site where users can create different subreddits based on specific topics of interest. One of the most popular subreddits related to DTC-GT is r/23andme, with more than 81,400 subscribers as of May 2022. In r/23andme, users discuss a wide range of topics related to genetic testing, including testing services, test results, explanations and interpretations, and share stories about what happened after undergoing testing (eg, health-related decisions) [8].

When r/23andme users share their results for discussion, instead of simply typing text, some users attach a screenshot of their DTC-GT result page (eg, the ancestry composition). Since Reddit is a virtual online community where users generally rely upon pseudonyms for communication, such screenshots of results typically do not contain a user's real name. Therefore, even when users share and discuss their DNA test results, this subreddit has historically been a community with a culture of anonymity.

However, in 2019, r/23andme users began attaching personal images to their posts. Figure 1 presents an example of a screenshot of a user's DTC-GT result page on the left, with the full-face image of this user on the right. This movement toward revealing one's face directly affects personal privacy [9,10]. Although these posts used pseudonyms, face image posting in online environments constitutes a knowing decision to give up one's privacy. Other users may utilize these face images to determine a user's identity, relying, in part, on the rapid development and deployment of modern face recognition [11] and identity detection systems [12]. This is a concern, because identity disclosure may lead to various negative consequences for individuals, including identity theft [13], discrimination [14], and threats to personal safety [15]. Since Reddit is a public platform, a user's posts and face images are readily accessible, making an identity disclosure attack feasible with little cost [16].

Figure 1. An example of a face image posted on the r/23andme subreddit. The report is shown together with a face image and testing results. The actual face and name are obscured for this publication; however, the data exist in the public domain.



Though users may be aware that revealing their face likely compromises their privacy, it is unclear why they choose to do so. Various investigations into behavioral psychology and economics show that some people waive their privacy rights in exchange for a service that they value [17]. Thus, we hypothesize that r/23andme users may receive more attention by publishing more personal information. This is supported by findings on other social platforms. For instance, including photos

XSL•FO
RenderX

with tweets on the Twitter platform can boost retweets by 35% [18]. Instagram photos with faces are 38% more likely to receive likes and 32% more likely to receive comments [19]. However, unlike Twitter or Instagram, the DTC-GT forum examined in this paper provides an anonymous environment for users to share and discuss sensitive personal genetic information. Thus, we sought to determine whether this forum supports the same privacy-service exchange hypothesis. To formally test our hypothesis, we investigated the following research questions: (1) What are the topics communicated in the natural language of posts with face images? (2) Is face image posting associated with the attention that a post receives?

To answer these questions, we collected posts from the r/23andme subreddit and categorized them into three types: (1) posts with only text, (2) posts with face images, and (3) posts with images not containing a face. We next measured the temporal posting trends regarding the type of post. Then, we applied topic modeling to compare the primary topics associated with types of post. Finally, we performed a regression analysis to infer the association between the attention that a post received, in terms of votes, comments, and whether the post contained a face image.

Methods

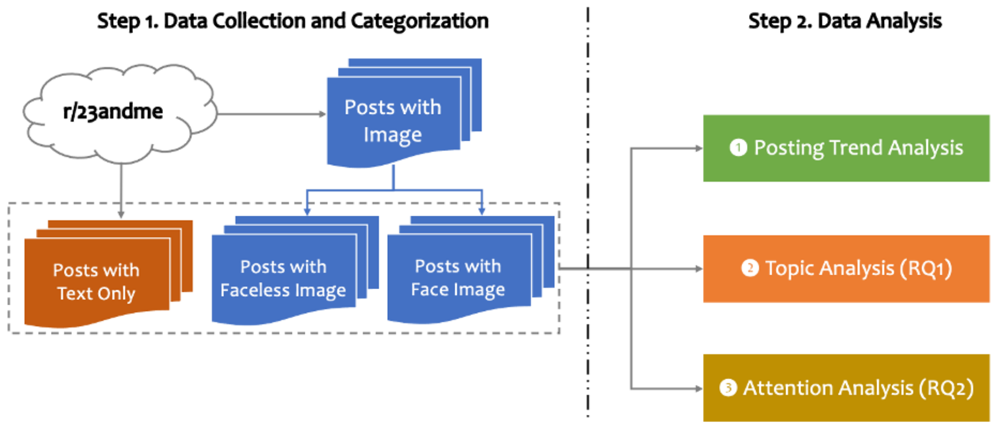
Ethics Considerations

This study involved only online posts that were openly accessible on Reddit. We have published the analysis results only in this paper, and any referenced posts or figures have been anonymized to protect the privacy of users.

Overview

Figure 2 provides an overview of the research pipeline, which had two primary steps. The first step involved data collection and categorization, in which we collected the posts on the r/23andme subreddit and extracted those with a face image using face recognition software. The second step focused on analysis. Specifically, we first conducted an exploratory analysis to investigate the temporal posting trends and then leveraged topic modeling to infer the themes communicated in these posts. Finally, we performed a regression analysis to determine whether including a face image in a post was associated with the attention it received. In this study, we characterized attention by the number of comments and the karma score that a post received from other online users. The karma score on Reddit is defined as the number of upvotes minus the number of downvotes, indicating the popularity of a post.

Figure 2. An overview of the research workflow for r/23andme post analysis. RQ: research question.



Data Collection and Categorization

To collect data from the r/23andme subreddit, we first gathered the IDs of all posts (ie, submissions) and comments using pushshift.io. We then applied the Python Reddit application programming interface wrapper package (version 6.3.1) to extract data from Reddit for each post ID. Specifically, we collected all posts and comments published on r/23andme between December 31, 2012, and January 31, 2020. Each collected post contained the following information: (1) author identifier, (2) post title, (3) post text body, (4) image URL (if there was an image in the post), (5) comments on the post, (6) post date, and (7) karma scores of the post and affiliated comments.

We downloaded the images from posts containing an image URL and applied the face-recognition Python package (version 1.3.0) [20] to classify images into (1) images with a face and (2) images without a face (ie, faceless images). To assess the accuracy of the face detection algorithm, we randomly selected 100 images from each group and manually examined the quality of classification. We found that 7 faceless images were classified as face images, indicating a false positive rate of 7% (7/100), while 2 face images were classified as faceless images, indicating a false negative rate of 2% (2/100). To achieve 100% precision, we manually reviewed all the images in the face group and relabeled the misclassified images. Due to a high true positive rate of 98% (98/100) and the large volume of the faceless images (3865), we did not perform a manual review step for the set of faceless images. As such, we categorized all of the collected posts into three types: (1) text-only posts; (2) posts with faceless images; and (3) posts with face images (such as the post in Figure 1), corresponding to 3 types of users.

Data Analysis

To describe face image posting behavior, we compared the face posts with the other two types of posts along three perspectives: (1) posting temporal trend, (2) post theme, and (3) the attention

XSL•FO
RenderX

that a post received from other users, in terms of the number of comments and karma score.

Topic Analysis

To examine the thematic differences between the three post types, we applied topic modeling [21] to the post title rather than the post body, because 41.1% (6404/15,596) of the posts had an empty text body. We first tokenized the data and removed all punctuation. Next, we lemmatized words into their base forms (eg, "walks" became "walk") using the nltk Python package (version 3.3). We also replaced personal pronouns, such as "we," "she," and "they," with the symbol "-PRON-," and replaced numbers with the word "datum." We then applied latent Dirichlet allocation (LDA) [22], as implemented in the gensim Python package (version 3.8.1), to extract topics. Since LDA is an unsupervised learning model, we calibrated the number of topics for the optimal model based on the coherence score, which measures the pairwise word semantic similarity in a topic. To do so, we ran LDA models with 2 to 20 topics (using a step size of 2) on the set of lemmatized words and selected the topic number that achieved the highest coherence score. Finally, to demonstrate the quality of topic modeling, we used *t*-distributed stochastic neighbor embedding [23] to cluster topics and displayed the results as a 2D representation (Figure S1 and Figure S2 in Multimedia Appendix 1).

Regression Analysis

We investigated two types of associations. First, we considered the association between an image post (with and without a face) and the attention it received. Second, we considered the association between a face post and the attention it received. Since the number of comments and the karma score are nonnegative count variables, we applied a negative binomial regression to infer the association [24].

Given that posts published earlier may be read by more readers and, thus, receive more comments and votes, we included the number of days a post had been published as a control variable. In addition, posts on different topics might receive different levels of attention. To reduce the effects of post topic, we incorporated the topic distribution of each post as an additional set of control variables. During model fitting, we dropped one topic (T_4, see below) to address collinearity.

Moreover, the activity level of users might affect the popularity of their posts. For example, posts from active users may receive more attention. To reduce the impact of user activity, we incorporated the number of posts and the number of comments of each user as an additional set of control variables. We utilized the implementation of negative binomial regression in the statsmodels Python package (version 0.11.1) to fit models for the karma score and the number of comments separately. We reported the features that achieved statistical significance at the *P*<.001 level.

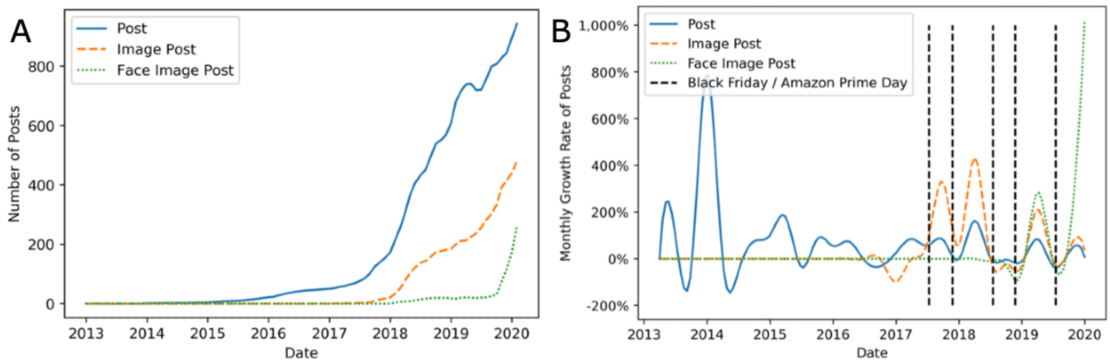
Results

We collected 15,596 posts and 188,843 comments, which were published by 20,883 users between December 31, 2012, and January 31, 2020. Among the collected posts, 24.8% (3818/15,596) contained faceless images, while 5.4% (849/15,596) contained face images.

Temporal Trends

In Figure 3A, the graph depicts the temporal post trend on a monthly basis. It can be seen that the r/23andme subreddit exhibited relatively low activity until 2017, after which the number of monthly posts grew rapidly. Image posts (with and without a face) became popular after 2018. In Figure 3B, the graph shows the quarterly growth rate of the number of posts. The green dotted line indicates that, since 2019, the number of face posts exhibited a rapid increase, with a growth rate that surpassed the growth rate of all posts (represented by the blue line) and image posts (represented by the orange dashed line). Notably, we find that posting rates for all 3 types of post increased rapidly after major promotions by 23andme (eg, as part of Black Friday and Amazon Prime Day, advertising events held by Amazon Inc), which is consistent with the findings of Yin et al [8].

Figure 3. Smoothed temporal trends of three types of post, including the number of posts published per month (A) and quarterly growth rate of posts (B).



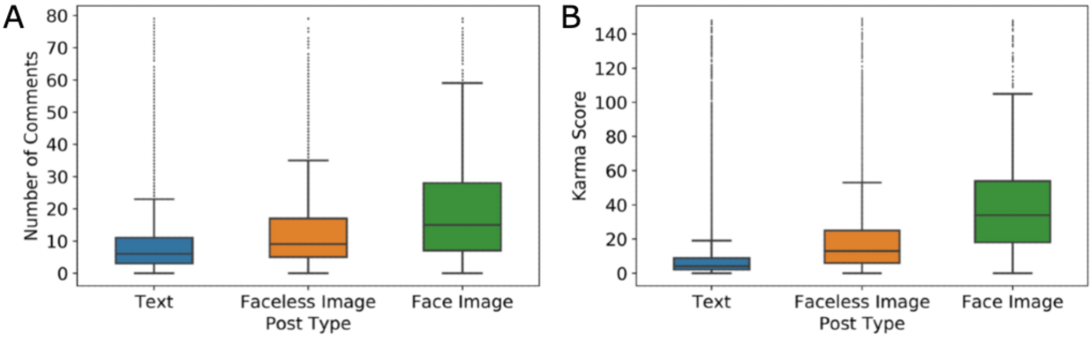
Attention to Posts

Figure 4A is a boxplot showing the number of comments per post for each post type. Face posts received the most comments, followed by posts not containing a face. The median number of comments for text-only posts was 6, but the median increased to 9 for posts with faceless images and 13 for posts with face images. Figure 4B is a boxplot showing the karma score by post type. Face posts received the highest median karma score (34), followed by posts with faceless images (median karma score

XSL•FO
RenderX

13). In contrast, the median karma score for text posts was only 4. One-way ANOVA tests for comments and karma scores indicated that the differences were statistically significant (*P*<.001).

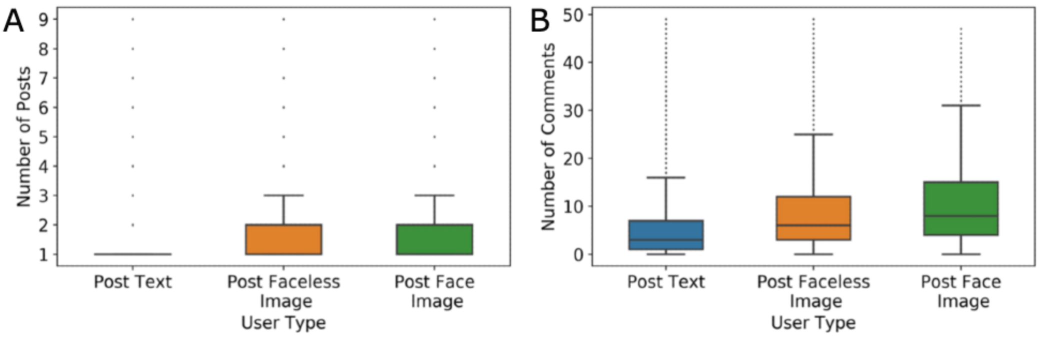
Figure 4. Attention to three types of posts. The number of comments per post (A) and karma score per post (B). For presentation purposes, we removed posts with more than 80 comments or karma scores greater than 150 (3% of the data). The entire data set is provided in Figure S3 and Figure S4 in Multimedia Appendix 1.



User Activity

We measured user activity in terms of the number of posts and comments. We found that 26.8% (2442/9114) of the users posted faceless images, while 8.5% (774/9114) posted face images. Figure 5A is a graph showing that the median number of posts for all 3 user types was 1. However, the third quartile of users who posted images (with or without a face) was 2. This suggests that, on average, authors who posted images (with or without a face) had more posts than authors who posted only text. The graph in Figure 5B depicts the number of comments posted for each user type. The users who posted face images wrote the most comments, with a median of 8. The median dropped to 6 for users who posted images not containing a face. For users who posted only text, the median number of comments was substantially lower, at 3. The results of 1-way ANOVA tests for the number of posts and the number of comments indicated that the differences were statistically significant (*P*<.001).

Figure 5. Number of posts per user (A) and number of comments per user (B) for users who posted (1) text only, (2) faceless images, and (3) face images. For presentation purposes, we removed users who published more than 10 posts or 50 comments, accounting for 4.4% of the total number of users. The entire data set is provided in Figure S3 and Figure S4 in Multimedia Appendix 1.



Topic Analysis

Table 1 shows the 10 inferred topics, their most relevant words, and the topic distribution (Figure S1 and Figure S2 in Multimedia Appendix 1 show details on the selection of the number of topics). The most relevant words were ranked based on their marginal distribution within a topic and displayed in descending order. The topic distribution was calculated as the percentage of posts belonging to the topic. Based on the relevant words and posts with the highest probability for each topic, we further grouped the 10 topics into three categories: (1) ancestry composition, (2) kinship and family discovery, and (3) general questions about genetic testing.

Ancestry composition included 4 topics: T_1, T_2, T_3, and T_4. Posts in this category focused on the presentation and discussion of ancestry composition testing results. The 4 topics captured ancestry information, which communicate a user's race, continental origin, and nationality. Textbox 1 shows example posts for each topic. Kinship finding and family discovery was communicated in T_5 and T_6. Specifically, T_5 communicated the discovery of ancestors and distinct relatives, where it can be seen that terms like "family" and "history" were often used. In T_6, words such as "find," "dad," and "siblings" show that this topic focused on findings relating to immediate family members. General questions related to DTC-GT were communicated in T_7, T_8, T_9, and T_{10}. Specifically, T_7 posts mainly asked about testing service progress. Words such as "time" and "wait" were highly weighted in this topic. T_8 posts were mainly comparisons of DTC-GT companies. There were mentions of companies, such as "MyHeritage," "23andme," and "WeGene." T_9 covered posts about understanding, or questions about, the test result report. T_{10} posts mainly discussed an upgrade to the genetic testing algorithm and the subsequent changes in testing results. Words such as "beta," "update," and "change" were highly weighted.

XSL•FO
RenderX

Figure 6 presents the topic distribution for each type of post. The 1-way ANOVA tests showed that there were statistically significant differences between the means of the 3 post types for all 10 topics (*P*<.001). Face posts were more likely to communicate ancestry composition (T_1, T_2, T_3, and T_4) and kinship and family discovery (T_5 and T_6), while text posts were more likely to be about general questions (T_7, T_8, and T_9). T_{10}, a topic about an algorithm upgrade by 23andMe, shows that faceless image posts were more likely to communicate this topic, followed by text posts and then face image posts. This may be because users tended to post screenshots of the results before and after the algorithm upgrade for easy comparison.

Table 1. The topics inferred from the r/23andme subreddit. The sample words are presented in descending order according to their relevance score within the topic.

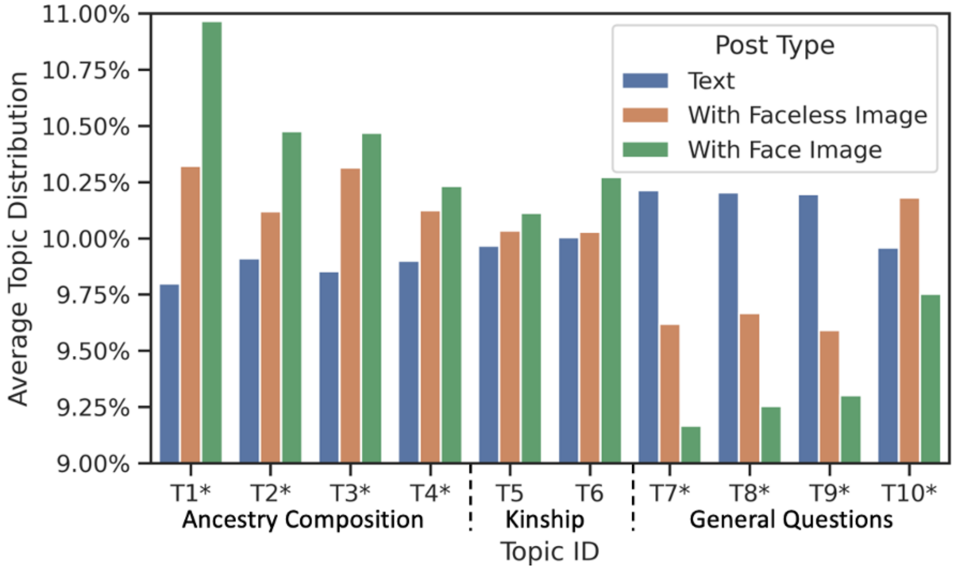
Category	Top-20 most relevant terms	Topic distribution
Ancestry composition		
Topic 1	European, -PRON-, result, Italian, Irish, British, surprise, Jewish, white, Chinese, broadly, bit, eastern, Ashkenazi, surprised, Scandinavian, give, eye, lot, surprising	11.6%
Topic 2	-PRON-, ancestry, German, guess, French, make, post, heritage, year, ethnicity, grandmother, common, grandparent, explain, mega-thread, feel, polish, Canadian, confused, wrong	7.9%
Topic 3	result, -PRON-, expect, finally, back, ancestor, interesting, pretty, AncestryDNA, bear, confidence, recent, location, Filipino, cool, guy, live, thought, Finnish, big	9.1%
Topic 4	American, Asian, African, native, Mexican, people, south, percentage, region, Neanderthal, gene, high, part, Spanish, unassigned, east, north, variant, trace, add	10.6%
Kinship and family discovery		
Topic 5	-PRON-, family, today, close, tree, understand, worth, info, don, trait, history, link, happen, picture, excited, love, list, connection, inherit, risk	6.5%
Topic 6	-PRON-, find, dad, half, mom, father, cousin, mother, side, sister, adopt, brother, great, sibling, grandfather, full, grandma, biological, aunt, figure	9.2%
General questions		
Topic 7	kit, long, time, extraction, wait, timeline, genetic, day, receive, sample, analysis, week, testing, step, send, batch, fail, information, work, stick	14.2%
Topic 8	andme, ancestry, datum, health, raw, accurate, GEDmatch, MyHeritage, good, DNA, upload, compare, site, comparison, land, data, service, difference, WeGene, interpret	11.0%
Topic 9	DNA, test, relative, question, parent, report, share, -PRON-, phase, show, generation, relate, computation, person, unexpected, noise, mystery, relationship, account, number	9.7%
Topic 10	result, update, beta, haplogroup, match, maternal, change, paternal, chromosome, map, mixed, chip, Puerto Rican, Korean, lose, comment, late, original, Romanian	10.2%

Textbox 1. Examples of posts for different topics.

- "So I'm a lot less British than I thought, and a lot more Swiss" (Topic 1).
- "Any guesses on my friend's ethnicity? He thinks he's French/German, English, and maybe some Slavic" (Topic 2).
- "Born and raised in Manila, grew up thinking I was 100% Filipino. A bit shocked at my results" (Topic 3).
- "Found out I am East Asian and Native American but I have northern Asian and Native American so high" (Topic 4).
- "Found out I have about a dozen cousins I didn't know about" (Topic 6).
- "My cousin did the DNA test and connected us to our great grandmother's family!" (Topic 5).
- "On my account apparently my mom and her twin sister are both my moms" (Topic 6).
- "Is my kit moving slow? It took 2 weeks to be marked as "arrived" after tracking showed it was delivered" (Topic 7).
- "23andMe vs WEGENE – uploaded 23andMe raw data to WEGENE and here are the differences" (Topic 8).
- "What is a likely relationship if the shared DNA is 1610 centimorgans across 80 segments?" (Topic 9).
- "Beta update v5.2 should now be available to all earlier chip (pre-V5) users, when opting into the Beta program" (Topic 10).

XSL•FO
RenderX

Figure 6. The prevalence of topics for each post type. The topics are arranged according to category. *P<.001 according to a 1-way ANOVA with post-hoc Tukey honestly significant difference tests for pairwise differences between the 3 post types for the topic.



Regression Analysis

Table 2 summarizes the results of the negative binomial regressions. *R* for image→comment and *R* for image→score indicate the association between the number of comments, karma score, and whether the post contained images, either faceless or with a face. Image posting exhibited statistically significant positive associations with both dependent variables, suggesting that image posts received more attention than text-only posts.

With respect to the *R* for face→comment and *R* for face→score tests, we selected 4717 image posts and assessed the association between the number of comments, karma score, and whether the image contained a face. Face image posting exhibited statistically significant positive associations with both dependent variables, which indicates that face posts received more attention than faceless posts. Comparing the *R* for image→comment and *R* for face→comment tests showed that posting a face image achieved a more positive impact on receiving comments. Comparing the *R* for image→score and *R* for face→score tests showed a similar result.

In addition, there were two notable findings with respect to the control variables. First, the log-transformed number of published days exhibited a negative association in the *R* for image→comment and *R* for image→score tests (β=–.09 for image→comment, β=–.26 for image→score, P<.001). Second, T_8 (the DTC-GT company comparison) had a negative association in all 4 tests (P<.001 for image→comment and face→comment, P=.003 for image→score, and P=.013 for face→score), while topic T_7 (asking about testing service progress) showed a negative association in *R* for image→score, *R* for face→score, and *R* for face→comment tests (P<.001 for image→score, P=.003 for face→score, and P=.04 for face→comment). The negative association between topics T_7, T_8, and face posting reinforce our previous finding that the topics in posts including a face were less likely to correspond to a general question about DTC-GT.

Table 2. Results of the regression analysis relating post type to comments and karma score. All associations were statistically significant (P<.001).

Negative binomial regression	Dependent variable	Independent variable	β	Z	SD	*P* value
R for image→comment	Number of comments	Posting image	.152	6.41	0.024	<.001
R for image→score	Karma score	Posting image	.618	12.35	0.050	<.001
R for face→comment	Number of comments	Posting face image	.451	10.21	0.044	<.001
R for face→score	Karma score	Posting face image	.760	9.64	0.079	<.001

Discussion

Principal Findings

This investigation made several notable findings. First, consistent with previous studies on other social platforms [18,19], we observed that posts with face images in the r/23andme subreddit received more attention than other posts. It is possible that the increase in attention drove the disclosure of personal information in this online environment. However, this is only a conjecture, as our investigation was not designed to be a causal analysis. Regardless of the motivation for face image posting, it is evident that this behavior has rapidly grown within this subreddit.

Second, the 10 inferred topics from the titles of r/23andme posts appeared to fall into three categories. Posts in the first category, which covered 4 out of 10 topics, focused on discussing users' ancestry composition. Notably, the topics in this category were associated with a higher rate of image and face image posting. It was further observed that users invoked their face images as proof (or counterexamples) of the genetic testing results. Posts

XSL•FO
RenderX

about kinship and family member discovery exhibited a moderate rate of face image sharing. When inspecting posts in this category, posts such as "finally find my half-sister," with a group photo of a reunion attached, were more prevalent than in other categories. Finally, posts asking general questions about genetic testing, which focused on comparisons between DTC-GT companies, the progress of testing result delivery, and upgrades to testing algorithms, exhibited the lowest rate of image sharing.

Third, counter to our expectation, we found that the number of days a post was published was negatively associated with a post's attention. One possible explanation for this result is that Reddit archives posts older than 6 months and no longer allows commenting on them. Thus, the number of comments and votes was limited for earlier posts. We further noticed that the topic related to general questions was negatively correlated with attention to a post.

Related Work

Natural language processing techniques have been applied to various health care applications [25]. Considering health care–related social media studies as an example, Liu et al [26] analyzed the association between weight loss progress and Reddit users' online interactions; Klein et al [27] relied upon Twitter data to identify potential cases of COVID-19 in the United States; and Ni et al [28] compared the attitudes of users of 4 different social platforms toward the "gene-edited babies" event. For DTC-GT, most investigations have focused on consumer motivations [29], health implications [30], and ethical implications [31], with only a handful considering the disclosure of test reports over social platforms [8,32,33]. Most previous studies that used social media data focused solely on mining knowledge from text. In this study, by taking image posting into consideration, we assess the behavior of personal image sharing on this DTC-GT forum.

This paper analyzes the association between face image sharing and attention paid to posts in an online setting; this setting may incentivize users to sacrifice their privacy in exchange for the benefit of a social response. This observation, however, does not imply that attention is undesirable in all cases, as several studies have shown that social engagement is beneficial to an individual's physical and mental health. For instance, in a large online breast cancer forum, Yin et al [34] found that the volume of online interchange was positively associated with patient treatment adherence. Pan et al [35] found that receiving replies could benefit online participants in depression forums. Naslund et al [36] analyzed the benefits and risks of using social media as a potentially viable platform for offering support intervention to persons with mental disorders. Thus, the perceived benefits an individual receives from a service typically outweigh the perceived privacy risks in the near term. Nevertheless, given that privacy concerns tend to be understood only later on [37], Reddit may wish to consider warning users about the potential negative consequences of their actions.

Limitations

Despite our findings, there are certain limitations to this work, which we believe serve as opportunities for future research. First, the face recognition package had an estimated 2% false negative rate, which means that approximately 76 of the 3865 face images (2%) were likely wrongly labeled as faceless images. These misclassified images might have influenced the accuracy of our findings, although not their overall direction. Second, most topics inferred from topic modeling were interpretable and intuitive, but topic T_{10} was difficult to interpret. As shown in Table 1, sample words for T_{10} conveyed different kinds of information: "Puerto Rican" and "Korean" are related to ancestry composition, whereas "late" and "lost" are evidence of asking about delivery progress. In this respect, newer topic modeling techniques [38-40] or language model–based topic modeling (eg, top2vec [41] and BERTopic [42]) may provide better insights into the semantics of posts on social platforms. Importantly, however, the quality of individual topics had little effect on our main conclusion, since the regression analysis (using the topic distribution as control variable; Table 2) and ANOVA test (without topic distribution; Figure 4) yielded the same finding—a statistically significant association between face image sharing on r/23andme and user engagement.

Conclusions

DTC-GT users are increasingly posting full-face images with their DTC-GT results on social platforms. In this study, we investigated the trend in this behavior in the r/23andme subreddit to obtain insight into potential underlying motivations. Our findings show that such behavior began in September 2019 and experienced rapid growth, with over 849 face-revealing posts by early 2020. Furthermore, our study suggests that posts including a face received, on average, 60% (5/8) more comments and 2.4 times higher karma scores than other posts. Posts that included face images were primarily about sharing and discussing ancestry composition and sharing family reunion photos with relatives discovered via DTC-GT. These findings verify our hypothesis that posting a personal image is associated with receiving more online attention, which is consistent with previous findings that people appear to be willing to give up their privacy (ie, their personal images) in exchange for a benefit (ie, attention from others). Based on this analysis, platform organizers and moderators might inform users about the risk of posting face images in a direct, explicit manner and make it clear that users' privacy may be compromised if personal images are disclosed.

Acknowledgments

YL, ZY, ZW, and CY proposed the research idea, which was finalized by BAM. YL and CN collected the data. YL and ZY designed and conducted the experiments. BAM and EWC provided advice on the data analysis. YL drafted the manuscript. EWC, ZY, BAM, YV, MK, and WX edited the final manuscript. All authors reviewed the final manuscript. This research was sponsored in part by the National Institutes of Health (grant RM1-HG009034, grant R01-HG006844, and grant U2COD023196).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 386 KB - infodemiology_v2i2e35702_app1.docx]

References

1. The Cost of Sequencing a Human Genome. National Human Genome Research Institute. URL: https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost [accessed 2022-07-02]
2. Bellcross C, Page P, Meaney-Delman D. Direct-to-consumer personal genome testing and cancer risk prediction. Cancer J 2012;18(4):293-302. [doi: 10.1097/PPO.0b013e3182610e38] [Medline: 22846729]
3. What is direct-to-consumer genetic testing? MedlinePlus. URL: https://ghr.nlm.nih.gov/primer/dtcgenetictesting/directtoconsumer [accessed 2022-07-02]
4. Regalado A. More than 26 million people have taken an at-home ancestry test. MIT Technology Review. URL: https://www.technologyreview.com/s/612880/more-than-26-million-people-have-taken-an-at-home-ancestry-test/ [accessed 2020-03-08]
5. McDermott M. 23andMe vs AncestryDNA. SmarterHobby. URL: https://www.smarterhobby.com/genealogy/23andme-vs-ancestry-dna/ [accessed 2021-11-27]
6. Ruhl GL, Hazel JW, Clayton EW, Malin BA. Public Attitudes Toward Direct to Consumer Genetic Testing. AMIA Annu Symp Proc 2019;2019:774-783 [FREE Full text] [Medline: 32308873]
7. Haeusermann T, Greshake B, Blasimme A, Irdam D, Richards M, Vayena E. Open sharing of genomic data: Who does it and why? PLoS One 2017 May 9;12(5):e0177158 [FREE Full text] [doi: 10.1371/journal.pone.0177158] [Medline: 28486511]
8. Yin Z, Song L, Clayton EW, Malin BA. Health and kinship matter: Learning about direct-to-consumer genetic testing user experiences via online discussions. PLoS One 2020 Sep 8;15(9):e0238644 [FREE Full text] [doi: 10.1371/journal.pone.0238644] [Medline: 32898148]
9. Xu H, Wang H, Stavrou A. Privacy risk assessment on online photos. 2008 Presented at: International Symposium on Recent Advances in Intrusion Detection; Nov 2, 2015; Berlin, Heidelberg p. 427-447. [doi: 10.1007/978-3-540-88313-5_3]
10. Venkatesaramani R, Malin BA, Vorobeychik Y. Re-identification of individuals in genomic datasets using public face images. Sci Adv 2021 Nov 19;7(47):eabg3296 [FREE Full text] [doi: 10.1126/sciadv.abg3296] [Medline: 34788101]
11. Masi I, Wu Y, Hassner T, Natarajan P. Deep Face Recognition: A Survey. 2018 Presented at: 31st SIBGRAPI Conference on Graphics, Patterns and Images; Oct 29-Nov 1, 2018; Parana, Brazil p. 471-478. [doi: 10.1109/sibgrapi.2018.00067]
12. Bäuml M, Tapaswi M, Stiefelhagen R. Semi-supervised Learning with Constraints for Person Identification in Multimedia Data. 2013 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; Jun 23-28, 2013; Portland, OR p. 3602-3609. [doi: 10.1109/cvpr.2013.462]
13. Irshad S, Soomro T. Identity theft and social media. Int J Comp Sci Net Sec 2018;18(1):43-55 [FREE Full text]
14. Acquisti A, Fong C. An Experiment in Hiring Discrimination via Online Social Networks. Manage Sci 2020 Mar;66(3):1005-1024. [doi: 10.1287/mnsc.2018.3269]
15. Nosko A, Wood E, Molema S. All about me: Disclosure in online social networking profiles: The case of FACEBOOK. Comput Hum Behav 2010 May;26(3):406-418. [doi: 10.1016/j.chb.2009.11.012]
16. Wan Z, Vorobeychik Y, Xia W, Clayton EW, Kantarcioglu M, Ganta R, et al. A game theoretic framework for analyzing re-identification risk. PLoS One 2015 Mar 25;10(3):e0120592 [FREE Full text] [doi: 10.1371/journal.pone.0120592] [Medline: 25807380]
17. Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. Science 2015 Jan 30;347(6221):509-514. [doi: 10.1126/science.aaa1465] [Medline: 25635091]
18. Maranga P. Social Photos Generate More Engagement: New Research Social Media Marketing. Social Media Examiner. URL: https://www.socialmediaexaminer.com/photos-generate-engagement-research/ [accessed 2020-03-25]
19. Bakhshi S, Shamma D, Gilbert E. Faces Engage Us: Photos with Faces Attract More Likes and Comments on Instagram. : Association for Computing Machiner; 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; Apr 26-May 1, 2014; Toronto, ON p. 965-974. [doi: 10.1145/2556288.2557403]
20. ageitgey/face_recognition: The world's simplest facial recognition api for Python and the command line. GitHub. URL: https://github.com/ageitgey/face_recognition [accessed 2020-03-19]
21. Mohr JW, Bogdanov P. Introduction—Topic models: What they are and why they matter. Poetics 2013 Dec;41(6):545-569. [doi: 10.1016/j.poetic.2013.10.001]
22. Blei D, Ng A, Jordan M. Latent Dirichlet Allocation. J Mach Learn Res 2003;3:1022 [FREE Full text] [doi: 10.5555/944919.944937]
23. van der Maaten L, Hinton G. Visualizing Data using t-SNE. J Mach Learn Res 2008;9:2579-2605 [FREE Full text]

XSL•FO
RenderX

24. Ver Hoef JM, Boveng PL. Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? Ecology 2007 Nov;88(11):2766-2772. [doi: 10.1890/07-0043.1] [Medline: 18051645]
25. Elbattah M, Arnaud E, Gignon M, Dequen G. The Role of Text Analytics in Healthcare: A Review of Recent Developments and Applications. 2021 Presented at: BIOSTEC 2021: 14th International Joint Conference on Biomedical Engineering Systems and Technologies; Feb 11-13, 2021; Vienna, Austria p. 825-832. [doi: 10.5220/0010414508250832]
26. Liu Y, Yin Z. Understanding Weight Loss via Online Discussions: Content Analysis of Reddit Posts Using Topic Modeling and Word Clustering Techniques. J Med Internet Res 2020 Jun 08;22(6):e13745 [FREE Full text] [doi: 10.2196/13745] [Medline: 32510460]
27. Klein AZ, Magge A, O'Connor K, Flores Amaro JI, Weissenbacher D, Gonzalez Hernandez G. Toward Using Twitter for Tracking COVID-19: A Natural Language Processing Pipeline and Exploratory Data Set. J Med Internet Res 2021 Jan 22;23(1):e25314 [FREE Full text] [doi: 10.2196/25314] [Medline: 33449904]
28. Ni C, Wan Z, Yan C, Liu Y, Clayton EW, Malin B, et al. The Public Perception of the #GeneEditedBabies Event Across Multiple Social Media Platforms: Observational Study. J Med Internet Res 2022 Mar 11;24(3):e31687 [FREE Full text] [doi: 10.2196/31687] [Medline: 35275077]
29. Roberts JS, Gornick MC, Carere DA, Uhlmann WR, Ruffin MT, Green RC. Direct-to-Consumer Genetic Testing: User Motivations, Decision Making, and Perceived Utility of Results. Public Health Genomics 2017;20(1):36-45 [FREE Full text] [doi: 10.1159/000455006] [Medline: 28068660]
30. European Society of Human Genetics. Statement of the ESHG on direct-to-consumer genetic testing for health-related purposes. Eur J Hum Genet 2010 Dec 25;18(12):1271-1273 [FREE Full text] [doi: 10.1038/ejhg.2010.129] [Medline: 20736974]
31. Clayton EW, Halverson CM, Sathe NA, Malin BA. A systematic literature review of individuals' perspectives on privacy and genetic information in the United States. PLoS One 2018 Oct 31;13(10):e0204417 [FREE Full text] [doi: 10.1371/journal.pone.0204417] [Medline: 30379944]
32. Olejnik L, Agnieszka K, Castelluccia C. I'm 2.8% Neanderthal - The Beginning of Genetic Exhibitionism? 2014 Presented at: Workshop on Genome Privacy; Jul 15, 2014; Amsterdam, Netherlands URL: https://hal.inria.fr/hal-01087696
33. Mittos A, Blackburn J, De Cristofaro E. Analyzing Twitter Discourse On Genetic Testing. arXiv Preprint posted online on April 20, 2018 [FREE Full text] [doi: 10.48550/arXiv.1801.09946]
34. Yin Z, Song L, Malin B. Reciprocity and its association with treatment adherence in an online breast cancer forum. 2017 Presented at: 30th IEEE International Symposium on Computer-Based Medical Systems (CBMS); Jun 22-24, 2017; Thessaloniki, Greece p. 618-623. [doi: 10.1109/cbms.2017.51]
35. Pan W, Feng B, Shen C. Examining Social Capital, Social Support, and Language Use in an Online Depression Forum: Social Network and Content Analysis. J Med Internet Res 2020 Jun 24;22(6):e17365 [FREE Full text] [doi: 10.2196/17365] [Medline: 32579125]
36. Naslund JA, Bondre A, Torous J, Aschbrenner KA. Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice. J Technol Behav Sci 2020 Sep 20;5(3):245-257 [FREE Full text] [doi: 10.1007/s41347-020-00134-x] [Medline: 33415185]
37. Acquisti A. Privacy in electronic commerce and the economics of immediate gratification. : Association for Computing Machinery; 2004 Presented at: 5th ACM conference on electronic commerce; May 17, 2004; New York, NY p. 21-29. [doi: 10.1145/988772.988777]
38. Lindstedt NC. Structural Topic Modeling For Social Scientists: A Brief Case Study with Social Movement Studies Literature, 2005–2017. Soc Curr 2019 May 02;6(4):307-318. [doi: 10.1177/2329496519846505]
39. Venkatesaramani R, Downey D, Malin B, Vorobeychik Y. A Semantic Cover Approach for Topic Modeling. In: Proceedings of the 8th Joint Conference on Lexical and Computational Semantics. 2019 Presented at: 8th Joint Conference on Lexical and Computational Semantics; Jun 6-7, 2019; Minneapolis, MN p. 92-102. [doi: 10.18653/v1/s19-1011]
40. Zhan Y, Liu R, Li Q, Leischow SJ, Zeng DD. Identifying Topics for E-Cigarette User-Generated Contents: A Case Study From Multiple Social Media Platforms. J Med Internet Res 2017 Jan 20;19(1):e24 [FREE Full text] [doi: 10.2196/jmir.5780] [Medline: 28108428]
41. Angelov D. Top2vec: Distributed representations of topics. arXiv Preprint posted online on August 19, 2020 [FREE Full text] [doi: 10.48550/arXiv.2008.09470]
42. Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv Preprint posted online on March 11, 2022 [FREE Full text] [doi: 10.48550/arXiv.2203.05794]

Abbreviations

DTC-GT: direct-to-consumer genetic testing
NLP: natural language processing
LDA: latent Dirichlet allocation

XSL•FO
RenderX

Edited by M Meacham; submitted 14.12.21; peer-reviewed by R Thompson, M Elbattah, W Ceron; comments to author 29.03.22; revised version received 24.05.22; accepted 22.06.22; published 03.08.22.

Please cite as:
Liu Y, Yin Z, Wan Z, Yan C, Xia W, Ni C, Clayton EW, Vorobeychik Y, Kantarcioglu M, Malin BA
Implicit Incentives Among Reddit Users to Prioritize Attention Over Privacy and Reveal Their Faces When Discussing Direct-to-Consumer Genetic Test Results: Topic and Attention Analysis
JMIR Infodemiology 2022;2(2):e35702
URL: https://infodemiology.jmir.org/2022/2/e35702
doi:10.2196/35702
PMID:37113452

©Yongtai Liu, Zhijun Yin, Zhiyu Wan, Chao Yan, Weiyi Xia, Congning Ni, Ellen Wright Clayton, Yevgeniy Vorobeychik, Murat Kantarcioglu, Bradley A Malin. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 03.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Physical Distancing and Social Media Use in Emerging Adults and Adults During the COVID-19 Pandemic: Large-scale Cross-sectional and Longitudinal Survey Study

Thabo van Woudenberg[1], PhD; Moniek Buijzen[1,2], PhD; Roy Hendrikx[3], PhD; Julia van Weert[4], PhD; Bas van den Putte[4], PhD; Floor Kroese[5,6], PhD; Martine Bouman[7], PhD; Marijn de Bruin[6,8], PhD; Mattijs Lambooij[3], PhD

[1]Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam, Rotterdam, Netherlands
[2]Behavioural Science Institute, Radboud University, Nijmegen, Netherlands
[3]Centre for Nutrition, Prevention and Health Services, National Institute for Public Health and the Environment, Bilthoven, Netherlands
[4]Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, Netherlands
[5]Social, Health and Organisational Psychology, University of Utrecht, Utrecht, Netherlands
[6]Corona Behavioural Unit, National Institute for Public Health and the Environment, Bilthoven, Netherlands
[7]Erasmus School of History, Culture and Communication, Erasmus University Rotterdam, Rotterdam, Netherlands
[8]Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands

Corresponding Author:
Thabo van Woudenberg, PhD
Erasmus School of Social and Behavioural Sciences
Erasmus University Rotterdam
Burgemeester Oudlaan 50
Rotterdam, 3062 PA
Netherlands
Phone: 31 10 408 2135
Email: vanwoudenberg@eur.nl

Abstract

Background: Although emerging adults play a role in the spread of COVID-19, they are less likely to develop severe symptoms after infection. Emerging adults' relatively high use of social media as a source of information raises concerns regarding COVID-19–related behavioral compliance (ie, physical distancing) in this age group.

Objective: This study aimed to investigate physical distancing among emerging adults in comparison with adults and examine the role of using social media for COVID-19 news and information in this regard. In addition, this study explored the relationship between physical distancing and using different social media platforms and sources.

Methods: The secondary data of a large-scale longitudinal national survey (N=123,848) between April and November 2020 were used. Participants indicated, ranging from 1 to 8 waves, how often they were successful in keeping a 1.5-m distance on a 7-point Likert scale. Participants aged between 18 and 24 years were considered emerging adults, and those aged >24 years were considered adults. In addition, a dummy variable was created to indicate per wave whether participants used social media for COVID-19 news and information. A subset of participants received follow-up questions to determine which platforms they used and what sources of news and information they had seen on social media. All preregistered hypotheses were tested with linear mixed-effects models and random intercept cross-lagged panel models.

Results: Emerging adults reported fewer physical distancing behaviors than adults ($\beta=-.08$, $t_{86,213.83}=-26.79$; $P<.001$). Moreover, emerging adults were more likely to use social media for COVID-19 news and information (b=2.48; odds ratio 11.93 [95% CI=9.72-14.65]; SE 0.11; Wald=23.66; $P<.001$), which mediated the association with physical distancing but only to a small extent (indirect effect: b=−0.03, 95% CI −0.04 to −0.02). Contrary to our hypothesis, the longitudinal random intercept cross-lagged panel model showed no evidence that physical distancing was not influenced by social media use in the previous wave. However, evidence indicated that social media use affects subsequent physical distancing behavior. Moreover, additional analyses showed that the use of most social media platforms (ie, YouTube, Facebook, and Instagram) and interpersonal communication were negatively associated with physical distancing, whereas other platforms (ie, LinkedIn and Twitter) and government messages had no or small positive associations with physical distancing.

XSL•FO
RenderX

Conclusions: In conclusion, we should be vigilant with regard to the physical distancing of emerging adults, but the study results did not indicate concerns regarding the role of social media for COVID-19 news and information. However, as the use of some social media platforms and sources showed negative associations with physical distancing, future studies should more carefully examine these factors to better understand the associations between social media use for news and information and behavioral interventions in times of crisis.

(JMIR Infodemiology 2022;2(2):e33713) doi:10.2196/33713

KEYWORDS

COVID-19; physical distancing; compliance; emerging adults; social media

Introduction

Background

In 2022, the COVID-19 pandemic is still ongoing in large parts of the world and, as of February, has been responsible for >381 million confirmed cases and >5.69 million deaths worldwide [1]. Given that most of the world's population has not been vaccinated yet, alternative precautionary measures are still essential to contain the spread of the COVID-19 infection. Therefore, many countries have adopted behavioral interventions, of which physical distancing is one of the most widely adopted, persistent, pragmatic, and effective policies [2]. However, the effectiveness of such strategies depends heavily on the compliance of the population with desired behaviors [3]. It is therefore important to study and understand compliance with governmental behavioral interventions, such as physical distancing during the COVID-19 pandemic to design future interventions most effectively.

In times of crisis, such as these, people tend to rely heavily on media to understand the situation and make informed decisions about related behavioral guidelines [4,5]. According to cultivation theory [6], content across the entire media landscape breeds a widespread meaning among the audience. The theory proposes that the more media-provided information the people consume, the greater the likelihood that their perceptions of reality align with that depicted in the media landscape. This cultivation process is driven by both mainstreaming and resonance, that is, different opinions and world viewpoints will move to align their opinions with the mediated content, and simultaneously, the mediated content becomes more relatable and relevant to media consumers. This means that people's perceptions and intentions will ultimately become similar to what is portrayed in the media landscape [6]. Therefore, the more the media emphasizes on the severity of the situation and the importance of physical distancing, the more likely it is that people will change their behavior.

Moreover, social cognitive theory [7] explains how a single media message can affect the behavior of people. This theory explains that people create cognitive schemas based on first-person experiences and observational learning. A large part of observational learning occurs through media exposure [8], meaning that people see and learn from others about COVID-19 and the counteractive measures via media exposure and adjust their perceptions and behavior accordingly [9]. This means that people learn how to behave from others portrayed in the media during the COVID-19 pandemic.

However, the nature and content of media messages in social media and mass media have different effects on people's perceptions and behavioral intentions during a crisis [10]. Although cultivation processes occur in both traditional media and social media [11], the nature and content of the message differ between the 2 forms of media. The immediacy of social media and the direct access to an unprecedented amount of content allows for less controlled and more fragmented view of the crisis [12-14]. Therefore, the process of mainstreaming and resonance is less likely to occur, and the importance of physical distancing will be less cultivated among social media users for COVID-19 news and information.

In addition, social media depicts more ambivalent messages about COVID-19 than traditional mass media does, contains more rumors or questionable information [12,15], is more subjective to algorithms that mediate and facilitate content promotion [16], and is more likely to only reach and circulate in subgroups of users in so-called echo chambers [17]. In general, people who use social media to inform themselves about COVID-19 will observe a broader range of ideas and behaviors on the web than those who only use traditional media. Therefore, it is less clear what normative behavior is during a crisis, and people are less likely to change their behaviors to comply with governmental behavioral interventions. This difference in behavioral change between social media users and nonusers has been observed in previous crises. For example, research on news consumption after the Great East Japan Earthquake in 2011 has demonstrated that mass media has a positive effect on people's perceptions of a crisis and the subsequent increased behavior change (ie, boosting civic communications, taking altruistic actions, and preparing for future crises). Social media showed only limited or no change in perceptions and behavioral intentions [10]. For the current crisis, this would mean that people who use social media to inform themselves about the crisis are less likely to change their behavior and, therefore, less likely to be physically distant from others.

This difference in news and information consumption and associated compliance with behavioral regulations is problematic when particular subgroups of the population rely more heavily on social media for news and information on COVID-19. In particular, young people differ substantially in news consumption compared with older generations. They are more attracted to social media as a source of news and information [18-20]; therefore, it seems likely that younger people also consume relatively more COVID-19 news and information via social media than adults do [21-23]. As a result, the subgroup

XSL•FO
RenderX

of young people, on average, would be less likely than adults to change their behavior and comply with behavioral regulations. Aside from a lower health risk when exposed to the coronavirus [24-26] and a stronger need to socialize with others [27] (A Orben, unpublished data, August 2020), this difference in the consumption of COVID-19 news and information might be important in understanding compliance of young people with behavioral regulations. That is, using social media for COVID-19 news and information might explain why young people are less often maintaining a physical distance from others.

A review of studies on protective behaviors during several pandemics before the COVID-19 crisis showed that older people have a higher chance of adopting relevant protective behaviors [28]. Contemporary research on COVID-19 corroborated this finding and showed that younger people engage less often in protective behaviors, such as physical distancing, than do older people. For example, a cross-sectional survey in the United States showed that adherence to distancing behaviors of young people aged between 18 and 24 years was considerably less than that of adults [29]. Similarly, other studies showed a linear increase in age with a range of protective behaviors, including physical distancing [30,31].

In this study, we are particularly interested in young people aged 18 years to their late 20s, termed *emerging adults* [32]. As these emerging adults grow as autonomous adults, they become more independent media consumers and less influenced by their parents. This is in contrast to the vast majority of children and adolescents who live with their parents and the associated influence of living with their parents on their media use [33,34]. A better understanding of the role that social media plays in compliance with behavioral interventions in emerging adults is valuable knowledge for governments, as this will help them better communicate behavioral regulations to all its citizens, boost the effectiveness of comparable behavioral interventions, and ultimately save lives.

This Study

This study investigated the differences between emerging adults and adults in terms of their physical distancing behavior while considering the role of using social media for COVID-19 news and information. On the basis of the theoretical framework and related empirical findings, we preregistered the following hypotheses: physical distancing is lower in emerging adults than in adults (H1), and the effect of age on physical distancing is mediated by the use of social media for COVID-19 news and information (H2). More specifically, we anticipated that age would negatively predict using social media for COVID-19 news and information (H2a) and using social media for COVID-19 news and information would negatively predicts physical distancing (H2b). This study further investigated the directionality of the association between physical distancing and social media use in a longitudinal sample.

In addition, to gain more insight into specific social media use, we performed exploratory research on a subsample of participants who were presented with an additional module of the questionnaire. These questions examined the use of different social media platforms and sources of messages consumed on social media. Specifically, these nonpreregistered analyses examined the association of physical distancing with (1) the most often used social media platforms (ie, Facebook, Twitter, Instagram, YouTube, and LinkedIn) and (2) the sources presented on the platform (ie, government, national news, regional news, personal communication, or another source).

Methods

Ethics Approval

We used secondary data from a large-scale national longitudinal study conducted by the Dutch National Institute for Public Health and the Environment. Participants provided informed consent before the start of the first survey. The data that we received did not contain any identifiable information. Therefore, this study did not require to be reviewed by an institutional review board. The study design, hypotheses, measured variables, and plan of analysis were preregistered before gaining access to the data and can be found on the Open Science Framework page of this study [35].

Participants and Procedure

Participants were recruited via 25 municipal health offices (Gemeenschappelijke Gezondheidsdienst) to participate in the national survey (N=124,580). Participants were asked to fill out 1 or more questionnaires during 8 waves of data collection between April and November 2020. During this period, COVID-19 was highly prevalent in the Netherlands, ranging from 0.47 to 57.87 daily new cases per 100,000 inhabitants, and various preventive measures were in effect. Some of the measures that were continuously communicated were to keep a physical distance from others (1.5 m), to not shake hands and to wash hands often, to sneeze and cough in the armpit, and to work from home as much as possible. Initially, participants received a questionnaire every 3 weeks, and after the fifth wave, the interval was increased to 6 weeks (Table 1).

For each wave, the survey was divided into 3 subcomponents, and each participant received 1 of these 3 subcomponents per wave. As a result, a subset of participants received questions relevant to this study. Also, the preregistered exclusion criteria were used to exclude participants aged <18 years and participants for which the control variables were missing. This resulted in an analytical sample of 123,848 adults aged >17 years (34.11% men) who participated in 1 wave (n=47,708, 38.5%) or multiple waves (n=76,140, 61.5%). The participants in the longitudinal sample participated in 2 and 8 waves (mean 5.36, SD 2.14). As this study used existing data, no a priori sample size calculation was performed. Given the sample size, we did not anticipate problems with the statistical power. For each analysis, the number of included participants and observations is reported.

XSL•FO
RenderX

Table 1. Number of participants and dates of measurements per wave.

Wave	Number of participants	Between dates
1	65,572	April 17, 2020, to April 24, 2020
2	52,847	May 7, 2020, to May 12, 2020
3	63,773	May 27, 2020, to June 1, 2020
4	50,200	June 17, 2020, to June 21, 2020
5	50,366	July 8, 2020, to July 12, 2020
6	61,361	August 19, 2020, to August 23, 2020
7	47,670	September 30, 2020, to October 4, 2020
8	63,989	November 11, 2020, to November 15, 2020

Measures

Physical Distancing

For each wave, participants first answered the question "In the past 7 days, how often were you with a group of four or more people with whom you do not live in 1 house? For example, at work, at the park, on the street with neighbors, or at a birthday" on a scale ranging from *never* (1) to *more than 20 times* (7). Participants who were, at least one time, with a group of 4 or more people in the last week were asked the follow-up question "In the past 7 days, how often were you successful in always keeping a physical distance of 1.5 meters from these people" and asked to respond on a Likert scale ranging from *never* (1) to *always* (7). The score on this scale for each wave was used as a measure of physical distancing. The higher the score on the variable, the more successful the participant was in maintaining physical distance in the past week ($mean_{grand}$ 4.34, SD_{grand} 1.58).

Age

The participants indicated the category according to their age group (Table 2). As only 0.54% (669/123,848) of the participants were in the eighth category (≥85 years), categories 7 and 8 were merged. Dummy coding was used to create a contrast between the emerging adults (n=6648) and older age categories (n=117,200). To further investigate differences between the age categories, reversed Helmert contrast coding was used to contrast the age category with all higher age categories combined, starting with the emerging adult category.

Table 2. Age in categories (N=123,848).

Answer	Age (years)	Label	Participants, n (%)
3	18-24	Emerging adults	6648 (5.37)
4	25-39	Early career	31,724 (25.62)
5	40-54	Midcareer	34,692 (28.01)
6	55-69	Late career	33,476 (27.03)
7-8	≥70	Retired	17,308 (13.98)

Social Media Use

Each wave, a subset of participants answered the question "In the past 7 days, which of these sources did you use to get information and news about the coronavirus?" Participants could respond by selecting 1 or more media sources from the given list. One such source was *social media*. A dummy variable *Social Media* was created to compare whether participants used social media (0.5, $n_{observations}$=33,941) or did not (−0.5, $n_{observations}$=81,008) for COVID-19 news and information per wave.

Social Media Platforms and Sources

In waves 2 and 4, a subset of the participants (n=18,047) received the module with more extensive questions regarding social media use. In these questions, participants indicated how many days of the past week they had used the following platforms for COVID-19 news and information: Facebook, Twitter, Instagram, YouTube, and LinkedIn. For each indicated social platform, participants were also asked to select 1 or multiple sources presented on the platform: government, national news, regional news, personal communication, or other sources.

Control Variables

To control for potential differences in physical distancing, all analyses were controlled for participant sex. In addition, the wave was added as a covariate to control for potential changes in behavior and context over time. As not all participants filled out the questions during the same wave, it is important to control this temporal context. During the measurements, the number of infected people was initially high and decreased during the summer but increased again after the fifth wave. In addition, the regulations changed regularly, and the overall sentiment might have changed as well. A linear wave variable would not reflect this trend; therefore, we have tested several other shapes that would fit the observed data [36]. The wave-transformed variable with the best fit to the observed data was selected. Specifically, the wave variable was centered on wave 5, and

XSL•FO
RenderX

the absolute values were used to create a v-shape. The standardized effect of the transformed variable was higher and explained more variance ($R^2_{marginal}$=0.031, β=0.18) than the linear wave variable ($R^2_{marginal}$=0.0002, β=−0.05).

Strategy of Analysis

We preregistered the intention to use Bayesian statistics to test the hypothesis. However, all analyses had to be performed on a secured remote desktop, and the possibilities of running extensive computations on this large data set were limited. Therefore, multivariate mixed effects models were run by using the *lme4* package [37] in R (R Foundation for Statistical Computing) [38]. SE, CIs, and *P* values were computed using the Satterthwaite approximation [39], and CIs not including 0 or *P*<.05 were considered statistically significant. Effect sizes were used to determine the direction and relative strength of the parameter, and parameter importance was determined based on the improved model fit.

In the mixed effects models, sex and wave were added as covariates, and random intercepts were included per participant. According to this hypothesis, the predictor was substituted for the variable of interest. To test the mediation for H2, a multilevel mediation model from the *mlma* package was used [40]. In addition, to determine the cross-lagged effects between physical distancing and using social media as a source, Random intercept cross-lagged panel models [41] were used to distinguish between-person (stable time-invariant traits) and within-person (in-person changes over time) associations. The cross-lagged paths were used to assess the directionality between using social media and physical distancing between current and subsequent waves while controlling for stability traits between waves and covariance within waves. All correlations at each wave, stability, and cross-lagged paths were restricted to be the same, resulting in 1 parameter estimate per path type.

In contrast to the preregistration, the weather conditions were not included as covariates because the exact dates of filling out the questionnaires were not included in the data set. We have tried to include the average weather conditions per wave, but this variable had too much collinearity with the wave variable, making the models unidentifiable. Moreover, the hypotheses on well-being are not reported in this paper because of an overlap with another group of researchers working with the same data set. The planned analyses are still part of the supportive materials for the Open Science Framework. Finally, 6 waves of data were available at the time of registration. Subsequently, 2 additional waves of data were gathered, which were added to the data set.

On top of the preregistered analyses, 2 exploratory analyses were performed on the subsample of the participants that received the module on the use of specific social media platforms (ie, Facebook, Twitter, Instagram, YouTube, and LinkedIn) and different sources that appear on these platforms (ie, governmental, national news, regional paper, personal post, or other sources). The 2 mixed effects models were specified similarly to the model to test the first hypothesis. In the first model, the number of days per week that participants used social media platforms for COVID-19 news and information were entered as predictors, and the age variable was treated as a covariate. In the second exploratory model, social media variables were again excluded, and dummy variables per source were used to determine whether participants were exposed to a specific source on social media.

Results

Physical Distancing

The linear mixed-effects model that was used to test the first hypothesis consisted of a random structure in the form of random intercepts per participant and a fixed structure explaining 4% of the variance in physical distancing (marginal R^2). Both fixed and random effects explained 50% of the variance (conditional R^2). The intraclass correlation coefficient of the random effect participant was 0.48, indicating that approximately half of the variance was explained by other observations on the outcome variable within the same participant.

The planned contrast indicated that emerging adults ($mean_{marginal}$ 3.48, $SE_{marginal}$ 0.03) maintained physical distance from others less often than the older participants ($mean_{marginal}$ 4.37, $SE_{marginal}$ 0.01; Figure 1).

The standardized effect size suggested that the effect of age was less important than that of the covariate wave (Table 3). However, the model fit and explained variance of the model (Akaike information criterion [AIC]=693,681 and Bayesian information criterion [BIC]=693,742) were better than the model fit without the emerging adult variable ($R^2_{marginal}$=0.03, AIC=694.394 and BIC=694.444; χ^2_1=714.9; *P*<.001). This indicated that although the effect of age could be considered small, the variable still contributed to explaining physical distancing behavior.

In a nonpreregistered additional analysis, we further investigated the differences in physical distancing between age categories. Therefore, the dichotomous *emerging adult* variable was substituted for multiple contrasts of the categorical age variables, as measured in the project. This variable increased the marginal R^2 of the model to 6% and indicated that with an increase in the age category, people practiced physical distancing more often (Figure 2). Together, these analyses provide support for hypothesis 1 that physical distancing is lower in emerging adults than in adults.

XSL•FO
RenderX

Figure 1. Physical distancing in emerging adults and adults over the eight waves.

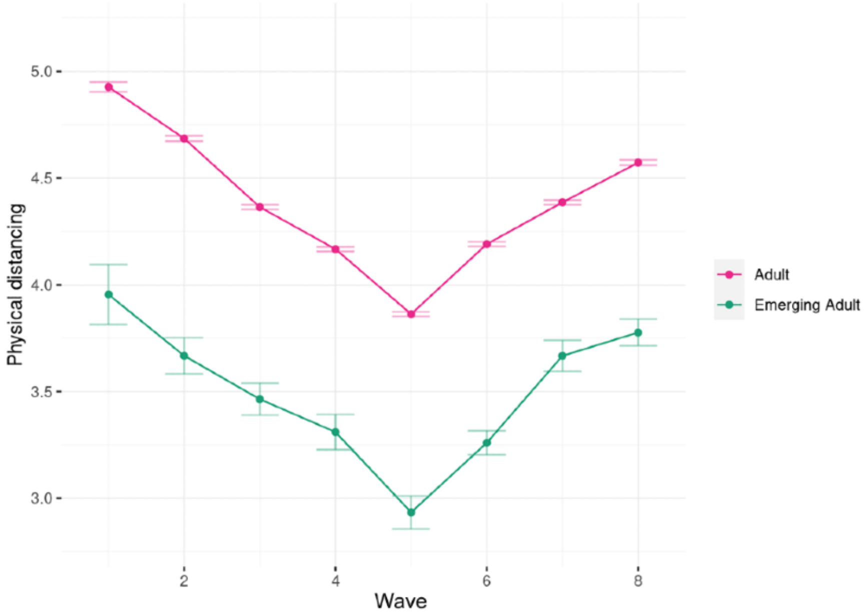
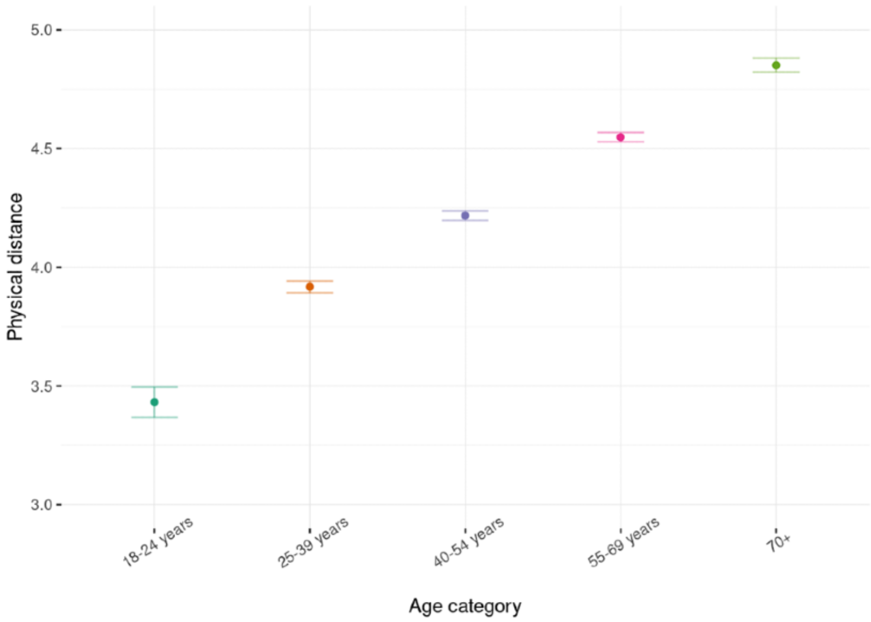


Table 3. Multivariate linear mixed-effects model predicting physical distancing behavior (n=70,629; number of observations=185,208; participant intraclass correlation coefficient=0.48; marginal R^2=0.04; conditional R^2=0.50).

Variable	B (SE)	95% CI	β	*t* test (*df*)	*P* value
Intercept	3.44 (0.02)	(3.41 to 3.48)	.00	194.46 (98,929.46)	<.001
Emerging adult	−0.89 (0.03)	(−0.96 to −0.82)	−.08	−26.79 (86,213.83)	<.001
Sex	0.12 (0.01)	(0.10 to 0.14)	.03	10.04 (65,587.82)	<.001
Wave	0.29 (0.00)	(0.29 to 0.30)	.18	96.81 (148,077.18)	<.001

Figure 2. Physical distancing per age category.



Using Social Media for COVID-19 News and Information

The second hypothesis investigated the role of using social media for COVID-19 news and information in physical distancing. The related social media use question was asked in waves 3, 5, and 8 in a subsample of participants (n=17,714, $n_{observations}$=38,423). A social media use dummy was added to the model used in H1 ($R^2_{marginal}$=0.03, $R^2_{conditional}$=0.48, intraclass correlation $coefficient_{participant}$=0.48). A significant

XSL•FO
RenderX

social media use parameter indicated that those who used social media ($mean_{marginal}$ 3.96, $SE_{marginal}$ 0.06) showed slightly less physical distancing behavior than those who did not use social media for COVID-19 news and information ($mean_{marginal}$ 4.03, $SE_{marginal}$ 0.06). In the same model, a significant emerging adult parameter indicated that after controlling for social media use, emerging adults kept physical distance less often than did adults (Table 4). Again, the standardized effect sizes for both social media use and emerging adults were small. An improved model fit indicated that both the social media use variable (full model: AIC=108,886, BIC=108,944; model without social media use predictor: AIC=108,900, BIC=108,950; χ^2_1=16.3; *P*<.001) and emerging adult variables (model without social emerging adult predictor: AIC=108,942, BIC=108,991; χ^2_1=57.7; *P*<.001) contributed to explaining physical distancing behavior.

Next, a mixed effects logistic regression model was used to test whether emerging adults were more likely to use social media for news and information on COVID-19. The model ($R^2_{marginal}$=0.05, $R^2_{conditional}$=0.76) showed a significant emerging adult parameter (B=2.48; SE 0.10; Wald=23.66; *P*≤.001). Emerging adolescents were 11.93 (95% CI 9.72-14.65) times more likely to use social media for COVID-19 news and information than adults did. This means that there is a stronger preference to use social media for COVID-19 news and information among emerging adults than among adults.

Finally, a mixed effects mediation model was used to dissociate the direct association between the emerging adult variable and physical distancing from the indirect association mediated by social media for COVID-19 news and information. The model showed that the total effect (β=−.91; 95% CI 1.06 to −0.77; *P*<.001), direct effect (β=−.88; 95% CI 1.04 to −0.74; *P*<.001), and indirect effect (β=−.03; 95% CI 0.04 to −0.02; *P*<.001) were all significant. However, the indirect effect was substantially smaller than the direct effect, and we concluded that there is a partial, but limited, mediating path of using social media for COVID-19 news and information. Therefore, using social media for COVID-19 news and information can only marginally explain why physical distancing is lower among emerging adults than among adults.

Table 4. Multivariate linear mixed-effects model predicting physical distancing behavior (n=17,714; number of observations=38,423; intraclass correlation coefficient of participants=0.47; marginal R^2=0.03; conditional R^2=0.48).

Variable	B (SE)	95% CI	β	*t* test (*df*)	*P* value
Intercept	3.56 (0.06)	(3.45 to 3.68)	.00	60.24 (13,774.81)	<.001
Social media use	−0.10 (0.02)	(−0.15 to −0.05)	−.02	−4.03 (28,459.80)	<.001
Emerging adult	−0.87 (0.11)	(−1.09 to −0.65)	−.06	−7.60 (12,566.76)	<.001
Sex	0.11 (0.03)	(0.06 to 0.17)	.03	3.94 (11,431.76)	<.001
Wave	0.27 (0.01)	(0.26 to 0.29)	.16	34.67 (21,764.25)	<.001

Determining Directionality

An additional analysis investigated the directionality of the effect of using social media on physical distancing and vice versa. This analysis used a subset of the participants (n=7325) in the last 4 waves because then the social media question was presented in 4 subsequent waves to the same participants. In addition, the sex and age groups of the participants were added as covariates in the model. The restricted random intercept cross-lagged panel model (χ^2_{35}=296.9; *P*≤.001; comparative fit index=0.987, Tucker–Lewis index=0.984, root mean square error of approximation=0.032, and standardized root mean square residual=0.025) showed a negative relationship between social media use and physical distancing (Table 5). This means that there was a small negative between-person association between physical distancing and using social media for COVID-19–related news and information. In addition, both stability paths were significant, indicating that the values of both variables were predicted by the value of the previous wave. Most interestingly, a small negative cross-lagged effect of physical distancing on social media use, but no effect of social media on physical distancing, was observed. This indicates that using social media for COVID-19 news and information did not affect physical distancing in the subsequent wave. By contrast, those who maintained physical distance less often were more likely to use social media as a source in the subsequent wave. However, the standardized effect size was very small.

Table 5. Random intercept cross-lagged panel model of physical distancing and social media (n=7324).

Variable	B (SE)	95% CI	β	*z* score	*P* value
W5 correlation	−0.01 (0.01)	(−0.02 to 0.00)	−.03	−2.21	.03
Distance → social media	0.00 (0.00)	(−0.01 to 0.00)	−.02	−2.14	.03
Social media → distance	−0.06 (0.04)	(−0.14 to 0.02)	−.02	−1.40	.16
Distance → distance	0.12 (0.01)	(0.10 to 0.14)	.12	10.94	<.001
Social media → social media	0.11 (0.01)	(0.09 to 0.14)	.11	9.9	<.001
Correlated change W6-8	0.00 (0.00)	(−0.01 to 0.01)	.00	0.33	.74
Between-person correlation	−0.01 (0.01)	(−0.03 to −0.00)	−.04	−2.54	.01

XSL•FO
RenderX

Differences in Social Media Platforms and Sources of Information on Social Media

In the last 2 analyses, we further explored the differences between several social media platforms and the sources that appear on these platforms for COVID-19 news and information. The explorations were performed in a subsample of participants who received the extensive social media module in waves 2 and 4 (n=9992 and $n_{observations}$=12,456). Facebook was the most frequently used platform (5274/12,456, 42.34%), whereas all other platforms were used between 15.7% (1995/12,456) and 11.03% (1374/12,456) of the time. When a social media platform was used, Facebook (n=5274; mean 4.91, SD 2.36), Instagram (n=1881; mean 4.49, SD 2.44), and Twitter (n=1786; mean 4.68, SD 2.38) were used for more than half of the days per week for COVID-19 news and information. LinkedIn (n=1955; mean 3.16, SD 2.12) and YouTube (n=1374; mean 2.81, SD 2.07) were used for fewer days per week for COVID-19 news and information.

The first linear mixed-effects model ($R^2_{marginal}$=0.07, $R^2_{conditional}$=0.89) investigated the association between physical distancing and the number of days per week during which different social media platforms were used for COVID-19 news and information (Table 6).

The results of the model showed that some platforms had no association or a slightly positive association with physical distancing (ie, Twitter and LinkedIn), whereas others had a negative association (ie, Facebook, Instagram, and YouTube; Figure 3).

Potentially, differences in associations emerged because various information sources were portrayed on different platforms. Therefore, we further investigated the sources of COVID-19 news and information on the social media platforms used by participants. Governmental (6511/12,456, 52.3%), national news (6429/12,456, 51.6%), and personal communication (7237/12,456, 58.1%) were the most common sources on social media platforms. Regional news (3036/12,456, 24.4%) and other sources (1454/12,456, 11.5%) were used less frequently for COVID-19 news and information.

In the second linear mixed-effects model, the social media platform variables were substituted for a dummy variable per source, contrasting seeing a source (0.5) on social media versus not seeing a source on social media (−0.5). The model ($R^2_{marginal}$=0.06, $R^2_{conditional}$=0.53) showed that being exposed to governmental sources had a distinctly small positive association, compared with the other sources that had no or a small negative association with physical distancing (Table 7). Together, these 2 exploratory analyses suggest that associations between physical distancing and using social media for COVID-19 news and information are less straightforward. Depending on the social media platform that people used and the sources they were exposed to on social media, the associations varied in effect size and direction.

Table 6. Multivariate linear mixed-effects model predicting physical distancing behavior (n=9992; number of observations=12,456; intraclass correlation coefficient of participants=0.48; marginal R^2=0.04; conditional R^2=0.50).

Variable	B (SE)	95% CI	β	*t* test (*df*)	*P* value
Intercept	4.01 (0.07)	(3.88 to 4.15)	.00	57.72 (9734.55)	<.001
Facebook	−0.04 (0.01)	(−0.05 to −0.03)	−.06	−6.06 (11,930.58)	<.001
Twitter	0.02 (0.01)	(0.00 to 0.04)	.02	2.17 (11,380.53)	.03
Instagram	−0.02 (0.01)	(−0.04 to 0.00)	−.02	−2.38 (12,429.39)	.02
YouTube	−0.09 (0.01)	(−0.12 to −0.07)	−.06	−6.36 (12,443.79)	<.001
LinkedIn	0.04 (0.01)	(0.01 to 0.06)	.03	2.94 (12,262.32)	.003
Sex (male)	0.12 (0.04)	(0.05 to 0.19)	.03	3.43 (9871.87)	<.001
Wave	−0.57 (0.03)	(−0.62 to −0.52)	−0.16	−20.88 (6389.34)	<.001
Emerging adult	−1.00 (0.14)	(−1.27 to −0.73)	−0.07	−7.36 (9631.78)	<.001

XSL•FO
RenderX

Figure 3. Associations between the number of days spent using different social media platforms and physical distancing.

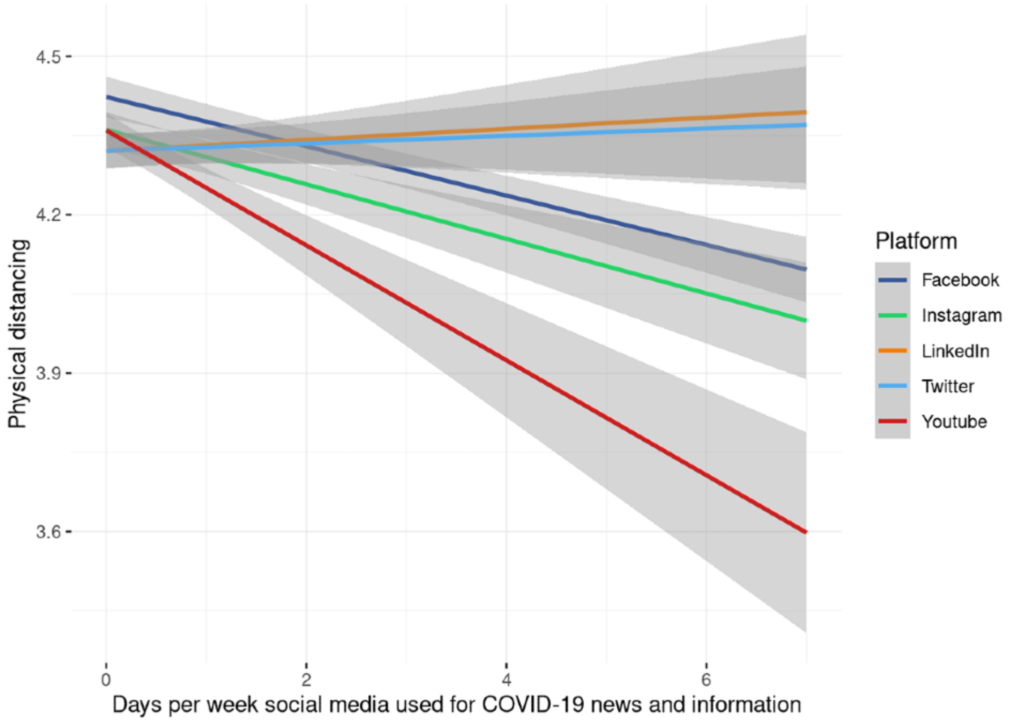


Table 7. Multivariate linear mixed-effects model of social media sources predicting physical distancing behavior (n=5986; number of observations=7221; intraclass correlation coefficient of participants=0.48; marginal R^2=0.04; conditional R^2=0.53).

Variable	B (SE)	95% CI	β	*t* test (*df*)	*P* value
Intercept	3.85 (0.08)	(3.68 to 4.01)	.00	46.39 (5994.70)	<.001
Government	0.10 (0.03)	(0.05 to 0.16)	.05	3.68 (7178.46)	<.001
National news	−0.05 (0.03)	(−0.11 to 0.01)	−.02	−1.73 (7106.50)	.08
Regional news	−0.04 (0.03)	(−0.11 to 0.03)	−.01	−1.14 (7077.41)	.25
Personal communication	−0.08 (0.02)	(−0.13 to −0.04)	−.05	−3.86 (7162.97)	<.001
Other	−0.10 (0.04)	(−0.18 to −0.03)	−.03	−2.74 (7085.96)	.006
Sex (male)	0.05 (0.05)	(−0.05 to 0.14)	.01	0.96 (5862.23)	.34
Wave	−0.59 (0.04)	(−0.66 to −0.52)	−.16	−16.35 (3694.48)	<.001
Emerging adult	−1.05 (0.15)	(−1.35 to −0.75)	−.09	−6.93 (5691.54)	<.001

Discussion

Principal Findings

This study investigated differences between emerging adults and adults in terms of physical distancing. Moreover, the role of using social media for COVID-19 news and information was investigated by examining its mediating role and to what extent different social media platforms and sources relate to physical distancing. These questions were addressed in a longitudinal panel study of a large sample of Dutch adults conducted between April and November 2020. On the basis of our findings, 3 main conclusions can be drawn.

First, our findings demonstrate that the physical distancing behavior is lower in the group of emerging adults than in the group of adults. Moreover, we believe that physical distancing increases with age as the exploration with nondichotomous age categories implies. This finding was in line with previous studies that reported that emerging adults or younger participants, on average, maintain physical distance less often than adults [29,31]. One potential explanation lies in psychosocial models such as the health belief model [42] and protection motivation theory [43]. Given the lower personal health risks for younger people, the perceived vulnerability, severity, and perceived benefits of physical distancing might be lower, whereas the cost and barriers to compliance would be higher, giving up more of social daily life [27,31].

The second conclusion is that using social media for news and information on COVID-19 is negatively related to physical distancing behavior, irrespective of age. Moreover, the emerging adults in our study were more likely to use social media for COVID-19 news and information, and social media played a small role in physical distancing behaviors. However, because the indirect relationship was trivially small compared with the

XSL•FO
RenderX

direct relationship, we consider social media use only as a very limited, not meaningful, explanation of emerging adults' lower physical distancing behavior.

Moreover, the longitudinal panel model showed no support for the direction of social media use, leading to lower physical distancing in the subsequent wave. Rather, the analysis showed a significant, albeit small, cross-lagged path between physical distancing and future social media use. A potential explanation is selective attention to COVID-19 news and information that affirms their current beliefs about COVID-19 and avoids media content that is dissonant with their behavior [44,45], similar to political news seeking [46]. This would mean that those who disagree with the prevailing measures turn away from other types of sources such as television, governmental websites, and newspapers when they seek information and news about the coronavirus, as the portrayed images are not in line with their beliefs. Therefore, using social media for COVID-19 news and information can be seen as a sign of noncompliance and not as a source of noncompliance to the prevailing measures. However, the observed effect in this study was so small that at this stage, we are not in a position to draw firm conclusions, and further investigation is warranted.

Finally, the explorations in this study suggest that social media use is not always bad for physical distancing behaviors, with some platforms showing small positive relations with physical distancing. Moreover, the types of sources portrayed in these social media messages seem to relate to physical distancing. We can draw tentative conclusions that users looking for COVID-19 news and information on LinkedIn and Twitter were more likely to adhere to physical distancing measures, albeit with relatively weak associations. Similarly, the use of social media posts from governmental sources was related to greater physical distancing, whereas web-based personal communication seemed to be related to less physical distancing. Overall, it should be noted that the strength of the observed associations between social media use and physical distancing was relatively low or even nonsignificant, such as for national and regional news sources on social media.

Strengths and Limitations

In the literature on COVID-19, an impressive number of studies have investigated the impact of the virus and its corresponding regulations. Some studies have focused on student samples but overlooked going beyond young people who attend postsecondary education or comparing this age group to adults. In this study, we had the opportunity to fill this gap by using a very large sample of emerging adults and adults. In addition, a substantive subset was part of a longitudinal sample, enabling us to investigate the relationships over time and sensitize the directionality of relationships. By using both multivariate mixed effects models and random intercept cross-lagged panel models, we were able to control for the clustering of data with each participant that responded multiple times and investigated the directionally of the studies association. Furthermore, open-science practices were used, in which the hypotheses and analyses were preregistered before the analyses were performed, and the scripts used are publicly available.

This research also has some limitations that must be considered when interpreting the findings, which can be addressed in future research. For example, a crude measurement of participants' age was used. As secondary data were used, this study had no control over the questions being asked or the data being stored. The survey was carried out with much attention paid to the privacy of the participants. To reduce the traceability of the participants, the age variable was measured in categories. We encourage these anonymization efforts, but they might have made the estimated parameters less precise, and another level of detail could have been achieved by having the exact age of the participants. Future large-scale projects could address this issue by creating synthetic data sets before analyzing the data to retain the privacy of the participants [47].

In addition, physical distancing was measured through retrospective self-report. Participants indicated in each wave how often they had maintained physical distance from others in the preceding 7 days. Considering all potential biases (eg, recall bias, primacy and recency bias, and social desirability), it is conceivable that the reported behavior deviates from the objective physical distancing behavior. However, we do not believe that the effects of potential biases may be different for different age groups. Related to this is the measurement of sources used for information and news on COVID-19. Participants responded by selecting several types of media from the provided list. The actual amount and specific content seen on social media or other sources could not be derived in this study. One way of obtaining more detailed information in large-scale studies would be to ask participants to donate logging data of the used social networking sites (eg, cookies or browser history) or ask participants to install a mobile sensing app to collect media use and physical distancing behaviors [48,49].

Finally, the size of the sample also warrants some caution in the context of null hypothesis significance testing because even tiny effects can reach the preregistered critical value of $P<.05$. As a result, the question arises of whether the significant effect is big enough to be concerned about. In our analyses, we used standardized effect sizes representing a 1 SD increase on the Likert scale measuring physical distancing. However, as the answers on the Likert scale do not form an absolute continuous scale, a quantifiable interpretation of the size of the significant effects is not straightforward. We have tried to indicate whether we deem the effect meaningful by examining an increased model fit of a particular variable. However, at the same time, the large sample size eliminates the argument of insufficient power to detect an effect and make a type 2 error. This gives us more confidence in deciding that when an effect is not statistically significant, it is highly likely to be absent. However, other arguments regarding why the hypotheses can be falsely rejected remain applicable to this study.

Conclusions

Our study indicates a substantive gap between emerging adults and adults in physical distance behavior during the COVID-19 pandemic and yet yields a nuanced view on emerging adulthood and the role of social media. Given the overall increase with age, we cannot make firm conclusions that the group of emerging adults should be seen as a particularly problematic

XSL•FO
RenderX

group in itself but rather that the older people become, the more often they comply with physical distancing measures. Moreover, although using social media for COVID-19 news and information is negatively related to physical distancing behavior, it does not seem to be an important factor in explaining why emerging adults comply less with the behavioral measures, nor does it lead to changes in physical distancing behavior over time. Finally, there are differences between the various social media platforms and sources, with some platforms and sources showing negative associations and other platforms showing positive to no associations with physical distancing. However, we should be cautious in assuming that these social media affect behaviors because they may very well be indicators of selective exposure to social media that match one's physical distancing behaviors.

Conflicts of Interest

None declared.

References

1. Dong E, Du H, Gardner L. An interactive Web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534 [FREE Full text] [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
2. Islam N, Sharp SJ, Chowell G, Shabnam S, Kawachi I, Lacey B, et al. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. BMJ 2020 Jul 15;370:m2743 [FREE Full text] [doi: 10.1136/bmj.m2743] [Medline: 32669358]
3. Seale H, Dyer CE, Abdi I, Rahman KM, Sun Y, Qureshi MO, et al. Improving the impact of non-pharmaceutical interventions during COVID-19: examining the factors that influence engagement and the impact on individuals. BMC Infect Dis 2020 Aug 17;20(1):607 [FREE Full text] [doi: 10.1186/s12879-020-05340-9] [Medline: 32807087]
4. Ball-Rokeach SJ. The origins of individual media-system dependency: a sociological framework. Commun Res 1985 Oct 1;12(4):485-510. [doi: 10.1177/009365085012004003]
5. Garfin DR, Silver RC, Holman EA. The novel coronavirus (COVID-2019) outbreak: amplification of public health consequences by media exposure. Health Psychol 2020 May;39(5):355-357 [FREE Full text] [doi: 10.1037/hea0000875] [Medline: 32202824]
6. Gerbner G. Cultivation analysis: an overview. Mass Commun Soc 1998 Jun;1(3-4):175-194. [doi: 10.1080/15205436.1998.9677855]
7. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ, USA: Prentice-Hall; 1986.
8. LaRose R. Social cognitive theories of media selection. In: Hartmann T, editor. Media Choice: A Theoretical and Empirical Overview. New York, NY, USA: Routledge; 2009:10-31.
9. Lin HC, Chang CM. What motivates health information exchange in social media? The roles of the social cognitive theory and perceived interactivity. Inf Manag 2018 Sep;55(6):771-780. [doi: 10.1016/j.im.2018.03.006]
10. Cheng JW, Mitomo H, Otsuka T, Jeon SY. Cultivation effects of mass and social media on perceptions and behavioural intentions in post-disaster recovery – the case of the 2011 Great East Japan Earthquake. Telemat Inform 2016 Aug;33(3):753-772. [doi: 10.1016/j.tele.2015.12.001]
11. Tang Z, Miller AS, Zhou Z, Warkentin M. Does government social media promote users' information security behavior towards COVID-19 scams? Cultivation effects and protective motivations. Gov Inf Q 2021 Apr;38(2):101572 [FREE Full text] [doi: 10.1016/j.giq.2021.101572] [Medline: 35719729]
12. 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]
13. Palen L, Vieweg S, Liu SB, Hughes AL. Crisis in a networked world: features of computer-mediated communication in the April 16, 2007, Virginia Tech Event. Soc Sci Comput Rev 2009 Apr 08;27(4):467-480. [doi: 10.1177/0894439309332302]
14. Wei YJ, McIntyre FS, Straub D. Does micro-blogging lead to a more positive attitude toward a brand?—a perspective of cultivation theory. J Promot Manag 2020 Jan 29;26(4):504-523. [doi: 10.1080/10496491.2020.1719957]
15. 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]
16. Kulshrestha J, Eslami M, Messias J, Zafar MB, Ghosh S, Gummadi KP, et al. Quantifying search bias: investigating sources of bias for political searches in social media. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW '17; February 25-March 1, 2017; Portland, OR, USA p. 417-432. [doi: 10.1145/2998181.2998321]
17. Jennings W, Stoker G, Willis H, Valgardsson V, Gaskell J, Devine D, et al. Lack of trust and social media echo chambers predict COVID-19 vaccine hesitancy. medRxiv 2021 Jan 27. [doi: 10.1101/2021.01.26.21250246]
18. Ghersetti M, Westlund O. Habits and generational media use. Journal Stud 2016 Nov 22;19(7):1039-1058. [doi: 10.1080/1461670x.2016.1254061]

XSL•FO
RenderX

19. Middaugh E. Teens, social media, and fake news. In: Journell W, editor. Unpacking Fake News: An Educator's Guide to Navigating the Media With Students. New York, NY, USA: Teachers College Press; 2019:42-59.
20. Westlund O, Ghersetti M. Modelling news media use. Journal Stud 2014 Jan 07;16(2):133-151. [doi: 10.1080/1461670x.2013.868139]
21. Dadaczynski K, Okan O, Messer M, Leung AY, Rosário R, Darlington E, et al. Digital health literacy and Web-based information-seeking behaviors of university students in Germany during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2021 Jan 15;23(1):e24097 [FREE Full text] [doi: 10.2196/24097] [Medline: 33395396]
22. Lisitsa E, Benjamin KS, Chun SK, Skalisky J, Hammond LE, Mezulis AH. Loneliness among young adults during COVID-19 pandemic: the mediational roles of social media use and social support seeking. J Soc Clin Psychol 2020 Dec;39(8):708-726. [doi: 10.1521/jscp.2020.39.8.708]
23. Wong FH, Liu T, Leung DK, Zhang AY, Au WS, Kwok WW, et al. Consuming information related to COVID-19 on social media among older adults and its association with anxiety, social trust in information, and covid-safe behaviors: cross-sectional telephone survey. J Med Internet Res 2021 Feb 11;23(2):e26570 [FREE Full text] [doi: 10.2196/26570] [Medline: 33523825]
24. Garbe L, Rau R, Toppe T. Influence of perceived threat of Covid-19 and HEXACO personality traits on toilet paper stockpiling. PLoS One 2020 Jun 12;15(6):e0234232 [FREE Full text] [doi: 10.1371/journal.pone.0234232] [Medline: 32530911]
25. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020 Apr 07;323(13):1239-1242. [doi: 10.1001/jama.2020.2648] [Medline: 32091533]
26. Zhang M. Estimation of differential occupational risk of COVID-19 by comparing risk factors with case data by occupational group. Am J Ind Med 2021 Jan;64(1):39-47 [FREE Full text] [doi: 10.1002/ajim.23199] [Medline: 33210336]
27. Crone EA, Dahl RE. Understanding adolescence as a period of social-affective engagement and goal flexibility. Nat Rev Neurosci 2012 Sep;13(9):636-650. [doi: 10.1038/nrn3313] [Medline: 22903221]
28. Bish A, Michie S. Demographic and attitudinal determinants of protective behaviours during a pandemic: a review. Br J Health Psychol 2010 Nov;15(Pt 4):797-824 [FREE Full text] [doi: 10.1348/135910710X485826] [Medline: 20109274]
29. Coroiu A, Moran C, Campbell T, Geller AC. Barriers and facilitators of adherence to social distancing recommendations during COVID-19 among a large international sample of adults. PLoS One 2020 Oct 7;15(10):e0239795 [FREE Full text] [doi: 10.1371/journal.pone.0239795] [Medline: 33027281]
30. Beeckman M, De Paepe A, Van Alboom M, Maes S, Wauters A, Baert F, et al. Adherence to the physical distancing measures during the COVID-19 pandemic: a HAPA-based perspective. Appl Psychol Health Well Being 2020 Dec;12(4):1224-1243. [doi: 10.1111/aphw.12242] [Medline: 33052008]
31. Kowalski RM, Black KJ. Protection motivation and the COVID-19 virus. Health Commun 2021 Jan;36(1):15-22. [doi: 10.1080/10410236.2020.1847448] [Medline: 33190547]
32. Arnett JJ. Emerging adulthood: a theory of development from the late teens through the twenties. Am Psychol 2000;55(5):469-480. [doi: 10.1037/0003-066x.55.5.469]
33. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design. Cambridge, MA, USA: Harvard University Press; 1979.
34. Chaffee SH, McLeod JM, Atkin CK. Parental influences on adolescent media use. Am Behav Sci 1971 Jan 1;14(3):323-340. [doi: 10.1177/000276427101400304]
35. Van Woudenberg T, Buijzen M, Lambooij M, Hendrikx R, van Weert J, van den Putte B. Physical distancing, media use and well being in (emerging) adults. Open Science Framework. 2022 Dec 2. URL: https://osf.io/9bygv [accessed 2022-07-22]
36. Van Woudenberg T, Buijzen M, Lambooij M, Hendrikx R, van Weert J, van den Putte B. Data beloning to physical distancing, media use and well being in (emerging) adults. Open Science Framework. 2022 Dec 2. URL: https://osf.io/ypa75 [accessed 2022-07-22]
37. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015 Oct 7;67(1):1-48. [doi: 10.18637/jss.v067.i01]
38. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2013. URL: http://www.R-project.org/ [accessed 2022-07-11]
39. Satterthwaite FE. An approximate distribution of estimates of variance components. Biometrics Bull 1946 Dec;2(6):110-114. [doi: 10.2307/3002019]
40. Yu Q, Li B. Third-variable effect analysis with multilevel additive models. PLoS One 2020 Oct 23;15(10):e0241072 [FREE Full text] [doi: 10.1371/journal.pone.0241072] [Medline: 33095796]
41. Hamaker EL, Kuiper RM, Grasman RP. A critique of the cross-lagged panel model. Psychol Methods 2015 Mar;20(1):102-116. [doi: 10.1037/a0038889] [Medline: 25822208]
42. Glanz K, Rimer BK, Viswanath K. Health Behavior: Theory, Research, and Practice. 5th edition. Hoboken, NJ, USA: John Wiley & Sons; 2015.
43. Rogers RW. A protection motivation theory of fear appeals and attitude change1. J Psychol 1975 Sep 02;91(1):93-114. [doi: 10.1080/00223980.1975.9915803] [Medline: 28136248]
44. Festinger L. A Theory of Cognitive Dissonance. Redwood City, CA, USA: Stanford University Press; 1957.

XSL•FO
RenderX

45. McQuail D. McQuail's Mass Communication Theory. 6th edition. Thousand Oaks, CA, USA: Sage Publications; 2010.
46. Messing S, Westwood SJ. Selective exposure in the age of social media. Commun Res 2012 Dec 31;41(8):1042-1063. [doi: 10.1177/0093650212466406]
47. Quintana DS. A synthetic dataset primer for the biobehavioural sciences to promote reproducibility and hypothesis generation. Elife 2020 Mar 11;9:e53275 [FREE Full text] [doi: 10.7554/eLife.53275] [Medline: 32159513]
48. Lind MN, Byrne ML, Wicks G, Smidt AM, Allen NB. The Effortless Assessment of Risk States (EARS) tool: an interpersonal approach to mobile sensing. JMIR Ment Health 2018 Aug 28;5(3):e10334 [FREE Full text] [doi: 10.2196/10334] [Medline: 30154072]
49. van Woudenberg TJ, Bevelander KE, Burk WJ, Smit CR, Buijs L, Buijzen M. Comparing the measurement of different social networks: peer nominations, online communication, and proximity data. Net Sci 2020 Jan 31;8(1):62-78. [doi: 10.1017/nws.2019.65]

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion

Edited by R Cuomo; submitted 20.09.21; peer-reviewed by T Lum, K Reuter; comments to author 14.01.22; revised version received 02.02.22; accepted 02.04.22; published 11.08.22.

Please cite as:
van Woudenberg T, Buijzen M, Hendrikx R, van Weert J, van den Putte B, Kroese F, Bouman M, de Bruin M, Lambooij M
Physical Distancing and Social Media Use in Emerging Adults and Adults During the COVID-19 Pandemic: Large-scale Cross-sectional and Longitudinal Survey Study
JMIR Infodemiology 2022;2(2):e33713
URL: https://infodemiology.jmir.org/2022/2/e33713
doi:10.2196/33713
PMID:35996459

©Thabo van Woudenberg, Moniek Buijzen, Roy Hendrikx, Julia van Weert, Bas van den Putte, Floor Kroese, Martine Bouman, Marijn de Bruin, Mattijs Lambooij. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 11.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Promoting Social Distancing and COVID-19 Vaccine Intentions to Mothers: Randomized Comparison of Information Sources in Social Media Messages

David Buller[1*], PhD; Barbara Walkosz[1*], PhD; Kimberly Henry[2*], PhD; W Gill Woodall[1*], PhD; Sherry Pagoto[3*], PhD; Julia Berteletti[1*], MSW; Alishia Kinsey[1*], BSc; Joseph Divito[3*], BSc; Katie Baker[4*], PhD; Joel Hillhouse[4*], PhD

[1]Klein Buendel, Inc, Golden, CO, United States

[2]Department of Psychology, Colorado State University, Fort Collins, CO, United States

[3]Department of Allied Health Sciences, University of Connecticut, Storrs, CT, United States

[4]Department of Community and Behavioral Health, East Tennessee State University, Johnson City, TN, United States

[*]all authors contributed equally

Corresponding Author:
David Buller, PhD
Klein Buendel, Inc
1667 Cole Blvd, Ste 220
Golden, CO, 80401
United States
Phone: 1 303 565 4321
Email: dbuller@kleinbuendel.com

Abstract

Background: Social media disseminated information and spread misinformation during the COVID-19 pandemic that affected prevention measures, including social distancing and vaccine acceptance.

Objective: In this study, we aimed to test the effect of a series of social media posts promoting COVID-19 nonpharmaceutical interventions (NPIs) and vaccine intentions and compare effects among 3 common types of information sources: government agency, near-peer parents, and news media.

Methods: A sample of mothers of teen daughters (N=303) recruited from a prior trial were enrolled in a 3 (information source) × 4 (assessment period) randomized factorial trial from January to March 2021 to evaluate the effects of information sources in a social media campaign addressing NPIs (ie, social distancing), COVID-19 vaccinations, media literacy, and mother–daughter communication about COVID-19. Mothers received 1 social media post per day in 3 randomly assigned Facebook private groups, Monday-Friday, covering all 4 topics each week, plus 1 additional post on a positive nonpandemic topic to promote engagement. Posts in the 3 groups had the same messages but differed by links to information from government agencies, near-peer parents, or news media in the post. Mothers reported on social distancing behavior and COVID-19 vaccine intentions for self and daughter, theoretic mediators, and covariates in baseline and 3-, 6-, and 9-week postrandomization assessments. Views, reactions, and comments related to each post were counted to measure engagement with the messages.

Results: Nearly all mothers (n=298, 98.3%) remained in the Facebook private groups throughout the 9-week trial period, and follow-up rates were high (n=276, 91.1%, completed the 3-week posttest; n=273, 90.1%, completed the 6-week posttest; n=275, 90.8%, completed the 9-week posttest; and n=244, 80.5%, completed all assessments). In intent-to-treat analyses, social distancing behavior by mothers (b=–0.10, 95% CI –0.12 to –0.08, P<.001) and daughters (b=–0.10, 95% CI –0.18 to –0.03, P<.001) decreased over time but vaccine intentions increased (mothers: b=0.34, 95% CI 0.19-0.49, P<.001; daughters: b=0.17, 95% CI 0.04-0.29, P=.01). Decrease in social distancing by daughters was greater in the near-peer source group (b=–0.04, 95% CI –0.07 to 0.00, P=.03) and lesser in the government agency group (b=0.05, 95% CI 0.02-0.09, P=.003). The higher perceived credibility of the assigned information source increased social distancing (mothers: b=0.29, 95% CI 0.09-0.49, P<.01; daughters: b=0.31, 95% CI 0.11-0.51, P<.01) and vaccine intentions (mothers: b=4.18, 95% CI 1.83-6.53, P<.001; daughters: b=3.36, 95% CI 1.67-5.04, P<.001). Mothers' intentions to vaccinate self may have increased when they considered the near-peer source to be not credible (b=–0.50, 95% CI –0.99 to –0.01, P=.05).

XSL•FO
RenderX

Conclusions: Decreasing case counts, relaxation of government restrictions, and vaccine distribution during the study may explain the decreased social distancing and increased vaccine intentions. When promoting COVID-19 prevention, campaign planners may be more effective when selecting information sources that audiences consider credible, as no source was more credible in general.

Trial Registration: ClinicalTrials.gov NCT02835807; https://clinicaltrials.gov/ct2/show/NCT02835807

(JMIR Infodemiology 2022;2(2):e36210) doi:10.2196/36210

KEYWORDS

social media; COVID-19; vaccination; nonpharmaceutical interventions; information source; misinformation; vaccine; public health; COVID-19 prevention; health promotion

Introduction

Background

To control the COVID-19 pandemic, the Centers for Disease Control and Prevention (CDC) has advised Americans to practice nonpharmaceutical interventions (NPIs; eg, social distancing, masking, and reduced group participation) and federal and state governments have mounted an unprecedented biomedical endeavor to develop and distribute vaccines [1-3]. NPIs are feasible, and social distancing and mask wearing reduce SARS-CoV-2 transmission [4-9]. Attention to prevention measures remains necessary because use of NPIs has declined and governments have relaxed restrictions [10-12]; even though vaccines are not universally accepted [13,14], individuals need to be revaccinated [15,16]; and groups that do not support vaccination are undermining confidence in COVID-19 vaccines [17,18].

In this study conducted from January to March 2021, we tested the impact of an intervention comprising social media posts promoting COVID-19 NPIs and vaccine intentions and compared 3 different types of information sources highlighted in the posts. In January 2021, COVID-19 case rates were high (7-day moving average=165,974 cases on January 25) [19] and NPIs were strongly recommended or mandated [20,21]. However, cases had declined substantially by March 2021 (7-day moving average=59,986 cases on March 26) [19] and some states were relaxing NPI advice and restrictions [20-22]. Two vaccines had been approved by January 2021 and a third in March 2021. Mass vaccination began during the intervention [22], but most states were still restricting vaccination to middle-age and older adults, with only 32% of American adults having received at least 1 dose at the end of March 2021 [23].

Role of Social Media in the COVID-19 Pandemic

Social media has played a large role in disseminating pandemic information, but it has also been used to spread misinformation [3,24], such as lack of severity of COVID-19, false virus transmission methods, ineffective prevention and diagnostic methods, unproven/pseudoscience treatments, risks from testing and face masks, and other conspiracy theories [25-28]. Misinformation has also spread about the COVID-19 vaccine, such as claims that vaccine safety was compromised by the rush to market, that the low risk from COVID-19 and effective prevention and treatment make vaccines less necessary, and that variation in the amount and length of effectiveness indicates vaccines are not useful [13,17]. Lower uptake of vaccines in general and lower COVID-19 vaccine intentions have been related to misinformation, unwarranted safety concerns, and conspiracies on social media, as has the practice of NPIs [29,30]. Thus, efforts are needed to promote COVID-19 prevention measures and correct misinformation on social media through fact checking and corrections, counternarratives, peer correction, coherence/credibility appeals, and digital and media literacy [31-38].

Impact of Sources for COVID-19 Information

The Extended Parallel Process Model (EPPM) of risk communication [39], an extension of protection motivation theory (PMT) [40,41], has explained mitigation behaviors in past pandemics, uptake of other vaccines [42-44], and COVID-19 pandemic responses [45]. It holds that the credibility of information sources influences the effectiveness of health messages [46]. High-credibility sources make it difficult for campaign audiences to derogate sources in order to decrease fear from risk information about COVID-19. In this way, messages from high-credibility sources motivate individuals to take actions that reduce risk with NPIs and vaccines.

We experimentally varied 3 types of sources, popular for information about the pandemic [47-49], that can vary in credibility (eg, trustworthiness and accuracy) in the social media posts on COVID-19: government agency, near-peer parents, and news media. Government health authorities are trusted sources of COVID-19 information for many (but not all) people [50,51], with nongovernmental content and unverifiable sources seen as less trustworthy, especially when posted on social media platforms [52,53]. A cross-sectional study of COVID-19 information sources found that attention to government sources is linked to greater COVID-19 knowledge [50]. Content shared on social media from (perceived) knowledgeable peers can have credibility and impact through identification processes based on similarity [54-57]. Peers (eg, friends, family, and work colleagues) have also been an often-used source of information about COVID-19, although they are not always as trusted as government and news media sources [48,51]. Consumers evaluate the credibility of both the source and message content of news media [58]. One study found that exposure to news media reduces conspiracy theories and misinformation beliefs regarding COVID-19 [59], but another reported that COVID-19 knowledge is lower among individuals who have greater trust in these sources [50]. The availability of a variety of information sources can elevate risk perceptions and fear; create information overload, anxiety, stress, and other negative psychological states; and possibly cause people to avoid information [45,47,48,60].

XSL•FO
RenderX

Hypothesis and Research Questions

This study was conducted with mothers of daughters aged 14-17 years who had participated in a previous trial on adolescent health. Mothers are an important audience for a COVID-19 prevention campaign because (1) mothers are often a primary decision maker for health and vaccination in families [61-63] and (2) parents use social media to track public health issues, share information, and seek advice [64]. The study tested the following primary hypothesis (H):

- H1: Mothers will report increased COVID-19 social distancing behaviors and vaccine intentions over the intervention period from baseline across 3 follow-up measures.

Posts also addressed theoretic antecedents of prevention behaviors prominent in the EPPM and social cognitive theory (SCT) [65]. In addition, whether mothers communicated with daughters about COVID-19 NPIs and vaccines was assessed because mother–daughter communication has influenced health behaviors of adolescent and young adult daughters in past research [66-68].

- H2: Mothers will report improved theoretic antecedents (perceived risk, self-efficacy, and response efficacy and cost) and mother–daughter communication about COVID-19 over the course of the intervention from baseline across 3 follow-up measures.

Analyses explored research questions asking whether the rate of change in social distancing, vaccine intentions, theoretic antecedents, and mother–daughter communication differed among the 3 types of information sources or by engagement with the social media messages.

Methods

Sample

Mothers were recruited to the study from a sample who had previously participated in a trial evaluating a social media campaign to prevent teen daughters from indoor tanning. In the original trial, mothers were recruited using community-based strategies (eg, schools, community events) and from the Qualtrics survey panel and met the following inclusion criteria: (1) having a daughter aged 14-17 years, (2) living in 1 of 34 states without a complete ban on indoor tanning (IT) by minors, (3) reading English, (4) having a Facebook account and logging in at least once per week, and (5) willing to “friend” the project’s community manager to join a private Facebook group. A detailed description of trial procedures has been published elsewhere [69,70]. In January 2021, 830 mothers were recontacted by email, invited to enroll in the current study that was described as a private group related to how mothers and daughters cope with the COVID-19 pandemic. Daughters were not enrolled in this study.

Experimental Design

Mothers were enrolled in a randomized pretest–posttest single-factor-design study with 4 assessments. After completing the baseline survey, mothers were randomly assigned to 1 of 3 experimental conditions that varied in the type of sources in the posts (government health agencies, near-peer parents, or news media) using a routine in Qualtrics survey software. Mothers “friended” the project community manager and were added to a Facebook private group for their assigned condition. As all mothers received experimental social media messages, they were blind to experimental manipulation of the information source. Study staff, other than the community manager and project manager, were blinded, too. The private groups prevented contamination between treatment groups while delivering the social media messages and made it possible to record engagement. Randomization controlled for background secular exposure to information in social media and other sources about COVID-19. Mothers received a series of Facebook posts for 9 weeks starting after randomization from January 25 to March 26, 2021. Each post contained text with a link to related information from 1 of the 3 types of sources. Mothers stayed in the groups for 9 weeks, completing online posttests at 3, 6, and 9 weeks postrandomization. After the intervention, 30 (9.9%) of 303 mothers were randomly selected to participate in focus groups, where they were asked what they liked most and least about the Facebook group and what they learned. A priori statistical power calculations via a Monte Carlo study in Mplus and with the *powerlmm* package [71] in R software (R Foundation for Statistical Computing) indicated that an initial sample size of 300 mothers (100 per condition) would have 0.90 power to detect a moderate-size rate of increase in vaccine intention (Cohen d=0.50). Retention was achieved by alerting mothers to upcoming posttests and compensating mothers for assessments (US $20 for baseline, US $10 for each posttest). Mothers also received 1 raffle entry for every survey completed in drawings for 20 US $100 gift cards after the final posttest.

Ethical Considerations

Mothers provided informed consent online before completing the baseline survey. The study procedures were approved by the Western Institutional Review Board (1-872442-1).

Intervention

The intervention contained 45 Facebook posts related to COVID-19 (5, 11.1%, per week) designed by the research team based on the EPPM [39] and SCT [65]. Posts addressed 4 topics: the 2 primary outcomes (NPIs and COVID-19 vaccination), digital and media literacy, and mother–daughter communication. These topics were rotated across weekdays by week to ensure that all topics had the same likelihood of being viewed. Posts on digital and media literacy were included to combat misinformation related to NPIs and vaccines by addressing source credibility, fact checking, lateral reading, sharing of posts with family/friends, social media algorithms, rebutting of misinformation, and deep fake videos [72-74]. Posts encouraged mothers to talk with teen daughters about the pandemic and promote prevention behavior [66-68] and sought to improve this communication by teaching skills, such as active listening, self-disclosure, empathy, and conflict management. Across these topics, posts addressed theoretic antecedents, including risk from COVID-19 (ie, severity and susceptibility), self-efficacy and response efficacy of NPIs and vaccination, descriptive norms for NPIs and vaccination, behavioral capability (knowledge of risks of COVID-19 and skills to

XSL•FO
RenderX

practice NPIs), and observational learning (stories about dangers of COVID-19 and skills related to NPIs, vaccination, and family communication). To increase mothers' engagement, posts encouraged mothers to react to (eg, like) and comment on posts, for example, by asking a question to solicit the mothers' own experience and opinions on a topic. Additionally, 12 posts provided study information or were aimed at engaging mothers with holiday plans, favorite books, family traditions, and recipes.

Each experimental post contained the same content in all 3 groups. The experimental manipulation of information sources was accomplished by linking each message in the posts to additional online content (eg, articles, blog posts, infographics, or videos) from either a government agency (eg, the CDC or the World Health Organization [WHO]), a near-peer parent, or news media. For the near-peer parent group, information was sourced primarily from Twitter, Instagram, Facebook, TikTok, and parenting blog posts or magazines. Near-peer parents were predominantly women. The term "near-peer" was used to reflect that these sources were similar to the participants, being obviously parents (although a few were female journalists, college professors, or nurses), and were selected to be close to the age of the sample (range 28-64 years, mean 42.7, SD 6.7). However, these sources were unlikely to be known personally by participants, as might be a "peer." For news media, content was sourced from 22 media organizations that focused on delivering news to the general public or a target public. Since individuals can differ in the credibility they assign to various news media, we selected content from news media that ranged from moderately conservative (eg, Fox News and the *New York Post*) to middle-of-the-road (eg, USA Today and *Newsweek*) to moderately liberal (eg, *Washington Post* and ABC), as ranked by All Sides Media Bias [77]. The research team confirmed that all links and content from information sources were accurate. Some of the content from the source was embedded in the experimental post (eg, infographic or screenshot), but a link was always provided to the information source.

Posts were developed by the investigators using an agile development process to reflect the rapidly changing pandemic information environment and ensure content was timely and relevant. Mothers (n=30, 9.9%) participated in virtual focus groups before and during the intervention to review and provide feedback on sample posts. Initially, 2 weeks of posts were prepared, after which new posts were developed weekly. All posts were reviewed by 4 of the investigators (authors DB, BW, WGW, SP), the project manager, and the community manager for readability, theoretical principles, accuracy, and information source prior to posting.

Posts were scheduled by the community manager. They appeared at 10:00 a.m. on Monday, Wednesday, and Friday and 7:00 p.m. on Tuesday and Thursday (1 post per day). Posting times were based on analytics from our prior study regarding the most popular times to view posts [69]. The initial post welcomed participants to the group, invited them to join in discussing the posts, and asked them to be respectful of other group members during discussions and to maintain the privacy of other participants when they communicated about content in the posts with family and friends outside the group. Posts on the 4 topics (NPIs, vaccination, digital and media literacy, and mother–daughter communication) appeared each week (1 post on each of the 3 topics and 2 posts on 1 topic in a week; topics with 2 posts were rotated across the weeks). On Wednesdays, an additional engagement post was published (n=12) to balance the seriousness of the pandemic topics and help keep mothers engaged. The community manager followed a protocol to monitor mothers' reactions and comments to each post and respond to any uncertainty or misinformation or requests for additional information from mothers. Responses had a respectful, empathy-driven, reflective-listening approach toward the mothers [76] that acknowledged the mothers' comments, advised them to follow local and national COVID-19 guidelines, and included links to government agencies, professional groups (eg, the American Diabetes Association), and news media.

Measures

All measures were self-reported by mothers and collected using Qualtrics survey software (see Multimedia Appendix 1).

Primary Outcomes

The primary outcomes, assessed at pretest and all posttests, were social distancing behaviors by self and daughters (self: α=.76 [baseline], .76 [week 3], .79 [week 6], .76 [week 9]; daughters: α=.76 [baseline], .72 [week 3], .78 [week 6], .78 [week 9]) [45,77,78] and mothers' intentions to vaccinate self and daughters for COVID-19 [79]. The vaccine intention questions were modified to use a 0-100 scale (0=definitely would not get the vaccine to 50=unsure whether I would get the vaccine to 100=definitely would get the vaccine) to maximize heterogeneity in responses and avoid forcing participants to choose among a finite set of categories. The intention scores were bimodal, so we divided responses into 5 categories based on the raw data plots: 1=0-20, 2=21-40, 3=41-60, 4=61-80, and 5=81-100. In the 9-week posttest, mothers were also asked whether they had received a COVID-19 vaccination; if vaccinated, mothers' vaccine intention was coded as 100.

Theoretic Antecedents

Theoretic antecedents from the EPPM and SCT were assessed, including perceived risk of COVID-19 (severity α=.86, susceptibility α=.72), self-efficacy for NPIs [45,80] and COVID-19 vaccination (α=.72-.73 [baseline], .59-.67 [week 3], .69-.67 [week 6], .58-.62 [week 9]) [81], and response efficacy (α=.91 [baseline], .92 [week 3], .80 [week 6], .89 [week 9]) response cost (α=.71 [baseline], .74 [week 3], .70 [week 6], .70 [week 9]) for COVID-19 NPIs [45].

Mother–Daughter Communication

Mother–daughter communication about COVID-19 was measured using a scale modified from the original trial [69,70], which asked whether they had discussed the 7 topics about COVID-19 with their daughters (α=.70 [baseline], .75 [week 3], .83 [week 6], .80 [week 9]).

Source Credibility

The credibility of the government agency, near-peer parents, and news media for COVID-19 information was assessed in 2 ways. At baseline, mothers rated the credibility of these 3 information sources on trustworthy, accurate, and bias (α=.79

XSL•FO
RenderX

[government], .76 [near-peer parent], .55 [news media]) [82]. In each posttest, mothers used these same items to rate 1-2 posts from their assigned group in the preceding 3 weeks (α=.60 [week 3], .64 [week 6], .63 [week 9, media literacy], .77 [week 9, mother–daughter communication]). Posts on social distancing (week 3), vaccination (week 6), media literacy (week 9), and mother–daughter communication (week 9) were presented at random.

Media Use

Mothers' media use was assessed at baseline. Mothers were asked about exposure to COVID-19 messages in the media (α=.91) [83]. They also reported the number of hours in a typical day they used any media to obtain news and information and used any media to inform themselves about COVID-19 [84]. Mothers completed measures on COVID-19 information overload (α=.76) and excessiveness (α=.60).

Mothers' Characteristics

Finally, individual differences among mothers on political leaning (conservative, middle-of-the-road, or liberal), history of COVID-19 infection (Do you believe you had COVID-19, and have you ever received a test to check for COVID-19 infection?) [78], vaccination antecedents (α=.82) [85], demographics (ie, race, Hispanic ethnicity, age, and education), urbanization of home county (from US Census), and health insurance status of self and daughter [86] were obtained from the original trial or the baseline survey.

Engagement With Social Media Messages

Engagement with the Facebook posts was recorded in 3 ways. Mothers' reactions (eg, like, love, wow, angry, and sad) and comments on all posts were extracted in the identified format using a customized program and counted. The number of views per post was recorded. Mothers reported whether they read posts on COVID-19, whether they felt connected to the group, and whether they shared/communicated about the posts on COVID-19 in the final posttest.

Acceptability of the Facebook Group

Finally, acceptability of the social media messages in the Facebook private group was evaluated in postintervention focus groups via 3 questions:

- What did you like most about the Facebook group?
- What did you like least about the Facebook group?
- What did you learn from the Facebook group?

Recordings of focus group discussions were reviewed and coded using a conventional content analysis protocol [87]. Two trained coders independently classified responses, and discussion was used to achieve consensus on disagreements. Interrater reliability was adequate (κ=0.78-0.87) [88]. We summarized the frequency of themes.

Statistical Analysis

Two sets of analyses were conducted to test the prespecified hypotheses and research questions. In the first set, a series of mixed effects growth models were used to model change in each of 4 primary outcomes (mothers' reports of social distancing behavior and vaccine intentions by self and daughters), 8 theoretic antecedent outcomes (perceived risk [severity and susceptibility], response efficacy and cost of NPIs, self-efficacy for NPIs and vaccination [self and daughters]), and mother–daughter communication in the hypotheses. Each outcome (measured over 4 occasions) was regressed on time (centered at 9 weeks), effect codes for treatment, and time-by-effect-code interactions. Random effects for the intercept and slope were included and specified to correlate. With the effect codes, estimates for the intercept (centered at week 9) and slope for time (rate of change in the outcome over time) represented the average of these estimates for the 3 conditions, rather than 1 single reference group as with dummy codes. An ordinal mixed effects model was fit for intentions to vaccinate, and a linear mixed effects model was fit for the other 13 outcomes. In the second set of analyses, 4 mixed effects models for social distancing behavior and vaccine intentions were examined to test the moderating effect of engagement with the social media feed to test the second research question. All models included all possible interactions between time, condition, and the moderator, with the treatment condition represented by effect codes. Therefore, simple effects for time and the moderators represented the average effect across the 3 conditions.

Next, a set of exploratory analyses were performed. Analyses fit mixed effects models to explore 4 additional possible moderators: baseline source credibility, COVID-19 media consumption, political leaning on social distancing behavior and vaccine intentions, and baseline vaccine intentions on follow-up vaccine intentions. Mothers' averaged interim credibility ratings of 4 posts from the Facebook private groups were examined as a moderator of treatment effects on social distancing behavior and vaccine intentions measured at week 9. A linear model was fit for social distancing behaviors and an ordinal regression model for intentions, regressing them on treatment (represented as 2 effect codes), post credibility, interaction of treatment and post credibility, baseline rating of credibility of the assigned treatment condition, and baseline rating of the outcome.

Results

Profile of the Sample

Overall, 303 mothers were enrolled (n=100, 33.0%, in the government agency group; n=99, 32.7%, in the near-peer parent group; n=104, 34.3%, in the news media group). Mothers were middle aged (range 28-64 years); well educated, with 160 (55.7%) completing college; and moderately affluent, with 150 (56.4%) having incomes over US $80,000 (see Tables 1-3). Nearly all were non-Hispanic White, because the original trial aimed at preventing indoor tanning. Mothers had diverse political leaning, and the majority lived in states with Republican governors. About 1 (22%) in 5 mothers believed that they had COVID-19 in the past, and nearly half (n=155, 51.3%) had been tested (n=30, 9.9%, had tested positive). At baseline, 199 (65.7%) of the participants lived in states with a mask mandate, and most states were limiting vaccination to older individuals (aged 46.1 years on average). There were no statistically

XSL•FO
RenderX

significant differences between the participants' characteristics by treatment group at baseline.

The retention of mothers was high. Nearly all mothers (n=298, 98.3%) remained in the Facebook private groups throughout the 9-week period (ie, did not actively "unfriend" themselves from the private group). Similarly, 276 (91.1%) completed the 3-week posttest, 273 (90.1%) completed the 6-week posttest, and 275 (90.8%) completed the 9-week posttest, while 244 (80.5%) completed all assessments; see the Consolidated Standards of Reporting Trials (CONSORT) diagram in Figure 1.

Mothers appeared to engage with the 57 messages posted to each Facebook private group. On average, mothers viewed over 35 posts (government mean 36.79 [SD 20.45], near-peer parents mean 37.30 [SD 8.99], news media mean 40.38 [SD 24.20]) and posted reactions or comments on over 10 of the posts (government mean 11.46 [SD 18.57], near-peer parents mean 10.23 [SD 16.51], news media mean 11.41 [SD 17.37]).

Table 1. Demographic characteristics of participants by treatment group.

Characteristics	Overall (N=303)	Treatment group		
		Government agency (n=100)	Near-peer parents (n=99)	News media (n=104)
Age (years), mean (SD)	42.8 (6.7)	42.7 (6.6)	42.8 (6.8)	42.8 (6.8)
Ethnicity, n (%)				
Hispanic	19 (6.3)	10 (10.0)	4 (4.0)	5 (4.8)
Non-Hispanic	284 (93.7)	90 (90.0)	95 (96.0)	99 (95.2)
Race, n (%)				
American Indian/Alaska Native	3 (1)	1(0.3)	1 (0.3)	1 (0.3)
Asian	4 (1.3)	0 (0)	4 (4.0)	0 (0)
Black/African American	23 (7.6)	7 (7)	8 (8.1)	8 (7.7)
White	264 (87.1)	90 (90)	83 (83.8)	91 (87.5)
Other	5 (1.7)	1 (1.0)	1 (1.0)	3 (2.9)
More than 1 race	4 (1.3)	2 (2.0)	1 (1.0)	1 (1.0)
Education, n (%)				
High school or less	22 (7.7)	6 (6.2)	5 (5.3)	11 (11.3)
Some education beyond high school	105 (36.6)	35 (36.5)	39 (41.5)	31 (32.0)
4-year college graduate	81 (28.2)	26 (27.1)	26 (27.7)	29 (29.9)
Postgraduate education	79 (27.5)	29 (30.2)	24 (25.5)	26 (26.8)
Total annual household income (US $), n (%)				
20,000 or less	13 (4.9)	2 (2.3)	6 (7.0)	5 (5.4)
20,001-40,000	32 (12.0)	12 (13.6)	7 (8.1)	13 (14.1)
40,001-60,000	38 (14.3)	9 (10.2)	15 (17.4)	14 (15.2)
60,001-80,000	33 (12.4)	16 (18.2)	7 (8.1)	10 (10.9)
80,001-100,000	49 (18.4)	19 (21.6)	14 (16.3)	16 (17.4)
More than 100,000	101 (38.0)	30 (34.1)	37 (43.0)	34 (37.0)

XSL•FO
RenderX

Table 2. COVID-19 prevention and history characteristics of participants by treatment group.

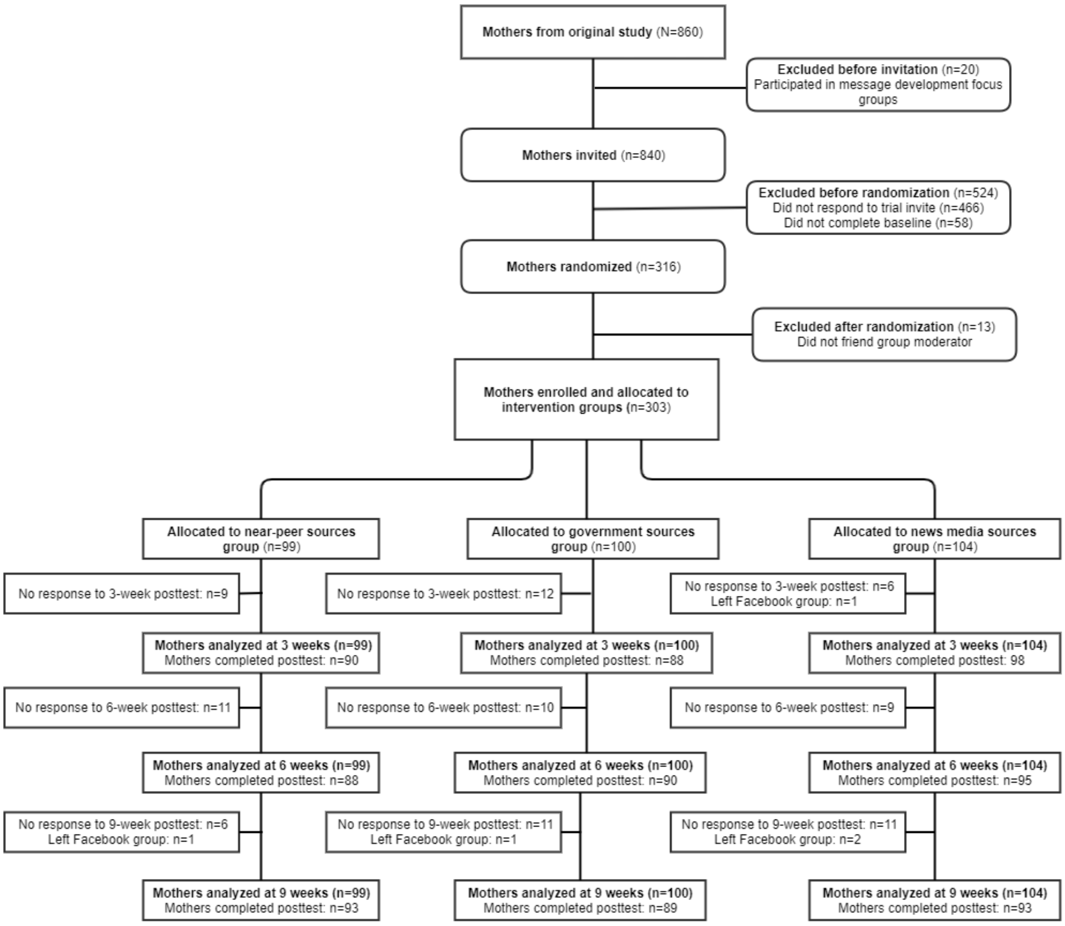
Characteristics	Overall (N=303)	Treatment group		
		Government agency (n=100)	Near-peer parents (n=99)	News media (n=104)
Statewide mask mandate in state of residence, n (%)				
Yes	199 (65.7)	71 (71.0)	57 (57.6)	71 (68.3)
No	104 (34.3)	29 (29.0)	42 (42.4)	33 (31.7)
Age eligibility for COVID-19 vaccine (years), mean (SD)	46.1 (17.7)	46.1 (17.7)	47.7 (17.2)	43.9 (18.2)
Have you ever received a test to check for COVID-19 infection?, n (%)				
Yes, tested positive	30 (9.9)	12 (12.0)	10 (10.2)	8 (7.7)
Yes, tested negative	123 (40.7)	46 (46.0)	40 (40.8)	37 (35.6)
Yes, still waiting for test results	2 (0.7)	1 (1.0)	0 (0.0)	1 (1.0)
No	147 (48.7)	41 (41.0)	48 (49.0)	58 (55.8)
Do you believe that you have had COVID-19?, n (%)				
Yes	67 (22.2)	25 (25.0)	22 (22.4)	20 (19.2)
No	197 (65.2)	63 (63.0)	63 (64.3)	71 (68.3)
I don’t know	38 (12.6)	12 (12.0)	13 (13.3)	13 (12.5)

Table 3. Political ideology characteristics of participants by treatment group.

Characteristics	Overall (N=303)	Treatment group		
		Government agency (n=100), n (%)	Near-peer parents (n=99), n (%)	News media (n=104), n (%)
Political leaning				
Conservative	72 (24.4)	25 (25.2)	25 (25.8)	22 (22.2)
Middle-of-the-road	148 (50.2)	54 (54.6)	48 (49.5)	46 (46.5)
Liberal	75 (25.4)	20 (20.2)	24 (24.7)	31 (31.3)
Political affiliation of governor of state of residence				
Democratic	115 (38.0)	44 (44.0)	31 (31.3)	40 (38.5)
Republican	188 (62.0)	56 (56.0)	68 (68.7)	64 (61.5)

XSL•FO
RenderX

Figure 1. CONSORT diagram for trial. CONSORT: Consolidated Standards of Reporting Trials.



Hypothesis 1 Test: Change in Social Distancing and Vaccine Intentions

At baseline, most mothers reported that they and their daughters were engaging in a moderate to high levels of social distancing (Table 4). Mothers' reports of social distancing by both themselves and daughters decreased over time when examining all 3 posttests relative to baseline (Table 5), disconfirming H1.

About half of the mothers had high vaccine intentions for themselves and their daughters, but up to one-quarter expressed low vaccine intentions (Table 4). Vaccine intentions for self and daughters increased over time (Table 5), supporting H1. However, vaccine intentions were bimodally distributed, with large groups of mothers consistently indicating low (<20.00 likelihood) and high (80.00 likelihood) intentions across all 4 time points. Thus, baseline vaccine intentions were split into 3 groups (low<20.00, moderate=20.00-79.00, and high80.00 likelihood) and tested as a moderator of change in the 5-level vaccine intention measure in the 3 posttests. There was a statistically significant improvement in vaccine intentions for self (b=0.76, 95% CI 0.31-1.21, *P*<.01) and daughters (b=0.48, 95% CI 0.06-0.89, *P*=.02) over time among mothers with moderate intentions at baseline. Likewise, there was a statistically significant increase in vaccine intention for self (b=9.21, 95% CI 6.60-11.82, *P*<.001) and daughters (b=5.51, 95% CI 3.78-7.23, *P*<.001) by the 9-week posttest among mothers with high baseline intentions. Mothers with low baseline vaccine intentions reported lower vaccine intention for self (b=–5.99, 95% CI –8.03 to –3.95, *P*<.001) and daughters (b=–4.83, 95% CI –6.69 to –2.97, *P*<.01) in the 9-week posttest.

XSL•FO
RenderX

Table 4. Percentage of mothers (N=303) reporting social distancing and vaccine intentions for themselves and daughters at baseline.

Ratings	Themselves, n (%)	Daughters, n (%)
Social distancing		
Low (rating=1.00-2.33)	12 (4.0)	8 (2.6)
Moderate (rating=2.34-3.66)	104 (34.3)	117 (38.7)
High (rating=2.67-5.00)	187 (61.7)	178 (58.7)
Intention to vaccinate		
Low (likelihood=0-20)	73 (24.5)	67 (22.6)
Moderate (likelihood=21-80)	73 (24.5)	94 (31.6)
High (likelihood=81-100)	152 (51.0)	136 (45.8)

Table 5. Results of regression analyses of a change in primary outcomes and theoretic mediators over time from baseline across 3-, 6-, and 9-week posttests.

	b	95% CI	*P* value
Social distancing			
Mother	–0.10	–0.12 to –0.08	<.001
Daughter	–0.10	–0.12 to –0.03	<.001
Intent to vaccinate			
Mother	0.34	0.19-0.49	<.001
Daughter	0.17	0.04-0.29	.01
Self-efficacy for NPIs[a]	0	–0.03 to 0.03	.96
Self-efficacy for vaccination			
Mother	0.08	0.05-0.12	<.001
Daughter	0.05	0.01-0.08	<.01
Response efficacy for NPIs	0.01	–0.02 to 0.03	.59
Response cost for NPIs	–0.03	–0.05 to 0.00	.02
Perceived risk			
Severity	0.04	0.01-0.07	.01
Susceptibility	–0.03	–0.06 to 0.00	.04
Mother–daughter communication	–0.02	–0.06 to 0.01	.16

[a]NPI: nonpharmaceutical intervention.

Hypothesis 2 Test: Change in Theoretic Antecedents and Mother–Daughter Communication

Several theoretic antecedents improved over time (Table 4), largely supporting H2. Specifically, self-efficacy for vaccination of self and daughters increased, and response costs for NPIs decreased. There was also some evidence that perceived risk increased over time, particularly with the severity of COVID-19 increasing over time; however, perceived susceptibility declined over time. By contrast, self-efficacy and response efficacy for NPIs did not change, nor did mother–daughter communication (Table 5), contrary to the hypothesis.

Differences Among Information Sources

Effect of Treatment Group

Only 1 outcome was moderated by the experimental manipulation of information sources. The decline in social distancing by daughters over time was greater when mothers were in the near-peer parents group (b=–0.04, 95% CI –0.07 to 0.00, *P*=.03) and lesser when mothers were in the government agency group (b=0.05, 95% CI 0.02-0.09, *P*=.003); see Table 6. Interactions between treatment group and time were not statistically significant for social distancing by mothers (near-peer parents: b=–0.01, 95% CI –0.03 to 0.02, *P*=.66; government agency: b=0.01, 95% CI –0.02 to 0.04, *P*=.51) and mother–daughter communication (near-peer parents: b=–0.03, 95% CI –0.08 to 0.02, *P*=.22; government agency: b=0.02, 95% CI –0.03 to 0.06, *P*=.51); see Table 7.

The information source moderated the improvement in mothers' own vaccine intentions in the analysis treating baseline vaccine intentions as a moderator. The increase in mothers' vaccine intentions among those who had high intentions at baseline was attenuated in the government agency source condition, both for change across all 3 posttests (b=–1.47, 95% CI –2.74 to –0.20,

XSL•FO
RenderX

P=.02) and at the 9-week posttest (b=–3.17, 95% CI –5.91 to –0.43, *P*=.02).

Table 6. Means (SD) of social distancing behavior and vaccine intention measures by treatment condition and time of assessment.

Outcome and source	Baseline	3-week posttest	6-week posttest	9-week posttest
Mothers' social distancing				
Government agency	3.90 (0.77)	3.80 (0.80)	3.72 (0.83)	3.62 (0.79)
Near-peer parents	3.87 (0.76)	3.74 (0.84)	3.67 (0.89)	3.56 (0.86)
News media	3.97 (0.68)	3.84 (0.76)	3.75 (0.83)	3.65 (0.86)
Daughters' social distancing				
Government agency	3.77 (0.70)	3.79 (0.74)	3.68 (0.75)	3.66 (0.75)
Near-peer parents	3.86 (0.72)	3.74 (0.71)	3.58 (0.83)	3.46 (0.87)
News media	3.98 (0.72)	3.82 (0.76)	3.68 (0.84)	3.64 (0.89)
Vaccine intentions for self				
Government agency	3.46 (1.78)	3.38 (1.80)	3.53 (1.75)	3.69 (1.68)
Near-peer parents	3.70 (1.64)	3.63 (1.71)	3.82 (1.64)	3.86 (1.65)
News media	3.70 (1.65)	3.66 (1.75)	3.76 (1.74)	3.80 (1.72)
Vaccine intentions for daughters				
Government agency	3.49 (1.71)	3.52 (1.72)	3.56 (1.66)	3.71 (1.61)
Near-peer parents	3.60 (1.59)	3.50 (1.69)	3.60 (1.63)	3.77 (1.62)
News media	3.66 (1.64)	3.66 (1.65)	3.75 (1.66)	3.74 (1.60)

XSL•FO
RenderX

Table 7. Means (SD) of secondary outcome measures by treatment condition and time of assessment.

Outcome and source	Baseline	3-week posttest	6-week posttest	9-week posttest
Perceived risk: severity				
Government agency	4.34 (0.85)	4.42 (0.80)	4.42 (0.88)	4.52 (0.72)
Near-peer parents	4.33 (0.89)	4.28 (0.99)	4.34 (0.79)	4.46 (0.80)
News media	4.36 (0.70)	4.49 (0.74)	4.45 (0.78)	4.44 (0.80)
Perceived risk: susceptibility				
Government agency	3.56 (0.86)	3.46 (0.99)	3.44 (0.90)	3.54 (0.96)
Near-peer parents	3.49 (0.96)	3.37 (0.98)	3.43 (0.81)	3.40 (0.92)
News media	3.50 (0.76)	3.56 (0.77)	3.42 (0.81)	3.28 (0.89)
Response efficacy of NPIs[a]				
Government agency	4.48 (0.66)	4.56 (0.68)	4.42 (0.63)	4.57 (0.62)
Near-peer parents	4.51 (0.76)	4.56 (0.74)	4.41 (0.76)	4.53 (0.70)
News media	4.55 (0.71)	4.54 (0.53)	4.68 (0.50)	4.54 (0.66)
Response cost for NPIs				
Government agency	4.43 (0.65)	4.45 (0.68)	4.40 (0.72)	4.39 (0.63)
Near-peer parents	4.49 (0.69)	4.45 (0.73)	4.41 (0.72)	4.42 (0.80)
News media	4.38 (0.78)	4.37 (0.80)	4.40 (0.78)	4.26 (0.90)
Self-efficacy for NPIs				
Government agency	4.35 (0.67)	4.40 (0.64)	4.34 (0.73)	4.35 (0.69)
Near-peer parents	4.28 (0.79)	4.22 (0.80)	4.27 (0.80)	4.22 (0.79)
News media	4.19 (0.84)	4.28 (0.80)	4.30 (0.80)	4.27 (0.82)
Self-efficacy for vaccinating mothers				
Government agency	3.88 (1.01)	3.95 (1.01)	4.03 (1.10)	4.19 (0.97)
Near-peer parents	4.15 (0.86)	4.07 (0.94)	4.19 (0.90)	4.32 (0.79)
News media	3.89 (1.10)	4.00 (1.03)	4.16 (1.03)	4.15 (0.97)
Self-efficacy for vaccinating daughters				
Government agency	3.83 (0.98)	3.80 (0.99)	3.98 (1.10)	4.06 (1.01)
Near-peer parents	4.02 (0.89)	3.95 (1.03)	4.06 (0.91)	4.05 (0.90)
News media	3.85 (1.04)	3.93 (1.04)	3.99 (1.04)	3.95 (1.03)
Mother–daughter communication about COVID-19				
Government agency	3.50 (0.86)	3.28 (0.98)	3.39 (1.08)	3.43 (1.06)
Near-peer parents	3.65 (0.82)	3.42 (0.97)	3.51 (1.06)	3.44 (1.08)
News media	3.62 (0.85)	3.45 (0.98)	3.50 (1.09)	3.57 (0.96)

[a]NPI: nonpharmaceutical intervention.

Moderation by Perceived Credibility of the Assigned Information Source

Approximately one-third of the mothers considered the assigned information source to be credible in general at baseline (government agency: n=100, 33.0%; near-peer parents: n=99, 32.7%; news media: n=104, 34.3%). Perceived credibility was associated with an increase in social distancing and vaccine intentions over time. Mothers who rated the assigned information source as credible reported greater social distancing for self (b=0.29, 95% CI 0.09-0.49, *P*<.01) and daughters (b=0.31, 95% CI 0.11-0.51, *P*<.01) and higher vaccine intentions for self (b=4.18, 95% CI 1.83-6.53, *P*<.001) and daughters (b=3.36, 95% CI 1.67-5.04, *P*<.001) at the 9-week posttest. However, these improvements in social distancing and vaccine intentions associated with source credibility were attenuated substantially in the near-peer parents condition (credibility × condition: social distancing, self: b=–0.41, 95% CI –0.68 to –0.14, *P*<.01 and daughters: b=–0.32, 95% CI –0.59 to –0.04, *P*=.02; vaccine intentions, self: b=–4.20, 95% CI –7.53 to –0.87, *P*=.01 and daughters: b=–2.85, 95% CI –5.12 to –0.58, *P*=.01). Moreover, mothers' intentions to vaccinate self may have

XSL•FO
RenderX

increased when they considered the near-peer parents to be not credible (b=–0.50, 95% CI –0.99 to –0.01, *P*=.05).

The higher perceived credibility of the individual posts rated during the intervention also predicted increased social distancing by daughters (b=0.23, 95% CI 0.04-0.42, *P*=.02) but not mothers (b=0.07, 95% CI –0.09 to 0.23, *P*=.37). It also was associated with greater vaccine intentions for self (b=1.09, 95% CI 0.27-1.91, *P*=.01) but not for daughters (b=0.63, 95% CI –0.09 to 1.35, *P*=.09). However, there were no significant interactions between the credibility of posts and information sources for social distancing for self (credibility × government agency: b=–0.05, 95% CI –0.26 to 0.16, *P*=.62; credibility × near-peer parents: b=0.04, 95% CI –0.20 to 0.29, *P*=.72) and for daughters (credibility × government agency: b=–0.16, 95% CI –0.41 to 0.08, *P*=.19; credibility × near-peer parents: b=0.06, 95% CI –0.22 to 0.35, *P*=.65) or vaccine intentions for self (credibility × government agency: b=0.20, 95% CI –0.84 to 1.23, *P*=.71; credibility × near-peer parents: b=0.42, 95% CI –0.87 to 1.71, *P*=.52) and for daughters (credibility × government agency: b=0.15, 95% CI –0.79 to 1.09, *P*=.75; credibility × near-peer parents: b=–0.52, 95% CI –1.60 to 0.57, *P*=.35).

Effects of Engagement With COVID-19 Social Media Messages

Two measures of exposure to the social media posts, number of views of the posts, and number of reactions and comments to the posts were tested as moderators of the intervention's effects on social distancing and vaccine intentions.

Social Distancing

The number of views of posts by participants did not influence their reports of social distancing by self or daughters, but reports of social distancing by daughters was higher among mothers who had more reactions and comments (b=0.01, 95% CI 0.01-0.01, *P*=.04). There was no evidence that engagement moderated differences among information sources (*P*>.05).

Vaccine Intentions

For views, the increase in vaccine intentions for self over time was attenuated when mothers viewed more posts across all conditions (b=–0.01, 95% CI –0.01 to –0.01, *P*=.01). This attenuation was stronger in the government agency group (self: b=–0.02, 95% CI –0.04 to 0.00, *P*<.001; daughters: b=–0.01, 95% CI –0.01 to –0.01, *P*=.01). By contrast, attenuation of the increase in vaccine intentions was less evident in mothers in the near-peer parents group who had more engagement (self: b=0.02, 95% CI 0.00-0.04, *P*<.01; daughters: b=0.02, 95% CI 0.00-0.04, *P*<.001). Engagement measured by reactions and comments did not affect changes in vaccine intentions (*P*>.05).

Moderation by Baseline Exposure to COVID-19 Media and Political Leaning

Potential moderation of change in social distancing and vaccine intentions by mothers' general exposure to media reporting on COVID-19 and political leaning at baseline was also examined.

Baseline COVID-19 Media Exposure

Baseline exposure to COVID-19 information in news media, averaged across 4 items, was similar across conditions on a 5-point scale (government agency mean 4.11, SD 0.88; near-peer parents mean 4.09, SD 0.91; news media mean 4.01, SD 0.82). Social distancing (self: b=0.46, 95% CI 0.36-0.56, *P*<.01; daughters: b=0.34, 95% CI 0.24-0.44, *P*<.01) and vaccine intentions (self: b=3.87, 95% CI 2.62-5.12, *P*<.001; daughters: b=2.80, 95% CI 1.93-3.66, *P*<.001) were higher at the 9-week posttest among mothers who reported more media exposure at baseline. However, baseline exposure did not affect differences by information source in either outcome.

Political Leaning

Political leaning was normally distributed among mothers within each condition (government agency: conservative n=25, 25.3%, moderate n=54, 54.6%, liberal n=20, 20.2%; near-peer parents: conservative n=25, 25.8%, moderate n=48, 49.5%, liberal n=24, 24.7%; news media: conservative n=22, 22.2%, moderate n=46, 46.5%, liberal n=31, 31.3%). Mothers reported increased social distancing (self: b=0.40, 95% CI 0.28-0.52, *P*<.001; daughters: b=0.31, 95% CI 0.19-0.42, *P*<.001) and vaccine intentions (self: b=3.16, 95% CI 1.49-4.82, *P*<.001; daughters: b=2.37, 95% CI 1.21-3.53, *P*<.001) over baseline at the 9-week posttest when they expressed a more liberal than conservative political leaning. Political leaning moderated differences by information source for reports of social distancing by daughters. Mothers who were more liberal and assigned to the near-peer parents group reported greater social distancing by daughters at the final posttest (b=0.19, 95% CI 0.01-0.37, *P*=.04), while more liberal mothers in the government agency group reported reduced social distancing at the final posttest (b=–0.25, 95% CI –0.43 to –0.07, *P*<.01). Political leaning did not show any other effects on vaccine intentions for self (near-peer parents: b=0.13, 95% CI –0.18 to 0.44, *P*=.43; government agency: b=–0.11, 95% CI –0.42 to 0.20, *P*=.50) or daughters (near-peer parents: b=–0.03, 95% CI –0.30 to 0.24, *P*=.85; government agency: b=0.20, 95% CI –0.09 to 0.49, *P*=.18).

Focus Group Results on Acceptability of the Social Media Messages

Of the 303 participants, 30 (9.9%) randomly selected participants (n=10, 33.3%, per treatment group) attended postintervention focus groups on reactions to the social media messages in the intervention. Coding of the 35 responses about what they liked most about the Facebook group (interrater reliability κ=0.82) revealed that the most common themes were a sense of community (n=15, 43%, responses) and program content or community manager (n=15, 43%, responses), followed by hearing opinions and perspectives that were different from the participants' (n=5, 14%). Of the 30 responses on what the participants liked least about the Facebook group, the most frequent theme was that they did not dislike any aspect of the program (n=14, 47%), followed by hearing opinions that they disagreed with or feeling fearful of offending people who might disagree (n=8, 27%; κ=0.78). A small number of participants (n=5,17%) said they did not remember any content (n=3, 10%, responses were classified as "other"; eg, wished other moms engaged more). Finally, of 39 responses about what they learned in the Facebook group, the mothers more commonly mentioned facts about the vaccine (n=14, 36%), followed by general facts about COVID-19 (n=5, 13%), media

XSL•FO
RenderX

literacy skills (n=5, 13%), and what other moms think about COVID-19 and vaccines (n=4, 10%; κ=0.87). A small number (n=5, 13%) said they had already heard all of the information in the messages, while a few (n=4, 10%) said they did not remember any of the content.

Discussion

Principal Findings

The results of this study must be interpreted within the context of the COVID-19 pandemic during the intervention. The relaxing of restrictions and ramping up of vaccination by March 2021 [22] may have made mothers feel that the risk from COVID-19 was diminishing, reflected in their lower perceived susceptibility to COVID-19 at 9 weeks. The EPPM asserts that health behavior is motivated by perceived risk [39,89], so this declining sense of susceptibility may have caused mothers and daughters to reduce their social distancing, a phenomenon seen in the H1N1 pandemic and other studies on COVID-19 [90-92]. Thus, these contextual factors may explain the failure to support our hypothesis of increased social distancing after the social media messages, which was seen in surveys [93,94]. By contrast, the expanding availability of the vaccine likely increased perceptions that mothers could get vaccinated, which produced greater self-efficacy for vaccination over time. This may have motivated stronger intentions to get vaccinated during the study. However, increased intentions appeared to occur mostly among mothers who had moderate-to-high intentions at baseline, while mothers with initially low intentions became more resistant over time.

The information source linked to the social media messages in the Facebook posts did not have a clear effect on mothers. Government sources may have attenuated the decline in social distancing mothers reported for daughters, while near-peer parents possibly amplified the decline. The government sources selected for the social media messages advocated for social distancing and thus rebutted local government decisions to relax restrictions. In a previous study, attention to government sources improved social distancing behaviors [50]. However, the near-peer parents may have increased participants' decisions to abandon social distancing, despite presenting messages supporting social distancing. It may be that other parents in the mothers' lives were strongly opposed to social distancing and hearing from "parents" in the social media posts made several mothers more aware of the parents' general opposition. By contrast, mothers with initially high intentions to get themselves vaccinated had weaker intentions at the end of the intervention period when receiving information from government sources. Their intentions could have declined because many of these mothers were vaccinated during the study, making intentions less relevant. Other studies have found that social media and online sources have limited impacts on perceptions related to COVID-19 prevention and sometimes result in lower knowledge [45,48,51]. Past research showed that in the United States, news media preferences affected COVID-19 knowledge and altered COVID-19 prevention behaviors, when comparing conservative news media outlets with outlets with more moderate or liberal political views [95,96]. We attempted to control these varying preferences by using randomization and linking to news media with different political perspectives from moderately liberal to moderately conservative. However, the heterogeneity of perceptions may have made it difficult to discern a consistent effect in the news media condition.

The intervention's social media messages seemed to affect mothers when they contained information sources that mothers considered credible, regardless of which source they received. Similarly, a recent study found that trust in specific sources of information on the pandemic results in higher COVID-19 health literacy [49]. Past research showed that risk communication must build trust in the government, medical organizations, and science to improve adherence to protection measures [97-99]. Consistent with the EPPM [39], information from high-credibility sources may make it more difficult to engage in fear control to reduce perceived severity, which increased during the intervention, through source derogation and dismissal. Instead, it may have motivated mothers to take steps to control the danger through social distancing and vaccinations, especially when perceived response costs declined.

The findings of this trial suggest that when using social media to improve COVID-19 prevention behaviors and vaccine uptake, campaign planners should, as a general strategy, select sources that recipients feel are trustworthy and accurate and construct messages that maintain these perceptions of high credibility. The sense of community cited by several mothers in follow-up interviews as something they liked about the private groups might have contributed to credibility, because goodwill toward others has been a dimension of source credibility [100]. In addition, mothers who liked the ability to hear perspectives different from their own may have seen the groups as a safe place to experience differing opinions, again expressing this sense of goodwill. Some mothers were hesitant to offend people who might disagree with their opinions, implying there may have been a norm of civility in the private groups that contributed to credibility as well. However, campaign planners need to avoid information overload, which has been associated with consuming certain sources, and a larger number of sources, which can cause recipients to actively avoid information [45,47,101-103].

The general conclusion that highly credible sources are most effective, however, may not always hold when considering near-peer parents as sources of information about COVID-19 (ie, parents in this case). In this study, mothers who felt near-peer parents were not credible initially may have been more influenced by the social media messages. It may be that mothers who generally considered near-peer parents to be less credible on COVID-19 may have found the near-peer parents included in the experimental posts to be more believable than they expected. Prior research has shown that individuals who argue for a position that they are not expected to hold are more influential, especially when the arguments are high quality [104]. In addition, a positive violation of expectations in persuasive messages can make individuals appear more credible and hence persuasive [105-108].

The finding that regardless of the information source, mothers' engagement with the social media messages in the Facebook

XSL•FO
RenderX

private groups was associated with an attenuated reduction in social distancing was consistent with other studies in which engagement improved social media's and other digital interventions' effectiveness [109-113]. However, engagement effects in this trial may have been limited by the generally high degree of exposure mothers had in all groups. Engaging with a social media intervention may be different from engaging with other forms of digital interventions, such as websites or online training. Reactions and comments are considered more involved engagement than just viewing posts, as the former represents conversation that may be more intrinsically engaging, while the latter is merely information consumption [114]. Viewing was more common than reacting and commenting in this study, and the 2 forms of engagement may have different motivations. Views may reflect information needs, while reaction or comments may fulfill social needs [115-117]. It is important to note that views, reactions, and comments are behavioral measures of online engagement, but researchers have recently argued that engagement is multidimensional and involves emotional and cognitive experiential processes that are better captured with self-reporting and other measures [118-120]. For example, mothers may have viewed a post and then discussed it with friends or family. Simply frequent, sustained online behavioral engagement may not capture the complex nature of engagement. There is a need to identify what constitutes effective engagement with social media [119,121].

Finally, 2 other contextual trends were apparent in the study results. Mothers who had paid more attention to COVID-19 information in the media prior to the study had higher social distancing and vaccine intentions by the final posttest. It may be that greater attention to the COVID-19 information environment provided mothers with more information that promoted COVID-19 prevention, including vaccine intentions. A recent study found that individuals with high perceived COVID-19 risk and greater prevention behaviors reported consuming information on COVID-19 from multiple sources [60]. Finally, there is ample evidence, including in this trial, that conservative political leaning is a major barrier to COVID-19 prevention [122,123]. This appears to be a robust tendency unaffected by different information sources.

Limitations and Strengths

The trial had some limitations. The design lacked a control group that did not receive messages on COVID-19, which made it challenging to determine whether the messages affected social distancing and vaccine intentions irrespective of the information source. The short duration of the intervention may have achieved only small effects. Although the sample was moderate in size and from a number of US states, generalizability was limited by enrolling mothers of teen daughters who may have been more attentive to the social media messages because they had elevated concerns about COVID-19 risks for their families. Whether individuals who are not parents would be affected in the same way is unknown. Mothers had already participated in a trial on other adolescent health topics, so the sample may have been biased to mothers with high interest in adolescent health. Most mothers were originally recruited from the Qualtrics survey panel, which tends to have a relatively high socioeconomic status, and nearly all mothers were non-Hispanic White because of the original trial's focus on indoor tanning. Although we varied the source of information contained in the posts, all posts were delivered through the Facebook platform, making it the primary source of the intervention and possibly undermining the experimental comparison. The multiple posttest measures may have introduced a testing effect (ie, reactivity) that increased the mothers' attention to the experimental messages because they knew they would be assessed every 3 weeks. All assessments were self-reporting, although many outcomes were intrapsychic processes (eg, perceptions, opinions, and intentions) measurable only through reports from mothers. We did use published scales, when available.

These limitations were offset somewhat by strengths of the study. Mothers were enrolled and pretested prior to the intervention, allowing for prospective tests of social media's effects, and were randomly assigned to 3 prominent sources of pandemic information, which improved the validity of these comparisons. A mixed methods approach was used to understand the impact of the social media messages on mothers. Finally, multiple posttests provided information on changes produced by the intervention over time.

Conclusion

There were several lessons learned to inform future trials using social media interventions. The group size of approximately 100 mothers was sufficient to achieve high viewership and active participation by group members over 9 weeks, although, as noted, the COVID-19 topic may have been generally interesting to them. Future studies should test how long engagement with a social media intervention can be sustained. In our parent trial with messages on general adolescent health topics, engagement declined over the first 6 months [70]. Participants were willing to remain in the group once they joined it, increasing the likelihood that the social media messages reached and affected them. Many large social media feeds are curated, and it required substantial time to manage the experimental Facebook groups, at least 10 hours a week by the community manager. The community manager played an important role in engaging participants by personalizing the experimental messages by highlighting that she was a mother and showing her picture.

Social media has been a source of information and misinformation even before the COVID-19 pandemic, but concerns over its role in the pandemic have been elevated as millions of Americans have been exposed to deceptive information, which some people can find believable [24,31,76,124,125]. Social media can affect vaccine-related decisions [126-128], and experts and researchers have called for efforts to correct information on social media [25,32,33,129]. In this context, the trial showed that a series of social media messages can be used to support pandemic responses when posts are based on health behavior change theories and information sources are tailored to the audiences' existing credibility beliefs.

XSL•FO
RenderX

Acknowledgments

The authors thank Christie Idiong and Haley Troy for their assistance with coding the focus group transcripts. This research was funded by the National Cancer Institute (Grant CA192652).

Conflicts of Interest

DB receives a salary from Klein Buendel, Inc., and his spouse is an owner of Klein Buendel. AK, BW, WGW, and JB receive a salary from Klein Buendel, Inc. SP, KH, KB, JD, and JH have no conflicts to declare.

Multimedia Appendix 1
Measurement scales.
[DOCX File , 18 KB - infodemiology_v2i2e36210_app1.docx]

References

1. Centers for Disease Control and Prevention (CDC). Coronavirus Disease 2019: Protect Yourself. 2021. URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2022-06-02]
2. Inglesby TV. Public health measures and the reproduction number of SARS-CoV-2. JAMA 2020 Jun 02;323(21):2186-2187. [doi: 10.1001/jama.2020.7878] [Medline: 32356869]
3. Schoch-Spana M, Brunson EK, Long R, Ruth A, Ravi SJ, Trotochaud M, et al. The public's role in COVID-19 vaccination: human-centered recommendations to enhance pandemic vaccine awareness, access, and acceptance in the United States. Vaccine 2021 Sep 24;39(40):6004-6012 [FREE Full text] [doi: 10.1016/j.vaccine.2020.10.059] [Medline: 33160755]
4. Qualls N, Levitt A, Kanade N, Wright-Jegede N, Dopson S, Biggerstaff M, CDC Community Mitigation Guidelines Work Group. Community mitigation guidelines to prevent pandemic influenza - United States, 2017. MMWR Recomm Rep 2017 Apr 21;66(1):1-34 [FREE Full text] [doi: 10.15585/mmwr.rr6601a1] [Medline: 28426646]
5. Teasdale E, Santer M, Geraghty AWA, Little P, Yardley L. Public perceptions of non-pharmaceutical interventions for reducing transmission of respiratory infection: systematic review and synthesis of qualitative studies. BMC Public Health 2014 Jun 11;14(1):589 [FREE Full text] [doi: 10.1186/1471-2458-14-589] [Medline: 24920395]
6. Eikenberry SE, Mancuso M, Iboi E, Phan T, Eikenberry K, Kuang Y, et al. To mask or not to mask: modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infect Dis Model 2020;5:293-308 [FREE Full text] [doi: 10.1016/j.idm.2020.04.001] [Medline: 32355904]
7. Delen D, Eryarsoy E, Davazdahemami B. No place like home: cross-national data analysis of the efficacy of social distancing during the COVID-19 pandemic. JMIR Public Health Surveill 2020 May 28;6(2):e19862 [FREE Full text] [doi: 10.2196/19862] [Medline: 32434145]
8. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. 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 Jun 01;395(10242):1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
9. Talic S, Shah S, Wild H, Gasevic D, Maharaj A, Ademi Z, et al. Effectiveness of public health measures in reducing the incidence of covid-19, SARS-CoV-2 transmission, and covid-19 mortality: systematic review and meta-analysis. BMJ 2021 Nov 17;375:e068302 [FREE Full text] [doi: 10.1136/bmj-2021-068302] [Medline: 34789505]
10. USC Center for Economic and Social Research. Understanding Coronavirus in America. 2021. URL: https://covid19pulse.usc.edu/ [accessed 2022-06-02]
11. Katz J, Sanger-Katz M, Quealy K. A Detailed Map of Who Is Wearing Masks in the U.S. 2020 Nov 11. URL: https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html [accessed 2022-08-01]
12. Zimet GD, Silverman RD, Fortenberry JD. Coronavirus disease 2019 and vaccination of children and adolescents: prospects and challenges. J Pediatr 2021 Apr;231:254-258 [FREE Full text] [doi: 10.1016/j.jpeds.2020.11.002] [Medline: 33161025]
13. Szilagyi PG, Thomas K, Shah MD, Vizueta N, Cui Y, Vangala S, et al. Likelihood of COVID-19 vaccination by subgroups across the US: post-election trends and disparities. Hum Vaccin Immunother 2021 Oct 03;17(10):3262-3267 [FREE Full text] [doi: 10.1080/21645515.2021.1929695] [Medline: 34170793]
14. Robinson E, Jones A, Lesser I, Daly M. International estimates of intended uptake and refusal of COVID-19 vaccines: a rapid systematic review and meta-analysis of large nationally representative samples. Vaccine 2021 Apr 08;39(15):2024-2034 [FREE Full text] [doi: 10.1016/j.vaccine.2021.02.005] [Medline: 33722411]
15. Centers for Disease Control and Prevention (CDC). CDC Extends Eligibility for COVID-19 Booster Shots. 2021. URL: https://www.cdc.gov/media/releases/2021/p1021-covid-booster.html [accessed 2022-06-02]
16. U.S. Food and Drug Administration. Coronavirus (COVID-19) Update: FDA Takes Additional Actions on the Use of a Booster Dose for COVID-19 Vaccines. 2021. URL: https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-takes-additional-actions-use-booster-dose-covid-19-vaccines [accessed 2022-06-02]
17. Johnson NF, Velásquez N, Restrepo NJ, Leahy R, Gabriel N, El Oud S, et al. The online competition between pro- and anti-vaccination views. Nature 2020 Jun 13;582(7811):230-233. [doi: 10.1038/s41586-020-2281-1] [Medline: 32499650]

XSL•FO
RenderX

18. Roose K. Get Ready for a Vaccine Information War. 2020. URL: https://www.nytimes.com/2020/05/13/technology/coronavirus-vaccine-disinformation.html [accessed 2022-08-01]
19. Centers for Disease Control and Prevention (CDC). COVID Data Tracker. URL: https://covid.cdc.gov/covid-data-tracker/#datatracker-home [accessed 2022-06-02]
20. Federal Reserve Bank of St. Louis. Timeline of Events Related to the COVID-19 Pandemic. URL: https://fraser.stlouisfed.org/timeline/covid-19-pandemic [accessed 2022-06-02]
21. New York Times. See Reopening Plans and Mask Mandates for all 50 States. URL: https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html [accessed 2000-06-02]
22. Centers for Disease Control and Prevention (CDC). COVID Data Tracker: Trends in Number of COVID-19 Cases and Deaths in the US Reported to CDC, by State/Territory. 2021. URL: https://covid.cdc.gov/covid-data-tracker/#trends_dailycases [accessed 2022-06-02]
23. Hamel L, Lopes L, Kearney A, Brodie M. Covid-19 Vaccine Monitor: March 2021. 2021. URL: https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-march-2021/ [accessed 2022-06-02]
24. Mitchell A, Jurkowitz M, Oliphant J, Shearer E. Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable. 2020. URL: https://www.journalism.org/2020/07/30/americans-who-mainly-get-their-news-on-social-media-are-less-engaged-less-knowledgeable/ [accessed 2022-06-02]
25. 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 12;38(2):143-149 [FREE Full text] [doi: 10.1111/hir.12320] [Medline: 32657000]
26. Satariano A, Alba D. Burning Cell Towers, Out of Baseless Fear They Spread the Virus. 2020. URL: https://www.nytimes.com/2020/04/10/technology/coronavirus-5g-uk.html [accessed 2022-06-02]
27. 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]
28. Larson HJ. A call to arms: helping family, friends and communities navigate the COVID-19 infodemic. Nat Rev Immunol 2020 Aug;20(8):449-450 [FREE Full text] [doi: 10.1038/s41577-020-0380-8] [Medline: 32616908]
29. Dunn AG, Surian D, Leask J, Dey A, Mandl KD, Coiera E. Mapping information exposure on social media to explain differences in HPV vaccine coverage in the United States. Vaccine 2017 May 25;35(23):3033-3040 [FREE Full text] [doi: 10.1016/j.vaccine.2017.04.060] [Medline: 28461067]
30. 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]
31. Kreps SE, Kriner D. Medical misinformation in the COVID-19 pandemic. SSRN J 2020:1-22 [FREE Full text] [doi: 10.2139/ssrn.3624510]
32. Chou WS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018 Dec 18;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
33. Yang YT, Broniatowski DA, Reiss DR. Government role in regulating vaccine misinformation on social media platforms. JAMA Pediatr 2019 Nov 01;173(11):1011-1012. [doi: 10.1001/jamapediatrics.2019.2838] [Medline: 31479099]
34. Bode L, Vraga EK. See something, say something: correction of global health misinformation on social media. Health Commun 2018 Sep;33(9):1131-1140. [doi: 10.1080/10410236.2017.1331312] [Medline: 28622038]
35. 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]
36. Walter N, Murphy ST. How to unring the bell: a meta-analytic approach to correction of misinformation. Commun Monogr 2018 May 15;85(3):423-441. [doi: 10.1080/03637751.2018.1467564]
37. Laura J, Tarunjose K, Angelina K, Mojisola O, Enisa S, Valentine V, et al. How storytelling can combat vaccine hesitancy: a transdisciplinary approach. Transdiscipl Insights 2018 Dec 15;2(1):92-103. [doi: 10.11116/tdi2018.2.4]
38. 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]
39. Witte K. Putting the fear back into fear appeals: the extended parallel process model. Commun Monogr 1992 Dec;59(4):329-349. [doi: 10.1080/03637759209376276]
40. Rogers RW. A protection motivation theory of fear appeals and attitude change. J Psychol 1975 Sep;91(1):93-114. [doi: 10.1080/00223980.1975.9915803] [Medline: 28136248]
41. Rogers R. Cognitive and physiological processes in fear appeals and attitude change: a revised theory of protection motivation. In: Cacioppo J, Petty R, editors. Social Psychophysiology. New York, NY: Guilford Press; 1983:153-176.
42. Woodall G, Starling R, Buller D, Kong A, Wheeler C. Beta-test and randomized trial results for GoHealthyGirls: a website for HPV vaccine adoption. 2014 Presented at: 29th Meeting of the International Papillomavirus Conference; August 20-25, 2014; Seattle, WA.
43. Woodall W. Digital interventions to improve HPV vaccine uptake: results and issues. 2018 Presented at: EUROGIN 2018 International Multidisciplinary HPV Congress; December 2-5, 2018; Lisbon, Portugal.

XSL•FO
RenderX

44. Starling R, Nodulman JA, Kong AS, Wheeler CM, Buller DB, Woodall WG. Usability testing of an HPV information website for parents and adolescents. Online J Commun Media Technol 2015 Oct;5(4):184-203 [FREE Full text] [Medline: 26594313]
45. Farooq A, Laato S, Islam AKMN. Impact of online information on self-isolation intention during the COVID-19 pandemic: cross-sectional study. J Med Internet Res 2020 May 06;22(5):e19128 [FREE Full text] [doi: 10.2196/19128] [Medline: 32330115]
46. Mutti-Packer S, Reid JL, Thrasher JF, Romer D, Fong GT, Gupta PC, et al. The role of negative affect and message credibility in perceived effectiveness of smokeless tobacco health warning labels in Navi Mumbai, India and Dhaka, Bangladesh: a moderated-mediation analysis. Addict Behav 2017 Oct;73:22-29 [FREE Full text] [doi: 10.1016/j.addbeh.2017.04.002] [Medline: 28441573]
47. Soroya SH, Farooq A, Mahmood K, Isoaho J, Zara S. From information seeking to information avoidance: understanding the health information behavior during a global health crisis. Inf Process Manag 2021 Mar;58(2):102440 [FREE Full text] [doi: 10.1016/j.ipm.2020.102440] [Medline: 33281273]
48. Wang P, Lu W, Ko N, Chen Y, Li D, Chang Y, et al. COVID-19-related information sources and the relationship with confidence in people coping with COVID-19: Facebook survey study in Taiwan. J Med Internet Res 2020 Jun 05;22(6):e20021 [FREE Full text] [doi: 10.2196/20021] [Medline: 32490839]
49. De Gani SM, Berger FMP, Guggiari E, Jaks R. Relation of corona-specific health literacy to use of and trust in information sources during the COVID-19 pandemic. BMC Public Health 2022 Jan 06;22(1):42 [FREE Full text] [doi: 10.1186/s12889-021-12271-w] [Medline: 34991525]
50. Fridman I, Lucas N, Henke D, Zigler CK. Association between public knowledge about COVID-19, trust in information sources, and adherence to social distancing: cross-sectional survey. JMIR Public Health Surveill 2020 Sep 15;6(3):e22060 [FREE Full text] [doi: 10.2196/22060] [Medline: 32930670]
51. Ali SH, Foreman J, Tozan Y, Capasso A, Jones AM, DiClemente RJ. Trends and predictors of COVID-19 information sources and their relationship with knowledge and beliefs related to the pandemic: nationwide cross-sectional study. JMIR Public Health Surveill 2020 Oct 08;6(4):e21071 [FREE Full text] [doi: 10.2196/21071] [Medline: 32936775]
52. Feldman D. Do People Trust the News about COVID-19?. 2020. URL: https://www.psychologytoday.com/us/blog/supersurvivors/202003/do-people-trust-the-news-about-covid-19 [accessed 2022-06-02]
53. McNeill A, Harris PR, Briggs P. Twitter influence on UK vaccination and antiviral uptake during the 2009 H1N1 pandemic. Front Public Health 2016;4:26 [FREE Full text] [doi: 10.3389/fpubh.2016.00026] [Medline: 26942174]
54. Walther JB, Pingree S, Hawkins RP, Buller DB. Attributes of interactive online health information systems. J Med Internet Res 2005 Jul 01;7(3):e33 [FREE Full text] [doi: 10.2196/jmir.7.3.e33] [Medline: 15998624]
55. Walther J, Tong S, DeAndrea D, Carr C, Van Der Heide B. A juxtaposition of social influences: web 2.0 and the interaction of mass, interpersonal, and peer sources online. In: Birchmeier Z, Dietz-Uhler B, Strasser G, editors. Strategic Uses of Social Technology: An Interactive Perspective of Social Psychology. Cambridge, England: Cambridge University Press; 2011.
56. 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]
57. Metzger M, Flanagin A. Credibility and trust of information in online environments: the use of cognitive heuristics. J Pragmat 2013 Dec;59:210-220 [FREE Full text] [doi: 10.1016/j.pragma.2013.07.012]
58. Metzger M, Flanagin A, Eyal K, Lemus D, Mccann R. Credibility in the 21st century: integrating perspectives on source, message, and media credibility in the contemporary media environment. In: Communication Yearbook. Mahwah, NJ: Lawrence Erlbaum; Jan 2003:293-335.
59. De Coninck D, Frissen T, Matthijs K, d'Haenens L, Lits G, Champagne-Poirier O, et al. Beliefs in conspiracy theories and misinformation about COVID-19: comparative perspectives on the role of anxiety, depression and exposure to and trust in information sources. Front Psychol 2021 Apr 16;12:646394 [FREE Full text] [doi: 10.3389/fpsyg.2021.646394] [Medline: 33935904]
60. Wang P, Chen Y, Chang Y, Wu C, Lu W, Yen C. Sources of COVID-19-related information in people with various levels of risk perception and preventive behaviors in Taiwan: a latent profile analysis. Int J Environ Res Public Health 2021 Feb 21;18(4):2091 [FREE Full text] [doi: 10.3390/ijerph18042091] [Medline: 33669977]
61. Zimet G, Panozzo C, Head K, Kornides M, Feemster K. Mothers are the primary decision-makers for adolescent HPV vaccination. Personal Communication 2022.
62. Ranji U, Salganicoff A. Data Note: Balancing on Shaky Ground: Women, Work and Family Health. 2014. URL: https://www.kff.org/womens-health-policy/issue-brief/data-note-balancing-on-shaky-ground-women-work-and-family-health/ [accessed 2022-06-02]
63. Employee Benefits Security Administration, U.S. Department of Labor. Fact Sheet: General Facts on Womand and Job Based Health. URL: https://www.dol.gov/sites/dolgov/files/ebsa/about-ebsa/our-activities/resource-center/fact-sheets/women-and-job-based-health.pdf [accessed 2022-06-02]

XSL•FO
RenderX

64. Hughes A. Using Social Media Platforms to Amplify Public Health Messages. An Examination of Tenets and Best Practices for Communicating with Key Audiences. 2010. URL: https://csic.georgetown.edu/wp-content/uploads/2016/12/public-health.pdf [accessed 2022-06-02]
65. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr 30;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
66. Stryker J, Lazovich D, Forster J, Emmons K, Sorensen G, Demierre M. Maternal/female caregiver influences on adolescent indoor tanning. J Adolesc Health 2004 Dec;35(6):528.e1. [doi: 10.1016/s1054-139x(04)00091-6]
67. Gore JS, Frederick H, Ramkissoon M. Mother-daughter communication and health: a cross-cultural comparison. Health Care Women Int 2018 Sep;39(9):994-1007. [doi: 10.1080/07399332.2018.1488852] [Medline: 30265834]
68. Fisher CL, Kastrinos A, Piemonte N, Canzona MR, Wolf B, Pipe T. Coping with breast cancer together: challenging topics for mothers and their adolescent-young adult (AYA) daughters. J Psychosoc Oncol 2021 Dec 13:1-14. [doi: 10.1080/07347332.2021.2005734] [Medline: 34898401]
69. Buller DB, Pagoto S, Baker K, Walkosz BJ, Hillhouse J, Henry KL, et al. Results of a social media campaign to prevent indoor tanning by teens: a randomized controlled trial. Prev Med Rep 2021 Jun;22:101382 [FREE Full text] [doi: 10.1016/j.pmedr.2021.101382] [Medline: 33996394]
70. Buller D, Pagoto S, Henry K, Baker K, Walkosz B, Hillhouse J, et al. Persisting effects of a social media campaign to prevent indoor tanning: a randomized trial. Cancer Epidemiol Biomark Prev 2022;31(4):885-892. [doi: 10.1158/1055-9965.epi-21-0059]
71. Magnusson K. Package 'powerlmm': Power Analysis for Longitudinal Multilevel Models. 2018 Aug 14. URL: https://mran.revolutionanalytics.com/snapshot/2020-04-25/web/packages/powerlmm/powerlmm.pdf [accessed 2022-08-01]
72. Tully M, Vraga EK, Bode L. Designing and testing news literacy messages for social media. Mass Commun Soc 2019 May 22;23(1):22-46. [doi: 10.1080/15205436.2019.1604970]
73. Austin EW, Austin BW, Willoughby JF, Amram O, Domgaard S. How media literacy and science media literacy predicted the adoption of protective behaviors amidst the COVID-19 pandemic. J Health Commun 2021 Apr 03;26(4):239-252. [doi: 10.1080/10810730.2021.1899345] [Medline: 33928871]
74. De AB. Gatekeeping misinformation with media literacy education. Knowledge Quest 2021;50(2):26-31.
75. AllSides. Media Bias Ratings. URL: https://www.allsides.com/media-bias/media-bias-ratings [accessed 2022-06-02]
76. 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]
77. Rubin GJ, Amlôt R, Page L, Wessely S. Public perceptions, anxiety, and behaviour change in relation to the swine flu outbreak: cross sectional telephone survey. BMJ 2009 Jul 02;339:b2651 [FREE Full text] [doi: 10.1136/bmj.b2651] [Medline: 19574308]
78. World Health Organization. COVID-19 Snapshot Monitoring (COSMO): Monitoring Knowledge, Risk Perceptions, Preventive Behaviours, and Public Trust in the Current Coronavirus Outbreak - WHO Standard Protocol. 2020. URL: https://doi.org/10.23668/PSYCHARCHIVES.2782 [accessed 2022-06-02]
79. Head KJ, Kasting ML, Sturm LA, Hartsock JA, Zimet GD. A national survey assessing SARS-CoV-2 vaccination intentions: implications for future public health communication efforts. Sci Commun 2020 Sep 23;42(5):698-723. [doi: 10.1177/1075547020960463]
80. Ling M, Kothe EJ, Mullan BA. Predicting intention to receive a seasonal influenza vaccination using protection motivation theory. Soc Sci Med 2019 Jul;233:87-92. [doi: 10.1016/j.socscimed.2019.06.002] [Medline: 31195194]
81. Lipschitz JM, Fernandez AC, Larson HE, Blaney CL, Meier KS, Redding CA, et al. Validation of decisional balance and self-efficacy measures for HPV vaccination in college women. Am J Health Promot 2013;27(5):299-307. [doi: 10.4278/ajhp.110606-QUAN-240] [Medline: 23402229]
82. McComas KA, Trumbo CW. Source credibility in environmental health-risk controversies: application of Meyer's credibility index. Risk Anal 2001 Jun;21(3):467-480. [doi: 10.1111/0272-4332.213126] [Medline: 11572427]
83. Liu M, Zhang H, Huang H. Media exposure to COVID-19 information, risk perception, social and geographical proximity, and self-rated anxiety in China. BMC Public Health 2020 Nov 04;20(1):1649 [FREE Full text] [doi: 10.1186/s12889-020-09761-8] [Medline: 33148201]
84. Bendau A, Petzold MB, Pyrkosch L, Mascarell Maricic L, Betzler F, Rogoll J, et al. Associations between COVID-19 related media consumption and symptoms of anxiety, depression and COVID-19 related fear in the general population in Germany. Eur Arch Psychiatry Clin Neurosci 2021 Mar 20;271(2):283-291 [FREE Full text] [doi: 10.1007/s00406-020-01171-6] [Medline: 32691135]
85. Betsch C, Schmid P, Heinemeier D, Korn L, Holtmann C, Böhm R. Beyond confidence: development of a measure assessing the 5C psychological antecedents of vaccination. PLoS One 2018;13(12):e0208601 [FREE Full text] [doi: 10.1371/journal.pone.0208601] [Medline: 30532274]
86. Child and Adolescent Health Measurement Initiative. 2018 National Survey of Children's Health: Guide to Topics and Questions. 2019. URL: https://www.childhealthdata.org/learn-about-the-nsch/topics_questions/2018-nsch-guide-to-topics-and-questions [accessed 2021-12-08]

XSL•FO
RenderX

87. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
88. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23-34 [FREE Full text] [doi: 10.20982/tqmp.08.1.p023] [Medline: 22833776]
89. Witte K. Fear control and danger control: a test of the extended parallel process model (EPPM). Commun Monogr 2009 Jun 02;61(2):113-134. [doi: 10.1080/03637759409376328]
90. Quinn SC, Kumar S, Freimuth VS, Kidwell K, Musa D. Public willingness to take a vaccine or drug under Emergency Use Authorization during the 2009 H1N1 pandemic. Biosecur Bioterror 2009 Sep;7(3):275-290 [FREE Full text] [doi: 10.1089/bsp.2009.0041] [Medline: 19775200]
91. Goldstein A, Clement S. 7 in 10 Americans Would Be Likely to Get a Coronavirus Vaccine, Post-ABC Poll Finds. 2020. URL: hhttps://tinyurl.com/mpp7895r [accessed 2022-06-02]
92. Li S, Feng B, Liao W, Pan W. Internet use, risk awareness, and demographic characteristics associated with engagement in preventive behaviors and testing: cross-sectional survey on COVID-19 in the United States. J Med Internet Res 2020 Jun 16;22(6):e19782 [FREE Full text] [doi: 10.2196/19782] [Medline: 32501801]
93. Crane MA, Shermock KM, Omer SB, Romley JA. Change in reported adherence to nonpharmaceutical interventions during the COVID-19 pandemic, April-November 2020. JAMA 2021 Mar 02;325(9):883-885 [FREE Full text] [doi: 10.1001/jama.2021.0286] [Medline: 33480971]
94. Jackson C, Newall M, Yi J. America's Reopening Shows Little Signs of Slowing Down. 2021. URL: https://www.ipsos.com/en-us/news-polls/axios-ipsos-coronavirus-index [accessed 2022-06-02]
95. Zhao E, Wu Q, Crimmins EM, Ailshire JA. Media trust and infection mitigating behaviours during the COVID-19 pandemic in the USA. BMJ Glob Health 2020 Oct;5(10):e003323 [FREE Full text] [doi: 10.1136/bmjgh-2020-003323] [Medline: 33037063]
96. Reisdorf B, Blank G, Bauer J, Cotten S, Robertson C, Knittel M. Information seeking patterns and COVID-19 in the United States. Journal of Quantitative Description: Digital Media 2021 Apr 26;1:1-38. [doi: 10.51685/jqd.2021.003]
97. Gilles I, Bangerter A, Clémence A, Green EGT, Krings F, Staerklé C, et al. Trust in medical organizations predicts pandemic (H1N1) 2009 vaccination behavior and perceived efficacy of protection measures in the Swiss public. Eur J Epidemiol 2011 Mar;26(3):203-210. [doi: 10.1007/s10654-011-9577-2] [Medline: 21476079]
98. Bish A, Michie S. Demographic and attitudinal determinants of protective behaviours during a pandemic: a review. Br J Health Psychol 2010 Nov;15(Pt 4):797-824 [FREE Full text] [doi: 10.1348/135910710X485826] [Medline: 20109274]
99. Pagliaro S, Sacchi S, Pacilli MG, Brambilla M, Lionetti F, Bettache K, et al. Trust predicts COVID-19 prescribed and discretionary behavioral intentions in 23 countries. PLoS One 2021;16(3):e0248334 [FREE Full text] [doi: 10.1371/journal.pone.0248334] [Medline: 33690672]
100. McCroskey JC, Teven JJ. Goodwill: a reexamination of the construct and its measurement. Commun Monogr 2009 Jun 02;66(1):90-103. [doi: 10.1080/03637759909376464]
101. Lee S, Kim K, Koh J. Antecedents of news consumers' perceived information overload and news consumption pattern in the USA. Int J Contents 2016 Sep 28;12(3):1-11 [FREE Full text] [doi: 10.5392/ijoc.2016.12.3.001]
102. Cao X, Sun J. Exploring the effect of overload on the discontinuous intention of social media users: an S-O-R perspective. Comput Hum Behav 2018 Apr;81:10-18 [FREE Full text] [doi: 10.1016/j.chb.2017.11.035]
103. Matthes J, Karsay K, Schmuck D, Stevic A. “Too much to handle”: impact of mobile social networking sites on information overload, depressive symptoms, and well-being. Comput Hum Behav 2020 Apr;105:106217 [FREE Full text] [doi: 10.1016/j.chb.2019.106217]
104. Baker SM, Petty RE. Majority and minority influence: source-position imbalance as a determinant of message scrutiny. J Pers Soc Psychol 1994;67(1):5-19. [doi: 10.1037/0022-3514.67.1.5]
105. Jensen ML, Averbeck JM, Zhang Z, Wright KB. Credibility of anonymous online product reviews: a language expectancy perspective. J Manage Inf Syst 2014 Dec 08;30(1):293-324. [doi: 10.2753/mis0742-1222300109]
106. Miller MD, Burgoon M. The relationship between violations of expectations and the induction of resistance to persuasion. Human Commun Res 1979 Jun;5(4):301-313. [doi: 10.1111/j.1468-2958.1979.tb00642.x]
107. Clementson D, Pascual-Ferrá P, Beatty M. When does a presidential candidate seem presidential and trustworthy? Campaign messages through the lens of language expectancy theory. Pres Stud Q 2016 Jul 28;46(3):592-617 [FREE Full text] [doi: 10.1111/psq.12299]
108. Siegel J, Burgoon J. Expectancy theory approaches to prevention: violating adolescent expectations to increase the effectiveness of public service announcements. In: Crano WD, Burgoon M, editors. Mass Media and Drug Prevetion: Classic and Contemporary Theories and Research. Mahwah, NJ: Lawrence Erlbaum Associates; 2002:163-186.
109. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
110. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]

XSL•FO
RenderX

111. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep 27;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
112. Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, Noakes M. Features predicting weight loss in overweight or obese participants in a web-based intervention: randomized trial. J Med Internet Res 2012 Dec 12;14(6):e173 [FREE Full text] [doi: 10.2196/jmir.2156] [Medline: 23234759]
113. Turner-McGrievy GM, Tate DF. Weight loss social support in 140 characters or less: use of an online social network in a remotely delivered weight loss intervention. Transl Behav Med 2013 Sep;3(3):287-294 [FREE Full text] [doi: 10.1007/s13142-012-0183-y] [Medline: 24073180]
114. Trunfio M, Rossi S. Conceptualising and measuring social media engagement: a systematic literature review. Ital J Mark 2021 Aug 11;2021(3):267-292. [doi: 10.1007/s43039-021-00035-8]
115. Muntinga DG, Moorman M, Smit EG. Introducing COBRAs. Int J Advert 2015 Jan 07;30(1):13-46. [doi: 10.2501/ija-30-1-013-046]
116. Khan ML. Social media engagement: what motivates user participation and consumption on YouTube? Comput Hum Behav 2017 Jan;66:236-247 [FREE Full text] [doi: 10.1016/j.chb.2016.09.024]
117. Fernandes T, Castro A. Understanding drivers and outcomes of lurking vs. posting engagement behaviours in social media-based brand communities. J Mar Manage 2020 Feb 10;36(7-8):660-681. [doi: 10.1080/0267257x.2020.1724179]
118. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun 13;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
119. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
120. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
121. Pagoto S, Waring M, Jake-Schoffman D, Goetz J, Michaels Z, Oleski J. What type of engagement predicts success in a Facebook weight loss group? 2018 Presented at: 51st Hawaii International Conference on System Sciences; January 3-6, 2018; Waikoloa Village, Hawaii. [doi: 10.24251/hicss.2018.419]
122. Agarwal R, Dugas M, Ramaprasad J, Luo J, Li G, Gao G. Socioeconomic privilege and political ideology are associated with racial disparity in COVID-19 vaccination. Proc Natl Acad Sci U S A 2021 Aug 17;118(33):e2107873118 [FREE Full text] [doi: 10.1073/pnas.2107873118] [Medline: 34326130]
123. Fridman A, Gershon R, Gneezy A. COVID-19 and vaccine hesitancy: a longitudinal study. PLoS One 2021 Apr 16;16(4):e0250123 [FREE Full text] [doi: 10.1371/journal.pone.0250123] [Medline: 33861765]
124. Sell T, Hosangadi D, Smith E, Trotochaud M, Vasudevan P, Gronvall G. National Priorities to Combat Misinformation and Disinformation for COVID-19 and Future Public Health Threats: a Call for a National Strategy. 2021. URL: https://tinyurl.com/2p86c7d7 [accessed 2022-06-02]
125. Kreps S, Kriner D. Good News and Bad News about COVID-19 Misinformation. URL: https://blogs.scientificamerican.com/observations/good-news-and-bad-news-about-covid-19-misinformation/ [accessed 2022-08-01]
126. Nan X, Madden K. HPV vaccine information in the blogosphere: how positive and negative blogs influence vaccine-related risk perceptions, attitudes, and behavioral intentions. Health Commun 2012;27(8):829-836. [doi: 10.1080/10410236.2012.661348] [Medline: 22452582]
127. Betsch C, Renkewitz F, Haase N. Effect of narrative reports about vaccine adverse events and bias-awareness disclaimers on vaccine decisions: a simulation of an online patient social network. Med Decis Making 2013 Jan;33(1):14-25. [doi: 10.1177/0272989X12452342] [Medline: 22875721]
128. McRee A, Reiter PL, Brewer NT. Parents' internet use for information about HPV vaccine. Vaccine 2012 May 28;30(25):3757-3762 [FREE Full text] [doi: 10.1016/j.vaccine.2011.11.113] [Medline: 22172505]
129. Lawes-Wickwar S, Ghio D, Tang MY, Keyworth C, Stanescu S, Westbrook J, et al. A rapid systematic review of public responses to health messages encouraging vaccination against infectious diseases in a pandemic or epidemic. Vaccines (Basel) 2021 Jan 20;9(2):72 [FREE Full text] [doi: 10.3390/vaccines9020072] [Medline: 33498395]

Abbreviations

CDC: Centers for Disease Control and Prevention
CONSORT: Consolidated Standards of Reporting Trials
EPPM: Extended Parallel Process Model
IT: indoor tanning
NPI: nonpharmaceutical intervention
SCT: social cognitive theory

XSL•FO
RenderX

Edited by L Leininger; submitted 05.01.22; peer-reviewed by C Latkin, A Farooq; comments to author 20.04.22; revised version received 28.06.22; accepted 26.07.22; published 23.08.22.

Please cite as:
Buller D, Walkosz B, Henry K, Woodall WG, Pagoto S, Berteletti J, Kinsey A, Divito J, Baker K, Hillhouse J
Promoting Social Distancing and COVID-19 Vaccine Intentions to Mothers: Randomized Comparison of Information Sources in Social Media Messages
JMIR Infodemiology 2022;2(2):e36210
URL: https://infodemiology.jmir.org/2022/2/e36210
doi:10.2196/36210
PMID:36039372

©David Buller, Barbara Walkosz, Kimberly Henry, W Gill Woodall, Sherry Pagoto, Julia Berteletti, Alishia Kinsey, Joseph Divito, Katie Baker, Joel Hillhouse. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 23.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Direct-to-Consumer Genetic Testing on Social Media: Topic Modeling and Sentiment Analysis of YouTube Users' Comments

Philipp A Toussaint[1,2], MSc; Maximilian Renner[1], MSc; Sebastian Lins[1], PhD; Scott Thiebes[1], MSc; Ali Sunyaev[1], PhD

[1]Department of Economics and Management, Karlsruhe Institute of Technology, Karlsruhe, Germany

[2]HIDSS4Health – Helmholtz Information and Data Science School for Health, Karlsruhe/Heidelberg, Germany

Corresponding Author:
Ali Sunyaev, PhD
Department of Economics and Management
Karlsruhe Institute of Technology
Kaiserstr. 89
Karlsruhe, 76133
Germany
Phone: 49 72160846037
Email: sunyaev@kit.edu

Abstract

Background: With direct-to-consumer (DTC) genetic testing enabling self-responsible access to novel information on ancestry, traits, or health, consumers often turn to social media for assistance and discussion. YouTube, the largest social media platform for videos, offers an abundance of DTC genetic testing–related videos. Nevertheless, user discourse in the comments sections of these videos is largely unexplored.

Objective: This study aims to address the lack of knowledge concerning user discourse in the comments sections of DTC genetic testing–related videos on YouTube by exploring topics discussed and users' attitudes toward these videos.

Methods: We employed a 3-step research approach. First, we collected metadata and comments of the 248 most viewed DTC genetic testing–related videos on YouTube. Second, we conducted topic modeling using word frequency analysis, bigram analysis, and structural topic modeling to identify topics discussed in the comments sections of those videos. Finally, we employed Bing (binary), National Research Council Canada (NRC) emotion, and 9-level sentiment analysis to identify users' attitudes toward these DTC genetic testing–related videos, as expressed in their comments.

Results: We collected 84,082 comments from the 248 most viewed DTC genetic testing–related YouTube videos. With topic modeling, we identified 6 prevailing topics on (1) general genetic testing, (2) ancestry testing, (3) relationship testing, (4) health and trait testing, (5) ethical concerns, and (6) YouTube video reaction. Further, our sentiment analysis indicates strong positive emotions (anticipation, joy, surprise, and trust) and a neutral-to-positive attitude toward DTC genetic testing–related videos.

Conclusions: With this study, we demonstrate how to identify users' attitudes on DTC genetic testing by examining topics and opinions based on YouTube video comments. Shedding light on user discourse on social media, our findings suggest that users are highly interested in DTC genetic testing and related social media content. Nonetheless, with this novel market constantly evolving, service providers, content providers, or regulatory authorities may still need to adapt their services to users' interests and desires.

(JMIR Infodemiology 2022;2(2):e38749) doi:10.2196/38749

KEYWORDS

direct-to-consumer genetic testing; health information; social media; YouTube; sentiment analysis; topic modeling; content analysis; online health information; user discourse; infodemiology

XSL•FO
RenderX

Introduction

Background and Objectives

Since the completion of the human genome project in 2003, dwindling genome sequencing costs and a rising interest in genomics among the general public have paved the way for direct-to-consumer (DTC) genetic testing [1]. Today, users can purchase DTC genetic tests via the internet for less than US $100 to gain genetic insights into their health, traits, heritage, and more without the involvement of health care professionals [2]. By providing users with such interesting and novel insights, DTC genetic testing markets are growing continuously. For example, North America's DTC genetic testing market alone accounted for 39% of an estimated global market value of US $1.5 billion in 2021. Moreover, with a projected annual growth rate of 15.3%, the DTC genetic testing market value is expected to triple in the next 8 years [3].

The uprise of DTC genetic testing and self-responsible genetics has also sparked countless ethical, social, technical, and legal issues [1]. For example, critics argue that DTC genetic testing lacks clinical validity and meaningful interpretation of test results, whereas service providers can make unregulated advertising and marketing claims, especially for health-related tests [1,2,4-7]. Indeed, consumers taking multiple DTC genetic tests found themselves receiving different results depending on the service provider [8]. Another concern often discussed by researchers and consumers is the potential sharing and reselling of genetic data (eg, to pharmaceutical companies) and the resulting implications on genetic privacy, including genetic data access to insurance companies, employers, law enforcement agencies, or malicious entities like hackers [9-14]. Although many consumers perceive these practices as unfair, low prices and potential genetic insights often outweigh the aforementioned concerns [15]. However, due to genetic similarity, these consequences may also apply to blood relatives who were not involved or did not consent to genetic testing [13,16]. This also ties in with media and research reporting that consumers in the United States use DTC genetic ancestry tests to prove their “genetic purity,” leading to instances of racism and genetic discrimination on social media [17,18].

With the increasing spread and availability of DTC genetic testing [2] and a general tendency in society to retrieve as well as discuss health information and health-related topics on the internet [19], it is by no means surprising that DTC genetic testing is a frequent and recent topic on many social media platforms [18,20,21]. In particular, YouTube, one of the largest social media platforms and the most comprehensive web-based video platform [22], serves as the first port of call for many internet users to discuss health information and DTC genetic testing in particular [23]. While YouTube can serve to share health information and experiences with a big audience for content creators (eg, consumers, service providers, health care professionals, or journalists), it also enables user discourse through textual comments below individual videos [24].

Understanding the topics, opinions, and attitudes discussed by the users can prove crucial for many stakeholders, as comments are the main form of user reaction and feedback on social media [23]. Service providers may gain, for instance, insights into consumer demands, whereas content creators may improve their videos by adjusting their content to meet user preferences. Moreover, with the ongoing debate on ethical and legal concerns toward DTC genetic testing [1,7], user opinions are of utmost importance to regulation authorities, politicians, and the industry in general. However, many stakeholders lack the means to extract the core themes discussed and attitudes expressed in the comments sections effectively and efficiently, given the sheer number of comments and manifold writing styles of users.

Extant research regarding DTC genetic testing on social media confirms this lack of understanding. Prior research focuses on microblogging services such as Twitter [25,26], Reddit [27], or 4chan [18] to investigate user discourse on DTC genetic testing and shows that we are still puzzled about users' interests and opinions toward DTC genetic testing. Inconsistent findings regarding which topics users discuss on different platforms (eg, ancestry testing on Twitter [25] and health testing on Reddit [27]) suggest that the DTC genetic testing discourse varies from platform to platform and must thus be investigated separately. Moreover, research has already shown the value of analyzing users' opinions and attitudes through user comments from select platforms for DTC genetic testing–related content. For instance, Mittos et al [18] have uncovered extensive use of hate speech on Twitter, whereas Basch et al [20] have identified the need for educational content about genetic testing on TikTok. Few studies have investigated information about DTC genetic testing on YouTube while primarily analyzing the multimedia information (ie, the content of the videos) [28-31] and overlooking the textual information provided by users' comments (see Multimedia Appendix 1 for a complete overview of research on DTC genetic testing on social media). Because most users do not actively produce YouTube videos but only consume them, we believe that analyzing the topics that users discuss in the YouTube comments sections provides a new perspective on the ongoing discussion regarding DTC genetic testing–related videos on social media platforms. Consequently, we ask the following research questions (RQs):

RQ1: What topics do YouTube users discuss in the comments sections of DTC genetic testing–related videos?

RQ2: What are users' attitudes toward DTC genetic testing–related videos, as expressed in their comments on YouTube?

To answer our RQs, we analyzed the 248 most viewed videos dealing with DTC genetics in a 3-step exploratory approach. First, we analyzed the selected videos regarding media type, genetic test purpose, and related health information. Second, we employed topic modeling to investigate user discourse in the comments sections of those videos. Third, we conducted a sentiment analysis unveiling users' attitudes toward the discussed topics and DTC genetic testing videos in general.

Through our study, we contribute to research and practice in several ways. As for research, we add to the literature on user attitudes toward DTC genetic testing by delineating topics and opinions discussed about these genetic tests. Further, we contribute to the research stream regarding health information on social media by showing that YouTube comments provide

XSL•FO
RenderX

valuable insights on user discourse on social media and demonstrate that DTC genetic testing and health information topics may generally vary from platform to platform. As for practice, our research may help providers of DTC genetic testing services and regulatory authorities gain further insights into user attitudes and consequently adapt or improve genetic testing services and regulations. As most videos are user-generated, our analysis of user discourse can provide valuable insights on the topics discussed in the comments sections of these videos, providing content creators with valuable information for improving their future DTC genetic testing–themed videos.

Health Information on Social Media Platforms

During the past decade, social media platforms have become increasingly attractive in the digital health sector as a means of communicating medical information [32]. In addition to accessing professional and nonprofessional medical information, users can also share their experiences and get in touch with each other [33]. Users already discuss various health topics like diabetes, medication and medication information, physical health, mental health, cancer, or more recently, COVID-19 on social media [19,34-38].

Consequently, information dissemination platforms (see Multimedia Appendix 1 for a detailed description of social media platform types), such as YouTube, have garnered interest from researchers to study various health care–related topics. For example, studies have investigated users' attitudes toward the effect of sleep-aiding music [24], users' preferences regarding treatment and symptoms of diabetes as well as the social culture pertaining to diabetes-related video clips [39], or public opinions and concerns about daily coverage of the COVID-19 crisis in Canada [23].

DTC Genetic Testing

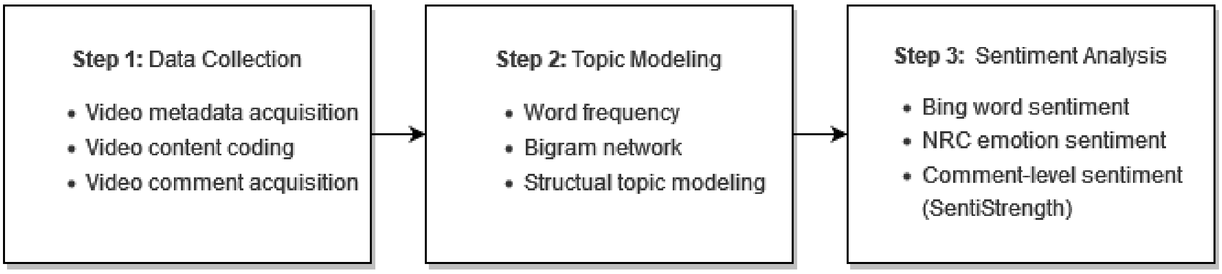
DTC genetic testing differs from traditional clinical genetic testing in that it is initiated by the consumers and does not require the direct interaction of consumers with health care professionals [2]. With the internet being the leading advertising and distribution channel, the DTC genetic testing service provider usually sends a DNA sample collection kit (eg, buccal swab or blood spot collection) to the consumers' homes for self-collection [5] or arranges for sample collection at a local laboratory [7]. Afterward, the service provider may perform various genetic analyses and then return the results directly to the consumers via the internet or mail [5]. Regarding DTC genetic testing, the consumers can choose the interpreter (ie, service provider) and the type and objective of the analysis of their genetic information (as opposed to a health care professional interpreting the genetic data). The most common types of testing services offered include ancestry tests (eg, AncestryDNA), nonmedical lifestyle tests (eg, FitnessGenes), relationship tests (eg, EasyDNA), and health tests (eg, 23andMe) [2]. Although DTC genetic testing provides consumers with novel and valuable information, it also has its downsides, such as consumers being responsible for managing and ensuring the security of their personal genetic information [1].

Methods

Research Approach

We employed a 3-step exploratory research approach to answer our RQs (see Figure 1). First, we performed comprehensive data collection by gathering DTC genetic testing–related videos on YouTube, including their comments, and coding the contents of these videos. Second, we performed topic modeling for the user discourse in the comments sections to reveal topics discussed in those comments (answering RQ1). Third, we analyzed users' attitudes toward DTC genetic testing videos using sentiment analysis (answering RQ2).

Figure 1. Overview of the 3-step research approach. NRC: National Research Council Canada.



Data Collection

We used the official YouTube application programming interface (API) to create a list of the most relevant DTC genetic testing–related videos on YouTube. With the region set to the United States (ie, the largest DTC genetic testing market), we queried the 300 most viewed video results for each of 6 different DTC genetic testing–related search terms (ie, direct to consumer genetic testing, home genetic testing, ancestry testing, DNA testing, genetic testing, and 23andMe). Thereafter, we combined the 1800 results from the 6 queries, removed duplicates, and sorted them by video views in descending order. We further excluded all videos with less than 50,000 views because they had very few comments per video (average of 61.2), with many having no comments (n=336).

Next, the remaining 468 videos were reviewed for relevance through iterative manual inspection by 2 researchers, with a third researcher breaking ties in case of differences. For this, our predefined exclusion criteria were as follows: (1) videos not focusing on DTC genetic testing, (2) videos focusing on

XSL•FO
RenderX

genetic testing of animals, (3) videos focusing on clinical prenatal genetic testing, (4) videos not in English, (5) live stream videos, (6) duplicate videos (ie, reuploads from different users), (7) videos commenting/reacting on videos (ie, showing the original video and adding commentary), or (8) videos with disabled ratings and comments sections (see Multimedia Appendix 2 for a detailed overview of the data collection process, including a rationale for each exclusion criterion). This resulted in a total of 250 relevant videos.

To gain insights on what topics the videos entailed, particularly the goal of the genetic test presented and the presentation type of the video, we coded the included videos according to their genetic test purpose and media type. For the genetic test purpose, we selected the most common test types suggested in the literature (ie, ancestry, traits, genetic predisposition, relationship, and other [2,7]). As for the media type, we adapted the categories used by Zhang et al [39] to our set of videos. Therefore, the categories were advertising, documentary, interview, news, user-generated video, and other. After the initial coding and comparison of 20 videos, 2 researchers conducted deductive coding of the remaining videos in parallel. In general, the agreement between both researchers was high, with the genetic test purpose and media type having Cohen κ values of 0.581 and 0.613, respectively. Differences in coding were discussed with a third author to break ties. This coding information allowed us to further analyze the comments regarding the contents of the videos and served as a base to evaluate the discussions in the comments.

With the final coded set of 250 videos in place, we again used the YouTube API to download each video's 500 most recent comments. This number was chosen due to the YouTube API download limitations while still allowing meaningful analysis. Among these, 80 videos had less than 500 comments, and 2 videos were no longer available, leaving us with 84,082 comments from 248 videos, which is a sufficient number for topic modeling and sentiment analysis [eg, 28,31,40,41].

Topic Modeling of Comments

To answer our first RQ, we employed topic modeling to identify common topics discussed by users in the comments sections of DTC genetic testing–related YouTube videos. Topic modeling is frequently used in medical informatics and related disciplines for text mining large data sets (such as comments or tweets) and deducing meaningful topics [23,37,38,40,41]. For our study, we used several topic modeling approaches, including word frequency, bigram correlations, and structural topic modeling, as described and recommended by Silge and Robinson [42]. Because they are some of the most common topic modeling methods and include different approaches [42-44], they are well suited for our exploratory study design. All analyses and visualizations were conducted using R (version 4.1.0, R Foundation for Statistical Computing) in RStudio (version 1.4.1106) and the tidytext package (version 0.3.2).

Before conducting any topic modeling, we first separated the comments into 1-word tokens (ie, comments were split into single words) and performed 2 essential data cleaning tasks. First, we used the SnowballC package to perform word stemming. This step was necessary to ensure that words with identical meanings (eg, plural or verb) were grouped together to allow for meaningful topic modeling. For each word stem, the most frequent word was used to represent its stem (eg, test represents test, tests, test's, and testing). Second, we removed common stop words with the stop word list included in the tidytext package. This list comprises 1149 common stop words such as the, of, or to. As these do not hold any topical information, removing stop words reduces the data set size and benefits topic accuracy [42].

With the cleansed word list in place, we first conducted a word frequency analysis by grouping, counting, and listing the words in descending order. This provides an overview of the most used words and can give a first insight into topics discussed most prominently (eg, “DNA” occurs 15,702 times and “test” 10,902 times).

Second, we created word bigrams. We created a frequency list of 2-word tokens, which are found by pairing every 2 consecutive words in each comment (eg, “DTC genetic testing” results in the bigrams “DTC genetic” and “genetic testing”). In contrast to the single word list, bigrams can be used to span a network with the number of occurrences indicating the weight of each bigram edge [42]. To allow for meaningful interpretation, we found that setting a minimum of 70 occurrences resulted in a comprehensible network. Lower values led to the inclusion of less interpretable and impactful bigrams while cluttering the network (eg, “grocery store,” “hey kelsey,” or “omg lol”).

Finally, we conducted structural topic modeling with the help of the stm package [43]. Structural topic modeling aims to group words from different documents (ie, comments) into topics based on their co-occurrences [43]. The stm package uses document-level covariate information to estimate topic models for a given number of topics. We estimated models ranging from 15 to 100 topics in increments of 5. We then compared these models in terms of best-practice metrics, such as held-out likelihood, lower bound, residuals, and semantic coherence [42,45].

Although there is no definite answer for the correct number of topics [43], after a manual review of these metrics and discussion among 3 researchers, we selected 50 as the appropriate number of topics. A more detailed description of the structural topic modeling process and metrics, as well as a comparison with the 45- and 55-topic model, can be found in Multimedia Appendix 3.

With the 50-topic model chosen, we sorted topics according to prevalence and within each topic, the words contributing to it in descending order. We then manually inspected the 50 most prevalent topics and their 10 most contributing words to deduce meaningful topics and categorized them according to their content. For this, we relied on our prior knowledge of DTC genetic testing as well as knowledge on the content of the videos that we gained during the video coding phase of the data collection step. All topic assignments were discussed among 3 researchers.

XSL•FO
RenderX

Sentiment Analysis of Comments

Because topic modeling can only help us identify topics discussed in the comments but not users' attitudes toward the videos, we next conducted word- and comment-level sentiment analyses to answer our second RQ. Sentiment analysis is a common tool to elicit people's opinions, sentiments, emotions, and attitudes from written language [46]. Although sentiment and attitude are near equivalents and often used synonymously, they do differ in the sense that sentiment is a more permanent disposition to react emotionally, cognitively, and conatively, whereas attitude is a disposition to react with belief, thought, feeling, and overt behavior as part of a larger sentiment [47]. In this sense, we can only deduce users' attitudes from a single YouTube comment and not their whole sentiment toward a certain topic.

Therefore, we decided to conduct 2 word-level sentiment analyses and 1 comment-level sentiment analysis to deduce users' attitudes. For the word-level sentiment, we again used the tidytext package, which entails typical word-level approaches that are well suited for a first exploratory overview [42]. We then followed an approach similar to that used by Mittos et al [18] for the comment-level analysis, who also performed sentiment analysis in the DTC genetic testing context.

Consequently, we first conducted a positive and negative sentiment analysis using the Bing lexicon, which consists of approximately 6800 words that are predefined and classified as either positive or negative [48]. Subsequently, we aggregated the sentiments by word and overall sentiment. Even though this method provides a good sentiment overview, the lexicon's limited number of words omits most topic-specific words.

We also used the National Research Council Canada (NRC) emotion lexicon to get a more detailed overview of users' sentiments toward DTC genetic testing [49]. This lexicon attributes 1 or multiple emotions to approximately 14,000 words (ie, a word may have more than 1 emotion), whereby the classification is also predefined. The emotions covered are anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. Similar to the Bing lexicon, we classified and aggregated all words by NRC sentiment. However, initial inspection revealed that the terms “black” and “white” were strongly associated with negative and positive emotions, respectively. Because it was likely that the overproportional use of these words in our data set was due to ancestry testing–related topics, and to avoid a strong association of ethnicity with emotions, we reran the analysis without them.

For the comment-level sentiment analysis, we used SentiStrength [50], a Java-based sentiment tool optimized for short social web text in English such as Twitter tweets or YouTube comments. The tool reports 2 predefined and experience-based sentiments for each document (ie, comment). First, a negative sentiment ranging from –1 (not negative) to –5 (extremely negative) and a second, positive sentiment ranging from 1 (not positive) to 5 (extremely positive). When combining both, we obtained a total sentiment score between –4 and +4. After calculating the sentiment score for each comment, we performed several analyses regarding sentiment as well as media type and test purpose.

Ethical Considerations

Ethics approval was not necessary for this study, as it did not directly involve human participants. All data used in this study (ie, videos and video comments) were publicly available on YouTube and accessible through the YouTube API at the time of retrieval. All results are only published in aggregated form, and single references are presented anonymously and without context to protect the privacy of the comments’ authors.

Results

Overview of Video Contents and Comments

We examined a total of 248 videos related to DTC genetic testing, collected on September 14, 2020, with a total of 30 videos from official company accounts (21 videos from 23andMe, 8 videos from Ancestry.com, and 1 video from MyHeritage). Based on the media type, these included 27 advertising-related videos, 14 documentaries, 16 interviews, 12 news, 174 user-generated videos, and 5 with other media types (mainly recordings of television shows such as The Late Show with Stephen Colbert or The Jim Jefferies Show/Comedy Central). Among the 248 videos, 194 videos address ancestry as a test purpose, 15 address trait testing, 9 address genetic predispositions, 19 address relationship testing, and 11 address other purposes (such as how to use a test kit or comparison/presentation of multiple genetic test purposes). In total, the videos had 724,574 comments on the day of video data aggregation. We collected the comments of the videos on January 3, 2021, focusing on the 500 most recent comments of each video (total number of comments=84,082). An overview of the video metadata, content, and comments is provided in Table 1.

XSL•FO
RenderX

Table 1. Overview of video metadata, content, and comments.

Video characteristic	Value
Number (N)	248
Date of collection	September 14, 2020
Media type (n)	
Advertising	27
Documentary	14
Interview	16
News	12
User-generated videos	174
Other	5
Test purpose addressed (n)	
Ancestry	194
Traits/characteristics	15
Genetic predisposition	9
Relationship	19
Other	11
Upload date	
Oldest	January 15, 2015
Newest	July 7, 2020
View count	
Minimum	52,802
Maximum	20,453,890
Average	1,158,064
Likes	
Minimum	0
Maximum	368,294
Average	22,114
Dislikes	
Minimum	0
Maximum	10,277
Average	813
Duration (minutes)	
Minimum	00:31
Maximum	34:23
Average	09:30
Comments	
Minimum	2
Maximum	24,523
Average	2922
Comment publication date	
Oldest	March 29, 2017
Newest	January 2, 2021

XSL•FO
RenderX

Topics of the DTC Genetic Testing Video Comments

Word frequency analysis using the comments on DTC genetic testing–related videos provides valuable insights into the topics discussed by users. DNA (n=15,702), test (n=10,902), and people (n=9259) are by far the most frequent terms, thus indicating that users indeed primarily discuss DTC genetic testing in their comments. Additionally, we identified many words referring to ancestry testing such as ancestry (n=5015), african (n=6268), or american (n=6139). Moreover, words such as family (n=5252), dad (n=2932), or parents (n=2228) can be attributed to relationship tests. Overall, the 100 most frequent words resemble the test purposes identified from the videos themselves as well as a general excitement for DTC genetic testing videos (eg, video, n=4794; love, n=4751). Table 2 provides an overview of the 20 most frequent words. Additionally, Multimedia Appendix 4 provides a word cloud and overview of the 100 most frequent words.

The bigram network of the comments provides a more fine-grained picture of the words used together often. Unlike the single word cloud, it allows us to see how multiple words are connected. Additionally, the arrows indicate in which order the words appear, whereas the shade of the edge represents the frequency of the word pair. Therefore, we can deduce possible topics discussed by users from the network.

As shown in Figure 2, we identified 5 main topics within the network. The largest topic we identified revolves around ancestry testing (blue cluster). Although the most indicative bigram is “ancestry DNA” (n=679), most bigrams in this topic describe a specific heritage such as “native american” (n=3255), “north african” (n=831), or “middle eastern” (n=756), further substantiating that users largely discuss ancestry results of genetic testing in the comments. The second-largest topic deals with trait testing (green cluster) and holds bigrams such as “blonde/brown/red hair” (n=203/n=72/n=41), “skin color” (n=131), or “blue eyes” (n=285). The third topic entails bigrams related to health testing (yellow cluster). Typical bigrams include “insurance companies” (n=121), “genetic makeup” (n=76), and “23andme test” (n=72). The last topic related to genetic testing indicates relationship testing (red cluster). It includes bigrams such as “identical twins” (n=231), “half sister” (n=124), or “biological parents” (n=74). We also identified 1 topic not specific to DTC genetic testing but YouTube as a platform in general (gray cluster). The bigrams found in this topic are parts of video URLs, for example, “https youtu.be” (n=246) or “www.youtube.com watch” (n=201). This indicates that users often share videos in the comments sections of videos, possibly on related topics.

Finally, we trained structural topic models, of which we selected the 50-topic model. Figure 3 shows the 20 most prevalent topics, including the 10 most important words for each topic of this model. The complete list of all 50 topics can be found in Multimedia Appendix 3. For a better overview of the topics discussed in the comments sections, we grouped these 20 topics into 6 categories, briefly described in the following:

Table 2. List of the 20 most frequent words obtained from comment analysis.

Rank	Word	Frequency (n)
1	dna	15,702
2	test	10,902
3	people	9259
4	african	6268
5	results	6178
6	american	6139
7	family	5252
8	european	5142
9	ancestry	5015
10	video	4794
11	love	4751
12	native	4665
13	white	4489
14	black	4203
15	lol	3469
16	asian	3276
17	irish	3177
18	mixed	2984
19	dad	2932
20	father	2782

XSL•FO
RenderX

Figure 2. Bigram network of 2-word tokens found in the comments of direct-to-consumer genetic testing–related videos on YouTube with a minimum of 70 occurrences. Colored legends indicate topic attribution.

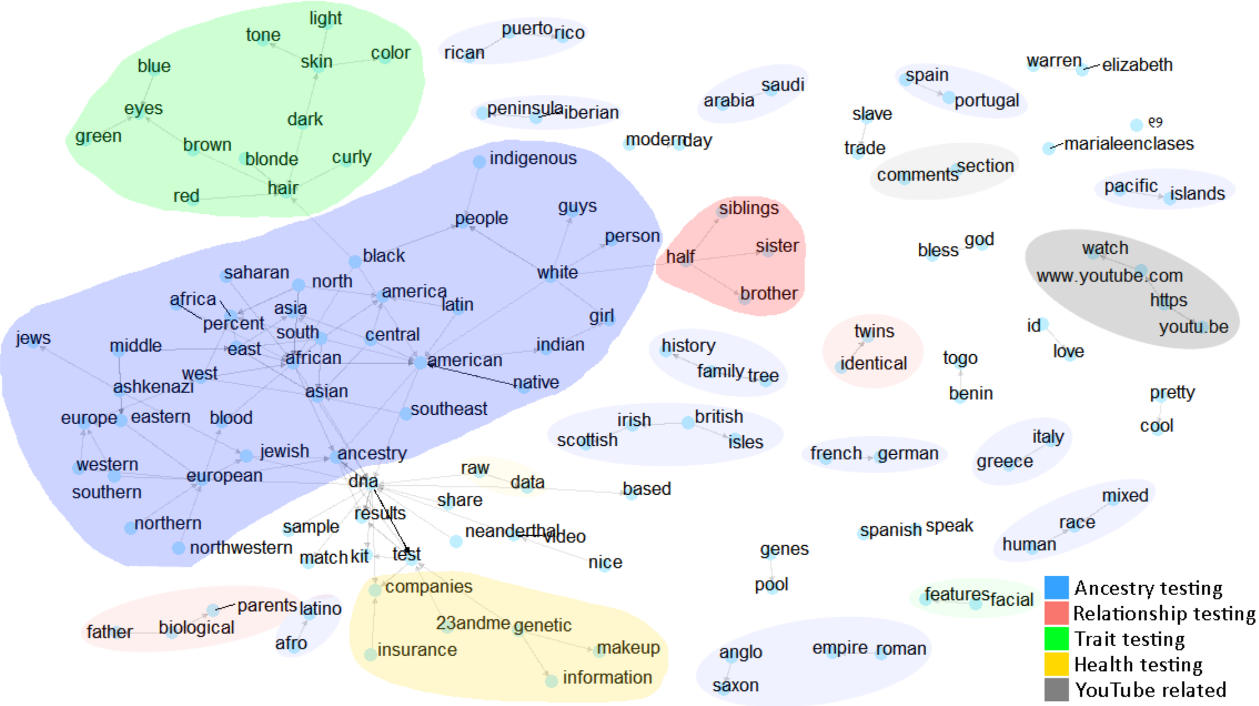


Figure 3. Top 20 topics and their 10 most indicative words from the 50-topic model. Colored legends indicate topic attribution.

General Genetic Testing

This topic group indicates a general interest in DTC genetic testing (eg, topics 16, 31, 49), entailing company names such as MyHeritage, AncestryDNA, or Ancestry.com and words of interest (eg, excited or expect). Moreover, topic 16 touches on the home collection (spit, tube) and financial (money) aspects of DTC genetic testing.

Ancestry Testing

In line with our previous findings, most topics are about the results of genetic ancestry testing. Topic 8 shows a general interest in ancestry testing by users. Topics 17, 26, 37, and 47 describe findings on heritage from a specific region, whereas topic 41 is about paternal and maternal ancestry. Additionally, topic 19 might indicate that users hope to find lost relatives through ancestry testing.

Relationship Testing

We also identified 3 topics about genetic relationship testing. Topics 34 and 48 deal with relationships between children such as identical twins, whereas topic 36 entails the aspects of adoption and genealogy (ie, searching for one's biological family).

Health and Trait Testing

Although less prevalent, health genetic testing and trait testing are also covered in the top 20 topics. Topic 44 focuses on health

XSL•FO
RenderX

information and data, whereas topic 28 entails words on traits such as hair or eye color.

Ethical Concerns

The 50-topic model also reveals some topics not contained in our previous findings. Topic 32 touches on instances of racism signified through words such as black, racist, or mad. Given the ongoing and complex debate toward instances of racism in the United States and the majority of DTC genetic testing revolving around ancestry and heritage, this could explain why this topic was found in the comments of these videos. Moreover, topic 22 deals with users' concerns regarding genetic testing and the government, with words such as lie, ad, or crime.

YouTube Video Reaction

In contrast to the previous findings, topics 18, 27, and 43 do not directly relate to genetic testing but rather entail reactions to the videos on YouTube (eg, love, awesome, watching, video, or channel). Further, users seem interested in personal stories (eg, amazing, story, or reaction).

Comparison of Topic Modeling Approaches and Identified Topics

Although the bigram network and structural topic modeling use different approaches, the majority of the identified topics are present in both methods. Both approaches show strong indications of ancestry testing, relationship testing, trait testing, and health testing topics. Moreover, both methods led to the deduction of a YouTube or YouTube video–related topic. Table 3 compares the topics covered by the bigram network and structural topic modeling and lists some of the most indicative bigrams and words for each method, respectively.

Table 3. Comparison of identified topics using the bigram network and structural topic modeling.

Topic	Bigram network	Structural topic modeling
General genetic testing	N/A[a]	Myheritage; ancestrydna; ancestrycom; excited; expect; spit; tube; money; genes; dna; genetic
Ancestry testing	Ancestry dna; native american; north african: middle eastern	Ancestry; african; american; native; irish; german; french; father; parents; race; mexican
Relationship testing	Identical twins; half sister; biological parents	Kids; cry; family; adopted; genealogy; lies
Trait testing	Blonde/brown/red hair; skin color; blue eyes	Hair; eyes; blonde; blue; red
Health testing	Insurance companies; genetic makeup; 23andme test	Companies; information; health; pay
Ethical concerns	N/A	Black; racist; claim; government; clone; crime; evidence
YouTube-related	https youtu.be; www.youtube.com watch	N/A
YouTube video reaction	N/A	Love; awesome; watching; video; channel; amazing; story; reaction

[a]N/A: not applicable.

Sentiments of DTC Genetic Testing Video Comments

Even though topic modeling can help unveil what users discuss in the comments sections, it does not provide insights into users' attitudes toward these topics. Therefore, conducting a Bing sentiment analysis can provide a first overview of the sentiment regarding words used in the comments sections. Figure 4 shows the 20 most used words with negative and positive sentiments. The results show that the most used positive words are used significantly more often. In fact, the first negative word, funny (n=864), is only the seventh most used word overall in the sentiment list. Moreover, the positive word love (n=4751) is used overproportionally, having more than twice as many occurrences as the second most used word, beautiful (n=1953). However, when observing all positively and negatively classified occurrences, we can identify more negative word uses (n=38,734) than positive ones (n=35,897).

Another type of sentiment analysis is the identification of emotions with the NRC lexicon. Our results show that the most frequent words representing positive emotions, namely anticipation, joy, surprise, and trust, have higher occurrences than the words expressing negative emotions, namely anger, fear, disgust, and sadness (see Figure 5). This finding is also supported by overall occurrences of positive word emotions (n=148,791) and negative word emotions (n=76,761). Love, the single most used word (n=4751), is associated with the emotion of joy, and the most frequent emotion is trust (n=54,814). In contrast, disgust (n=15,541) has the least word occurrences.

The comment-level sentiment analysis provides insights into user attitudes as well as attitudes toward DTC genetic testing videos and their respective content (ie, test purpose and media type). Although the SentiStrength sentiment can vary on a scale of –4 to 4, the average sentiment score of all comments is 0.32, meaning slightly positive. This is also reflected by almost half of all the comments (n=36,804) having a neutral sentiment (ie, 0). Grouping comment sentiment by video shows that the lowest sentiment score per video comments section is –0.62, whereas the highest is 1.33. Overall, only 30 of the 248 inspected videos have a negative sentiment, indicating an overall positive attitude toward DTC genetic testing videos.

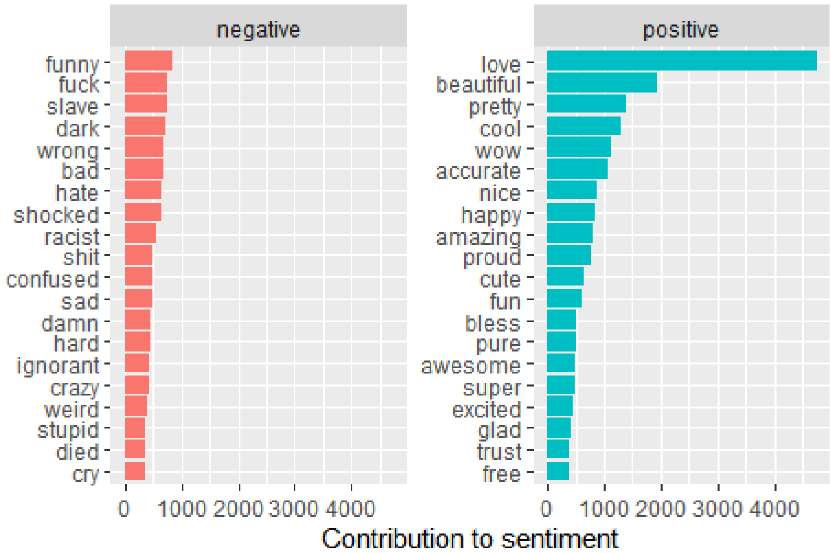
When comparing comment sentiment regarding the test purpose of the videos, our results show that from the comments with a sentiment score of 4, 91.6% (230/251) are in the comments sections of videos about ancestry testing (most frequent test purpose), whereas for comments with a sentiment score of –4, ancestry testing videos only account for 67.9% (76/112). In

XSL•FO
RenderX

contrast, only 1.6% (4/251) of the comments with a sentiment score of 4 are in the responses to a video dealing with relationship testing. However, this increases to 17% (19/112) for comments with a sentiment score of –4. As shown in Figure 6 (left), videos with an ancestry test purpose seem to evoke more positive user comments, whereas this is the opposite for relationship test videos.

The analysis of comment sentiment regarding media type unveils that user-generated videos account for the most significant number of positive comments with 91.6% (230/251) for a sentiment score of 4. On the contrary, for a sentiment score of –4, user-generated videos only account for 60.7% (68/112) of the comments. Consequently, as shown in Figure 6 (right), user-generated videos tend to evoke the most positive attitude toward their video content. This is in contrast to the media types advertising, documentary, and interview; all of these show an increase in the number of comments with decreasing sentiment values. For example, the number of comments for the media type documentary increases from 2% (5/251) with a sentiment score of 4 to 15.2% (17/112) with a sentiment score of –4. Therefore, advertisements, documentaries, and interviews may evoke more negative responses than user-generated videos.

Figure 4. Bing sentiment by most frequent words for negative and positive sentiments.



XSL•FO
RenderX

Figure 5. National Research Council Canada (NRC) sentiment by most frequent words for the emotions anger, anticipation, disgust, fear, joy, sadness, surprise, and trust.

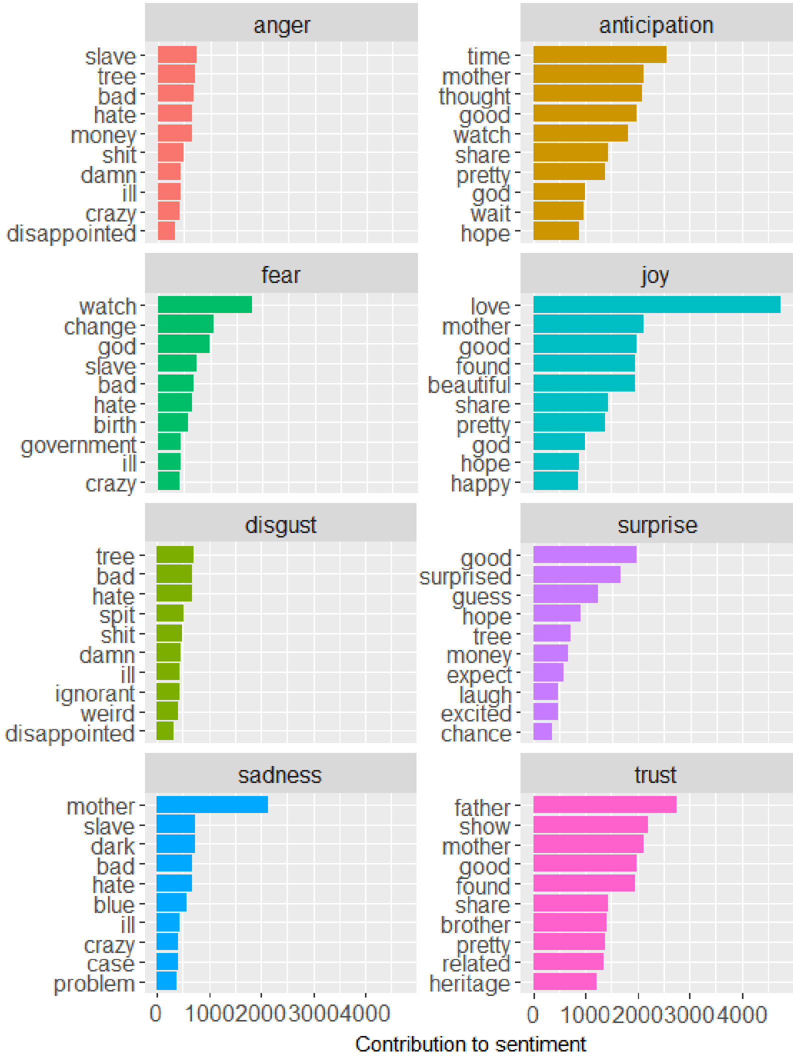
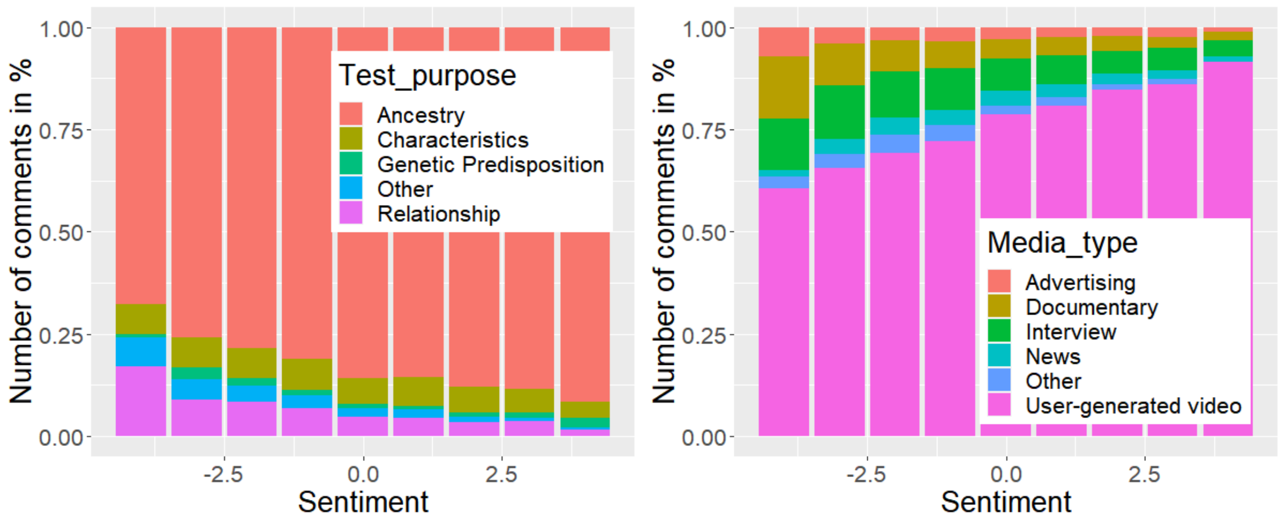


Figure 6. Spreads for test purpose (left) and media type (right) by sentiment.



XSL•FO
RenderX

Discussion

Principal Findings

Our analysis of user comments on DTC genetic testing–related YouTube videos yields several valuable findings. The test purposes found in the videos largely resemble the most common genetic test purposes, with most videos talking about ancestry or relationship testing and fewer about trait and health testing. This finding is in line with previous research on YouTube videos related to DTC genetic testing [28,31] and social media in general [20,21,25]. Nonetheless, in contrast to our study, Yin et al [27] found in their collected Reddit data set that relationship and health testing were more often mentioned than ancestry testing. Although Mittos et al [18] do not report the same finding for their Reddit data set, this may indicate that users of different social media platforms have other interests regarding DTC genetic testing. Another possible explanation for this could be that platform suggestion algorithms differ and may hence propose distinct content to users depending on the platform. Thus, discourses on the respective platforms should be investigated individually before assuming DTC genetic testing–related findings to be true across multiple platforms.

Moreover, most topics found with the bigram network and structural topic modeling can be attributed to common DTC genetic testing purposes. This indicates that user discourse revolves around the contents of the videos and DTC genetic testing. In line with previous research, we also identified topics dealing with general genetic testing and users' interest in and excitement for DTC genetic testing [18,51].

Besides, research has also shown instances of racism regarding ancestry testing on Twitter [18], which we also identified as a topic in the video comments. Even though it is unclear whether these comments relate directly to the content of the respective video or are in the replies to other comments, the identified topics largely revolve around racism and discrimination against African Americans and Native Americans. However, our results did not show any specific topics on the educational content of DTC genetic testing. Considering that consumers in the United States continue to use DTC ancestry tests to prove their “genetic purity” and discriminate against marginalized ethnic groups such as the aforementioned ones, especially on social media [17,18], research has called for more educational content and scientific explanations about DTC genetic testing [20,21]. Despite finding some videos expressing concerns toward DTC genetic testing (eg, documentaries), the majority of the videos seem to fail to highlight the advantages as well as the disadvantages and risks of DTC genetic testing. Hence, the discussions in the comments section may also largely neglect these aspects.

Sentiment analysis revealed that users have more negative attitudes toward the content of advertisements, news, or documentary videos compared to user-generated videos on DTC genetic testing. Although this finding could be explained through some media types being more thought-provoking (eg, documentaries covering disadvantages and risks of DTC genetic testing or news covering stories of genetic discrimination), another explanation might be that user-generated videos are often produced by single creators often trying to engage more with their YouTube community (eg, through specific content or active discussion in the comments sections) than, for example, a news broadcaster or DTC genetic testing service provider. Hence, this may result in a more positive user attitude. This assumption is further supported by our findings on YouTube-related and YouTube video reaction topics. On the one hand, these findings once again indicate that users discuss and respond to the content discussed in the respective videos, and on the other hand, they suggest a more complex discussion between content creators and their community (eg, through expressing enjoyment of content or including links to further YouTube videos). It should be noted that the revealed user attitudes on DTC genetic testing videos do not necessarily reflect user attitudes toward DTC genetic testing in general. However, as our topic modeling results suggest that user comments largely revolve around DTC genetic testing, it is likely that users' attitudes toward DTC genetic testing videos also reflect their attitudes toward DTC genetic testing to some degree. This notion is further supported by the finding that videos discussing the disadvantages and risks of DTC genetic testing tend to have more negative user attitudes. Comparable results on user attitudes toward DTC genetic testing were also found for Twitter and related textual platforms [18,25,51], thereby strengthening this assumption.

Similar to DTC genetic testing–related Reddit posts [41], we found that user emotions toward DTC genetic testing videos expressed through the comments are mainly positive. The NRC sentiment and comment-level sentiment analyses also indicate a clear tendency toward a positive user attitude. This may be explained by the majority of videos being user-generated ones and aforementioned higher community engagement of content creators. Previous research on user sentiment toward Twitter tweets also shows a positive sentiment toward DTC genetic testing [51]. However, Mittos et al [18] found that most tweets only have a sentiment score of 0 or 1. In line with previous research [21,51], these less positive emotions and attitudes could indicate that although users are generally interested in DTC genetic testing, they still have reservations regarding this new technology. These reservations are mirrored in the results of the NRC sentiment analysis that highlighted fear as the most prominent negative attitude toward DTC genetic testing, whereas trust was the most prominent positive attitude. These reservations toward DTC genetic testing were also highlighted in prior research [7].

Implications for Research and Practice

This study conveys several implications for research and practice. As for research, we contribute to the literature on user attitudes toward DTC genetic testing by investigating topics and opinions discussed about these genetic tests. We examined the 248 most viewed DTC genetic testing videos on YouTube in terms of their content (ie, test purpose, media type) and analyzed users' attitudes in the form of their comments. Further, we contribute to research regarding health information on social media by showing that YouTube comments provide valuable insights into user discourse on social media. This study suggests that video content and user comments are co-dependent and should therefore be investigated together. To this end, we

XSL•FO
RenderX

provide new insights into the discourse on genetic testing on YouTube by showing that the discourse in the comments primarily revolves around the content of the videos. Our research indicates that the discourse on YouTube may differ from that on other social media platforms, and hence, a detailed and differentiated consideration of the different platforms may be necessary. We further contribute to knowledge regarding user behavior on social media by examining users' attitudes and emotions toward DTC genetic testing videos on YouTube.

As for practice, our research offers important implications for DTC genetic testing service providers, content creators, and regulatory authorities regarding user attitudes, which may help adapt or improve genetic testing services, multimedia content, or regulations. Similar to the study of Lee et al [21] involving Twitter, our identified topics indicate a lack of educational information about DTC genetic testing in YouTube videos. Further, sentiment analysis shows that users have more negative attitudes toward advertisements, news, or documentary videos and prefer user-generated content on DTC genetic testing. Hence, authorities could consider working with content creators to promote user education on DTC genetic testing. Finally, our topic modeling indicates instances of racism, especially regarding ancestry testing. Service providers and authorities should be aware of this and ensure genetic testing is not used for discrimination. Therefore, we suggest that it may be helpful to flag videos with high numbers of negative comments, including racism or anxiety, and provide further information regarding DTC genetic testing via banners or other visual cues, similar to those used on many platforms for content related to COVID-19 [52].

Limitations and Future Research

The limitations of this study are as follows. First, we only considered a limited number of videos and comments. Even though we attempted to include an appropriate sample by saturating the videos and comments using metrics such as views and number of comments, examining all the initially identified videos (n=1325) and comments could provide further insight, particularly concerning topic modeling and sentiment analysis. Second, we limited our YouTube API queries to the United States because the related DTC genetic testing market is the most evolved there. However, other regions with striving markets, such as Asia [30], could offer further insights into user discourse and should therefore be investigated in future research. Third, because there is no way to determine the optimal number of topics [42], we concentrated on models in increments of 5, selecting the 50-topic model. Although adjacent models tend to have many similar topics, it is possible that we did not identify a vital topic covered in a different solution. Future research could also attempt using different topic modeling methods and larger sample sizes to unveil a more fine-grained view of the topics discussed. Fourth, despite covering several sentiment lexicons, they may have been limited with respect to words associated with a sentiment (eg, Bing sentiment), and research should further investigate YouTube comment sentiment to gain deeper insight into user attitudes. It should also be pointed out that the generic association of words with sentiment values and emotions could omit or alter some findings in specific contexts such as DTC genetic testing. However, we tried to minimize this effect by using different approaches and content-specific modifications such as removing the words "white" and "black" from the NRC sentiment analysis, as these were used overproportionally. Finally, although this study investigated videos spanning from 2015 to 2020, we did not specifically focus on whether or how user discourse and attitudes might have changed over time. Because we only collected the 500 most recent comments, the majority of these can be dated to 2021. However, the DTC genetic testing market has and continues to evolve and change rapidly [1,2,7,14]. Future research should thus consider a temporal analysis of DTC genetic testing videos and comments to investigate if the market changes also affected user discourse and attitudes.

Conclusions

This study examined 248 DTC genetic testing videos and 84,082 comments on YouTube to investigate user discourse. To this end, we employed topic modeling and identified 6 prevailing topics discussed among users, which largely revolve around the test purposes mentioned within those videos, such as ancestry or relationship testing. Further, we conducted sentiment analysis, showing that users have positive emotions, as indicated by the NRC sentiments of anticipation, joy, surprise, and trust, and a generally neutral-to-positive attitude toward DTC genetic testing expressed through words such as love, beautiful, pretty, and cool as well as a positive attitude toward DTC genetic testing–related videos on YouTube in general. Through this study, we show how users' attitudes toward DTC genetic testing can be determined by analyzing topics and opinions based on YouTube video comments. Our findings show that users are highly interested in DTC genetic testing and related social media content. Nonetheless, with this novel market still evolving, service providers, content providers, or regulatory authorities may need to adapt their services to users' interests and desires.

Acknowledgments

The present contribution is supported by the Helmholtz Association under the joint research school "HIDSS4Health – Helmholtz Information and Data Science School for Health." We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Direct-to-consumer genetic testing on social media.
[PDF File (Adobe PDF File), 61 KB - infodemiology_v2i2e38749_app1.pdf]

Multimedia Appendix 2
Data collection process.
[PDF File (Adobe PDF File), 97 KB - infodemiology_v2i2e38749_app2.pdf]

Multimedia Appendix 3
Structural topic modeling.
[PDF File (Adobe PDF File), 321 KB - infodemiology_v2i2e38749_app3.pdf]

Multimedia Appendix 4
Word frequency analysis results.
[PDF File (Adobe PDF File), 122 KB - infodemiology_v2i2e38749_app4.pdf]

References

1. Allyse MA, Robinson DH, Ferber MJ, Sharp RR. Direct-to-Consumer Testing 2.0: emerging models of direct-to-consumer genetic testing. Mayo Clin Proc 2018 Jan;93(1):113-120. [doi: 10.1016/j.mayocp.2017.11.001] [Medline: 29304915]
2. Ramos E, Weissman SM. The dawn of consumer-directed testing. Am J Med Genet C Semin Med Genet 2018 Mar;178(1):89-97. [doi: 10.1002/ajmg.c.31603] [Medline: 29512889]
3. Ugalmugle S, Swain R. Direct-to-consumer (DTC) genetic testing market size projections 2022 - 2028. Global Market Insights. 2022 Apr. URL: https://www.gminsights.com/industry-analysis/direct-to-consumer-dtc-genetic-testing-market [accessed 2022-09-02]
4. Haga SB, Willard HF. Letting the genome out of the bottle. N Engl J Med 2008 May;358(20):2184. [doi: 10.1056/NEJMc086053] [Medline: 18480218]
5. Hudson K, Javitt G, Burke W, Byers P, American Society of Human Genetics Social Issues Committee. ASHG statement* on direct-to-consumer genetic testing in the United States. Obstet Gynecol 2007 Dec;110(6):1392-1395. [doi: 10.1097/01.AOG.0000292086.98514.8b] [Medline: 18055737]
6. Hunter DJ, Khoury MJ, Drazen JM. Letting the genome out of the bottle--will we get our wish? N Engl J Med 2008 Jan;358(2):105-107. [doi: 10.1056/NEJMp0708162] [Medline: 18184955]
7. Thiebes S, Toussaint PA, Ju J, Ahn J, Lyytinen K, Sunyaev A. Valuable genomes: taxonomy and archetypes of business models in direct-to-consumer genetic testing. J Med Internet Res 2020 Jan;22(1):e14890 [FREE Full text] [doi: 10.2196/14890] [Medline: 31961329]
8. Saey TH. What I actually learned about my family after trying 5 DNA ancestry tests. ScienceNews. 2018 Jun. URL: https://www.sciencenews.org/article/family-dna-ancestry-tests-review-comparison [accessed 2022-09-02]
9. Briscoe F, Ajunwa I, Gaddis A, McCormick J. Evolving public views on the value of one's DNA and expectations for genomic database governance: results from a national survey. PLoS One 2020 Mar;15(3):e0229044 [FREE Full text] [doi: 10.1371/journal.pone.0229044] [Medline: 32160204]
10. Majumder MA, Cook-Deegan R, McGuire AL. Beyond our borders? Public resistance to global genomic data sharing. PLoS Biol 2016 Nov;14(11):e2000206 [FREE Full text] [doi: 10.1371/journal.pbio.2000206] [Medline: 27806054]
11. Majumder MA, Guerrini CJ, McGuire AL. Direct-to-consumer genetic testing: value and risk. Annu Rev Med 2021 Jan;72:151-166. [doi: 10.1146/annurev-med-070119-114727] [Medline: 32735764]
12. Raz AE, Niemiec E, Howard HC, Sterckx S, Cockbain J, Prainsack B. Transparency, consent and trust in the use of customers' data by an online genetic testing company: an exploratory survey among 23andMe users. New Genet Soc 2020 May;39(4):459-482. [doi: 10.1080/14636778.2020.1755636]
13. Thiebes S, Lyytinen K, Sunyaev A. Sharing is about caring? Motivating and discouraging factors in sharing individual genomic data. In: ICIS 2017 Proceedings. 2017 Presented at: Thirty Eighth International Conference on Information Systems; December 10-13, 2017; Seoul, South Korea URL: https://aisel.aisnet.org/icis2017/IT-and-Healthcare/Presentations/15/
14. Beyene M, Toussaint PA, Thiebes S, Schlesner M, Brors B, Sunyaev A. A scoping review of distributed ledger technology in genomics: thematic analysis and directions for future research. J Am Med Inform Assoc 2022 Jul;29(8):1433-1444. [doi: 10.1093/jamia/ocac077] [Medline: 35595301]
15. Toussaint PA, Thiebes S, Schmidt-Kraepelin M, Sunyaev A. Perceived fairness of direct-to-consumer genetic testing business models. Electron Mark 2022 Jul:1-18 [FREE Full text] [doi: 10.1007/s12525-022-00571-x] [Medline: 35874303]
16. Weidman J, Aurite W, Grossklags J. On sharing intentions, and personal and interdependent privacy considerations for genetic data: a vignette study. IEEE/ACM Trans Comput Biol Bioinform 2019 Jul;16(4):1349-1361. [doi: 10.1109/TCBB.2018.2854785] [Medline: 30004885]

XSL•FO
RenderX

17. Boodman E. White nationalists are flocking to genetic ancestry tests--with surprising results. Scientific American. 2017 Aug. URL: https://www.scientificamerican.com/article/white-nationalists-are-flocking-to-genetic-ancestry-tests-with-surprising-results/ [accessed 2022-09-02]
18. Mittos A, Zannettou S, Blackburn J, Cristofaro ED. Analyzing genetic testing discourse on the web through the lens of Twitter, Reddit, and 4chan. ACM Trans Web 2020 Sep;14(4):1-38. [doi: 10.1145/3404994]
19. 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]
20. Basch CH, Fera J, Quinones N. A content analysis of direct-to-consumer DNA testing on TikTok. J Community Genet 2021 Jul;12(3):489-492 [FREE Full text] [doi: 10.1007/s12687-021-00526-3] [Medline: 33860464]
21. Lee NM, Abitbol A, VanDyke MS. Science communication meets consumer relations: an analysis of Twitter use by 23andMe. Sci Commun 2020 Apr;42(2):244-264. [doi: 10.1177/1075547020914906]
22. McLachlan S. 23 YouTube stats that matter to marketers in 2022. Hootsuite. 2022 Feb. URL: https://blog.hootsuite.com/youtube-stats-marketers/ [accessed 2022-09-02]
23. Zheng C, Xue J, Sun Y, Zhu T. Public opinions and concerns regarding the Canadian prime minister's daily COVID-19 briefing: longitudinal study of YouTube comments using machine learning techniques. J Med Internet Res 2021 Feb;23(2):e23957 [FREE Full text] [doi: 10.2196/23957] [Medline: 33544690]
24. Eke R, Li T, Bond K, Ho A, Graves L. Viewing trends and users' perceptions of the effect of sleep-aiding music on YouTube: quantification and thematic content analysis. J Med Internet Res 2020 Aug;22(8):e15697 [FREE Full text] [doi: 10.2196/15697] [Medline: 32831182]
25. Mittos A, Blackburn J, Cristofaro ED. "23andMe confirms: I'm super white" - analyzing Twitter discourse on genetic testing. ArXiv. Preprint posted online on Jan 30, 2018 2021 [FREE Full text] [doi: 10.48550/arXiv.1801.09946]
26. Roberts MC, Allen CG, Andersen BL. The FDA authorization of direct-to-consumer genetic testing for three pathogenic variants: a Twitter analysis of the public's response. JAMIA Open 2019 Dec;2(4):411-415 [FREE Full text] [doi: 10.1093/jamiaopen/ooz037] [Medline: 32025636]
27. Yin Z, Song L, Clayton EW, Malin BA. Health and kinship matter: learning about direct-to-consumer genetic testing user experiences via online discussions. PLoS One 2020 Sep;15(9):e0238644 [FREE Full text] [doi: 10.1371/journal.pone.0238644] [Medline: 32898148]
28. Basch CH, Hillyer GC, Wahrman MZ, Garcia P, Basch CE. DNA testing information on YouTube: inadequate advice can mislead and harm the public. J Genet Couns 2021 Jun;30(3):785-792. [doi: 10.1002/jgc4.1375] [Medline: 33382174]
29. Harris A, Kelly SE, Wyatt S. Autobiologies on YouTube: narratives of direct-to-consumer genetic testing. New Genet Soc 2014 Mar;33(1):60-78 [FREE Full text] [doi: 10.1080/14636778.2014.884456] [Medline: 24772003]
30. Kragh-Furbo M, Tutton R. Spitting images: remaking saliva as a promissory substance. New Genet Soc 2017 May;36(2):159-185. [doi: 10.1080/14636778.2017.1320943]
31. Marcon AR, Rachul C, Caulfield T. The consumer representation of DNA ancestry testing on YouTube. New Genetics and Society 2020 Aug;40(2):133-154. [doi: 10.1080/14636778.2020.1799343]
32. 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;22(8):e19996 [FREE Full text] [doi: 10.2196/19996] [Medline: 32750004]
33. Bennett GG, Glasgow RE. The delivery of public health interventions via the internet: actualizing their potential. Annu Rev Public Health 2009 Apr;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
34. Rezaallah B, Lewis DJ, Pierce C, Zeilhofer H, Berg B. Social media surveillance of multiple sclerosis medications used during pregnancy and breastfeeding: content analysis. J Med Internet Res 2019 Aug;21(8):e13003 [FREE Full text] [doi: 10.2196/13003] [Medline: 31392963]
35. Robinson P, Turk D, Jilka S, Cella M. Measuring attitudes towards mental health using social media: investigating stigma and trivialisation. Soc Psychiatry Psychiatr Epidemiol 2019 Jan;54(1):51-58 [FREE Full text] [doi: 10.1007/s00127-018-1571-5] [Medline: 30069754]
36. Vraga EK, Stefanidis A, Lamprianidis G, Croitoru A, Crooks AT, Delamater PL, et al. Cancer and social media: a comparison of traffic about breast cancer, prostate cancer, and other reproductive cancers on Twitter and Instagram. J Health Commun 2018 Jan;23(2):181-189. [doi: 10.1080/10810730.2017.1421730] [Medline: 29313761]
37. Ren F, Kang X, Quan C. Examining accumulated emotional traits in suicide blogs with an emotion topic model. IEEE J Biomed Health Inform 2016 Sep;20(5):1384-1396. [doi: 10.1109/JBHI.2015.2459683] [Medline: 26208372]
38. Jelodar H, Wang Y, Orji R, Huang S. Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach. IEEE J Biomed Health Inform 2020 Oct;24(10):2733-2742. [doi: 10.1109/JBHI.2020.3001216] [Medline: 32750931]
39. Zhang J, Zheng Z, Wang Y, Zhu Y. A study on users' preference towards diabetes-related video clips on YouTube. BMC Med Inform Decis Mak 2020 Feb;20(1):43 [FREE Full text] [doi: 10.1186/s12911-020-1035-1] [Medline: 32111208]
40. Venkatesaramani R, Downey D, Malin B, Vorobeychik Y. A semantic cover approach for topic modeling. In: Proceedings of the Eighth Joint Conference on Lexical and Computational Semantics (*SEM 2019)). 2019 Presented at: Eighth Joint

XSL•FO
RenderX

Conference on Lexical and Computational Semantics; June 6-7, 2019; Minneapolis, MN p. 92-102. [doi: 10.18653/v1/s19-1011]
41. Jiang J, Adam M, Benlian A. ECIS 2021 Research Papers. 2021 Presented at: 29th European Conference on Information Systems; June 14-16, 2021; Marrakech, Morocco.
42. Silge J, Robinson D. Text Mining With R: A Tidy Approach. Sebastopol, CA: O'Reilly Media; 2017.
43. Roberts ME, Stewart BM, Tingley D. stm: An R package for structural topic models. J Stat Soft 2019 Oct;91(2):1-40. [doi: 10.18637/jss.v091.i02]
44. Taddy M. On estimation and selection for topic models. In: Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics, PMLR 22. 2012 Presented at: Fifteenth International Conference on Artificial Intelligence and Statistics; April 21-23, 2012; La Palma, Canary Islands p. 1184-1193. [doi: 10.1109/inmic.2016.7840071]
45. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. In: Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing. 2011 Presented at: Conference on Empirical Methods in Natural Language Processing; July 27-31, 2011; Edinburgh, Scotland p. 262-272.
46. Liu B. Sentiment analysis and opinion mining. Synth Lect Hum Lang Technol 2012 May;5(1):1-167. [doi: 10.2200/s00416ed1v01y201204hlt016]
47. Cattell RB. Sentiment or attitude? The core of a terminology problem in personality research. J Pers 1940 Sep;9(1):6-17. [doi: 10.1111/j.1467-6494.1940.tb02192.x]
48. Hu M, Liu B. Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2004 Presented at: KDD04: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 22-25, 2004; Seattle, WA p. 168-177. [doi: 10.1145/1014052.1014073]
49. Mohammad SM, Turney PD. Crowdsourcing a word-emotion association lexicon. Comput Intell 2013 Aug;29(3):436-465. [doi: 10.1111/j.1467-8640.2012.00460.x]
50. Thelwall M, Buckley K, Paltoglou G, Cai D, Kappas A. Sentiment strength detection in short informal text. J Am Soc Inf Sci 2010 Dec;61(12):2544-2558. [doi: 10.1002/asi.21416]
51. Chow-White P, Struve S, Lusoli A, Lesage F, Saraf N, Oldring A. ‘Warren Buffet is my cousin’: shaping public understanding of big data biotechnology, direct-to-consumer genomics, and 23andMe on Twitter. Inf Commun Soc 2017 Feb;21(3):448-464. [doi: 10.1080/1369118x.2017.1285951]
52. Nunziato DC. Misinformation mayhem: social media platforms' efforts to combat medical and political misinformation. GW Law Faculty Publications & Other Works 2020 Aug:1-68 [FREE Full text]

Abbreviations

API: application programming interface
DTC: direct-to-consumer
NRC: National Research Council Canada
RQ: research question

Edited by A Mavragani; submitted 14.04.22; peer-reviewed by A Pal, B Zimmermann, S Persky; comments to author 09.05.22; revised version received 15.07.22; accepted 18.08.22; published 15.09.22.

Please cite as:
Toussaint PA, Renner M, Lins S, Thiebes S, Sunyaev A
Direct-to-Consumer Genetic Testing on Social Media: Topic Modeling and Sentiment Analysis of YouTube Users' Comments
JMIR Infodemiology 2022;2(2):e38749
URL: https://infodemiology.jmir.org/2022/2/e38749
doi:10.2196/38749
PMID:37113449

©Philipp A Toussaint, Maximilian Renner, Sebastian Lins, Scott Thiebes, Ali Sunyaev. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 15.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Perspectives of the COVID-19 Pandemic on Reddit: Comparative Natural Language Processing Study of the United States, the United Kingdom, Canada, and Australia

Mengke Hu[1], PhD; Mike Conway[1,2,3], PhD

[1]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

[2]School of Computing & Information Systems, University of Melbourne, Carlton, Australia

[3]Centre for Digital Transformation of Health, University of Melbourne, Carlton, Australia

Corresponding Author:
Mengke Hu, PhD
Department of Biomedical Informatics
University of Utah
6301 S Madrid St
Salt Lake City, UT, 84121
United States
Phone: 1 2159150417
Email: mengke.hu@utah.edu

Abstract

Background: Since COVID-19 was declared a pandemic by the World Health Organization on March 11, 2020, the disease has had an unprecedented impact worldwide. Social media such as Reddit can serve as a resource for enhancing situational awareness, particularly regarding monitoring public attitudes and behavior during the crisis. Insights gained can then be utilized to better understand public attitudes and behaviors during the COVID-19 crisis, and to support communication and health-promotion messaging.

Objective: The aim of this study was to compare public attitudes toward the 2020-2021 COVID-19 pandemic across four predominantly English-speaking countries (the United States, the United Kingdom, Canada, and Australia) using data derived from the social media platform Reddit.

Methods: We utilized a topic modeling natural language processing method (more specifically latent Dirichlet allocation). Topic modeling is a popular unsupervised learning technique that can be used to automatically infer topics (ie, semantically related categories) from a large corpus of text. We derived our data from six country-specific, COVID-19–related subreddits (r/CoronavirusAustralia, r/CoronavirusDownunder, r/CoronavirusCanada, r/CanadaCoronavirus, r/CoronavirusUK, and r/coronavirusus). We used topic modeling methods to investigate and compare topics of concern for each country.

Results: Our consolidated Reddit data set consisted of 84,229 initiating posts and 1,094,853 associated comments collected between February and November 2020 for the United States, the United Kingdom, Canada, and Australia. The volume of posting in COVID-19–related subreddits declined consistently across all four countries during the study period (February 2020 to November 2020). During lockdown events, the volume of posts peaked. The UK and Australian subreddits contained much more evidence-based policy discussion than the US or Canadian subreddits.

Conclusions: This study provides evidence to support the contention that there are key differences between salient topics discussed across the four countries on the Reddit platform. Further, our approach indicates that Reddit data have the potential to provide insights not readily apparent in survey-based approaches.

(JMIR Infodemiology 2022;2(2):e36941) doi:10.2196/36941

KEYWORDS

COVID-19; social media; natural language processing; Reddit

XSL•FO
RenderX

Introduction

In December 2019, several cases of respiratory disease were reported in Wuhan City, China [1]. This respiratory disease, ultimately named COVID-19, was caused by a novel coronavirus identified as SARS-CoV-2. COVID-19 is a highly contagious infection, typically spread through respiratory droplets or by contact [2]. In the period since COVID-19 was declared a pandemic by the World Health Organization (WHO) on March 11, 2020, the disease has had an unprecedented impact worldwide, with, as of June 13, 2022, more than 540 million confirmed cases and 6.3 million deaths [3]. The number of people who have died because of the COVID-19 pandemic could be roughly three times higher than official figures suggest, according to a new analysis [4].

To suppress the transmission of COVID-19, governments have enforced several waves of border shutdowns, travel restrictions, quarantine, and other nonpharmaceutical interventions such as mask mandates, limiting public activities, and restricting travel [5-7], sparking fears of social unrest, educational disruption, and economic crisis [8]. The scientific uncertainties regarding the virus and its transmission have created a volatile political and social environment [9,10]. These concerns are exacerbated by the dynamic nature of the virus, with new variants emerging over time [10,11], creating uncertainty regarding the projected course of the pandemic and impacts on policy. Further, the advent of COVID-19 has been associated with a marked deterioration in population-level mental health issues, especially for vulnerable populations such as college students and pregnant women [12-14].

Traditional surveillance systems, including those utilized by the US Centers for Disease Control and Prevention and the European Influenza Surveillance Scheme, rely on both virologic and clinical data, and publish data once per week, typically with a 1-2–week reporting lag [15]. Survey data have also been leveraged to investigate the spread of COVID-19 in the community. In particular, ecological momentary assessment has proven to be a valuable research tool [16]. Further, the peer-reviewed scientific literature and preprint data are popular data sources to study the impact of COVID-19.

Social media such as Reddit, Twitter, Facebook, Weibo, and others provide a readily available source of abundant, organic, publicly accessible first-person narratives [17-25], which can serve as data sets for identifying outbreaks and providing situational awareness. Even more important during the COVID-19 pandemic, social media data provide a means of better understanding public attitudes and behaviors during a crisis to support communication and health-promotion messaging, especially in situations in which survey data are not readily available [15,26].

During lockdown events, social media platforms have—through their individual users—provided informational support and online access to services for pregnant women to obtain prenatal care services, such as consulting and scheduling necessary appointments [27]. Similarly, Weibo posts have proved useful in investigating public attitudes toward COVID-19 vaccination in China [28,29]. Alternative data sources such as Reddit are especially valuable in situations where traditional survey data are limited. For example, Reddit has been employed to study the impact of the pandemic on disordered eating behaviors [30].

Topic modeling, a popular statistical unsupervised machine-learning technique, has been widely used for discovering the underlying themes that occur in collections of health-related texts [31]. Because of its utility in facilitating the analysis of large-scale document collections, useful results have been obtained in areas such as biological/biomedical text mining; clinical informatics; and information extraction from other text data sources, including government reports, newspaper articles, and scientific journals [32]. Social media data such as Reddit are frequently used in conjunction with topic modeling methods to explore public concerns, attitudes, and policies. For example, Zhang et al [33] identified eight popular topics using Chinese social media platforms that served to characterize the COVID-19 infodemic, including conspiracy theories, government response, preventive action, new cases, transmission routes, origin and nomenclature, vaccines and medicines, and symptoms and detection. Topic modeling has also been used to examine COVID-19–related concerns across different countries [34]. Categorizing posts by topic modeling technique such as latent Dirichlet allocation (LDA) [35], perhaps the most popular topic modeling method, has been used extensively to analyze sentiments and concerns during the COVID-19 crisis [1,10,19,20,36-41], especially in the context of large social media data sets. Topic modeling with LDA has also demonstrated utility in discovering themes from combined data sets, such as combining news articles and tweets in Brazil to study the impact of COVID-19 [42]. LDA has also been used to study sentiment variations over time [10,43-46]. In particular, as COVID-19 vaccine–related issues received increasing public attention, LDA was employed to study the changes in people's opinions toward COVID-19 vaccination, discovering that public attitudes became more favorable over time [47,48]. However, whether the topics identified are interpretable typically requires qualitative evaluation [49,50].

Reddit is one of the most popular social media platforms with over 430 million active users and 1.2 million subreddits (ie, topic-focused subforums) as of May 2020, with over 70% of its user base coming from English-speaking countries [51,52]. Some subreddits have clear descriptions regarding locations (eg, *r/CoronavirusUK*, *r/CanadaCoronavirus*), which enables a more targeted analysis of users from different countries [43].

In this work, we employed Reddit data from six geographically specific COVID-19–related subreddits representing four English-speaking countries, the United States, the United Kingdom, Canada, and Australia, to investigate (1) whether there were key differences between salient topics discussed across the four countries and (2) whether Reddit data have the potential to provide insights not readily apparent in survey-based approaches. In general, LDA topic modeling was applied to each country-specific Reddit data set. We trained multiple topic models for each country consisting of a different number of topics and manually inspected each model to find the optimal model for each country (ie, the model that generated the most coherent and least redundant topics). We further compared the summarized topics for each country based on each country's

XSL•FO
RenderX

model, and mapped them to four common topic categories (ie, metacategories). Finally, longitudinal topic trends were examined to identify trends in the common topic categories, which were then mapped to the COVID-19 events for each country.

Methods

Data Collection and Preprocessing

As Reddit data do not generally include geolocation information, we collected data from the six most popular subreddits (topical forums on Reddit) related to the United States, the United Kingdom, Canada, and Australia (*r/CoronavirusUK*, *r/coronavirusus*, *r/CoronavirusCanada*, *r/CanadaCoronavirus*, *r/CoronavirusAustralia*, *r/CoronavirusDownunder*), as shown in Table 1.

Data were collected using the pushshift.io [52] application programming interface (API), a service that archives Reddit data to its online database in real time. We employed the pushshift.io API to harvest COVID-19–related data, as previous work has indicated that this approach yields a more complete data set than alternative methods (eg, the PRAW API) [53]. However, in the data collection process, we noticed that pushshift.io failed to identify all of the new updates, including deleted comments [54]. To ensure we collected the most complete data set possible, we recollected the data over the same time frame after 3 months and consolidated the new and old data sets to gain a more complete data set.

The consolidated Reddit data set consisted of 84,229 initiating posts and 1,094,853 associated comments collected between February and November 2020 derived from the six subreddits shown in Table 1. These subreddits are related to a specific country according to the subreddit description. For example, *r/CanadaCoronavirus* is used primarily by Canadians to discuss the COVID-19 crisis. Among all the country-specific COVID-19 subreddits, the six subreddits we chose have the largest number of members (>8000), which means they are the most active and popular geographically specific COVID-19–related subreddits available. As users typically present their own experiences in the initiating posts [55], with subsequent comments frequently subject to off-topic discussion, we restricted our topic study to only initiating posts. Given that Reddit does not provide user-level geolocation information, we regarded the fact that a Reddit user posted in a country-specific subreddit as a proxy for their location in that country.

To build the corpus for each country, we organized the submissions from the six subreddits shown in Table 1. For example, to build an Australia data set, we extracted all text data (the title section and the description section) from the submissions of *r/CoronavirusAustralia* and *r/CoronavirusDownunder*. We then automatically identified URLs and email addresses, which were removed from the texts of submissions to simplify the subsequent topic modeling process. To remove the stop words (ie, common English function words such as “the,” “of,” and “it”), we first used the Natural Language Toolkit (NLTK 3.3 for Python 2.7) [56] to initialize the stop-words list. The stop-words list was then further augmented using the Essential Word List (a lexicon originally developed for language learning and testing) [57]. Subsequently, the text data from submissions were tokenized (ie, the string *Let’s go!* was tokenized into the list “let,” “’s,” “go,” “!”) and lemmatized (ie, the string *I was reading the paper* was broken down into the list “I,” “be,” “read,” “the,” “paper”) using the Python SpaCy 2.2.1 package [58] to convert various forms of words (eg, *cough*, *coughing*) into a canonical form (eg, *cough*).

Table 1. Subreddit information on July 31, 2021.

Country	Subreddit	Number of members	Date subreddit created
United Kingdom	*r/CoronavirusUK*	92,600	February 11, 2020
United States	*r/coronavirusus*	141,000	February 12, 2020
Canada	*r/CoronavirusCanada*	9000	February 12, 2020
Canada	*r/CanadaCoronavirus*	67,300	March 1, 2020
Australia	*r/CoronavirusAustralia*	10,800	February 21, 2020
Australia	*r/CoronavirusDownunder*	90,300	February 23, 2020

Topic Modeling and Common Topic Annotation

We used the topic modeling technique to compare the broad themes emerging from the United States, the United Kingdom, Canada, and Australia. The general procedure is described in Figure 1. Specifically, we adopted a generative probabilistic modeling algorithm, LDA, which models documents as random mixtures over topics, where each topic is characterized as a distribution of words [35].

We trained multiple topic models (consisting of 10, 15, and 20 topics) for each of the four countries using the LDA implementation in the Gensim 3.8.3 [59] toolkit. Under each model, we summarized the topics according to the topic keywords. We then manually checked if the topics overlapped or were redundant. We found that topics thematically overlapped when the model contained fewer than 10 topics, while the topics were redundant when the model had more than 20 topics. Thus, we chose 10, 15, and 20 topics to train the models for further manual examination.

For each topic model, the most characteristic keywords associated with each of the thematic topics were manually examined, focusing specifically on the posts that were particularly representative according to the contribution probability of those topics to determine which model best characterized the data set. In the process of manual identification of topics, we noticed that the models for the four countries had

XSL•FO
RenderX

different optimal numbers of coherent, nonoverlapping topics. Further, some models contained topics idiosyncratic to that country (ie, they did not appear in the models of other countries). For example, the "mental health" topic in the UK topic model did not appear in the US topic model. To compare and contrast the common themes among the four countries, we consolidated these various topics into four common topic categories. Topics and their mappings to the common topic categories are listed in Table 2.

Figure 1. Procedure for topic training and mapping to the common topic categories.

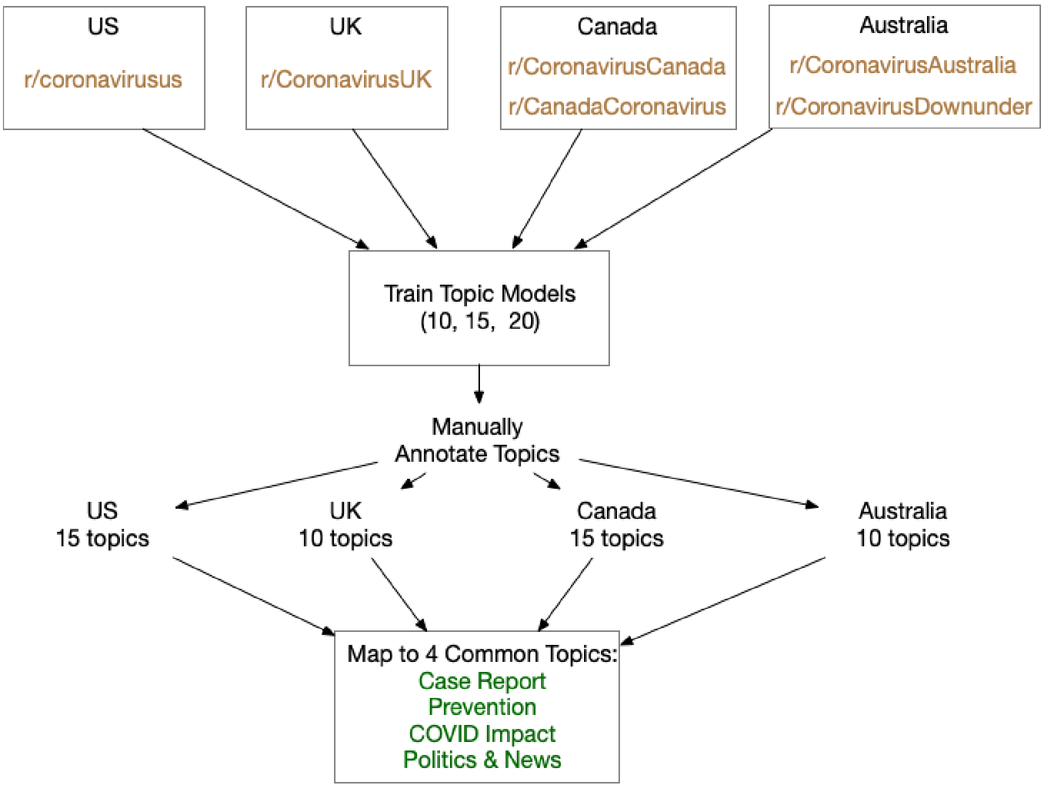


Table 2. Consolidated topics common across all four countries.

Common topic	Topics
COVID Impact	work, finance, education, travel restriction, social distancing
COVID Prevention	mask wearing, hand washing, transmission risk
Case Report	case report, report of interaction with hospital
Policy & News	policy announcement, news, question and answer

Common Topic Prevalence in the United States, the United Kingdom, Canada, and Australia

The prevalence of common topics for the United States, the United Kingdom, Canada, and Australia was studied by first finding the "document-topic" for each post. The document-topic refers to a topic that is the major constituent (according to the contribution probability) of a given document [60], which can be used to study the proportion of a specific topic for each country-related data set. As the topics and their distributions vary among the US, UK, Canada, and Australia data sets, the document-topics were analyzed separately based on each country-related data set. To find document-topics for each country, we needed to first find the threshold probability to identify the major topics. Specifically, for each country-related data set, if the topic probability for a certain document was above the threshold, this topic was deemed to be one of the major constituents for this document. Practically, document-topics were not uniformly distributed (ie, some documents contain more than one while some contain no document-topic). To evenly address each country-related data set, we iteratively tested different candidate probability values until the number of document-topics was close to the number of documents in that country-related data set. More precisely, from the topic models we trained for each country, we have: (1) a set of topics, (2) a list of words (we used 40 words) associated with each topic ranked by their contribution probability to that topic, and (3) a list of documents (submission posts) with estimates of the proportion of each topic. To find the threshold, whenever we set up the threshold probability for testing, we counted the number of document-topics for each submission and summed them for all submissions until the total number of document-topics was close to the number of submissions in that country-related data set. The reason for carrying out this process was to help ensure that the document-topics accurately covered the topics of all submissions

XSL•FO
RenderX

in the Reddit data set, thus maximizing the proportion of content that was represented [60]. We repeated this process until finding the threshold for each country.

Using the document-topic threshold for each country, we identified the proportion of each topic by first calculating the number of posts whose topic probability was above the threshold, and then dividing this number by the total number of posts to obtain the topic proportion. The proportion of the common topic categories was determined by summing the proportion of the topics that belonged to each common topics category.

Common Topic Trend in Reddit and COVID-19 Event Timeline

With the document-topic threshold for each country, we also calculated the number of submissions on a specific common topic category for each week, before counting the weekly volume of submissions on each common topic category, to plot the common topic trend for each country from February to November in 2020. We also mapped the COVID-19 event timeline from the WHO [61] and Think Global Health [62] to our Reddit data trend plot for comparison.

Ethical Considerations

We restricted our analysis to publicly available discussion content and the University of Utah's Institutional Review Board exempted the study procedure and data from ethical review (IRB_00076188) under Exemption 2 as defined in the United States' Code of Federal Regulations (CFR), 45 CFR 46.101(b).

Results

Corpus Characteristics

Our COVID-19 Reddit data set comprises 10 months of discussions (February 2020 to November 2020), which covers the main early COVID-19 events, including the initial outbreak and subsequent lockdowns in the United States, the United Kingdom, Canada, and Australia. During this time, 103,180 unique users posted some 84,229 submissions and their associated 1,094,853 comments. Table 3 summarizes the numbers of unique users, submissions, and associated comments for each subreddit.

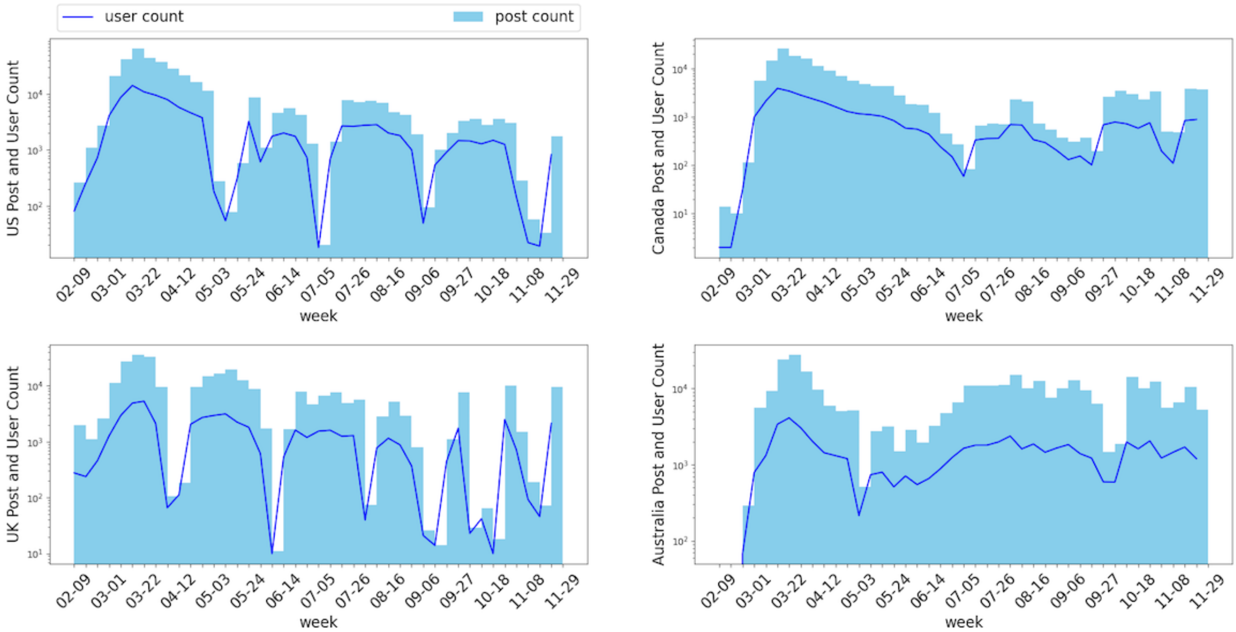
To further study the behavior of posting on Reddit, we summarized the weekly post volume and user volume for each country, as shown in Figure 2. We found that the user volume is consistent with the post volume, which indicates that posts are created by organic Reddit users rather than by a "water army" [63] of paid posters. Thus, the post data we used for analysis can be considered to be reflective of subreddit users' genuine opinions and behavior during the COVID-19 crisis. We also noticed that for the four countries, the highest volume peak appeared between February and April 2020 when the first wave of lockdowns were enforced. Moreover, the post volume and user volume decreased over time.

Table 3. Reddit COVID-19 data from February 2020 to November 2020.

Country data set	Subreddit	Unique users, n	Submissions, n
United Kingdom	r/CoronavirusUK	20,482	17,350
United States	r/coronavirusus	55,380	35,885
Canada	r/CoronavirusCanada	4061	4625
Canada	r/CanadaCoronavirus	10,420	9670
Australia	r/CoronavirusAustralia	3114	2359
Australia	r/CoronavirusDownunder	15,537	14,340

XSL•FO
RenderX

Figure 2. Post and user volume by week (log scale).



Results from Topic Modeling With Common Topic Annotation

After manually examining the topic models (10, 15, 20 topics) for the United States, the United Kingdom, Canada, and Australia, we qualitatively identified the most coherent model, as well as the threshold of the document-topics for each country-related data set, as shown in Table 4. The reason we chose the model manually instead of using automated methods (eg, LDA coherent score) is due to the limitation of topic model interpretation [49].

Table 4. Manually selected topic models for the United States, the United Kingdom, Canada, and Australia, and the associated document-topic thresholds.

Country	Chosen model	Threshold for document-topics
United States	15-topics model	0.19881
United Kingdom	10-topics mode	0.24
Canada	15-topics model	0.15864215
Australia	10-topics model	0.18434

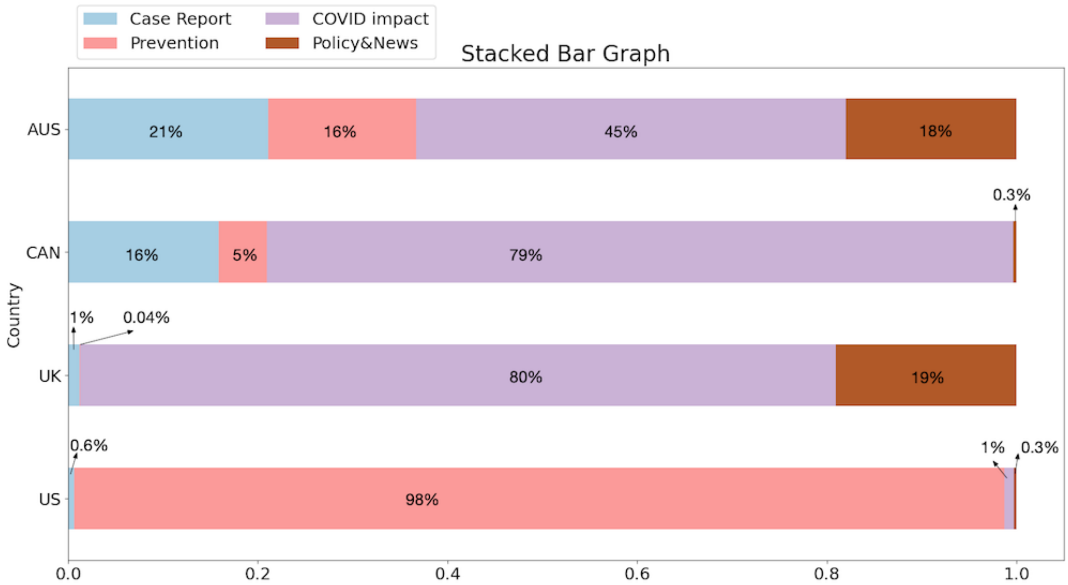
Common Topic Prevalence in the United States, the United Kingdom, Canada, and Australia

For each topic in each model, we mapped that topic to four common topics (described in Table 2) and calculated the number of documents for each topic according to the thresholds shown in Table 4. The document proportion for each topic for each country is presented in Figure 3. The detailed calculations for generating Figure 3 are presented in Multimedia Appendix 1.

We found that the majority of the US posts focused on COVID-19 prevention strategies, whereas the posts in the United Kingdom, Canada, and Australia were more focused on the impacts of COVID-19, including education, finance, and potentially limited availability of food.

XSL•FO
RenderX

Figure 3. Topic proportions for the United States (US), the United Kingdom (UK), Canada (CAN), and Australia (AUS).

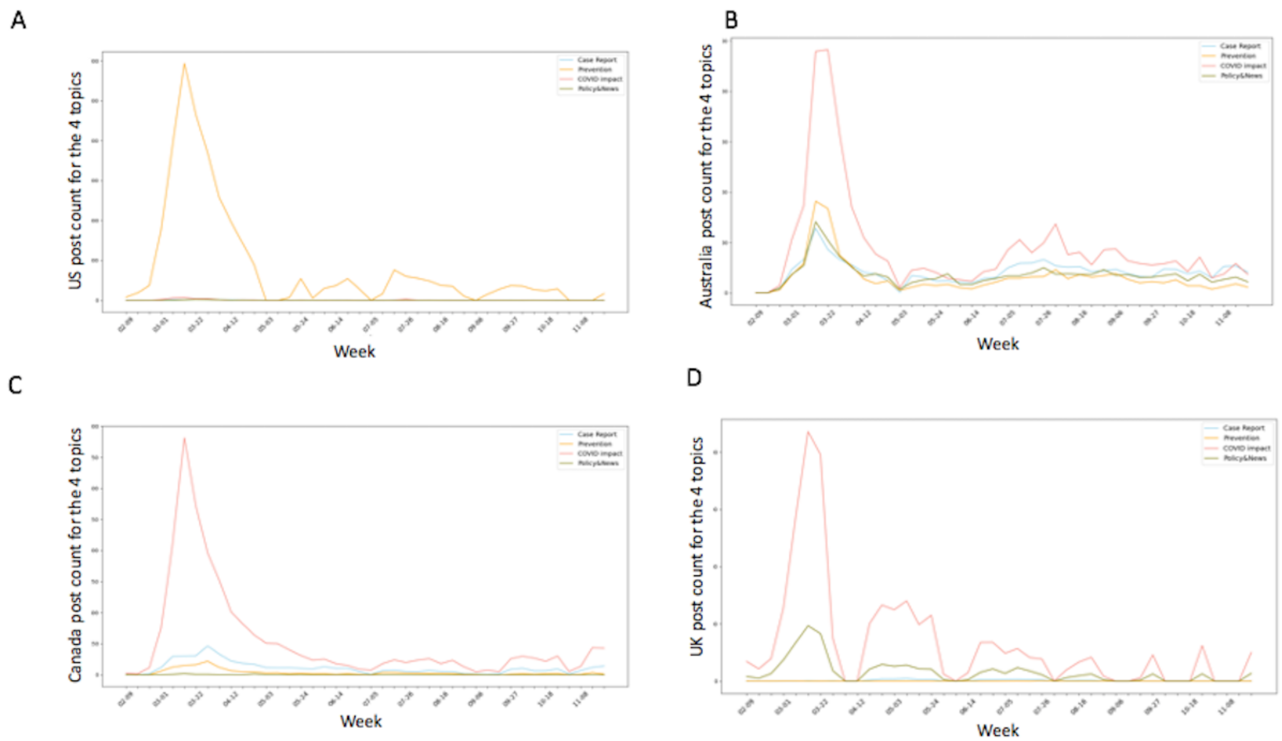


Common Topic Trend in Reddit and COVID-19 Event Timeline

In visualizing the identified topic model, we also summarize the topic trends for the United States, the United Kingdom, Canada, and Australia in Figure 4.

In both Figure 2 and Figure 4, it can be observed that all countries experienced an early peak in posting activity. The user volume plot in Figure 2 and trends in Figure 4 imply that users post more during lockdown events. For all four countries, the post volume and user volume reached a peak in March 2020. In the same month, all of these countries announced lockdown or travel restriction policies. This increase in posts may reflect a combination of public fear and concern regarding the virus, and the fact that many individuals found themselves confined to their homes, with abundant time to access social media. A list of salient pandemic-related events is shown in Table 5.

Figure 4. Topic weekly trend for (A) the United States, (B) Australia, (C) Canada, and (D) the United Kingdom. Higher-resolution version of this figure is available in Multimedia Appendix 2.



XSL•FO
RenderX

Table 5. Timeline of COVID-19 lockdown events in 2020 [62].

Date	Event
March 11	United Kingdom lockdown; United States announces level 3 travel advisory
March 18	United States and Canada suspend nonessential travel between the two countries
March 23	United Kingdom lockdown
March 24	Australia bans all overseas travel
April 18	United States: protests of the country’s lockdown
June 24	United States: increase in case rates in 26 states since easing lockdown restrictions
July 3	United Kingdom announces an end to travel restrictions except for the United States
July 4	Melbourne, Australia tightens restrictions on 12 suburbs
September 5	Australia extends its hard lockdown until the end of September
October 12	United Kingdom announces new lockdown rules

Discussion

Principal Findings

In this work, we applied topic modeling and visualization techniques to compare perspectives on events related to the COVID-19 pandemic for the United States, the United Kingdom, Canada, and Australia, and investigated the impact of COVID-19 events from February to November 2020.

Post Volume Variation for the COVID-19 Reddit Data Set

As shown in Figure 2, we observed that the post volume and user volume gradually decreased over the 10-month study period. We also observed that an early peak appeared during February and April 2020, which was the critical period for fighting the spread of COVID-19 in the United States, the United Kingdom, Canada, and Australia. One potential reason for the decline in post volume is that some users may avoid social media since they experienced increased anxiety from COVID-19–related news and discussions, and sought to protect their mental health [64]. Another reason is that users may become habituated to the “new normal,” which is identified as the acceptance phase after the authorities imposed social distancing measures [65]. In this stage, Aiello et al [65] found that people were more open to find solutions to continue social interaction; for example, the number of visits to parks and outdoor spaces increased. Hence, users posted less content in COVID-19–related subreddits to seek physical social support.

From the large volume of posts, we can see that Reddit supports the collection of a large volume of data that can provide insights into population attitudes and behavior. Previous studies have demonstrated that the analysis of public behavior and attitudes can help public health agencies and policymakers cope effectively in times of crisis [1].

Topic Variation Among the United States, the United Kingdom, Canada, and Australia

The common topics shown in Table 2 varied among the four countries. As shown in Figure 3, we found that in the United States, the majority of the posts focused on COVID-19 prevention, with only a small portion of posts directly discussing COVID-19–related policies. For the United Kingdom, Canada, and Australia, the majority of posts focused on the impact of COVID-19, including job loss, food insecurity, and feelings of anxiety. Especially for the United Kingdom and Australia, users’ concerns—at least as expressed in these subreddits—focused on the impact of COVID-19 and government policies. At the beginning of the pandemic, a core concern among the Reddit-using population centered on effective COVID-19 prevention strategies due to the scientific uncertainties regarding how the virus was transmitted [9,10]. The social impact of COVID-19 is also a leading topic, which is consistent with the fact that the COVID-19 crisis poses huge psychological pressure and is associated with mental health issues [12-14].

As shown in Figure 4, we found that the totality of topic-related posts reached a peak in March when all four countries announced a lockdown and enforced travel restrictions (see Table 5 for a summary of lockdown events). Especially in March 2020, when the COVID-19 outbreak started and governments enforced border shutdowns, travel restrictions, and quarantine [5-7], people’s topics focused on the impact of COVID-19, including education and economic disruptions [8].

Limitations

The work reported in this paper is not without limitations. COVID-19–related subreddits are still relatively new, with most of them initiated in February 2020. In the early stages of the COVID-19 pandemic, a considerable volume of COVID-19–related rumors spread [66] making Reddit data less reliable for the purposes of monitoring the outbreak, but useful for monitoring disinformation and public concerns. Additionally, Reddit has known sociodemographic biases. For example, the service is more popular in urban and suburban areas than in rural areas [67].

Topic modeling with LDA has a number of limitations, especially with respect to assessing topic quality. We noticed two problems when we manually checked the topic models: (1) very similar posts (eg, COVID-19 case report) may be assigned to different topics and (2) very simple posts (eg, lockdown announcement) may correlate to many topics. Similar problems were discovered by Xu et al [68] when analyzing clinical data.

XSL•FO
RenderX

Another issue in this work is related to the completeness of the Reddit data we collected via the pushshift.io API [52]. Although pushshift.io allows collecting a large amount of historical data from Reddit and yields a more complete data set than alternative methods (eg, the PRAW API) [53], it failed to identify all new updates, including deleted comments [54]. Even though we recollected the data to make it more complete, the Reddit data we curated may still be missing data.

A further limitation is related to the differences in culture associated with different subreddits. As Reddit data do not in general include geolocation information, we collected data from the six most popular COVID-19 subreddits related to the United States, the United Kingdom, Canada, and Australia. We examined the posts and noticed that most users are local people (ie, users from *r/CanadaCoronavirus* are mostly Canadians). Thus, the subreddits not only reflect people's opinions but also the culture differences in the four countries. For example, people in the United Kingdom concentrate on discussing politics or COVID-19–related breaking news. Thus, the leading topic, politics-related policies, in *r/CoronavirusUK* does not fully reflect people's concerns related to COVID-19, as it may simply reflect people's discussion habit in the United Kingdom. Therefore, the differences in topics may not fully reflect people's opinion toward COVID-19 in the United States, the United Kingdom, Canada, and Australia.

Finally, in this work we did not explicitly consider the demographic characteristics (eg, age, socioeconomic status, gender [52,69]) of Reddit users across the four countries and how these characteristics may differ.

Conclusion

In this work, we used Reddit data to examine variations in people's concerns during the COVID-19 crisis in the United States, the United Kingdom, Canada, and Australia. We found that people posted more on Reddit during lockdown events, and people's concerns differ among the four countries. Further, this work provides evidence to support the contention that there are key differences between salient topics discussed across the four countries on the Reddit platform. Further, our approach indicates that Reddit data have the potential to provide insights not readily apparent in survey-based approaches.

Acknowledgments

The research reported in this publication was partially supported by the "Special Emphasis: Emerging COVID-19/SARS-CoV-2 Research" seed grant program from University of Utah. The content is solely the responsibility of the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Calculation for the topic bar chart in Figure 3.
[DOCX File , 14 KB - infodemiology_v2i2e36941_app1.docx]

Multimedia Appendix 2
Higher resolution version of Figure 4. Topic weekly trend for (A) the United States, (B) Australia, (C) Canada, and (D) the United Kingdom.
[PNG File , 1124 KB - infodemiology_v2i2e36941_app2.png]

References

1. Wang J, Zhou Y, Zhang W, Evans R, Zhu C. Concerns expressed by Chinese social media users during the COVID-19 pandemic: content analysis of Sina Weibo microblogging data. J Med Internet Res 2020 Nov 26;22(11):e22152 [FREE Full text] [doi: 10.2196/22152] [Medline: 33151894]
2. Ather A, Patel B, Ruparel NB, Diogenes A, Hargreaves KM. Coronavirus disease 19 (COVID-19): implications for clinical dental care. J Endod 2020 May;46(5):584-595 [FREE Full text] [doi: 10.1016/j.joen.2020.03.008] [Medline: 32273156]
3. Coronavirus cases. Worldometer. URL: https://www.worldometers.info/coronavirus/ [accessed 2022-09-19]
4. Adam D. COVID's true death toll: much higher than official records. Nature 2022 Mar;603(7902):562. [doi: 10.1038/d41586-022-00708-0] [Medline: 35277684]
5. Coronavirus: travel restrictions, border shutdowns by country. AlJazeera. 2020 Jun 03. URL: https://www.aljazeera.com/news/2020/6/3/coronavirus-travel-restrictions-border-shutdowns-by-country [accessed 2022-09-19]
6. Betsch C, Wieler LH, Habersaat K, COSMO group. Monitoring behavioural insights related to COVID-19. Lancet 2020 Apr 18;395(10232):1255-1256 [FREE Full text] [doi: 10.1016/S0140-6736(20)30729-7] [Medline: 32247323]
7. Nicola M, Alsafi Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. Int J Surg 2020 Jun;78:185-193 [FREE Full text] [doi: 10.1016/j.ijsu.2020.04.018] [Medline: 32305533]
8. Buck T, Chazan G, Arnold M, Cookson C. Coronavirus declared a pandemic as fears of economic crisis mount. Financial Times. 2020 Mar 11. URL: https://www.ft.com/content/d72f1e54-6396-11ea-b3f3-fe4680ea68b5 [accessed 2022-09-19]

XSL•FO
RenderX

9. Werron T, Ringel L. Pandemic practices, part one. How to turn “Living Through the COVID-19 Pandemic” into a heuristic tool for sociological theorizing. Sociologica 2020;14:55-72. [doi: 10.6092/issn.1971-8853/11172]
10. Doogan C, Buntine W, Linger H, Brunt S. Public perceptions and attitudes toward COVID-19 nonpharmaceutical interventions across six countries: a topic modeling analysis of Twitter data. J Med Internet Res 2020 Sep 03;22(9):e21419 [FREE Full text] [doi: 10.2196/21419] [Medline: 32784190]
11. SARS-CoV-2 variant classifications and definitions. Centers for Disease Control and Prevention. 2022 Apr 06. URL: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html [accessed 2022-09-19]
12. Wang X, Hegde S, Son C, Keller B, Smith A, Sasangohar F. Investigating mental health of US college students during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2020 Sep 17;22(9):e22817 [FREE Full text] [doi: 10.2196/22817] [Medline: 32897868]
13. Son C, Hegde S, Smith A, Wang X, Sasangohar F. Effects of COVID-19 on college students' mental health in the United States: interview survey study. J Med Internet Res 2020 Sep 03;22(9):e21279 [FREE Full text] [doi: 10.2196/21279] [Medline: 32805704]
14. Zhang X, Liu J, Han N, Yin J. Social media use, unhealthy lifestyles, and the risk of miscarriage among pregnant women during the COVID-19 pandemic: prospective observational study. JMIR Public Health Surveill 2021 Jan 05;7(1):e25241 [FREE Full text] [doi: 10.2196/25241] [Medline: 33293263]
15. Ginsberg J, Mohebbi M, 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. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. J Med Internet Res 2020 Jun 17;22(6):e20185 [FREE Full text] [doi: 10.2196/20185] [Medline: 32519963]
17. Klein AZ, Magge A, O'Connor K, Flores Amaro JI, Weissenbacher D, Gonzalez Hernandez G. Toward using Twitter for tracking COVID-19: a natural language processing pipeline and exploratory data set. J Med Internet Res 2021 Jan 22;23(1):e25314 [FREE Full text] [doi: 10.2196/25314] [Medline: 33449904]
18. Ahmad AR, Murad HR. The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: online questionnaire study. J Med Internet Res 2020 May 19;22(5):e19556 [FREE Full text] [doi: 10.2196/19556] [Medline: 32369026]
19. 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]
20. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978 [FREE Full text] [doi: 10.2196/21978] [Medline: 33108310]
21. Foufi V, Timakum T, Gaudet-Blavignac C, Lovis C, Song M. Mining of textual health information from Reddit: analysis of chronic diseases with extracted entities and their relations. J Med Internet Res 2019 Jun 13;21(6):e12876 [FREE Full text] [doi: 10.2196/12876] [Medline: 31199327]
22. Paul MJ, Dredze M. Social Monitoring for Public Health. In: Synthesis Lectures on Information Concepts, Retrieval, and Services. San Rafael, CA: Morgan & Claypool Publishers; 2018.
23. Park A, Conway M. Longitudinal changes in psychological states in online health community members: understanding the long-term effects of participating in an online depression Community. J Med Internet Res 2017 Mar 20;19(3):e71 [FREE Full text] [doi: 10.2196/jmir.6826] [Medline: 28320692]
24. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J Med Internet Res 2017 Jun 29;19(6):e228 [FREE Full text] [doi: 10.2196/jmir.7215] [Medline: 28663166]
25. Conway M, O'Connor D. Social media, big data, and mental health: current advances and ethical implications. Curr Opin Psychol 2016 Jun;9:77-82 [FREE Full text] [doi: 10.1016/j.copsyc.2016.01.004] [Medline: 27042689]
26. Katz M, Nandi N. Social media and medical education in the context of the COVID-19 pandemic: scoping review. JMIR Med Educ 2021 Apr 12;7(2):e25892 [FREE Full text] [doi: 10.2196/25892] [Medline: 33755578]
27. Yang X, Song B, Wu A, Mo PKH, Di JL, Wang Q, et al. Social, cognitive, and eHealth mechanisms of COVID-19-related lockdown and mandatory quarantine that potentially affect the mental health of pregnant women in China: cross-sectional survey study. J Med Internet Res 2021 Jan 22;23(1):e24495 [FREE Full text] [doi: 10.2196/24495] [Medline: 33302251]
28. Yin F, Wu Z, Xia X, Ji M, Wang Y, Hu Z. Unfolding the determinants of COVID-19 vaccine acceptance in China. J Med Internet Res 2021 Jan 15;23(1):e26089 [FREE Full text] [doi: 10.2196/26089] [Medline: 33400682]
29. 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 Mar 16;23(3):e25977 [FREE Full text] [doi: 10.2196/25977] [Medline: 33651709]
30. Nutley SK, Falise AM, Henderson R, Apostolou V, Mathews CA, Striley CW. Impact of the COVID-19 pandemic on disordered eating behavior: qualitative analysis of social media posts. JMIR Ment Health 2021 Jan 27;8(1):e26011 [FREE Full text] [doi: 10.2196/26011] [Medline: 33465035]
31. Blei D. Probabilistic topic models. Commun ACM 2012 Apr;55(4):77-84. [doi: 10.1145/2133806.2133826]

XSL•FO
RenderX

32. Liu L, Tang L, Dong W, Yao S, Zhou W. An overview of topic modeling and its current applications in bioinformatics. Springerplus 2016;5(1):1608 [FREE Full text] [doi: 10.1186/s40064-016-3252-8] [Medline: 27652181]
33. Zhang S, Pian W, Ma F, Ni Z, Liu Y. Characterizing the COVID-19 infodemic on Chinese social media: exploratory study. JMIR Public Health Surveill 2021 Feb 05;7(2):e26090 [FREE Full text] [doi: 10.2196/26090] [Medline: 33460391]
34. 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 Apr 05;5(4):e23593 [FREE Full text] [doi: 10.2196/23593] [Medline: 33750736]
35. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Machine Learn Res 2003;3:993-1022 [FREE Full text] [doi: 10.5555/944919.944937]
36. Jang H, Rempel E, Roth D, Carenini G, Janjua NZ. Tracking COVID-19 discourse on Twitter in North America: infodemiology study using topic modeling and aspect-based sentiment analysis. J Med Internet Res 2021 Feb 10;23(2):e25431 [FREE Full text] [doi: 10.2196/25431] [Medline: 33497352]
37. McQuillan L, McAweeney E, Bargar A, Ruch A. Cultural convergence: insights into the behavior of misinformation networks on Twitter. arXiv. 2020 Jul 07. URL: https://arxiv.org/abs/2007.03443 [accessed 2022-09-19]
38. Koh JX, Liew TM. How loneliness is talked about in social media during COVID-19 pandemic: text mining of 4,492 Twitter feeds. J Psychiatr Res 2022 Jan;145:317-324 [FREE Full text] [doi: 10.1016/j.jpsychires.2020.11.015] [Medline: 33190839]
39. 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]
40. 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]
41. Adikari A, Nawaratne R, De Silva D, Ranasinghe S, Alahakoon O, Alahakoon D. Emotions of COVID-19: content analysis of self-reported information using artificial intelligence. J Med Internet Res 2021 Apr 30;23(4):e27341 [FREE Full text] [doi: 10.2196/27341] [Medline: 33819167]
42. 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 Feb 10;7(2):e24585 [FREE Full text] [doi: 10.2196/24585] [Medline: 33480853]
43. Biester L, Matton K, Rajendran J, Mower E, Mihalcea R. Quantifying the effects of COVID-19 on mental health support forums. arXiv. 2020 Sep 08. URL: https://arxiv.org/abs/2009.04008 [accessed 2022-09-19]
44. 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]
45. Wang X, Zou C, Xie Z, Li D. Public opinions towards COVID-19 in California and New York on Twitter. medRxiv. 2020 Jul 14. URL: https://www.medrxiv.org/content/10.1101/2020.07.12.20151936v1 [accessed 2022-09-19]
46. Stokes DC, Andy A, Guntuku SC, Ungar LH, Merchant RM. Public priorities and concerns regarding COVID-19 in an online discussion forum: longitudinal topic modeling. J Gen Intern Med 2020 Jul;35(7):2244-2247 [FREE Full text] [doi: 10.1007/s11606-020-05889-w] [Medline: 32399912]
47. 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]
48. Kwok SWH, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: machine learning Analysis. J Med Internet Res 2021 May 19;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: 33886492]
49. Chang J, Boyd-Graber J, Gerrish S, Wang C, Blei DM. Reading tea leaves: how humans interpret topic models. 2009 Presented at: NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems; December 7-9, 2009; Vancouver, BC p. 288-296 URL: https://dl.acm.org/doi/10.5555/2984093.2984126
50. Jang H, Rempel E, Carenini G, Janjua N. Exploratory analysis of COVID-19 related tweets in North America to inform public health institutes. 2020 Presented at: 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020; December 2020; online. [doi: 10.18653/v1/2020.nlpcovid19-2.18]
51. Aggarwal J, Rabinovich E, Stevenson S. Exploration of gender differences in COVID-19 discourse on Reddit. 2020 Presented at: 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020; December 2020; online.
52. Reddit. pushshift. URL: https://pushshift.io/reddit/ [accessed 2022-09-19]
53. Gaffney D, Matias JN. Caveat emptor, computational social science: Large-scale missing data in a widely-published Reddit corpus. PLoS One 2018 Jul 6;13(7):e0200162 [FREE Full text] [doi: 10.1371/journal.pone.0200162] [Medline: 29979741]
54. How many redditors delete their posts? Reddit. URL: https://www.reddit.com/r/pushshift/comments/ikpxrf/how_many_redditors_delete_their_posts/ [accessed 2022-09-19]
55. MacLean D, Gupta S, Lembke A, Manning C, Heer J. Forum77: An analysis of an online health forum dedicated to addiction recovery. 2015 Presented at: CSCW '15: Proceedings of the 18th ACM Conference on Computer Supported Cooperative

XSL•FO
RenderX

Work & Social Computing; March 14-18, 2015; Vancouver, BC URL: https://doi.org/10.1145/2675133.2675146 [doi: 10.1145/2675133.2675146]
56. Natural Language Toolkit. URL: https://www.nltk.org/ [accessed 2022-09-19]
57. Essential Word List. URL: https://www.edu.uwo.ca/faculty-profiles/docs/other/webb/essential-word-list.pdf [accessed 2022-09-19]
58. Library architecture. spaCy Industrial-Strength Natural Language Processing. URL: https://spacy.io/api [accessed 2022-09-19]
59. Latent Dirichlet Allocation. GENSIM. URL: https://radimrehurek.com/gensim/models/ldamodel.html [accessed 2022-09-19]
60. Chen AT, Zhu S, Conway M. What online communities can tell us about electronic cigarettes and Hookah use: a study using text mining and visualization techniques. J Med Internet Res 2015 Sep 29;17(9):e220 [FREE Full text] [doi: 10.2196/jmir.4517] [Medline: 26420469]
61. Timeline: WHO's COVID-19 response. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline/ [accessed 2022-09-19]
62. Kantis C, Kiernan S, Bardi JS, Posner L. UPDATED: Timeline of the Coronavirus. Think Global Health. 2022 Sep 16. URL: https://www.thinkglobalhealth.org/article/updated-timeline-coronavirus [accessed 2022-09-19]
63. Chen C, Wu K, Srinivasan V, Zhang X. Battling the internet water army: detection of hidden paid posters. 2013 Presented at: ASONAM '13: Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining; August 25-28, 2013; Niagara, ON. [doi: 10.1145/2492517.2492637]
64. 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]
65. Aiello L, Quercia D, Zhou K, Constantinides M, Šćepanović S, Joglekar S. How epidemic psychology works on Twitter: evolution of responses to the COVID-19 pandemic in the U.S. Humanit Soc Sci Commun 2021 Jul 23;8(1):179. [doi: 10.1057/s41599-021-00861-3]
66. 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. [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
67. 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]
68. Xu X, Jin T, Wei Z, Wang J. Incorporating topic assignment constraint and topic correlation limitation into clinical goal discovering for clinical pathway mining. J Healthc Eng 2017;2017:5208072. [doi: 10.1155/2017/5208072] [Medline: 29065617]
69. Zhang C, Xu S, Li Z, Hu S. Understanding concerns, sentiments, and disparities among population groups during the COVID-19 pandemic via Twitter data mining: large-scale cross-sectional study. J Med Internet Res 2021 Mar 05;23(3):e26482 [FREE Full text] [doi: 10.2196/26482] [Medline: 33617460]

Abbreviations

API: application programming interface
LDA: latent Dirichlet allocation
WHO: World Health Organization

Edited by M Meacham; submitted 31.01.22; peer-reviewed by A Rovetta, J Li; comments to author 06.06.22; revised version received 13.08.22; accepted 15.09.22; published 27.09.22.

Please cite as:
Hu M, Conway M
Perspectives of the COVID-19 Pandemic on Reddit: Comparative Natural Language Processing Study of the United States, the United Kingdom, Canada, and Australia
JMIR Infodemiology 2022;2(2):e36941
URL: https://infodemiology.jmir.org/2022/2/e36941
doi:10.2196/36941
PMID:36196144

©Mengke Hu, Mike Conway. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 27.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a

XSL•FO
RenderX

link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Monitoring Mentions of COVID-19 Vaccine Side Effects on Japanese and Indonesian Twitter: Infodemiological Study

Kiki Ferawati[1], MStat; Kongmeng Liew[1], PhD; Eiji Aramaki[1], PhD; Shoko Wakamiya[1], PhD

Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

Corresponding Author:
Shoko Wakamiya, PhD
Graduate School of Science and Technology
Nara Institute of Science and Technology
8916-5, Takayama-cho
Ikoma, 630-0192
Japan
Phone: 81 743725250
Email: wakamiya@is.naist.jp

Abstract

Background: The year 2021 was marked by vaccinations against COVID-19, which spurred wider discussion among the general population, with some in favor and some against vaccination. Twitter, a popular social media platform, was instrumental in providing information about the COVID-19 vaccine and has been effective in observing public reactions. We focused on tweets from Japan and Indonesia, 2 countries with a large Twitter-using population, where concerns about side effects were consistently stated as a strong reason for vaccine hesitancy.

Objective: This study aimed to investigate how Twitter was used to report vaccine-related side effects and to compare the mentions of these side effects from 2 messenger RNA (mRNA) vaccine types developed by Pfizer and Moderna, in Japan and Indonesia.

Methods: We obtained tweet data from Twitter using Japanese and Indonesian keywords related to COVID-19 vaccines and their side effects from January 1, 2021, to December 31, 2021. We then removed users with a high frequency of tweets and merged the tweets from multiple users as a single sentence to focus on user-level analysis, resulting in a total of 214,165 users (Japan) and 12,289 users (Indonesia). Then, we filtered the data to select tweets mentioning Pfizer or Moderna only and removed tweets mentioning both. We compared the side effect counts to the public reports released by Pfizer and Moderna. Afterward, logistic regression models were used to compare the side effects for the Pfizer and Moderna vaccines for each country.

Results: We observed some differences in the ratio of side effects between the public reports and tweets. Specifically, fever was mentioned much more frequently in tweets than would be expected based on the public reports. We also observed differences in side effects reported between Pfizer and Moderna vaccines from Japan and Indonesia, with more side effects reported for the Pfizer vaccine in Japanese tweets and more side effects with the Moderna vaccine reported in Indonesian tweets.

Conclusions: We note the possible consequences of vaccine side effect surveillance on Twitter and information dissemination, in that fever appears to be over-represented. This could be due to fever possibly having a higher severity or measurability, and further implications are discussed.

(JMIR Infodemiology 2022;2(2):e39504) doi:10.2196/39504

KEYWORDS

COVID-19; vaccine; COVID-19 vaccine; Pfizer; Moderna; vaccine side effects; side effects; Twitter; logistic regression

Introduction

Background

Vaccinations have been proposed as one of the solutions to contain and end the COVID-19 pandemic [1-3]. Prior to their widespread deployment, an early Twitter poll suggested that public sentiment toward vaccinations was mostly positive, with many individuals indicating that they would seek vaccination, despite ongoing concerns about the safety of the vaccines [4]. However, these concerns persisted in the public eye, including issues like safety, the unusually quick development of vaccines, and possible side effects after administration [5]. These were also observed in vaccine-related search trends in early 2021,

XSL•FO
RenderX

when resulting side effects were identified as a significant area of concern [6].

The vaccine rollout evoked a diverse set of reactions from the general population, be it for or against vaccination. Bonnevie et al [7] investigated vaccine acceptance before the pandemic and in the middle of pandemic in 2020 and found stronger vaccine opposition on Twitter during the latter period. The trend appeared reversed in the study conducted by Lyu et al [8], who studied public perception and reactions toward COVID-19 vaccinations on Twitter through topic modeling and sentiment analysis and found that early discussions about vaccines stemmed from the development stage of the vaccines, and public sentiment leaned toward a positive outlook later on.

As one of the most popular social media platforms in use today, Twitter has been widely utilized as a source for research in COVID-19 infodemiology (see [9]), building on an extant body of literature on epidemic surveillance via that platform. For example, in the case of influenza, Twitter has been used as a detection tool to estimate individual diagnoses [10] and as social surveillance, functioning as an early warning tool for outbreak detection [11]. Despite the usefulness of Twitter for epidemiological surveillance studies, there are limitations, such as the spread of misinformation during the early periods of the COVID-19 pandemic [12] that may bias these data. In this paper, we proposed that Twitter can also be used in a similar fashion to monitor side effects from COVID-19 vaccination, by focusing on the Pfizer and Moderna vaccines in Japan and Indonesia.

Attitudes Toward COVID-19 Vaccination

Most studies on COVID-19 vaccination and Twitter have focused on general collective attitudes toward vaccination. Marcec and Likic [13] applied lexicon-based sentiment analysis to English tweets mentioning AstraZeneca, Pfizer, and Moderna and found that the sentiment for Pfizer and Moderna was generally more positive than that for AstraZeneca. Sattar and Arifuzzaman [14] analyzed tweets related to public sentiment about COVID-19 vaccination awareness and found strong positive sentiments despite the side effects of the vaccine. Kwok et al [15] used latent Dirichlet allocation topic modeling to identify topics in tweets related to COVID-19 vaccination in Australia, applied sentiment analysis to the tweets, and found that counts of tweets with positive sentiment were only slightly larger than counts of tweets with negative sentiment, thereby raising concerns over widespread vaccine acceptance. Yet, most of these studies were with Western and English samples, and there has been a considerable lack of similar research for tweets in non-English languages.

In this paper, we focused on tweets concerning widely available messenger RNA (mRNA) vaccines in Japan and Indonesia, 2 island countries located in the Pacific that have a large Twitter-using population (top 10 in terms of Twitter users globally [16]). As the Japanese and Indonesian languages are largely ubiquitously spoken in their respective countries [17,18], it provides for a relatively controlled environment to observe patterns unique to each country. This allows for a contextualization of Twitter usage to the wider society for added interpretations and behavioral analyses, especially in this pandemic.

The year 2021 saw the adoption of COVID-19 vaccinations on a global scale. Vaccination for health care workers in Japan started in February 2021, and vaccinations for the elderly started in April 2021 [19]. The vaccination rate started picking up quickly in June 2021 and continued to rise until over 80% were fully vaccinated by the end of 2021 [20]. Meanwhile, vaccination in Indonesia started in January 2021, with health care workers as a priority, followed by the elderly and public officers, and finally for the general public [21]. Although the early vaccination campaign used the Coronavac and AstraZeneca vaccines, in August 2021, the Moderna and Pfizer vaccines started to be administered for vaccinations and boosters in the country. By the end of 2021, 46.7% of the population of Indonesia were fully vaccinated.

One key context behind vaccine hesitancy identified in both cultures is the role of side effects. In Japan, concerns about adverse side effects were arguably the main reason for vaccine hesitancy, alongside other factors like gender, living arrangements, economic status, and psychological issues [22]. Vaccine hesitancy was also found to be significantly more frequent in the younger generation than in the older generation [23]. Meanwhile, in Indonesia, concerns about vaccine safety, distrust toward the vaccine, and concerns about side effects were identified as a few common reasons for vaccine hesitancy [24]. As Twitter has been used to identify symptoms of COVID-19 [25], we proposed for this paper that it can also be utilized to examine side effects of COVID-19 vaccination. Moreover, by comparing side effect counts from Twitter with rates reported in phase 3 clinical trials of the Pfizer and Moderna vaccines, we can observe patterns in information dissemination on side effects that may be unique to Twitter (eg, are side effects overrepresented, appropriately, or underrepresented when mentioned on Twitter). These results may then determine the usefulness of Twitter for vaccine side effect monitoring or alternatively illustrate misinformation biases that are present on the platform.

Finally, we examined if there were differences in side effect reporting between the Pfizer and Moderna vaccines found in tweets by country. This is because publicly available research concerning vaccine (and maker-specific) side effects in both countries is still limited, and to our knowledge, only one study by Kitagawa et al [26] compared the side effects of the Moderna and Pfizer vaccines available in Japan through self-reported data. Analyses were conducted separately for Japan and Indonesia.

Methods

Data

Tweet Collection

To get a general sense of public opinions for the vaccination campaigns in Japan and Indonesia, tweets in Japanese (ja) and Indonesian (id; based on Twitter's language filter) were collected for the whole of 2021, from January 1, 2021, to December 31, 2021 (UTC). The search query comprised keywords for vaccines and side effects and excluded retweets (see Table 1). Of all the vaccines used in both countries, we

XSL•FO
RenderX

limited the query to Moderna and Pfizer because these 2 vaccines were used in both countries. Although AstraZeneca's vaccine was also used in both countries, it was much less common in Japan and was represented by a variety of names in the public sphere, so relevant tweets were even more difficult to obtain.

All data were obtained using the Python Twarc library (version 2.8.3) for Academic Research Access in Twitter [27]. As limitations on tweet quota restricted access to vast amounts of Japanese tweets, we were unable to obtain tweets that mentioned "vaccine" only or "side effects" only. Consequently, the vaccine keywords used for Japanese and Indonesian tweet scraping included the vaccine type ("Pfizer" and "Moderna"), and the side effect keywords followed the symptoms described by the Centers for Disease Control and Prevention in English: tiredness, headache, muscle pain, chills, fever, and nausea [28]. Similarly, a list of symptoms was also available on the Indonesian government's official webpage for effect after COVID-19 vaccination, which is abbreviated as *kipi* (in Indonesian: *kejadian ikutan pasca imunisasi*) [29]. Corresponding sources from the Japanese government also mentioned the same symptoms, excluding nausea but including diarrhea [30]. We decided to exclude diarrhea since it was not listed on the Indonesian source and the detailed prevalence is not available in Moderna public reports. However, we decided to retain nausea in this study because of the availability of the corresponding statistics, and it was also referenced by some research about vaccine side effects in Japan [26,31]. This list of symptoms was then translated into Japanese and Indonesian, with additional keywords added from synonyms (see Table 1).

Table 1. List of keywords related to COVID-19 vaccines.

Terms[a]	Keywords (delimited by commas)	
	Japanese	Indonesian
Vaccine		
Vaccine-related	ファイザー, モデルナ	pfizer, moderna
Side effects		
Side effects	副反応	efek, kipi
Tiredness	疲労, 疲れ, 倦怠感, だるい, だるさ	lelah, capai, capek, pegal, lemas, letih
Headache	頭痛, 頭が痛い	pusing, sakit kepala
Muscle pain	筋肉痛	nyeri otot
Chills	寒気, 悪寒, さむけ	meriang, menggigil
Fever	熱, 高熱, 微熱, 発熱, 熱が高い, 熱があった, 熱がある, 熱が出た, 熱風邪	demam, panas
Nausea	悪心, 吐, 嘔吐, おう吐, 気分悪	mual

[a]Translated into English.

Public Report Data

The comparison percentages listed in this paper were obtained from publicly available reports (press releases) published by Pfizer [32] and Moderna [33]. For Pfizer, these included data from participants aged 16 years to 55 years, and for Moderna, these included data from participants aged 18 years to 64 years. Both were collected for 7 days after the vaccination and classified as systemic adverse reactions.

Preprocessing of Tweet Data

For tweets from both languages, the initial preprocessing steps were removing usernames and web links. Afterward, for Japanese tweets, we removed emojis and special characters (such as Japanese punctuation). Tweets were then tokenized using mecab-ipadic-NEologd [34-36], which reduced terms into their simplest forms to facilitate further analyses. All keyword filtering was done using full-width characters. For Indonesian tweets, all characters were set into lowercase, and non-ASCII characters were removed.

To assess the 2 vaccines separately, we filtered tweets to select tweets with the term Pfizer or Moderna only. Tweets mentioning both vaccines were removed and excluded from the analyses. Next, we defined user accounts with more than 10 tweets in our data as "high frequency users." We removed these high frequency users to avoid having data biased by excessive tweet counts from the same individual. We then grouped tweets by user account, focusing on user-level analyses. If a user had more than one tweet, they were merged into a single sentence for the analyses. This was to reduce bias arising from the same individual tweeting their side effects multiple times over different tweets.

Following that, tweets with the term "Pfizer" were coded as 1, while tweets with the term "Moderna" were coded as 0. As mentioned earlier and considering the respective timelines for vaccinations, we excluded tweets mentioning both types (Pfizer and Moderna) and filtered them out at this stage. The resulting variable from this step served as the outcome variable for the logistic regression analysis.

The sample of filtered tweets is shown in Textbox 1. The tweet samples were paraphrased due to Twitter's privacy policy. The Tweet ID for the data set processed in this study is available in Multimedia Appendix 1 (Japanese tweets) and Multimedia Appendix 2 (Indonesian tweets). In each filtered tweet, we

XSL•FO
RenderX

applied word matching of the specified keywords to the merged tweets. If the word for side effect was present, then the column was marked as “1,” and if it was not, then it was marked as “0.” There were 7 predictor variables in total: effect, tiredness, headache, muscle pain, chills, fever, and nausea. The presence of each respective side effect was checked using exact word matching. For example, based on the Japanese tweet in Textbox 1, the columns for fever and headache would be marked as 1, while the rest of the side effects would be 0. Mentions of “pain” were not classified as “muscle pain,” so it was marked as 0.

Textbox 1. Examples of filtered tweets.

Japanese tweet: こんばんは皆様。2回目のファイザーワクチンの接種後、夜から次の日にかけて、微熱と頭痛と接種部位の痛みを感じます。若い人の方が熱が出やすいと思われます。＃ファイザー＃コロナワクチン (Good evening everyone. After the second dose of Pfizer vaccine, I feel a slight fever, headache and pain at the vaccination site during the night and the next day. It seems that younger people are more prone to fever. #pfizer #coronavaccine)

Indonesian tweet: *efek dosis kedua vaksin moderna membuat aku menangis karena sakit demam, menggigil, badan terasa nyeri dan sakit kepala* (side effect of second dose of Moderna made me cry in pain with fever, chills, sore and headache)

Logistic Regression Analysis

Logistic regression analysis is a statistical method to analyze associations with a binary outcome variable [37]. In this study, we conducted separate logistic regression models by country: Japan and Indonesia. The outcome variable was the vaccine type (Pfizer or Moderna), and the predictor variables were the identified side effects: side effect, tiredness, headache, muscle pain, chills, fever, and nausea (from the tweets). We then examined the likelihood of a specific side effect for each vaccine type by using the odds ratio obtained from the model parameter. A significance level of 5% was used to construct the confidence interval for the odds ratios. The model was evaluated using the Nagelkerke R^2 [38]. All statistical analyses were conducted using SPSS version 28.0.1 (IBM Corp, Armonk, NY). All variables included in the analyses were binary.

Ethical Considerations

This study did not require participants to be involved in any physical or mental intervention. As this research did not use personally identifiable information, it was exempt from institutional review board approval in accordance with the Ethical Guidelines for Medical and Health Research Involving Human Subjects stipulated by the Japanese national government.

Results

Comparisons With Public Report Data (Press Releases) From Clinical Trials

The final data set used in this research included 286,887 Japanese tweets from 214,165 users and 14,484 Indonesian tweets from 12,289 users. Table 2 shows the final tweet count after merging and the detailed breakdown for each side effect. For the final data set, the mean number of tweets per user for Japanese was less than the mean number of tweets per user for Indonesian, as shown in Table 2. However, since we aggregated tweets by user, we focused on user counts from merged tweets for subsequent analyses. The proportions of tweets about the Pfizer vaccine and Moderna vaccine in the Indonesian data set were 58.80% and 41.20%, respectively, with more tweets mentioning Moderna. For the Japanese data, 98.47% of the overall set of tweets mentioned Pfizer. The proportions of Pfizer and Moderna shots administered in Japan at the end of 2021 were 79.85% and 20.08%, respectively, and 0.08% for others (AstraZeneca) [39]. We were unable to access comparable statistics for Indonesia.

Table 2. Counts from the tweets.

	Japanese			Indonesian		
	Pfizer	Moderna	Total	Pfizer	Moderna	Total
Tweets, n	283,530	3357	286,887	5684	8800	14,484
Number of tweets per user, mean (SD)	1.344 (0.890)	1.028 (0.221)	1.339 (0.885)	1.122 (0.449)	1.218 (0.606)	1.178 (0.549)
Individual users (after merging tweets), n	210,899	3266	214,165	5063	7226	12,289
Individual users who mentioned side effects from any vaccine, n						
Side effects	101,794	1224	103,018	1535	4628	6163
Tiredness	39,724	406	40,130	375	356	731
Headache	34,878	398	35,276	457	858	1315
Muscle pain	25,167	339	25,506	9	37	46
Chills	9361	143	9504	248	780	1028
Fever	131,172	1993	133,165	2089	2386	4475
Nausea	5334	49	5383	105	396	501

XSL•FO
RenderX

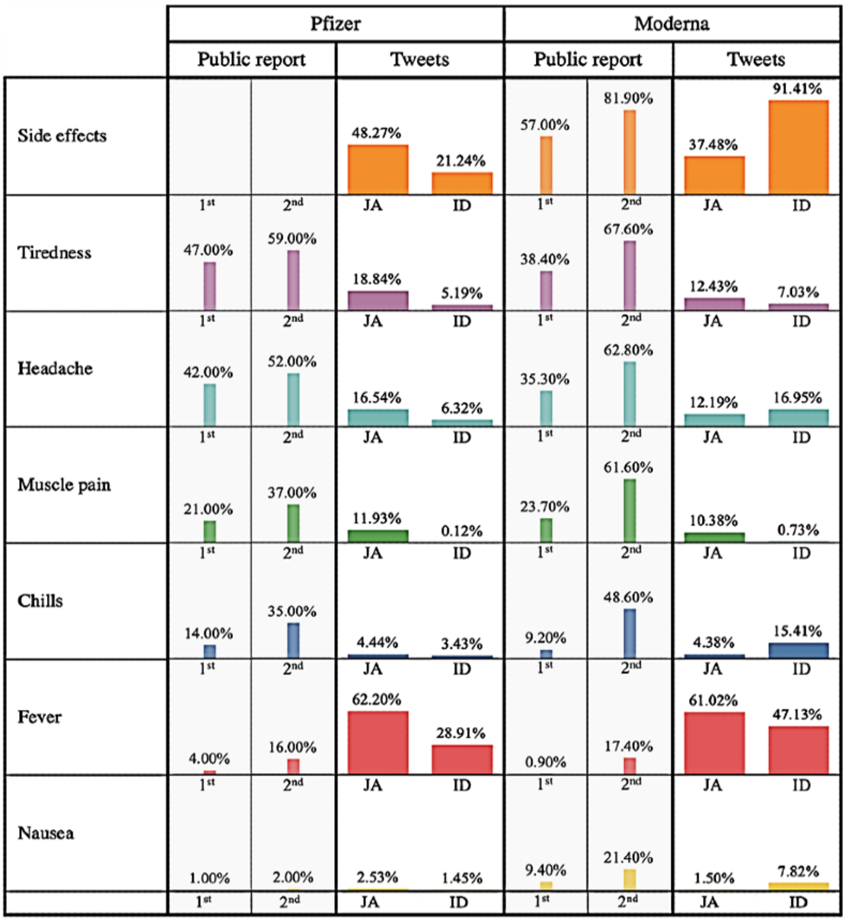
Figure 1 displays the side effects reported in press releases by Pfizer and Moderna, followed by the percentages of tweets observed in our study. The figure contains the side effects obtained from the word matching in tweets and the comparison with public report data to illustrate the ability of tweet data to capture the side effects. There appeared to be a difference between the percentage of side effects reported by press releases and that reported in tweet data and a slight difference between the percentages in Japanese and Indonesian tweets.

In the public reports, side effects were more frequently observed with the second dose than with the first dose. However, we lacked comparable information from the tweets to similarly differentiate side effects between the first dose and the second dose in our data set. Except for fever, we noticed that all side effects were reported more frequently in the public reports (higher percentages) than in the tweets. A radar graph showing the comparison of each side effect for the first dose and second dose in the public reports versus those obtained from the tweets is available in Multimedia Appendix 3, with the value plotted corresponding to the side effect counts in Table 2 and percentages from the values shown in Figure 1.

Looking into the tweet comparisons for Japanese and Indonesian tweets, for both Pfizer and Moderna, several side effects such as tiredness, muscle pain, and fever were reported at higher percentages in Japanese tweets than in Indonesian tweets. The percentages of headache, chills, and nausea were also different, with higher percentages in Indonesian tweets. Regardless of vaccine type, fever was by far the most reported side effect in tweets.

The percentages of all the side effects in Japanese tweets for the Pfizer vaccine were slightly higher than those for the Moderna vaccine, even when the total number of tweets for the 2 were notably different. On the other hand, in Indonesian tweets, the percentages of side effects with the Moderna vaccine were higher. We also noticed that the percentage reported for muscle pain was really small, which was probably caused by an inappropriate word choice used to represent this type of pain. For example, not many users may have been able to locate the exact part of the body from which the pain originated. A more general term "pain" may have been more suitable to represent this side effect than the specific term "muscle pain."

Figure 1. Percentage of side effects experienced after vaccination, as obtained from public reports by Pfizer and Moderna, compared with the percentages found in our Twitter sample. The overall percentage of general side effects was not mentioned in Pfizer's report, and the percentage of side effects in tweets was calculated from the number of tweets for specific side effects divided by the total number of filtered tweets. 1st: first vaccine dose; 2nd: second vaccine dose; ID: Indonesian tweets; JA: Japanese tweets.



Logistic Regression Analysis of Vaccine Side Effects

We then compared Twitter mentions of side effects for the Pfizer and Moderna vaccines, with side effects as predictors and vaccine type (Pfizer or Moderna) as the outcome variable. Most of the predictor variables were significant (Table 3), suggesting that reported side effects differed significantly between Pfizer and Moderna vaccines. However, we note the high statistical

XSL•FO
RenderX

power resulting from a large sample size may have affected the calculation of statistical tests and the respective *P* value. We first report results for the Japanese data, followed by the Indonesian data separately.

The Nagelkerke R^2 for Japanese tweets was 1.2%. In interpreting the model, we found that the odds of the umbrella term “side effect” appearing in a tweet about the Pfizer vaccine was about 1.907 times more. For specific terms (ie, muscle pain, fever, headache, and nausea), the odds ratios were close to each other: They were 1.338, 1.357, 1.362, and 1.458 times more likely, respectively, to be mentioned in Pfizer tweets, suggesting that those terms were more frequent in tweets about the Pfizer vaccine than in tweets about the Moderna vaccine. However, only chills had a small odds ratio. This is different from past results reported by Kitagawa et al [26], who compared the prevalence of the side effects of both vaccine types through a questionnaire study conducted in Japan and found that users receiving the Moderna vaccine reported more side effects than those receiving the Pfizer vaccine.

Indonesian tweets showed a different result, as other than tiredness and fever, all other side effects appeared less likely to be mentioned in tweets about the Pfizer vaccine and more likely to be mentioned in tweets about the Moderna vaccine. The Nagelkerke R^2 for the model for the Indonesian tweets was 17.4%. However, we noticed that tiredness appeared significantly more often in tweets about the Pfizer vaccine than in tweets about the Moderna vaccine. A closer look at the 95% CI for fever, which contained a value of 1, suggested that there may be little difference in those 2 side effects between the Pfizer and Moderna vaccines.

Table 3. Logistic regression analysis results for Japanese and Indonesian tweets.

Variable	Coefficient	Standard error	*P* value	Odds ratio	95% CI
Japanese Tweets					
(Intercept)	3.551	0.043	<.001	34.831	—[a]
Side effects	0.646	0.043	<.001	1.907	1.755-2.073
Tiredness	0.489	0.054	<.001	1.631	1.468-1.814
Headache	0.309	0.055	<.001	1.362	1.223-1.517
Muscle pain	0.291	0.059	<.001	1.338	1.191-1.503
Chills	–0.101	0.087	.25	0.904	0.762-1.072
Fever	0.305	0.042	<.001	1.357	1.250-1.473
Nausea	0.377	0.146	.01	1.458	1.095-1.941
Indonesian Tweets					
(Intercept)	0.455	0.036	<.001	1.576	—
Side effects	–1.461	0.042	<.001	0.232	0.214-0.252
Tiredness	0.238	0.084	.004	1.269	1.077-1.495
Headache	–0.294	0.068	<.001	0.745	0.653-0.851
Muscle pain	–0.98	0.398	.01	0.375	0.172-0.818
Chills	–1.029	0.080	<.001	0.358	0.306-0.418
Fever	0.003	0.043	.94	1.003	0.922-1.092
Nausea	–0.821	0.120	<.001	0.440	0.348-0.557

[a]Not applicable.

Discussion

Principal Findings

Our results highlight a large gap between expressions of side effects on Twitter and percentages reported in public press releases: For most of the side effects, we found that the percentages of Twitter users who reported them were far lower than the percentages reported in the public reports. Our first result is focused on the descriptive comparison of counts reported on Twitter and counts of observation in public press releases. Considering that several studies make use of symptom reporting on Twitter for epidemiological surveillance, our study shows that, at least for vaccine side effect surveillance, we may be at risk of overrepresenting “fever” relative to other (milder) side effects. Although our study did not explicitly examine the reasons behind this phenomenon, we speculate on a few possible explanations.

First, this could reflect a difference in how lay people and health professionals perceive and talk about vaccinations [40]. Our study may be relying too much on relatively specialized terminology (eg, muscle pain) that may not be the most salient term accessible to the broader lay population, at least in Japan and Indonesia. Hence, the difference in focus and word choice between lay people and professionals may result in different expressions used to describe side effects experienced after vaccination. Given that our study used more “professional”

XSL•FO
RenderX

terms, our results for side effects like muscle pain or chills could thus have been an underrepresentation of available tweets.

Second, the higher reporting rates for fever could be due to its ease of measurement by lay members of the public. Thermometers are widely available and widely used, and there are general conventions (thresholds) for determining if a person has a fever. On the other hand, other side effects such as chills, headaches, and tiredness sometimes might not have clear, objective thresholds and measurement methods that are of common knowledge to the lay person. With the ambiguity and subjectivity around these side effects, Twitter users may hesitate to update their statuses, especially compared with fever, which comes with a clear and objective threshold. Accordingly, Twitter users who may experience more than one side effect may then decide to only report the clearer, more observable one.

Finally, another possible reason could be the age difference between people observed in the studies (public reports) and Twitter users who share their experiences in their tweets. In a survey conducted by Statista, close to 80% of Japanese respondents aged 20 years to 29 years reported using the microblogging and social networking service Twitter. Although this suggests that the penetration rate among Japanese youths was also on a high level, it was much less widely used by older age groups [41]. A past study also suggested that systemic incidence of side effects from the Pfizer vaccine was significantly higher in young participants than in older adults [31], a finding that was also previously extended to the Japanese context [42]. The rate of Twitter users in the country also shows that there are people who did not use Twitter, which means they will not share their side effects through tweets. As the incidence of the vaccine side effects is higher in older age groups while the Twitter penetration rate is lower, this might also influence the number of side effects that can possibly be found in Twitter.

Our findings from Japanese and Indonesian tweets were also different from reported (vaccine) side effects in English tweets in the United States, where soreness, fatigue, and headache were listed as the top 3 side effects for the Pfizer and Moderna vaccines [43]. One probable reason could be cultural differences in how people express themselves on Twitter, which might stem from their respective cultural background and habits. People in collectivist cultures (like Japan and Indonesia) may be less open and active on social media, as compared with individuals in individualistic cultures (like the United States) [44]. Accordingly, users in the United States could be reporting their symptoms with more detail and frequency on social media, whereas Japanese and Indonesian users may be “saving” their posts for worse side effects (ie, fever). In any case, we suggest that future studies on infodemiological surveillance of vaccine side effects may consider focusing primarily on fever-related keywords in these countries.

Regardless of interpretation, our study appears to suggest that “fever,” as a subjectively stronger side effect of vaccination, is discussed disproportionately more on Twitter in Japan and Indonesia. One possible consequence could be in the echo chamber effect on Twitter [45], which could contribute to vaccine hesitancy or other aversive behaviors. As an illustration, due to this disproportionate reporting, consider Marie, a Twitter user, who is currently considering vaccination. She may observe that many users on her Twitter feed discuss their experiences with fever as a side effect of vaccination, which could lead to a perceived overrepresentation of fever risks that may dissuade her from receiving the vaccine. In contrast, if tweets had discussed side effects in a more representative manner, Marie would have had an accurate representation of the risks and may not have been discouraged from vaccination for this reason. We note again that side effects were a strong reason for vaccine hesitancy in Japan [22] and postulate that this overrepresentation of strong side effects on Twitter may have had a role to play in contributing toward hesitancy, although follow-up research is needed to test this hypothesis.

Limitations

Although we limited our investigation to only Moderna and Pfizer vaccines, these received relatively late approval in Indonesia, and we did not examine tweets on side effects from other vaccine makers (eg, Sinovac, AstraZeneca). Consequently, some of the discourse surrounding vaccines and their side effects was not captured in the earlier tweets. Second, some tweets also elaborated on side effects without mentioning the specific vaccine type received. Third, the search query for tweets was limited to the specified keywords (Table 1) and did not include other possible words not listed in the table, and we did not consider the positive or negative sentiment expressed by the tweets. Last, we focused only on tweets that mentioned one type of vaccine only and removed tweets mentioning both. Nevertheless, there was little chance of people getting both Pfizer and Moderna vaccinations administered in the observed period, and tweets mentioning both mainly referred to news articles and related discussions, not the actual side effects experienced by the public.

Finally, we did not control for negation in tweets. However, we sampled 100 tweets for each vaccine and language. Of all sampled tweets, we focused our observation on fever, which was the most frequent side effect found in our Twitter data set and found that only a minority of tweets contained negation. Based on a manual inspection of the sample, we found that, for Japanese tweets, negation (for fever) was observed in 15 of 63 tweets mentioning fever, or 23.80% of relevant tweets about the Pfizer vaccine. By doing the same process, we obtained 32.76% negation in Moderna tweets. Meanwhile, in Indonesian tweets, the negation for fever was 21.43% for the Pfizer vaccine and 16.67% for the Moderna vaccine. Negation is a difficult challenge in Twitter analyses, as there are many ways to express negation and may not necessarily be easily filtered out through designated negation words [46]. Although negation handling may improve the final results, it did not appear to hinder the utility of tweets in displaying consistent patterns as real-world (unfiltered) data in past research [47]. Considering that the observed percentage of negation for both vaccines in each language was similar in our random sample of tweets, we decided to retain all tweets for these analyses.

We also lacked the means to verify whether the tweets were from a personal or nonpersonal (eg, corporate) account and whether the said individual behind the account actually received vaccination and follow-up confirmations of reported side effects.

XSL•FO
RenderX

Finally, we also lacked sufficient information about whether the side effects were from the first, second, or third dose of the vaccination, as we were limited to the side effects shared in the tweets by the users that matched the language filter of our Twitter API query without any deeper demographic or contextual information.

Conclusions

We found that fever was the most prevalent side effect reported in Japanese and Indonesian tweets, and this may be a reflection of bias on social media toward reporting severe or measurable side effects (like fever). Furthermore, in examining side effects from different vaccine makers, we found that Twitter yielded inconsistent information from Japan and Indonesia, in that side effects were reported relatively more in tweets about the Pfizer vaccine in Japan but more in tweets about the Moderna vaccine in Indonesia. As such, given the inconsistencies and gaps in findings from Twitter and the vaccine press releases, we present cautious optimism that Twitter can prove useful for infodemiological surveillance for vaccine side effects that is best suited for detecting prevalences of fever symptoms in Japanese and Indonesian populations.

Acknowledgments

This work was supported by Japan Science and Technology Agency (JST) Strategic International Research Cooperative Program (SICORP) Grant Number JPMJSC2107, Japan and Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP22K12041.

Authors' Contributions

KF collected the data and performed the analysis, with the assistance of KL. KF and KL wrote the manuscript, and SW and EA provided critical comments. All authors contributed to the conceptualization of the study and to the initial study design. EA and SW supervised the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tweet ID for Japanese tweets.

[TXT File , 5603 KB - infodemiology_v2i2e39504_app1.txt]

Multimedia Appendix 2

Tweet ID for Indonesian tweets.

[TXT File , 283 KB - infodemiology_v2i2e39504_app2.txt]

Multimedia Appendix 3

Radar graph for Pfizer (A) and Moderna (B), showing comparisons of the percentages of each side effect.

[PNG File , 213 KB - infodemiology_v2i2e39504_app3.png]

References

1. Dhama K, Sharun K, Tiwari R, Dadar M, Malik Y, Singh K, et al. COVID-19, an emerging coronavirus infection: advances and prospects in designing and developing vaccines, immunotherapeutics, and therapeutics. Hum Vaccin Immunother 2020 Jun 02;16(6):1232-1238 [FREE Full text] [doi: 10.1080/21645515.2020.1735227] [Medline: 32186952]
2. Fadda M, Albanese E, Suggs LS. When a COVID-19 vaccine is ready, will we all be ready for it? Int J Public Health 2020 Jul;65(6):711-712 [FREE Full text] [doi: 10.1007/s00038-020-01404-4] [Medline: 32529534]
3. Yamey G, Schäferhoff M, Hatchett R, Pate M, Zhao F, McDade KK. Ensuring global access to COVID-19 vaccines. Lancet 2020 May 02;395(10234):1405-1406 [FREE Full text] [doi: 10.1016/S0140-6736(20)30763-7] [Medline: 32243778]
4. Eibensteiner F, Ritschl V, Nawaz FA, Fazel SS, Tsagkaris C, Kulnik ST, et al. People's willingness to vaccinate against COVID-19 despite their safety concerns: Twitter poll analysis. J Med Internet Res 2021 Apr 29;23(4):e28973 [FREE Full text] [doi: 10.2196/28973] [Medline: 33872185]
5. Latkin CA, Dayton L, Yi G, Konstantopoulos A, Boodram B. Trust in a COVID-19 vaccine in the U.S.: A social-ecological perspective. Soc Sci Med 2021 Feb;270:113684 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113684] [Medline: 33485008]
6. An L, Russell DM, Mihalcea R, Bacon E, Huffman S, Resnicow K. Online search behavior related to COVID-19 vaccines: Infodemiology study. JMIR Infodemiology 2021;1(1):e32127 [FREE Full text] [doi: 10.2196/32127] [Medline: 34841200]
7. Bonnevie E, Gallegos-Jeffrey A, Goldbarg J, Byrd B, Smyser J. Quantifying the rise of vaccine opposition on Twitter during the COVID-19 pandemic. Journal of Communication in Healthcare 2020 Dec 15;14(1):12-19. [doi: 10.1080/17538068.2020.1858222]

XSL•FO
RenderX

8. 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]
9. Banda JM, Tekumalla R, Wang G, Yu J, Liu T, Ding Y, et al. A large-scale COVID-19 Twitter chatter dataset for open scientific research -- an international collaboration. ArXiv 2020 Apr 07:1. [Medline: 32550247]
10. Wakamiya S, Kawai Y, Aramaki E. Twitter-based influenza detection after flu peak via tweets with indirect information: Text mining study. JMIR Public Health Surveill 2018 Sep 25;4(3):e65 [FREE Full text] [doi: 10.2196/publichealth.8627] [Medline: 30274968]
11. Yousefinaghani S, Dara R, Poljak Z, Bernardo TM, Sharif S. The assessment of Twitter's potential for outbreak detection: Avian influenza case study. Sci Rep 2019 Dec 03;9(1):18147 [FREE Full text] [doi: 10.1038/s41598-019-54388-4] [Medline: 31796768]
12. 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]
13. Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J 2022 Jul;98(1161):544-550 [FREE Full text] [doi: 10.1136/postgradmedj-2021-140685] [Medline: 34373343]
14. Sattar NS, Arifuzzaman S. COVID-19 vaccination awareness and aftermath: Public sentiment analysis on Twitter data and vaccinated population prediction in the USA. Applied Sciences 2021 Jun 30;11(13):6128. [doi: 10.3390/app11136128]
15. Kwok SWH, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis. J Med Internet Res 2021 May 19;23(5):e26953 [FREE Full text] [doi: 10.2196/26953] [Medline: 33886492]
16. Dixon S. Countries with the most Twitter users 2022. Statista. 2022 Jul 27. URL: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ [accessed 2022-09-23]
17. Bernier A. An Introduction To The Japanese Language. Babbel Magazine. 2021 Jul 28. URL: https://www.babbel.com/en/magazine/guide-to-japanese-language [accessed 2022-09-23]
18. Devlin TM. How Many People Speak Indonesian, And Where Is It Spoken? Babbel Magazine. 2021 May 11. URL: https://www.babbel.com/en/magazine/how-many-people-speak-indonesian-where-is-it-spoken [accessed 2022-09-23]
19. Wang S. Japan begins vaccine rollout on Wednesday. CNN. 2021 Feb 17. URL: https://edition.cnn.com/world/live-news/coronavirus-pandemic-vaccine-updates-02-16-21/h_45ac6dc81faf4905a94fb4c8ab93092a [accessed 2022-09-23]
20. Hasell J, Mathieu E, Beltekian D, Macdonald B, Giattino C, Ortiz-Ospina E, et al. A cross-country database of COVID-19 testing. Sci Data 2020 Oct 08;7(1):345 [FREE Full text] [doi: 10.1038/s41597-020-00688-8] [Medline: 33033256]
21. Jiao C, Aditya A. Indonesia Begins COVID-19 Vaccine Rollout – and Jokowi Is First to Get Jab. Time. 2021 Jan 13. URL: https://time.com/5929160/indonesia-vaccine-rollout-jokowi/ [accessed 2022-09-23]
22. Okubo R, Yoshioka T, Ohfuji S, Matsuo T, Tabuchi T. COVID-19 vaccine hesitancy and its associated factors in Japan. Vaccines (Basel) 2021 Jun 17;9(6):662. [doi: 10.3390/vaccines9060662] [Medline: 34204465]
23. Khan MSR, Watanapongvanich S, Kadoya Y. COVID-19 vaccine hesitancy among the younger generation in Japan. Int J Environ Res Public Health 2021 Nov 07;18(21):11702 [FREE Full text] [doi: 10.3390/ijerph182111702] [Medline: 34770216]
24. Indonesian Ministry of Health, National Immunization Technical Advisory Group (NITAG), UNICEF, World Health Organization (WHO). COVID-19 Vaccine Acceptance Survey in Indonesia. 2020 Nov. URL: https://covid19.go.id/storage/app/media/Hasil%20Kajian/2020/November/vaccine-acceptance-survey-en-12-11-2020final.pdf [accessed 2022-09-23]
25. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: Sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978 [FREE Full text] [doi: 10.2196/21978] [Medline: 33108310]
26. Kitagawa H, Kaiki Y, Sugiyama A, Nagashima S, Kurisu A, Nomura T, et al. Adverse reactions to the BNT162b2 and mRNA-1273 mRNA COVID-19 vaccines in Japan. J Infect Chemother 2022 Apr;28(4):576-581 [FREE Full text] [doi: 10.1016/j.jiac.2021.12.034] [Medline: 35058126]
27. twarc. URL: https://twarc-project.readthedocs.io/en/latest/ [accessed 2022-09-23]
28. What to Expect after Getting a COVID-19 Vaccine. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/expect/after.html [accessed 2022-09-23]
29. Information about side effects or reaction after COVID-19 vaccination. kipi. URL: https://kipi.covid19.go.id/ [accessed 2022-09-23]
30. COVID-19 Vaccines. Ministry of Health, Labour and Welfare Japan. URL: https://www.mhlw.go.jp/stf/covid-19/vaccine.html [accessed 2022-09-23]
31. Izumo T, Kuse N, Awano N, Tone M, Sakamoto K, Takada K, et al. Side effects and antibody titer transition of the BNT162b2 messenger ribonucleic acid coronavirus disease 2019 vaccine in Japan. Respir Investig 2021 Sep;59(5):635-642 [FREE Full text] [doi: 10.1016/j.resinv.2021.06.003] [Medline: 34210623]
32. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, 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]

XSL•FO
RenderX

33. Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, COVE Study Group. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. N Engl J Med 2021 Feb 04;384(5):403-416 [FREE Full text] [doi: 10.1056/NEJMoa2035389] [Medline: 33378609]
34. Sato T, Hashimoto T, Okumura M. Implementation of word-spacing dictionary mecab-ipadic-NEologd and its effective use in information retrieval. 2017 Presented at: The 23rd Annual Conference of the Association for Natural Language Processing; March 13-17, 2017; Tsukuba, Ibaraki, Japan URL: https://www.anlp.jp/proceedings/annual_meeting/2017/pdf_dir/B6-1.pdf
35. Sato T. Neologism dictionary based on the language resources on the Web for Mecab. neologd / mecab-ipadic-neologd. 2015. URL: https://github.com/neologd/mecab-ipadic-neologd [accessed 2022-09-23]
36. Sato T, Hashimoto T, Manabu O. Operation of a word segmentation dictionary generation system called NEologd (in Japanese). 2016 Presented at: Information Processing Society of Japan, Special Interest Group on Natural Language Processing (IPSJ-SIGNL) Information Processing Society of Japan; 20-22 December 2016; Tokyo p. NL-229-NL-215.
37. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression. Hoboken, NJ: John Wiley & Sons; 2013.
38. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78(3):691-692. [doi: 10.1093/biomet/78.3.691]
39. COVID-19 Vaccines. Prime Minister's Office of Japan. URL: https://www.kantei.go.jp/jp/headline/kansensho/vaccine.html [accessed 2022-09-23]
40. Luo C, Ji K, Tang Y, Du Z. Exploring the expression differences between professionals and laypeople toward the COVID-19 vaccine: text mining approach. J Med Internet Res 2021 Aug 27;23(8):e30715 [FREE Full text] [doi: 10.2196/30715] [Medline: 34346885]
41. Twitter's penetration rate in Japan FY 2021, by age group. Statista. 2022 Sep 07. URL: https://www.statista.com/statistics/1077538/japan-twitter-penetration-rate-by-age-group/ [accessed 2022-09-23]
42. Saita M, Yan Y, Ito K, Sasano H, Seyama K, Naito T. Reactogenicity following two doses of the BNT162b2 mRNA COVID-19 vaccine: Real-world evidence from healthcare workers in Japan. J Infect Chemother 2022 Jan;28(1):116-119 [FREE Full text] [doi: 10.1016/j.jiac.2021.09.009] [Medline: 34580011]
43. 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 Jan 11;10(1):103 [FREE Full text] [doi: 10.3390/vaccines10010103] [Medline: 35062764]
44. Jackson LA, Wang J. Cultural differences in social networking site use: A comparative study of China and the United States. Computers in Human Behavior 2013 May;29(3):910-921. [doi: 10.1016/j.chb.2012.11.024]
45. Colleoni E, Rozza A, Arvidsson A. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. J Commun 2014 Mar 19;64(2):317-332. [doi: 10.1111/jcom.12084]
46. Kharde VA, Sonawane S. Sentiment analysis of Twitter data: A survey of techniques. IJCA 2016 Apr 15;139(11):5-15. [doi: 10.5120/ijca2016908625]
47. Aramaki E, Maskawa S, Morita M. Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter. 2011 Presented at: Conference on Empirical Methods in Natural Language Processing; July 27-31, 2011; Edinburgh, United Kingdom.

Abbreviations

JSPS: Japan Society for the Promotion of Science
JST: Japan Science and Technology Agency
mRNA: messenger RNA
SICORP: Strategic International Research Cooperative Program

Edited by T Mackey; submitted 12.05.22; peer-reviewed by S Kwok, J Banda, S Matsuda; comments to author 24.06.22; revised version received 29.07.22; accepted 19.09.22; published 04.10.22.

Please cite as:
Ferawati K, Liew K, Aramaki E, Wakamiya S
Monitoring Mentions of COVID-19 Vaccine Side Effects on Japanese and Indonesian Twitter: Infodemiological Study
JMIR Infodemiology 2022;2(2):e39504
URL: https://infodemiology.jmir.org/2022/2/e39504
doi:10.2196/39504
PMID:36277140

©Kiki Ferawati, Kongmeng Liew, Eiji Aramaki, Shoko Wakamiya. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 04.10.2022. This is an open-access article distributed under the terms of the Creative Commons

XSL•FO
RenderX

Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Investigating COVID-19 Vaccine Communication and Misinformation on TikTok: Cross-sectional Study

Katherine van Kampen[1*], BHSc; Jeremi Laski[2*], MSc; Gabrielle Herman[1], BSc; Teresa M Chan[3,4,5,6], MD, MHPE

[1]Michael G DeGroote School of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[2]College of Medicine, Central Michigan University, Mount Pleasant, MI, United States
[3]Division of Emergency Medicine, Department of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[4]Office of Continuing Professional Development, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[5]McMaster Education Research, Innovation, and Theory, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[6]Division of Education & Innovation, Department of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[*]these authors contributed equally

Corresponding Author:
Teresa M Chan, MD, MHPE
Division of Emergency Medicine, Department of Medicine
Faculty of Health Sciences
McMaster University
100 Main St W, Room 5003
Hamilton, ON, L8P 1H6
Canada
Phone: 1 905 525 9140
Email: teresa.chan@medportal.ca

Abstract

Background: The COVID-19 pandemic has highlighted the need for reliable information, especially around vaccines. Vaccine hesitancy is a growing concern and a great threat to broader public health. The prevalence of social media within our daily lives emphasizes the importance of accurately analyzing how health information is being disseminated to the public. TikTok is of particular interest, as it is an emerging social media platform that young adults may be increasingly using to access health information.

Objective: The objective of this study was to examine and describe the content within the top 100 TikToks trending with the hashtag #covidvaccine.

Methods: The top 250 most viewed TikToks with the hashtag #covidvaccine were batch downloaded on July 1, 2021, with their respective metadata. Each TikTok was subsequently viewed and encoded by 2 independent reviewers. Coding continued until 100 TikToks could be included based on language and content. Descriptive features were recorded including health care professional (HCP) status of creator, verification of HCP status, genre, and misinformation addressed. Primary inclusion criteria were any TikToks in English with discussion of a COVID-19 vaccine.

Results: Of 102 videos included, the median number of plays was 1,700,000, with median shares of 9224 and 62,200 followers. Upon analysis, 14.7% (15/102) of TikToks included HCPs, of which 80% (12/102) could be verified via social media or regulatory body search; 100% (15/15) of HCP-created TikToks supported vaccine use, and overall, 81.3% (83/102) of all TikToks (created by either a layperson or an HCP) supported vaccine use.

Conclusions: As the pandemic continues, vaccine hesitancy poses a threat to lifting restrictions, and discovering reasons for this hesitancy is important to public health measures. This study summarizes the discourse around vaccine use on TikTok. Importantly, it opens a frank discussion about the necessity to incorporate new social media platforms into medical education, so we might ensure our trainees are ready to engage with patients on novel platforms.

(JMIR Infodemiology 2022;2(2):e38316) doi:10.2196/38316

KEYWORDS

TikTok; COVID-19 vaccines; vaccinations; misinformation; COVID-19; Infodemiology; social media; health information; content analysis; vaccine hesitancy; public health; web-based health information

XSL•FO
RenderX

Introduction

Social media has become a prominent vehicle for educating both learners and the public. Learners and young physicians are increasingly savvy with these technologies [1,2], engaging as influencers and gaining outsized influence over young people [3].

Although the rapid development and emergency approval of multiple vaccines is something to be celebrated, vaccine hesitancy and misinformation remain significant obstacles to global vaccination. Vaccine hesitancy has been noted as one of the greatest threats to global health by the World Health Organization in 2019 [4-6]. In particular, this is evident by lower vaccination rates in some countries such as the United States [4]. In comparison, other G7 countries have higher percentages of their citizens receiving at least one dose [7]. This misinformation may stem from social media use. A total of 82% of Americans use social media, and many may use it for health information [8]. Social media, including Twitter, Facebook, Instagram, and TikTok, among others, has fueled rumors, hoaxes, misinformation, as well as disinformation [9].

Misinformation occurs when incorrect information is unintentionally propagated [10]. Even more worryingly, the use of targeted disinformation, where medical facts are intentionally falsified, can propagate distrust of public health measures, such as mask wearing or vaccination [9,11]. Social media platforms have increasingly faced more pressure from both citizens and regulators alike to combat this disinformation [12]. Nevertheless, these platforms continue to be ongoing sources of both misinformation and disinformation, revealing a need to understand the vaccine discourse on these platforms [13]. A recent study by Griffith et al [14] explored some of the etiology of vaccine hesitancy by analyzing over 500 Twitter tweets containing COVID-19 vaccine hesitancy content. Several overarching themes related to vaccine hesitancy were identified that included concerns of safety, lack of knowledge about the vaccine, mistrust of the medical community, confusing messages from authority figures, and mistrust of vaccine companies [14].

TikTok is the twin of "Douyin"—the Chinese short video app, originally known as "Musical.ly"—later rebranded as TikTok to a western audience [15]. Founded in 2018, TikTok is a growing social media platform in which users upload short videos under 120 seconds. Users interact with the platform typically by the "For You" page, which is meant to algorithmically present videos in which the user may be interested [11,16]. Gaining incredible popularity, about 1 in 6 people in the United States are current TikTok users [17]. However, despite its popularity, TikTok's algorithm has come under criticism for perpetuating misinformation. A report by NewsGuard [18] found that TikTok accounts that spread vaccination misinformation and antivaccination sentiments were being viewed by children as young as age 9, although the app technically only allows users over the age of 13 to use it. Prior to TikTok's revision of their algorithm, interacting with a single video containing false medical information could modify the "For You" page to be populated with similarly oriented vaccine hesitancy and COVID-19 misinformation content [19].

Given the vast implications of perpetuating medical misinformation during the onset of the COVID-19 pandemic, past research has sought to explore TikTok's role in vaccine misinformation. A study from the end of 2020 [3] found more TikToks overall that discouraged vaccination; however, those encouraging vaccination gained more traffic. TikToks pertaining to vaccination typically included humor or parody, with parodies of adverse reactions gaining higher view counts [3,11]. Furthermore, a small number of TikToks included health care professionals (HCPs), and a few TikToks conveyed medical education [11]. TikToks pertaining to vaccination seem to be created by a majority of non-HCP creators, and vaccine hesitancy prevails as a common theme on the platform. It is imperative that social media platforms be analyzed to reveal public attitudes toward vaccination and allow for more targeted public health campaigns [11,20].

To close the gap between public perceptions and the science behind vaccines, there is certainly an avenue for engaging learners and providing them with tools to engage with the public more robustly [21]. Instead of engaging in financial gain via social media stardom [1], increased efforts to formalize social media use and communication skills and incorporate them into medical school curricula may be of great benefit to our communities. However, to do so, it is imperative that we have a firm handle on what the current state of web-based communications are for physicians and other HCPs on platforms like TikTok.

Our cross-sectional study seeks to examine trends and attitudes toward COVID-19 vaccines by analyzing the most viewed TikToks with the hashtag #covidvaccine in July 2021 and specify which of these were generated by physicians and other HCPs.

Methods

Cross-sectional Study

We conducted a cross-sectional study of published TikToks with the hashtag #covidvaccine to characterize the discourse regarding vaccine use on the platform and to explore HCPs presence on the app regarding vaccine use. Furthermore, HCPs were identified and validated (ie, through regulatory bodies), which helped to show HPCs at what fields can better understand the sentiments of the general population, especially in regard to misinformation being spread.

Data Extraction

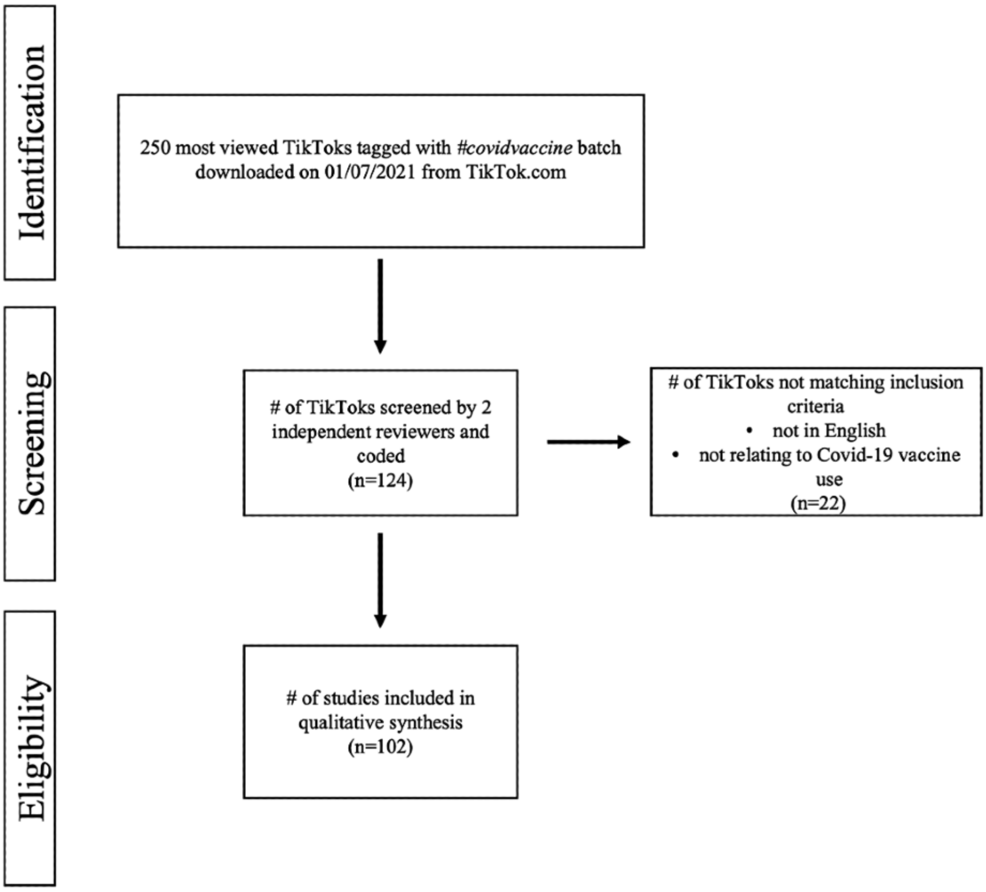
The most viewed TikToks with the hashtag #covidvaccine were batch downloaded using the open source TikTokApi Python wrapper [22] on July 1, 2021, with their respective metadata (ie, number of views, likes, shares, comments, author followers, and hashtags; Multimedia Appendix 1). TikToks were subsequently reviewed by 2 authors (JL and KvK) and encoded or categorized deductively (Multimedia Appendix 2); discrepancies between the 2 reviewers were resolved by the third author (GH).

XSL•FO
RenderX

Inclusion Criteria and Descriptive Coding of TikToks

Primary inclusion criteria were any TikToks in English with discussion of a COVID-19 vaccine (either positive, negative, or neutral). Inclusion criteria were purposefully left as broad as possible to encompass as many TikToks that would refer to the COVID-19 vaccine and could be potentially viewed by a general TikTok user in the future. Exclusion criteria were TikToks in languages other than English and those not relating to COVID-19 vaccine or vaccine use (Figure 1). Review continued until we reached approximately 100 appropriate TikToks for analysis, for a total of 124 videos reviewed and 102 eligible. Descriptive features that were additionally recorded if possible included the following: number of people in the video, country of origin, HCP status of creator, verification of HCP status, type of medium (eg, dance, commentary, storytelling, question and answer, responding to comments, silent video with visuals, satire, skit, stitch, or other), scientific validity of claims evaluated at the time of TikTok creation, and if COVID-19 vaccine misinformation is either referred to or combatted. The agreement of the coded data (ie, whether to include or to exclude it) between the 2 reviewers (JL and KvK) was calculated (κ=0.64, 95% CI 0.63-0.64). Due to the ambiguous nature of the content and messaging of numerous TikToks, agreement of the coded data often required the input of the third reviewer, with consensus being reached following discussion between all 3 reviewers on whether the content was related to the COVID-19 vaccine. Indeed, in our preliminary investigation, we found that several TikToks met our broad inclusion criteria, but they used the hashtag #covidvaccine likely as a method of generating traffic to their TikTok, without actually mentioning any pertinent content related to vaccination.

Figure 1. Schematic workflow with inclusion and exclusion criteria.



Data Visualization

The coding information was combined with the metadata to generate descriptive statistics and graphs. The data extraction and visualization workflow can be found in the study's GitHub repository [23].

Ethical Considerations

This study only analyzed publicly available data from existing data sets, and results do not contain any identifiable information that is not already in the public domain or are presented in aggregate.

Results

Overall Metrics

Of the 102 coded TikToks, 19 (18.6%) contained vaccine-hesitant messaging, whereas 83 (81.3%) were provaccine. Median plays between these two groups were 290,000 and 160,000, respectively. Of note, many of the provaccine TikToks were simply people recording themselves receiving the vaccine or recording their experience and symptoms post vaccination. Other broad themes noted in the provaccine category were people celebrating vaccines as a measure to ending lockdowns, encouraging others to vaccinate

XSL•FO
RenderX

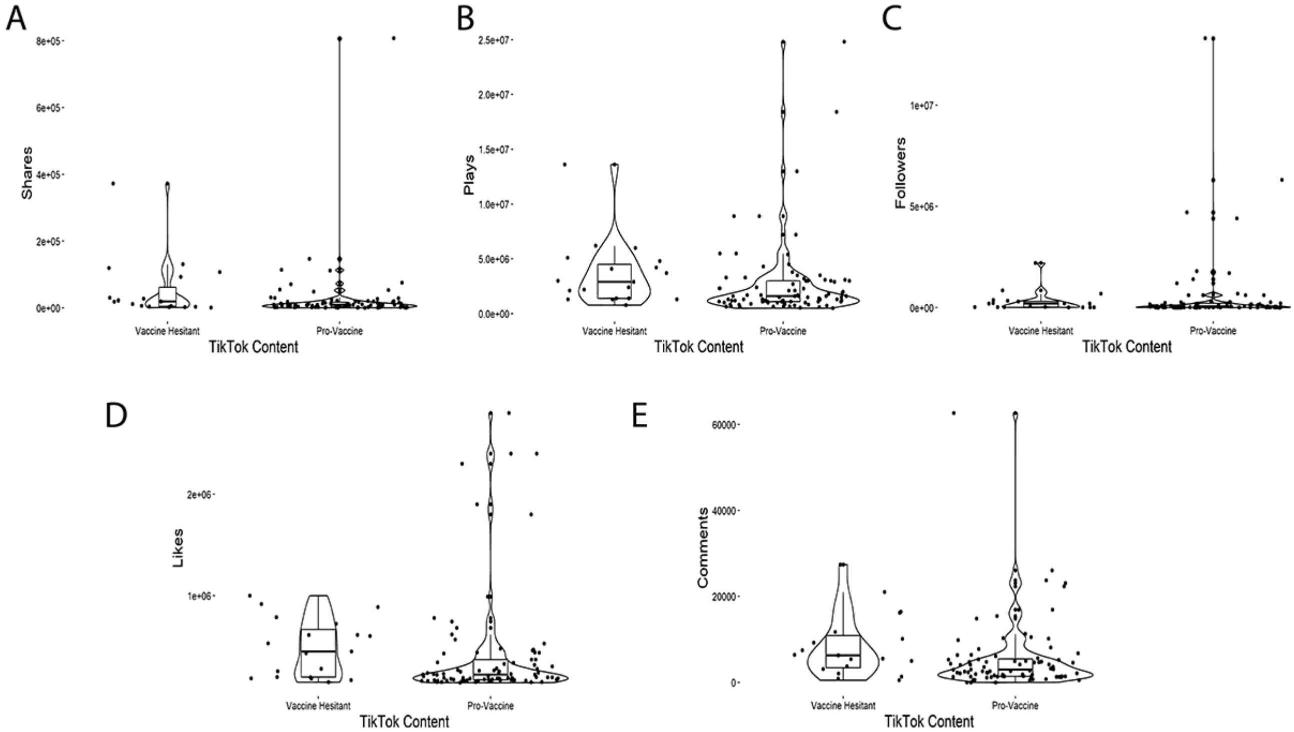
themselves. Vaccine-hesitant TikToks generated higher median comments, shares, and author followers than provaccine TikToks (Table 1; Figure 2). Interestingly, a relatively low number (n=15, 14.7%) of total TikToks were attributed to HCP creators. These HCP TikToks, however, all contained provaccine content. Furthermore, when comparing TikToks created by either layperson or HCP creators, HCP TikToks had higher median plays, comments, shares, and author followers (Table 2; Figure 3). One particular HCP TikTok creator, Dr Noc, is of particular interest, as he is the only creator to have more than one (n=4) TikToks that fall into the top 102 TikTok category for the month of July 2021. We additionally investigated TikTok retention 4 months after our original analysis, on November 29, 2021. During this period, we assessed the number of TikToks that still remained on the internet and were viewable to the general public. Out of a total of 102 original TikToks, 94 (92.1%) still remained active. All removed TikToks (n=8, 7.9%) were from separate provaccine content creators.

Table 1. General metrics of characteristics for both vaccine-hesitant and provaccine individual TikToks.

Characteristics	Vaccine hesitant (N=19)[a]	Provaccine (N=83)[a]
Plays, median (IQR)	2,900,000 (1,400,000-4,500,000)	1,600,000 (1,100,000-3,000,000)
Likes, median (IQR)	447,800 (194,400-666,450)	220,600 (168,250-369,200)
Comments, median (IQR)	6253 (3402-10,900)	2963 (1408-5480)
Shares, median (IQR)	18,500 (4448-61,100)	8986 (2636-16,800)
Followers, median (IQR)	191,400 (17,150-312,150)	55,550 (7690-210,200)
Health care expert, n (%)	0 (0)	15 (18)
TikTok still present as of November 29, 2021, n (%)	19 (100)	75 (90)

[a]TikToks were categorized as created by either a layperson or health care expert.

Figure 2. Violin plot depiction of individual TikTok metrics stratified as either vaccine hesitant (n=19) or pro-vaccine (n=83) as presented within Table 1.

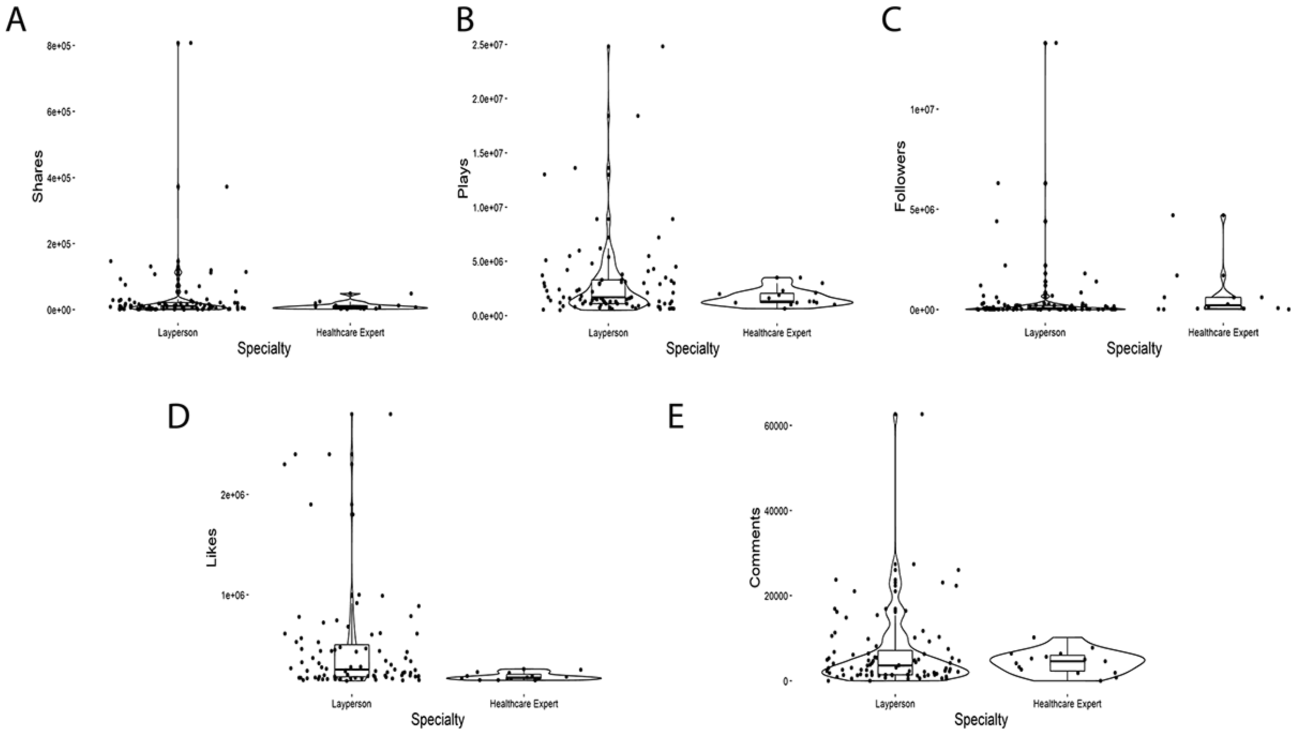


XSL•FO
RenderX

Table 2. General metrics of individual TikTok characteristics created by laypeople or health care experts. Of note, 4 of the 15 health care expert–created TikToks are from the same user, Dr Noc.

Characteristics	Layperson (N=87)	Health care expert (N=15)
Supporting vaccine, n (%)	68 (78)	15 (100)
Plays, median (IQR)	1,700,000 (1,100,000-3,300,000)	1,300,000 (1,200,000-2,050,000)
Likes, median (IQR)	252,600 (180,800-501,900)	173,100 (162,250-205,100)
Comments, median (IQR)	3545 (1408-7108)	4562 (2268-5998)
Shares, median (IQR)	10,300 (3034-20,750)	6885 (3142-12,200)
Followers, median (IQR)	53,000 (7804-198,300)	209,000 (44,400-610,200)
TikTok still present as of November 29, 2021, n (%)	81 (93)	13 (87)

Figure 3. Violin plot depiction of individual TikTok metrics stratified by TikTok creator, either layperson (n=87) or health care expert (n=15), as presented within Table 2 (N=102).



Vaccine-Hesitant Content Misinformation Analysis

Beyond general TikTok metrics, we performed content analysis to identify certain perceptions and misinformation associated with TikToks (n=19, 18.6%) containing vaccine-hesitant content. Most of vaccine-hesitant TikToks (n=10, 53%) did not voice any particular vaccine-hesitant themes other than that the person chose not to get vaccinated. From the remaining (n=9, 47%) vaccine-hesitant TikToks, several vaccine-hesitant sentiments were noted as follows:

- We do not know the long-term side effects
- The vaccine injects you with a microchip
- The vaccine makes you magnetic

From these 9 vaccine-hesitant TikToks, 5 (55%) TikToks were listed as the individual creator's "Top Liked" video; 8 of the 9 (88%) vaccine-hesitant TikToks pertained in some way to parodying or alluding to the vaccines causing neurological side effects that included dystonia or dysarthria; 2 out of the 9 (22%) vaccine-hesitant TikToks alleged that the vaccine injects you with a microchip that may make the individual magnetic.

HCP Creator Verification

When attempting to verify HCP status of all (n=15) HCP-related TikToks, 12 (80%) Tik Toks were able to be attributed to a verified HCP through assessing medical professional registries, professional or academic institutions, and social media verification blue check marks (on TikTok and Instagram; Table 3). We were unable to verify 3 (20%) HCP-related TikToks, as either the creators purely self-reported HCP status or the TikTok account was deleted with no potential for follow-up investigation.

XSL•FO
RenderX

Table 3. Analysis of the specialty of the health care expert TikTok creators who were subsequently verified by JL and KvK. Total specialty number (n=12) does not align with total number of health care professional–created TikToks (n=15), as one creator within the research scientist category (Dr Noc) created multiple (n=3) TikToks within our analysis.

Specialty	Creators in each field, n	creators verified, n/N (%)
Physician	6	5/6 (83)
Nurse	2	2/2 (100)
Doula	1	0/1 (0)
Pharmacist	1	0/1 (0)
Phlebotomist	1	1/1 (100)
Research scientist (Immunology)	1	1/1 (100)

Discussion

Our study characterized the content on TikTok during the summer of 2021 by analyzing TikToks tagged under the hashtag #covidvaccine. The results allow us to draw some conclusions regarding attitudes prevalent on TikTok during this time. Most of TikToks were supportive of vaccination, though the vaccine-hesitant content garnered more likes, shares, and views. HCPs represented a small portion of creators and all created provaccine content. Generally, vaccine-hesitant content reflected fears about side effects of the vaccine that were unfounded, such as magnetism.

Over 80% (n=83) of TikToks included in the study contained provaccine sentiments (Table 1). In comparison, Basch et al [3], who analyzed the same hashtag in March 2021, found only 36% of videos encouraging vaccination. As Basch et al did not code support for vaccines binarily (coding encompassed vaccine support, genre, and claims) [3], it is difficult to directly compare our study's findings with their prior research. However, given the change over time, it's suggestive that there was some increase in provaccine sentiments between the months of March and July 2021. This could be influenced by world events such as the increased distribution of vaccination around the globe. Between the months of March and July 2021, the number of individuals fully vaccinated against COVID-19 rose from 30.11 million to 159.79 million [24]. As vaccines became more available to the general population, more people may have posted about getting the vaccine [25]. In fact, many of our coded provaccine TikToks contained people recording themselves receiving their vaccination. It is also possible TikTok's misinformation management algorithm may have changed their system for flagging and removing inaccurate videos. Currently, TikTok claims to combat medical misinformation by banning antivax advertisements, and it directs users to the World Health Organization's website for COVID-19– related information [26]. They also claim to be removing TikToks containing misinformation within 24 hours [26]. TikTok USA promoted vaccine use on the platform, using #VaccinatedFor, a hashtag for users to share their reasons for being vaccinated [27]. Ultimately, the rise in provaccine TikToks is likely multifactorial. Some of the contributing factors may include the increasing proportion of vaccinated individuals, improved TikTok algorithm management for removal of misinformation, and increasing provaccine social outreach campaigns.

Nevertheless, even with the improvements in promoting vaccine use, none of the 9 TikToks containing misinformation from our original analysis were removed from the TikTok platform by November 2021. This suggests that TikTok may not be fully successful with their misinformation policy and is still struggling with detecting misinformation on the platform, with certain vaccine-hesitant content remaining on the app for more than 5 months. It is important to note that the deletion of provaccine content TikToks may have been due to a variety of reasons, such as the user leaving the TikTok app, or more worryingly, due to harassment over provaccine sentiments [28]. A possible reason for the great popularity of vaccine-hesitant TikToks may be the inclusion of misinformation that has a broad shock appeal [29]. We noted several TikToks that alluded to the vaccines' side effect of making individuals magnetic. Although content on TikTok may generally be more provaccine, it cannot be ignored that the smaller portion of vaccine-hesitant videos gaining higher traffic represents a dangerous avenue for the spread of misinformation on TikTok.

Our analysis found that HCP-created TikToks accounted for 15% (n=15) of the total 102 most popular TikToks, and only 6% (n=6) were posted by physicians. This is only a slight increase from Southwick et al's findings of 4% of individuals who were posting vaccine content on TikTok self-reporting as HCPs [11]. The small percentage of HCP-created popular TikToks suggests that HCPs can further use this platform to disseminate accurate medical information to a broad audience. However, with the advent of the medical influencer [1], it must be ensured that information provided by HCPs is correct and not biased by financial incentives. Although many current medical students have used social media for both their personal and professional lives, many have not received formalized social media training on how to disseminate information correctly beyond maintaining a professional image. Indeed, guidelines for maintaining a professional image have been created by both the American and Canadian Medical Associations [30,31]; however, these recommendations do not provide a guide on how to create new content that has educational value or helps combat misinformation. Social media platforms will continue to keep growing and gaining new followers, regardless of whether the health care community participates or not. As such, it is imperative that health care programs, residencies, and medical schools offer training to providers who choose to engage in medical education to broaden the reach of HCPs on social media.

XSL•FO
RenderX

Limitations

As this was a cross-sectional study, there are inherent limitations to interpreting trends found on TikTok for the month of July. The TikToks deemed viral at the time may not be viral currently, and this may change the viewership metrics. Future studies may benefit from comparing several cross-sectional studies and perform content analysis on how trends change over time as more vaccinations are rolled out globally. Our data extraction is also limited by TikTok's algorithm, which is known to show users content related to their interests. Although the extracted TikToks were highly viewed, it is difficult to determine whether the views originated from people being recommended content by the "For You" page algorithm or whether individuals specifically searched out the #covidvaccine hashtag. Due to the small sample size of HCP content creators, it is difficult to draw conclusions on what makes an HCP creator reach a broad audience.

Future Studies

Further studies should work to continue to characterize HCP content to gain an understanding of how HCPs can better combat misinformation. Examination of more than one hashtag could better categorize the growing field of vaccine-related content. Furthermore, comparing the TikToks at two different time points could better depict the ever-changing discourse of vaccine use on TikTok. Future studies should seek to understand the underlying causes that allow TikToks with blatant misinformation to succeed on the app.

Conclusions

Given the 3 billion views of content about #covidvaccine [32], TikTok is clearly a platform where vaccine discourse is taking place. Although most of the content is provaccine, the smaller proportion of vaccine-hesitant content continues to receive more traffic in likes, shares, and comments, indicating that misinformation is still being engaged with on this platform. Encouragingly though, HCPs can play a role in curbing misinformation by posting provaccine content and establishing a larger presence on the app. Using TikTok and other social media responsibly is imperative to how health information will be spread around the globe. Studying these trends must be continued to understand how the world perceives medical information and how HCPs can improve trust in science and vaccines.

Acknowledgments

The authors would like to thank Dr Yusuf Yilmaz for his invaluable help with citation management and TMC's research group for their constructive feedback throughout the research process. TMC has received funding for her work about social media for education and knowledge translation from the Physician Services Incorporated (PSI) foundation in Ontario, Canada.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The most viewed TikToks with the hashtag #covidvaccine downloaded.

[XLSX File (Microsoft Excel File), 34 KB - infodemiology_v2i2e38316_app1.xlsx]

Multimedia Appendix 2

TikToks reviewed by 2 authors (JL and KvK) encoded or categorized.

[XLSX File (Microsoft Excel File), 41 KB - infodemiology_v2i2e38316_app2.xlsx]

References

1. Dugdale L, Braswell H. Instacash: the ethics of leveraging medical student status on social media. Acad Med 2021 Apr 01;96(4):507-511. [doi: 10.1097/ACM.0000000000003885] [Medline: 33298692]
2. Abbas N, Ojha U. Not just a medical student: delivering medical education through a short video series on social media. JMIR Med Educ 2019 May 06;5(1):e11971 [FREE Full text] [doi: 10.2196/11971] [Medline: 31066690]
3. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. Hum Vaccin Immunother 2021 Mar 25:1-5. [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
4. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: an overview. Hum Vaccin Immunother 2013 Aug 27;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
5. World HO. Ten threats to global health in 2019. World Health Organization. 2019. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2021-12-19]
6. Safo SA. Exploring barriers and facilitators to COVID-19 vaccine uptake among Black and Latinx communities. JAMA Netw Open 2021 Oct 01;4(10):e2129675 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.29675] [Medline: 34643724]
7. Alyssa L. The U.S. is falling to lowest vaccination rate of G7 nations. The New York Times. 2021 Sep 11. URL: https://www.nytimes.com/2021/09/11/world/asia/us-vaccination-rate-low.html [accessed 2021-12-19]

XSL•FO
RenderX

8. Percentage of U.S. population who currently use any social media from 2008 to 2021. Statista. 2021. URL: https://www.statista.com/statistics/273476/percentage-of-us-population-with-a-social-network-profile/ [accessed 2021-12-19]
9. 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]
10. Swire-Thompson B, Lazer D. Public health and online misinformation: challenges and recommendations. Annu Rev Public Health 2020 Apr 02;41(1):433-451 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094127] [Medline: 31874069]
11. Southwick L, Guntuku SC, Klinger EV, Seltzer E, McCalpin HJ, Merchant RM. Characterizing COVID-19 content posted to TikTok: public sentiment and response during the first phase of the COVID-19 pandemic. J Adolesc Health 2021 Aug;69(2):234-241 [FREE Full text] [doi: 10.1016/j.jadohealth.2021.05.010] [Medline: 34167883]
12. Lapowsky I. New report shows kids see COVID-19 misinfo on TikTok in minutes. Protocol. 2021 Sep 22. URL: https://www.protocol.com/policy/tiktok-newsguard-covid [accessed 2021-12-19]
13. Burki T. Vaccine misinformation and social media. Lancet Digit Health 2019 Oct;1(6):e258-e259. [doi: 10.1016/S2589-7500(19)30136-0]
14. 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]
15. Kaye DBV, Chen X, Zeng J. The co-evolution of two Chinese mobile short video apps: parallel platformization of Douyin and TikTok. Mob Media Commun 2020 Aug 30;9(2):229-253. [doi: 10.1177/2050157920952120]
16. Li Y, Guan M, Hammond P, Berrey L. 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]
17. TikTok revenue and usage statistics (2022). Business of Apps. 2022 Aug 19. URL: https://www.businessofapps.com/data/tik-tok-statistics/ [accessed 2022-08-22]
18. Grierson J, Milmo D, Farah H. Revealed: anti-vaccine TikTok videos being viewed by children as young as nine. The Guardian. 2021 Oct 08. URL: https://www.theguardian.com/technology/2021/oct/08/revealed-anti-vaccine-tiktok-videos-viewed-children-as-young-as-nine-covid [accessed 2021-12-19]
19. Olivia L, Abbie R. TikTok's algorithm is amplifying COVID-19 and vaccine misinformation. Media Matters for America. 2021 Aug 18. URL: https://www.mediamatters.org/tiktok/tiktoks-algorithm-amplifying-covid-19-and-vaccine-misinformation [accessed 2021-12-19]
20. Zenone M, Ow N, Barbic S. TikTok and public health: a proposed research agenda. BMJ Glob Health 2021 Nov 24;6(11):e007648 [FREE Full text] [doi: 10.1136/bmjgh-2021-007648] [Medline: 34819326]
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;111(3):389-391 [FREE Full text] [doi: 10.17269/s41997-020-00343-0] [Medline: 32519085]
22. DavidTeather/TikTok-Api. GitHub. 2021 Dec 05. URL: https://github.com/davidteather/tiktok-api [accessed 2021-12-05]
23. Gabiherman-tiktok-vaccine-study. GitHub. 2021 Nov 12. URL: https://github.com/gabiherman/tiktok-vaccine-study [accessed 2021-12-05]
24. Ritchie H, Mathieu E, Rodés-Guirao L. Coronavirus (COVID-19) vaccinations. Our World in Data. 2020 Mar 05. URL: https://ourworldindata.org/covid-vaccinations [accessed 2021-12-19]
25. Anderson R. COVID-19 vaccine selfies taking social media by storm. CTV News. 2021 Apr 10. URL: https://www.ctvnews.ca/health/coronavirus/covid-19-vaccine-selfies-taking-social-media-by-storm-1.5382518 [accessed 2021-12-19]
26. COVID-19-safety center. TikTok. URL: https://www.tiktok.com/safety/en-us/covid-19/ [accessed 2021-12-19]
27. Supporting vaccine education on TikTok. TikTok. 2021 Apr 22. URL: https://newsroom.tiktok.com/en-us/supporting-vaccine-education-on-tiktok [accessed 2022-01-04]
28. Perrett C. Medical experts on TikTok are paying the price for stepping up against a 'ridiculous amount of myths' like taking Vitamin C for at-home abortions. Insider. 2021 May 26. URL: https://www.businessinsider.com/tiktok-misinfo-doctors-pay-the-price-for-addressing-2021-4 [accessed 2021-12-19]
29. 'These are false': Chicago's top doc debunks viral vaccine magnet social media videos. NBC Chicago. 2021 Jun 11. URL: https://www.nbcchicago.com/news/coronavirus/these-are-false-chicagos-top-doc-debunks-viral-vaccine-magnet-social-media-videos/2529790/ [accessed 2021-12-19]
30. Kind T. Professional guidelines for social media use: a starting point. AMA J Ethics 2015 May 01;17(5):441-447 [FREE Full text] [doi: 10.1001/journalofethics.2015.17.5.nlit1-1505] [Medline: 25986088]
31. Social media and Canadian physicians: issues and rules of engagement. DOC Player. URL: https://docplayer.net/25751070-Social-media-and-canadian-physicians-issues-and-rules-of-engagement.html [accessed 2022-01-30]
32. #COVIDVaccine videoes. TikTok. URL: https://www.tiktok.com/tag/covidvaccine [accessed 2022-08-22]

Abbreviations

HCP: health care professional

XSL•FO
RenderX

Edited by T Mackey; submitted 28.03.22; peer-reviewed by D Goergen, M Nali; comments to author 15.08.22; revised version received 30.08.22; accepted 01.10.22; published 25.10.22.

Please cite as:
van Kampen K, Laski J, Herman G, Chan TM
Investigating COVID-19 Vaccine Communication and Misinformation on TikTok: Cross-sectional Study
JMIR Infodemiology 2022;2(2):e38316
URL: https://infodemiology.jmir.org/2022/2/e38316
doi:10.2196/38316
PMID:36338548

©Katherine van Kampen, Jeremi Laski, Gabrielle Herman, Teresa M Chan. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 25.10.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

COVID-19 Health Beliefs Regarding Mask Wearing and Vaccinations on Twitter: Deep Learning Approach

Si Yang Ke[1*], BS; E Shannon Neeley-Tass[1*], PhD; Michael Barnes[2*], PhD; Carl L Hanson[2*], PhD; Christophe Giraud-Carrier[3*], PhD; Quinn Snell[3*], PhD

[1]Department of Statistics, Brigham Young University, Provo, UT, United States
[2]Department of Public Health, Brigham Young University, Provo, UT, United States
[3]Computer Science Department, Brigham Young University, Provo, UT, United States
[*]all authors contributed equally

Corresponding Author:
Quinn Snell, PhD
Computer Science Department
Brigham Young University
3361 TMCB
Provo, UT, 84602
United States
Phone: 1 801 422 5098
Email: snell@cs.byu.edu

Abstract

Background: Amid the global COVID-19 pandemic, a worldwide infodemic also emerged with large amounts of COVID-19–related information and misinformation spreading through social media channels. Various organizations, including the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC), and other prominent individuals issued high-profile advice on preventing the further spread of COVID-19.

Objective: The purpose of this study is to leverage machine learning and Twitter data from the pandemic period to explore health beliefs regarding mask wearing and vaccines and the influence of high-profile cues to action.

Methods: A total of 646,885,238 COVID-19–related English tweets were filtered, creating a mask-wearing data set and a vaccine data set. Researchers manually categorized a training sample of 3500 tweets for each data set according to their relevance to Health Belief Model (HBM) constructs and used coded tweets to train machine learning models for classifying each tweet in the data sets.

Results: In total, 5 models were trained for both the mask-related and vaccine-related data sets using the XLNet transformer model, with each model achieving at least 81% classification accuracy. Health beliefs regarding perceived benefits and barriers were most pronounced for both mask wearing and immunization; however, the strength of those beliefs appeared to vary in response to high-profile cues to action.

Conclusions: During both the COVID-19 pandemic and the infodemic, health beliefs related to perceived benefits and barriers observed through Twitter using a big data machine learning approach varied over time and in response to high-profile cues to action from prominent organizations and individuals.

(JMIR Infodemiology 2022;2(2):e37861) doi:10.2196/37861

KEYWORDS

COVID-19; Health Belief Model; deep learning; mask; vaccination; machine learning; vaccine data set; Twitter; content analysis; infodemic; infodemiology; misinformation; health belief

Introduction

On January 30, 2020, the World Health Organization (WHO) declared the Chinese outbreak of SARS-CoV-2 (ie, COVID-19) to be a public health emergency of international concern [1]. The following day, the United States Department of Health and Human Services (HHS) secretary declared a US public health emergency to respond to COVID-19 [1]. The president of the United States signed a "Proclamation on Suspension of Entry as Immigrants and Nonimmigrants of Persons Who Pose a Risk

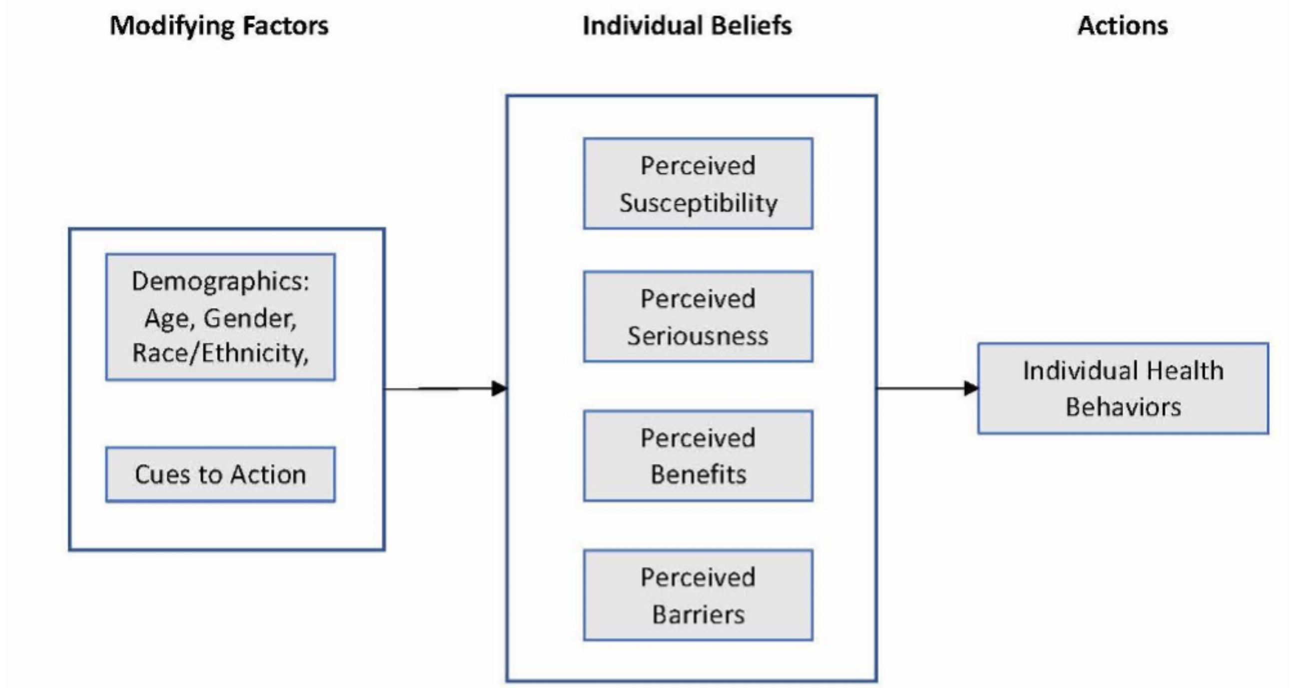
XSL•FO
RenderX

of Transmitting 2019 Novel Coronavirus," limiting entry into the United States of persons who traveled to mainland China. Subsequently, on March 11, 2020, WHO declared COVID-19 to be a global pandemic.

With the emergence of a global pandemic came another concern, the emergence of a worldwide infodemic. As it pertained to COVID-19, WHO described the infodemic as an overabundance of information and misinformation related to the COVID-19 pandemic that led to mistrust of health authorities and hampered public health efforts [2]. With the growth in social media use, information about COVID-19 spread quickly, necessitating infodemic management or the need to manage false and misleading information in such a way that would reduce the impact on health behaviors [2]. Greater attention is being paid to sources of COVID-19 information, especially low-credibility sources responsible for spreading COVID-19 misinformation through social media channels [3]. As such, several researchers have begun to address methods for fighting the COVID-19 infodemic and acknowledge the influential role of social media [4-7]. The continuous monitoring and analysis of social media information (infodemiology) has been heralded as a critical tool for understanding the influence of social media and combating misinformation [5]. Although social media data are not specifically designed for public health purposes, they are a valuable and accessible resource for public health surveillance purposes [8]. For example, topic modeling of Twitter posts has been used in understanding topics and sentiments related to COVID-19 in general [9-11], face masks [12,13], and vaccine discussions [14]. Although simply monitoring social media information can provide valuable insight into COVID-19 information/misinformation, understanding the influence on health beliefs and behaviors during an infodemic can assist public health in better managing information during health emergencies, such as COVID-19, through risk communication [14]. In addition, less is known about the influence of higher-credibility sources of information, such as prevention guidelines coming from WHO and the Centers for Disease Control and Prevention (CDC).

The Health Belief Model (HBM) was developed to explain how beliefs impact health decisions [15]. The theory posits that people engage in health-related behaviors based on (1) their perception of the health condition (eg, COVID-19), (2) their perception of the advantages and disadvantages of the health behavior (eg, mask wearing or receiving a vaccine), and (3) cues to action or stimuli that encourage them to participate in the behavior (eg, health organization recommendations). This theory consists of 5 main elements: perceived susceptibility, perceived severity, perceived benefits, perceived barriers, and cues to action (Figure 1). The model has been successfully used to assess health beliefs on social media regarding physical distancing during the COVID-19 pandemic [16], Zika virus [17], and the human papillomavirus vaccine [18]. Although traditional polling methods require a substantial number of resources and have limitations in assessing public health beliefs (eg, difficulty reaching a large-scale population in large geographic areas and tracking changes in real time), social media has provided millions of people, worldwide, a chance to voluntarily and continuously express their thoughts and opinions on issues that they deem important [18]. Although 1 study used machine learning of Twitter posts to monitor health beliefs regarding COVID-19, health care treatments, and the influence of various external cues to action [19], no identified study has used this methodology to explore the HBM regarding important COVID-19–related behavioral outcomes—mask wearing and vaccinations.

Figure 1. Health Belief Model (HBM).



XSL•FO
RenderX

Research has demonstrated that health organizations, physicians, and the media during COVID-19 represent important HBM cues to action [20]. WHO and the CDC have both issued COVID-19 prevention recommendations; however, recommendations have evolved over time. For example, in the first official advice document regarding the need for and usage of masks, WHO stated that a medical mask is not required for healthy individuals, as no evidence was available on its usefulness to protect nonsick persons [21]. Later, in the updated mask guidelines, the previous advice was modified, and the general public was encouraged to wear masks [22]. Similarly, the CDC initially asserted that the wearing of face masks was an unnecessary public health tool, but a short while later, it issued new guidelines advising people to wear face coverings in public settings where social distancing was difficult [23]. Understanding the individual beliefs regarding COVID-19–preventive behaviors in response to various cues to action from these high-credibility sources is crucial toward helping manage an infodemic. The fallout from these types of COVID-19 shifts in prevention guidelines have created controversy among many sources.

Generally, US guidelines emerging from national, state, and local public health organizations received prominent attention. Lessons have been learned from global and national guidelines, including the following: (1) Travel restriction delays allowed citizens traveling from high-risk areas to pass freely through airports without screening; (2) quarantine delays in high-risk areas allowed potentially infected individuals to spread the infection; (3) public misinformation allowed racism, incorrect public precautions, and unprecedented fear surrounding COVID-19, allowing rumors, speculation, and misinformation to spread; and (4) emergency announcement regarding the outbreak severity was delayed and not widely broadcast for a month when WHO declared the public health emergency of international concern [24]. Additionally, the WHO guidelines came under political scrutiny by the US president, who blamed WHO for delays and dysfunctions to investigate early cases of COVID-19 and suspended WHO funding [25].

The purpose of this study is to investigate health beliefs and cues to action for mask wearing and vaccination using machine learning of COVID-19–related Twitter posts. External cues to action from prominent pandemic declarations (eg, WHO and the CDC) regarding mask wearing and vaccines and prominent examples of displayed preventive behaviors (eg, presidential mask wearing) were explored for possible influence on health beliefs, as explained by HBM constructs. Although these prominent events could not be studied for a cause-effect relationship, given the surveillance approach of this study, observing the prominent events along with the ongoing Twitter posts may help provide clues to their potential effect on cues to action. This unique approach is an important way to begin exploring how infodemics may influence cues to action. Tables 1 and 2 show the HBM constructs for masks and vaccines, respectively. Findings from this study revealed that cues to action are associated with increased conversations around the perceived health beliefs about mask wearing and vaccinations.

Table 1. HBM[a] constructs related to COVID-19 and face coverings.

Construct	Definition
Perceived susceptibility	Assessment of the likelihood or risk of contracting COVID-19; increased likelihood of contracting the disease (eg, increased/decreased prevalence, high/low number of COVID-19 cases)
Perceived severity	Assessment of the perceived seriousness and consequences of contracting COVID-19 (eg, hospitalization, death, mortality, disability)
Perceived benefits	Comments mentioning the benefits of masks or face coverings to reduce the transmission of COVID-19 or the removal of barriers (eg, promotion of mask or face coverings)
Perceived barriers	Comments mentioning the difficulties, challenges, and negative effects of masks and face coverings or the perceived ineffectiveness of masks and face coverings (eg, negative reports of masks or face coverings)

[a]HBM: Health Belief Model.

Table 2. HBM[a] constructs related to COVID-19 and vaccines.

Construct	Definition
Perceived susceptibility	Assessment of one's likelihood or risk of contracting COVID-19 if not vaccinated; references increased/decreased prevalence, high/low number of cases, and high/low risk/chance/probability
Perceived severity	Assessment of the seriousness of COVID-19 and the major consequences that contracting COVID-19 would have on one's life, such as hospitalization, death, mortality, or disability
Perceived benefits	Assessment of the benefits of COVID-19 vaccines or being vaccinated against COVID-19; the removal of barriers (see the Perceived Barriers section for more information); positive opinion
Perceived barriers	Assessment of the barriers to COVID-19 vaccination, including difficulties, challenges, conspiracies, negative effects, dangers, and perceived ineffectiveness; the removal of benefits; negative opinion

[a]HBM: Health Belief Model.

XSL•FO
RenderX

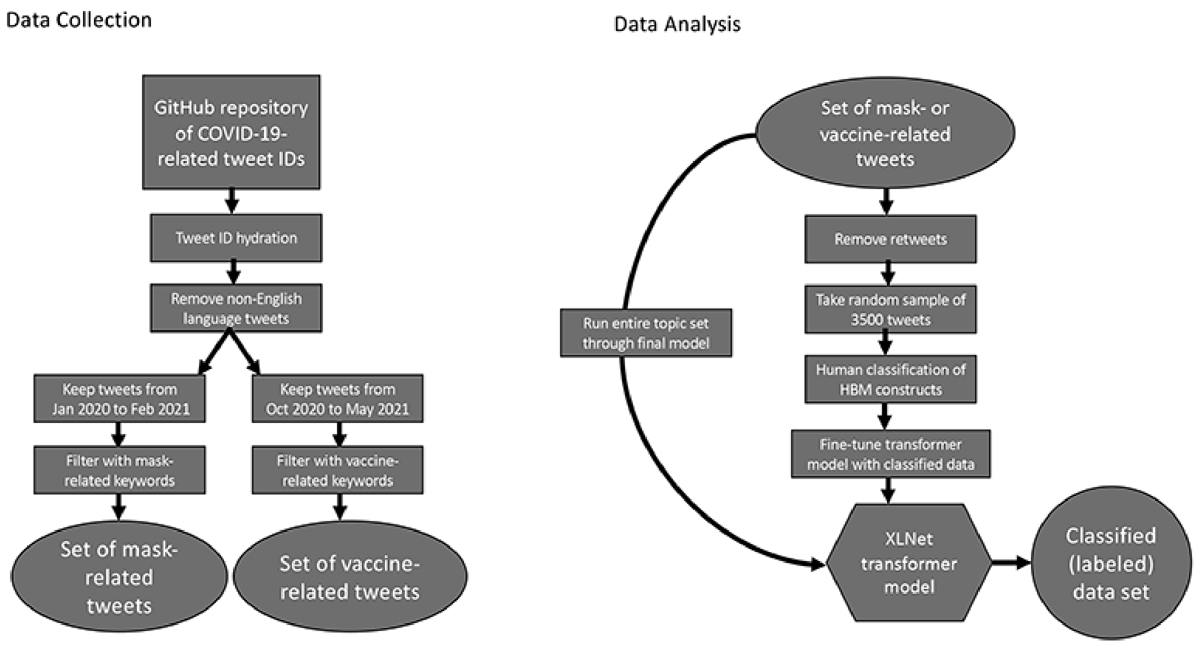
Methods

Data Collection

For this paper, a large, publicly available data set of COVID-19–related tweets was used [26,27]. Since Twitter's terms of service only allow the tweet IDs to be publicly available, the authors hydrated the tweet IDs with their own Twitter developer accounts.

The diagram on the left in Figure 2 outlines the data collection process. Since the transformer models used are pretrained and can handle rather raw information, only minor preprocessing of the data was necessary. Non-English tweets were excluded, and all text was converted to lowercase. Tweets were then filtered by date, and an iterative process was used to filter the tweets by keywords. Research students came up with initial lists of keywords for both mask-related and vaccine-related tweets. The lists were reviewed by the extended research group and modified, as appropriate. For example, the first list included the keyword "face," which picked up a lot of tweets talking about "Facebook." Hence, the keyword "face" was changed to "face " (with a space after the "e"). The keyword lists may not be perfect and may cause the inclusion of tweets outside of the scope of face masks or vaccines related to COVID-19. However, they do produce a more relevant population than the entire corpus, and together with the hand labeling, it was felt that the model would be able to identify HBM-related tweets.

Figure 2. Data collection and analysis procedures. HBM: Health Belief Model.



The set of mask-related tweets was created by filtering tweets from January 2020 through January 2021 with the following keywords: "mask," "face cover," "facemask," "cloth cover," "cover your face," "face covering," "maskup," and "face." Likewise, the set of vaccine-related tweets was created by filtering tweets from October 2020 through November 2021 with the following keywords: "vaccine," "antiva," "anti-va," "vax," "shot," "inoculat," "needle," "booster," "pfizer," "biontech," "immun," "mrna," "trials," "moderna," "novacax," "astrazeneca," "johnson," "sanofi," and "glaxosmithkline." As expected, vaccine-related Twitter conversation came later than mask-related conversation, which is reflected in the different start dates of the data collection. As to the end dates, their selection was simply a matter of choice based on the authors' hypotheses. Upon examination of the mask-related data and resultant graphs, the hypothesized correlations were clearly exhibited, and it was felt that additional data would not significantly change these results (the HBM is all about beliefs affecting behavior), so data collection for the mask-related tweets was interrupted (similarly for the vaccine-related tweets, but for the relevant, later date range).

The final data set statistics are as follows: (1) 1.8 TB on disk, (2) 646,885,238 total tweets, (3) 59,724,507 mask-related tweets, and (4) 113,542,400 vaccine-related tweets.

Data Analysis

Once the topic-related tweets sets were created, they were classified according to the HBM constructs. The classification process was performed separately for each topic-related set, but the process was the same. A random sample of 3500 tweets was selected for manual labeling. No retweets were included in these sets to avoid biasing the models, since retweets would cause repetition of content. Hence, the labeled data consisted of unique tweets, and only these were used in the subsequent model-building phase. Three independent reviewers manually classified the sample according to their relevance to the HBM constructs. Every tweet in the sample was classified with either a positive or a negative label for each of the 4 HBM constructs (ie, perceived susceptibility, perceived severity, perceived benefits, and perceived barriers). Next, a tweet that was labeled positive for at least 1 of the 4 constructs was also classified as HBM related. Tables 1 and 2 show the criteria for a positive label for each of the 4 HBM constructs for the mask- and vaccine-related tweets, respectively. These criteria were defined

XSL•FO
RenderX

using constructs from similar work. Compiled together, each construct is assumed to impact the likelihood for persons taking (or not taking) action. The reviewers classified the tweets independently but compared classifications together after labeling the first 100 tweets and again after the first 500 tweets to resolve conflicts. During these calibration meetings, the reviewers came together and examined samples of tweets that differed in classifications and came to consensus based on the criteria defined in Tables 1 and 2. The final label for each tweet was the label that received the majority vote (ie, at least 2 votes out of 3 from the reviewers).

Note that each tweet could be labeled as belonging to more than 1 of the 4 HBM constructs as 1 part of a tweet could fall into 1 HBM category and another part of the tweet could fall into a different HBM category. As an example, consider the following tweet from the set of mask-related tweets:

> *Buying reusable masks is generally a waste of time and hurts healthcare workers (hello) who need to use them in everyday care during the flu season. Is the corona virus scary? The weirdo incubation time and severity/rapid spread is honestly wild.*

This tweet was labeled for both perceived barriers and perceived susceptibility because the user first advocated not buying reusable masks (perceived barriers) and then proceeded to comment on how fast the virus spreads (perceived susceptibility).

Once labeled, each set of 3500 tweets (mask related and vaccine related) could be used for model-building purposes. A random stratified 2450/1050 (70%/30%) split was applied to create the training and test sets. Each model was trained exclusively on the training data (n=2450, 70%). The only hyperparameter considered was the dropout rate, which did not significantly change the results. The test data (n=1050, 30%) were then used to assess the quality of the models, and the results presented here are based on those data alone. Following their construction, the models were applied to label all tweets, as a real system would indeed be expected to label both original tweets as well as all retweets. The diagram on the right of Figure 2 illustrates the steps taken to process the data from the GitHub repository through the creation of 2 topic-related tweet sets and finally through the classification process.

State-of-the-art bidirectional transformer models [28] were used, combined with custom classification layers. Three different pretrained transformer models were considered and fine-tuned: bidirectional encoder representations from transformers (BERT) [29], a distilled version of bidirectional encoder representations from transformers (DistilBERT; more memory efficient), and XLNet [30]. Three additional simpler models were also included for comparison, namely logistic regression, RepresentationNet (a vanilla version of BERT with a custom classification network), and a bidirectional gated recurrent unit (BiGRU) network [31]. All models were implemented in Python with the *pytorch* library. Each model's predictive ability was evaluated by the area under the receiver operating characteristic (AUROC) curve, accuracy, precision, recall, and F_1 score. In addition, the final model size (pytorch binary format) and a measure of performance in the form of the number of tweets the model can classify per second were also computed.

After classifying all the tweets using the transformer model, the tweets were separated into calendar weeks and counted. HBM-positive label percentages were computed by taking raw counts of HBM-positive labels divided by the total raw count of COVID-19–related tweets filtered for mask- or vaccine-related keywords, respectively. Potential linear relationships between HBM label percentages by week and COVID-19–related statistics from the corresponding weeks, such as US confirmed case counts, US COVID-19 death counts, and US COVID-19 vaccine doses administered, were also investigated using scatter plots, Spearman correlation matrices, simple linear regression models, and regression with added quadratic terms.

Ethical Considerations

Ethical approval was not needed, since the study only analyzed publicly available data from existing data sets, and results do not contain any identifiable information and are represented in aggregate only.

Results

Manual Labeling and Interrater Reliability

For the mask-related data set, interrater reliability coefficients, measured by Gwet AC1, were 0.784, 0.762, 0.957, and 0.938 for perceived benefits, perceived barriers, perceived severity, and perceived susceptibility, respectively. Interrater reliability coefficients for the vaccine-related data set, measured by Gwet AC1, were 0.825, 0.814, 0.937, and 0.915, respectively, for the 4 HBM constructs in the same order as above. These interrater reliability coefficients are all interpreted as substantial agreement to almost perfect agreement according to the Landis-Koch benchmarking scale [32]. Compared to other interrater reliability calculations, Gwet AC1 is more stable than traditional κ coefficient calculations [33].

For the mask-related data set, 2135 (61%) tweets were manually labeled as related to the HBM model. For the vaccine-related data set, 1330 (38%) tweets were labeled as related to the HBM model.

Machine Learning Model

Table 3 reports on the AUROC, accuracy, precision, recall, and F_1 score for each model. It also includes the model size, as well as the number of tweets the model can classify per second. All 3 pretrained transformer models outperformed the simpler models. Among the transformer models, the XLNet transformer model (with a custom dense 3-layer classification network and a dropout rate set to 0.25) was clearly the best model for this task.

To provide a finer-grained analysis of performance, Figure 3 displays the complete AUROC curve for each of the models. The graph confirms the superiority of the transformer models. Furthermore, it shows that the XLNet transformer generally dominates, and at worse is on par with, the other transformer models.

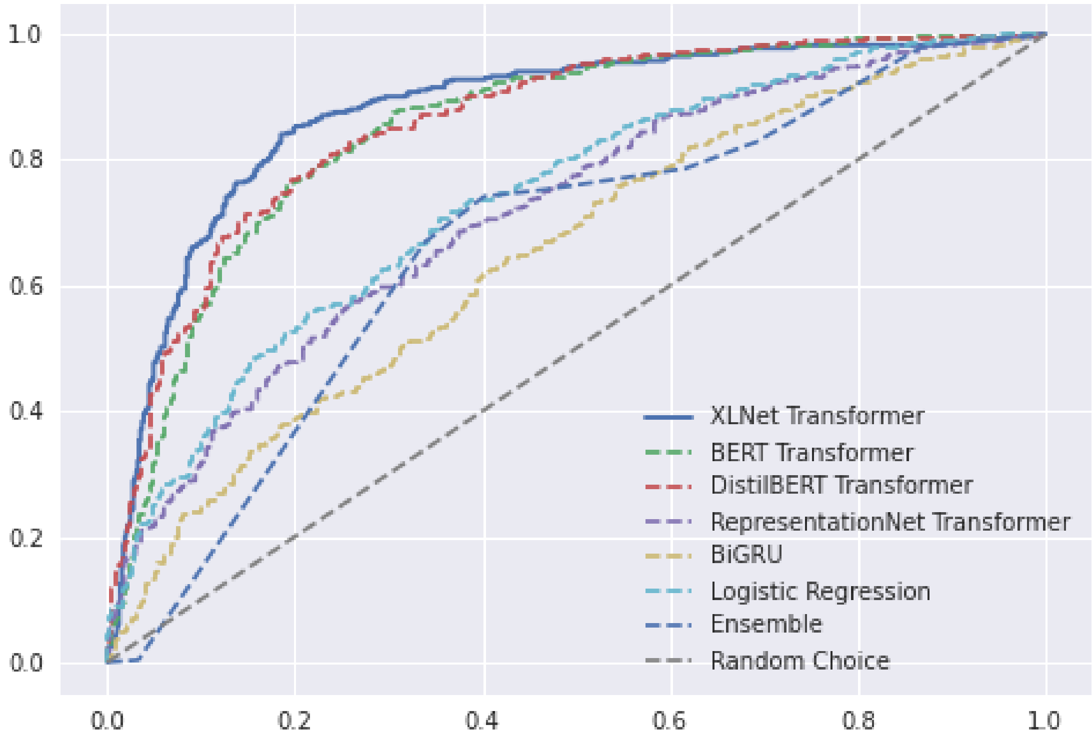
XSL•FO
RenderX

Table 3. Model evaluation metrics.

Model	AUROC[a]	Accuracy	Precision	Recall	F_1 score	Size (MB)	Evaluations/second
XLNet	0.878	0.824	0.775	0.761	0.768	467	234.8
BERT[b]	0.850	0.786	0.685	0.760	0.721	440	321.1
DistilBERT[c]	0.858	0.795	0.760	0.692	0.725	261	563.5
RepresentationNet	0.717	0.735	0.644	0.648	0.646	484	316.0
BiGRU[d]	0.737	0.752	0.665	0.678	0.672	27	473.9
Logistic regression	0.566	0.651	0.584	0.229	0.329	0.0016	15,925.3

[a]AUROC: area under the receiver operating characteristic.
[b]BERT: bidirectional encoder representations from transformers.
[c]DistilBERT: distilled version of bidirectional encoder representations from transformers.
[d]BiGRU: bidirectional gated recurrent unit.

Figure 3. AUROC curves. AUROC: area under the receiver operating characteristic; BERT: bidirectional encoder representations from transformers; BiGRU: bidirectional gated recurrent unit; DistilBERT: distilled version of bidirectional encoder representations from transformers.



It is worth noting the strength of transformer models, especially pretrained ones. Indeed, it is quite remarkable that XLNet achieved over 82% accuracy with only 3500 training data points. By contrast, logistic regression achieved only about 65% accuracy. It is likely that a larger training corpus would improve the performance of logistic regression, but transformer models can do rather well with relatively small corpora. It is also of note that although XLNet is better across predictive evaluation metrics, its memory footprint is significantly larger, and its speed of execution is slower. In cases of low memory, the DistilBERT model may be a viable alternative at the slight cost of some performance loss in classification. Here, the XLNet model was preferred because it fit well within available memory constraints.

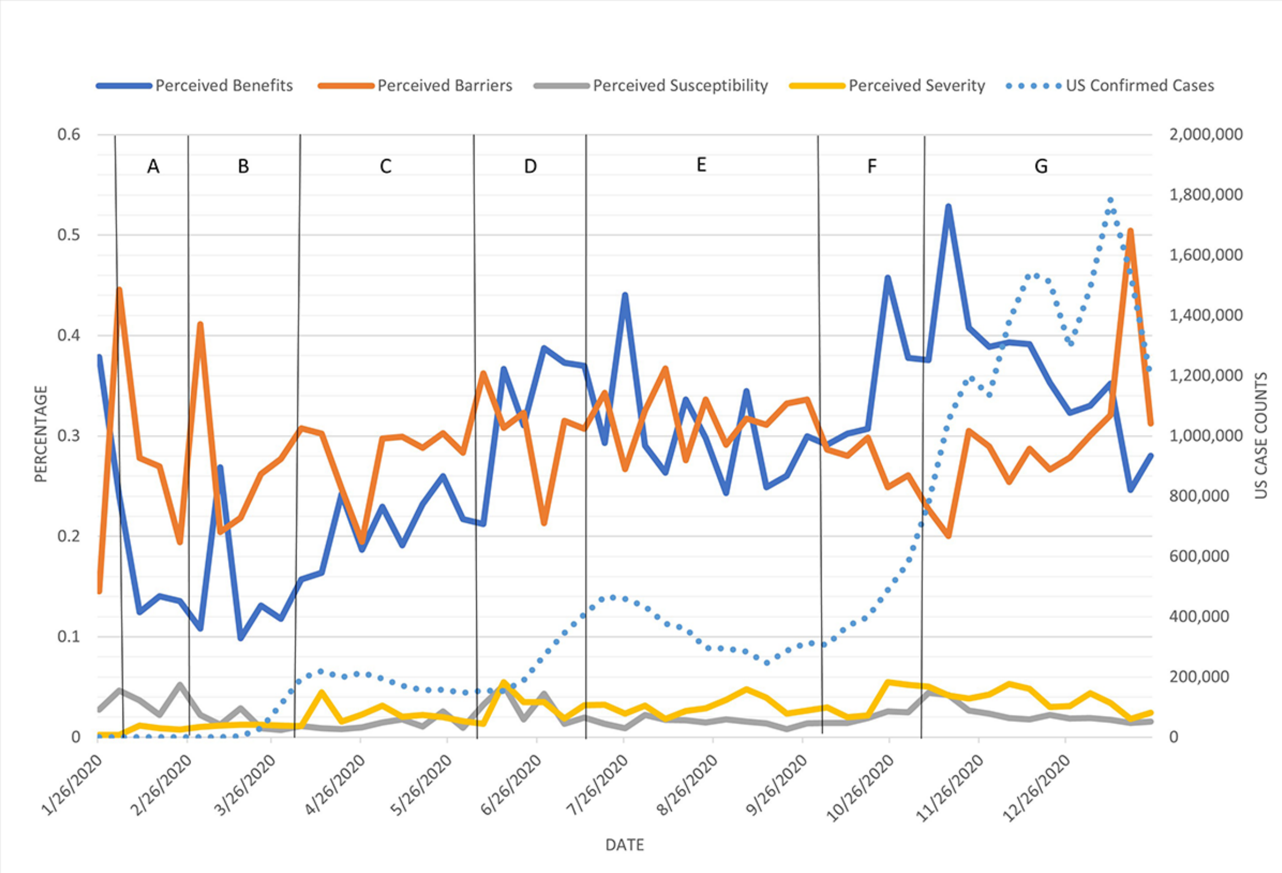
Using the XLNet transformer model, 5 models were trained to classify the tweets for HBM relatedness and for the 4 HBM constructs (in the order of perceived benefits, perceived barriers, perceived severity, and perceived susceptibility). The 5 models each achieved over 81% classification accuracy (81%, 97%, 96%, 85%, 82%, respectively) on test data for the mask-related tweets and over 79% classification accuracy (82%, 81%, 86%, 79%, 85%, respectively) on test data for the vaccine-related tweets. Again, the pretrained transformers proved effective at embedding the tweets, making it simple to train the custom classification networks.

The HBM-positive label percentages are plotted by week in Figure 4 for mask-related tweets and Figures 5 and 6 for vaccine-related tweets. In Figures 4-6, the HBM-positive label percentages were computed by dividing raw counts of HBM

XSL•FO
RenderX

labels by the total raw counts of COVID-19–related tweets filtered for mask- or vaccine-related keywords, respectively. Due to the nature of these calculations, the analysis and comparisons synthesized from Figures 4-6 focused on the direction of change in individual HBM labels rather than on the magnitude of change. Moreover, the magnitude of individual HBM label percentages were not compared across HBM label categories.

Figure 4. US COVID-19 case counts and each HBM scale across time for mask-related tweets. There was a statistically significant correlation between case counts and perceived benefits. Important events corresponding to the time intervals are listed in Table 4. HBM: Health Belief Model.



XSL•FO
RenderX

Figure 5. US COVID-19 case counts and each HBM scale across time for vaccine-related tweets. There was a statistically significant correlation between case counts and perceived barriers. Important events corresponding to the time intervals are listed in Table 5. HBM: Health Belief Model.

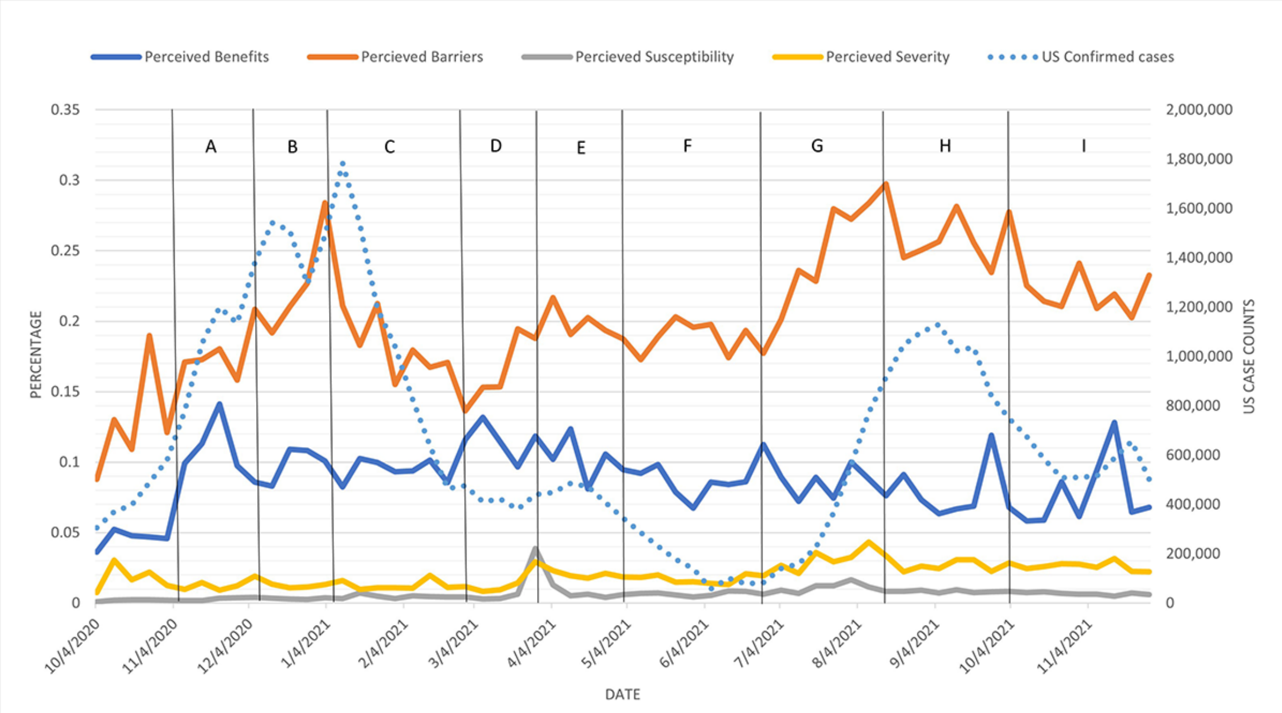
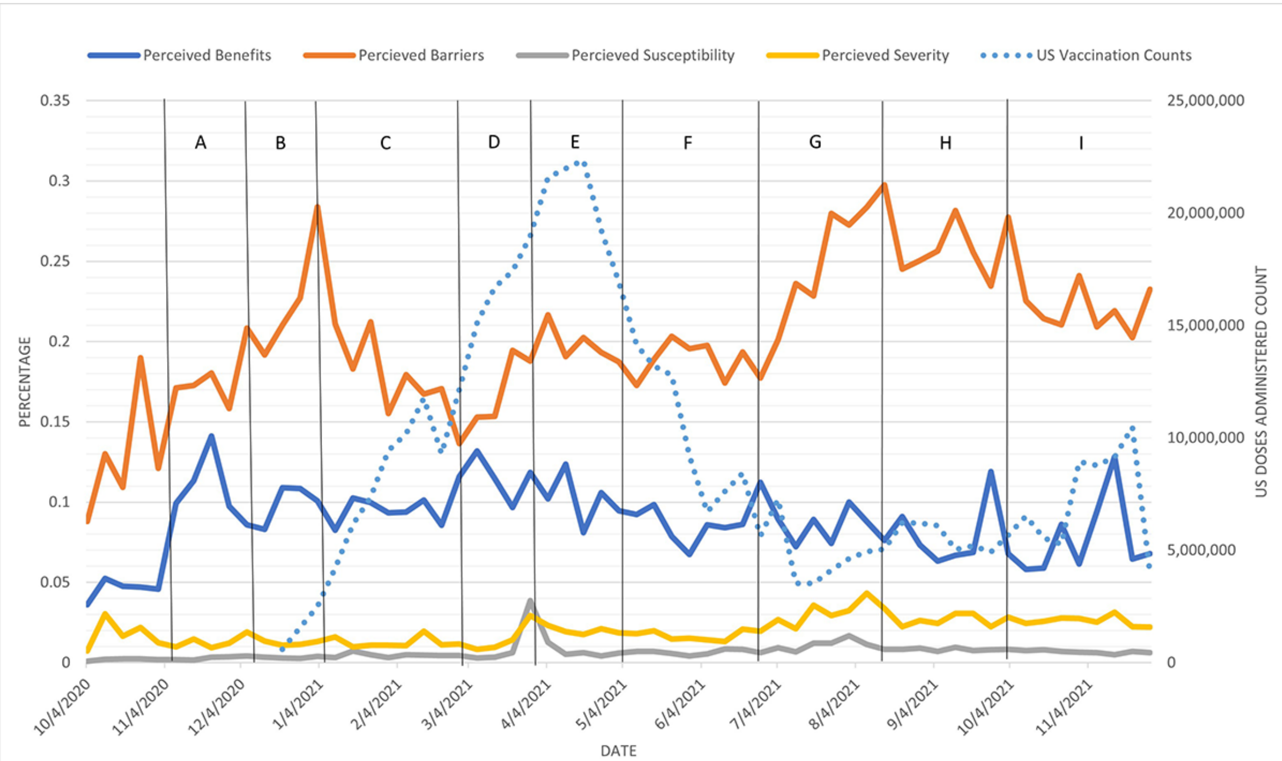


Figure 6. US COVID-19 vaccine doses administered and each HBM scale across time for vaccine-related tweets. There was a statistically significant negative correlation between vaccination counts and perceived barriers. Important events corresponding to the time intervals are listed in Table 5. HBM: Health Belief Model.



XSL•FO
RenderX

Table 4. Mask-related events and cues to action.

Dates	Cues to action
Time frame A	
January 21, 2020	The CDC[a] confirms the first COVID-19 case.
January 31, 2020	WHO[b] declares a global emergency.
Time frame B	
February 27, 2020	WHO issues interim guidance on masks.
March 8, 2020	This National Institutes of Health's (NIH) Dr Fauci says face masks are not fully effective.
March 11, 2020	WHO declares a pandemic.
March 13, 2020	President Trump declares an emergency.
Time frame C	
April 3, 2020	The CDC recommends wearing face masks in public.
April 10, 2020	The first state (New Jersey) mandates face masks in public.
April 16, 2020	The White House begins a formal discussion to open the economy.
June 3, 2020	The surgeon general (HHS[c]) asks Americans to stop buying face masks.
Time frame D	
June 5, 2020	WHO recommends masks for areas with community transmission.
June 17, 2020	WHO halts hydroxychloroquine production.
July 6, 2020	WHO is asked by scientists to revise guidelines to acknowledge airborne transmission.
July 9, 2020	WHO declares that COVID-19 is airborne-transmissible.
Time frame E	
July 12, 2020	President Trump is seen wearing a mask in public for the first time.
August 13, 2020	Presidential candidate Joe Biden calls for a 3-month mask mandate.
August 17, 2020	The United States declares COVID-19 as the third-leading cause of death.
August 28, 2020	The first US reinfection case is found.
Time frame F	
September 16, 2020	President Trump releases a vaccine distribution plan.
September 21, 2020	The CDC withdraws guidance saying COVID-19 is airborne-transmissible.
October 2, 2020	President Trump and the First Lady are diagnosed with COVID-19 and the president is hospitalized.
October 15, 2020	WHO declares conclusive evidence that hydroxychloroquine is ineffective.
Time frame G	
November 4, 2020	The United States reports an unprecedented 100,000 cases in 1 day.
December 14, 2020	WHO reports the first of the COVID-19 variants (in the United Kingdom).

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

[b]WHO: World Health Organization.

[c]HHS: Department of Health and Human Services.

XSL•FO
RenderX

Table 5. Vaccine-related events and cues to action.

Dates	Cues to action
Time frame A	
November 4, 2020	Joe Biden is elected the 46th President of the United States over Donald Trump.
November 9, 2020	Pfizer publishes vaccine results.
November 16, 2020	Moderna reveals vaccine efficacy results.
Time frame B	
December 4, 2020	President Biden asks Americans to commit to 100 days of wearing masks, his first act as president-elect.
December 11, 2020	The FDA[a] approves the emergency use of the Pfizer vaccine.
December 18, 2020	The FDA approves the emergency use of the Moderna vaccine.
December 21, 2020	President Biden get the first dose of the vaccine.
December 31, 2020	The United States falls short of the goal to give 20 million vaccinations by year-end (2.8 million).
Time frame C	
January 4, 2021	The White House says more individuals will be vaccinates using reserve supplies.
January 6, 2021	The HHS[b] provides US $22 billion to fund testing and vaccine distribution.
January 7, 2021	The CDC[c] says COVID-19 vaccine benefits outweigh allergic reaction risks for Pfizer or Moderna.
January 12, 2021	The CDC and the HHS update vaccine allocation to release all available doses.
February 4, 2021	The FDA begins considering the J&J[d] vaccine.
February 26, 2021	The Kaiser Family Foundation (KFF) poll shows vaccine acceptance increases among Americans.
February 27, 2021	The FDA approves the J&J vaccine with emergency use authorization.
March 1, 2021	The former president and First Lady urge followers to get vaccinated.
Time frame D	
March 4, 2021	The COVID-19 UK variant does not affect vaccine efficacy.
March 8, 2021	The CDC releases guidance on safe activities for fully vaccinated.
March 11, 2021	The Public Broadcasting Service (PBS) poll says nearly half of Republican men will not get COVID-19 vaccines.
March 11, 2021	President Biden pushes for expanded vaccine eligibility to all adults aged 18 years and older by May 1.
March 15, 2021	The White House unveils an expansive public relations vaccine confidence campaign.
Time frame E	
April 2, 2021	The CDC expands travel guidelines for those fully vaccinated.
April 6, 2021	The COVID-19 variant is detected in all 50 states.
April 13, 2021	The CDC and the FDA recommend pausing the J&J vaccine.
April 23, 2021	The CDC lifts the pause on the J&J vaccine.
April 27, 2021	The CDC eases mask restrictions for fully vaccinated individuals.
Time frame F	
May 4, 2021	The FDA prepares to authorize the Pfizer vaccine in adolescents.
May 4, 2021	President Biden announces a new goal of having 70%, or 160 million, American adults with at least 1 dose of a COVID-19 vaccine by July 4.
May 10, 2021	The Pfizer/BioNTech vaccine is approved for adolescents.
May 11, 2021	WHO[e] declares the Delta variant a variant of concern.
May 24, 2021	Heart problems are investigated in vaccinated Teens.
June 1, 2021	Employers can require a COVID-19 vaccine.
June 3, 2021	The Biden administration announces a National Month of Action with the goal of immunizing at least 70% Americans by July 4.

XSL•FO
RenderX

Dates	Cues to action
June 10, 2021	President Biden is set to announce the purchase of 500 million doses of Pfizer's COVID-19 vaccine to be donated to the rest of the world.
June 25, 2021	*Associated Press* reports: Almost all COVID-19 deaths recorded in the United States are among those who are not vaccinated.
Time frame G	
July 9, 2021	Pfizer says it will pursue booster shots.
July 12, 2021	The FDA warns that the COVID-19 J&J vaccine can lead to increased risk of Guillain-Barré syndrome.
July 13, 2021	HHS officials say fully vaccinated individuals do not need a COVID-19 vaccine booster shot.
July 21, 2021	New research published in the *New England Journal of Medicine* shows the Pfizer vaccine is not as effective against the Delta variant.
July 29, 2021	President Biden calls for federal worker vaccines.
August 3, 2021	Around 70% of Americans are vaccinated.
Time frame H	
August 12, 2021	The CDC recommends pregnant people get vaccinated.
August 13, 2021	A booster shot is endorsed for those immunocompromised.
August 18, 2021	US health officials announce a plan for booster shots for the general public.
August 23, 2021	The Pfizer/BioNTech vaccine gains full FDA approval.
August 24, 2021	Several large organizations issue vaccine mandates for workers.
September 9, 2021	President Biden announces all companies with over 100 employees must mandate COVID-19 vaccinations.
September 10, 2021	Los Angeles schools mandate vaccines.
September 17, 2021	The FDA committee votes against boosters for the general public.
September 20, 2021	Pfizer says its vaccine is safe and effective in children.
September 22, 2021	The FDA authorizes the Pfizer booster.
Time frame I	
October 15, 2021	The United States opens to vaccinated travelers.
October 25, 2021	The Moderna dose is effective in children.
October 29, 2021	The FDA authorizes the Pfizer vaccine for children.
November 1, 2021	The FDA investigates Moderna vaccine adverse effects.
November 5, 2021	At least 2 groups file lawsuits against implementation of President Biden's vaccine mandate.
November 13, 2021	The appeals court affirms hold on the employer vaccine mandate.
November 19, 2021	The FDA approves a vaccine booster for all adults.

[a]FDA: Food and Drug Administration.

[b]HHS: Department of Health and Human Services.

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

[d]J&J: Johnson & Johnson.

[e]WHO: World Health Organization.

Mask-Related Tweet Results

Figure 4 displays HBM label percentages. The U.S. confirmed cases by week were added to the secondary y-axis and letters within vertical lines represent a selection of events that could potentially help interpret the fluctuations in the figure from the U.S. government and related sources from the American Journal of Managed Care. Case count data came from the New York Times COVID-19 data GitHub repository [34].

Perceived benefits and perceived barriers frequently corresponded to cues for action, such as official government or public health guidance or policies (Figure 4). Perceived benefits for mask wearing trended with the rise and fall of the US COVID-19 case counts for the majority of the 2020 calendar year until mid-November and early January when vaccines were available. This observation was supported by a statistically significant (P<.001) Spearman correlation of 0.686. The lowest levels of perceived benefits occurred between February and mid-March, where barriers to mask wearing overshadowed benefits. Benefits first emerged above barriers after the

XSL•FO
RenderX

pandemic was declared by WHO and when WHO recommended mask wearing for health care workers and sick individuals, but mask wearing was not sustained. Mask wearing emerged again more steadily when WHO declared COVID-19 was airborne-transmissible and when the CDC began recommending masks in early April. Benefits emerged most quickly after WHO changed its stance on mask wearing. The highest levels of perceived barriers were from February to mid-March. According to the top 5 retweets categorized for perceived barriers from February 24 to March 1, the "stop buying masks, save it for the health care workers" idea was the content of 4 of the top 5 most retweeted retweets. The top retweets correspond well with the initial stance of WHO regarding mask wearing for the general public [21] and the elevated perceived barriers to the action shown in Figure 4. Although WHO did not change its official stance on public mask wearing until June 5, the CDC first recommended face masks on April 3, and the first state (New Jersey) began a mask mandate on April 10. Perceived benefits gradually trended upward and peaked to correspond with each of the peaks in US case counts in late July and again in October and November and exceeded perceived barriers during much of that time. Perceived benefits trended with case counts until November around the US presidential election and in anticipation of the US Thanksgiving holiday. The lowest perceived barriers and the most significant divergence from benefits occurred as the COVID-19 case counts mounted toward their highest level in late November 2020. Table 4 lists a few of the major events and cues to action during the time frame in Figure 4.

Perceived severity and perceived susceptibility possessed a much lower percentage of HBM influence compared to perceived benefits and perceived barriers. The overall rate of perceived severity tended to be slightly higher than that of perceived susceptibility except at the beginning of the pandemic.

Vaccine-Related Tweet Results

Figure 5 displays the HBM label percentages (raw count of HBM labels divided by the total raw count of COVID-19–related tweets filtered for vaccines). US confirmed cases by week were added to the secondary y axis, and letters within vertical lines represent a selection of events that could potentially help interpret the fluctuations in the figures from the US government and related sources from the *American Journal of Managed Care* [35,36]. In addition, US confirmed case count data were collected from the *New York Times* COVID-19 data GitHub repository [34]. Figure 6 looks at how US vaccine doses administered trended with the HBM label percentages. US vaccine data were collected from the Bloomberg Covid-19 Vaccine Tracker Open Data GitHub repository [37].

Perceived barriers and US case counts had a statistically significant (*P*=.03), positive Spearman correlation of 0.28. This means that as US case counts increased, perceived barriers also increased. Perceived barriers and US vaccination counts had a statistically significant (*P*<.001), negative Spearman correlation of –0.67. This means that there was an inverse relationship between US vaccination counts and perceived barriers. The displayed findings do not reflect magnitude but simply identify the relative volume of tweets for each of the HBM constructs.

Perceived barriers for vaccination consistently remained higher than perceived benefits, but they did not appear to trend inversely as expected. Like the face mask findings, perceived susceptibility and perceived severity accounted for a noticeably lower percentage than perceived benefits and perceived barriers. Unlike the face mask findings, perceived barriers had no crossovers with perceived benefits (Figures 5 and 6). Analysis of US case counts tended to trend with perceived barriers throughout fall 2020 through winter 2021. The percentage of perceived barriers was steadily higher than US case counts in 2021. The initial peak of perceived barriers to vaccination occurred at the beginning of the availability and distribution of vaccines. Still, it then steadily declined until early March 2021, when the CDC released guidance that safe activities were available for fully vaccinated individuals (March 8) and when President Biden pushed for expanded vaccination eligibility for all adults aged 18 years and older (March 11). The highest level of perceived barriers to vaccine acceptance peaked in late summer and early fall 2021 around the time that schools resumed operations. This rapid incline seemed to correspond with the greatest decline in US vaccination counts (Figure 6). The second greatest rise in barriers occurred early during the vaccination process from late December 2020 to early January 2021. Perceived benefits did not rise until late October to early November 2020, when vaccine efficacy was established, and then trended and held relatively steady in 2021.

Vaccination distribution rates continued to climb steadily from mid-March until mid-April 2021, but perceived barriers also continued to climb up and down. Overall vaccination counts dropped steadily in mid-April, near the time when the CDC and the Food and Drug Administration (FDA) recommended pausing the Johnson & Johnson (J&J) vaccine (April 13) and continued despite those restrictions being lifted 10 days later (April 23). The vaccination drop continued to plummet even though the FDA formally expanded the availability of vaccines for adolescents aged 12-15 years in early May (May 4). At that time, barriers again climbed upward through June 1. An inverse correlation between vaccine doses administered and perceived barriers to obtaining vaccination was observed. That is, as doses increased, perceived barriers decreased. During this inverse correlation period, WHO declared the COVID-19 Delta variant a concern (May 11), heart problems were noted in vaccinated teenagers (May 24), the government allowed employers to require the COVID-19 vaccine (June 1), and research in the *New England Journal of Medicine* indicated the Pfizer vaccine was not as effective (July 21). From June through September 2021, the conversation on perceived barriers for vaccines continued to climb, while the case counts also climbed, and doses administered declined. Table 5 lists some of the major events and cues to action during the time frame in Figures 5 and 6.

Discussion

Principal Findings

With COVID-19 also came the emergence of a worldwide infodemic, or the overabundance of information and misinformation about COVID-19. A central informational theme

XSL•FO
RenderX

for COVID-19 in 2020 in public health revolved around controlling its spread using social distancing and face mask wearing and included ramping up vaccines that could be quickly distributed. A continuing theme in 2021 for COVID-19 involved efforts to end the pandemic through vaccine distribution and continuing some level of preventative measures (eg, stay-at-home orders, mask mandates, and capacity limits in certain businesses or other settings). However, social media played an influential role in creating an infodemic [21]. These social media influences are closely related to the political and personal reactions to COVID-19 from early in the US pandemic and appear to have continued throughout [38]. Although some social media research during the infodemic pointed to the influence of low-credibility content regarding COVID-19 [3], less has been done to understand the influence of higher-credibility content largely from prominent sources on COVID-19 health beliefs. This study used social media to investigate the COVID-19–related Twitter posts in the United States to understand health beliefs related to mask wearing and vaccinations in the midst of an infodemic. We also explored external cues to action from prominent pandemic declarations (eg, WHO and the CDC) regarding mask wearing and vaccines and notable examples of displaying preventive behaviors (eg, presidential mask wearing) and their possible association with health beliefs, as explained by the HBM constructs. Understanding the influence of social media information on COVID-19 health beliefs and preventative behaviors is important for information management during emergencies.

First, health beliefs relating to the perceived benefits of and perceived barriers to mask wearing appeared to be most influenced by external cues to action. Furthermore, they tended to inversely weakly mirror each other over time. That is, as more perceived benefits of mask wearing were discussed, fewer perceived barriers were discussed. These findings are consistent with other HBM research, particularly among student pharmacists [39]. Although previous research on COVID-19 mask-wearing beliefs pointed to the influence of perceived severity [40], in this study, the perceived susceptibility to and perceived severity of COVID-19 were much less prominent. This remained fairly consistent over time despite numerous high-profile cues to actions, such as major announcements, COVID-19 case counts, or COVID-19 deaths. This contrasts with early assumptions about what health beliefs influence face mask wearing in context with the HBM [41]. Regardless, the perceptions of disease susceptibility and severity seemed muted in this study, presumably because the worldwide pandemic may have appeared obviously relevant for most people in most places. Another possible explanation is that the immediacy of an ongoing and rapidly changing pandemic tended not to discuss the severity and seriousness of getting the disease but seemed to emphasize the importance of behaviors (ie, face masks and vaccines) and taking action.

Perceived benefits of and perceived barriers to mask wearing varied over time with sometimes dramatic swings that appeared to align with specific major cues to action. The most important or consistent cue to action was the rise and fall of total US case counts from March through December 2020 with perceived benefits. This observation is reinforced by the fact that the perceived benefits of mask wearing occurred before formal mask wearing was recommended by WHO (June 5). The benefits of mask wearing continued despite the CDC and other sources asking that masks not be hoarded so they could be available for health care workers and other individuals who were sick.

The second-most important cues to action included how the timing of certain messaging cues may have prompted temporary perceptions of the benefits of mask wearing. For example, perceived benefits peaked around the same time that WHO recommended masks in June 2020, WHO asked scientists to revise guidelines to acknowledge airborne transmission in mid-July, and President Trump was seen wearing a mask for the first time in mid-July [35]. By contrast, other kinds of messaging cues to action may have prompted a rise in perceived barriers to mask wearing when contradictory messages from WHO and the CDC urged people to not use masks early in the pandemic or messages were released about vaccines' immediate availability at the end of 2020. Thus, the greatest cues to action for mask wearing identified from these data were primarily the rise in US case counts followed by episodic messaging that promoted the health belief that mask wearing was important. Such observed cues to action do not reflect a cause-and-effect relationship, but findings from this study are verified by what others have recently suggested [42].

Second, health belief findings regarding vaccinations from this study demonstrated several important implications for cues to action and the 4 HBM constructs. Similar to mask wearing, health beliefs relating to perceived benefits of and perceived barriers to vaccination appeared to be influenced by specific cues to action and often inversely mirrored each other over time, while not being as pronounced as mask-wearing perceptions. Twitter conversations regarding perceived vaccine barriers generally increased up until January 1 before declining and flattening. These trends tended to precede and mirror US confirmed COVID-19 case counts and death counts, suggesting the potential influence of several high-profile announcements (cues to action) during those peak times. Although previous HBM survey research in Malaysia indicated perceived barriers, such as vaccine efficacy, safety, affordability, and side effects [19], findings from this study appear to be linked more with vaccine access. Perceived barriers among COVID-19 vaccine Twitter conversations began to increase sharply when emergency use began (December 11), only peaking after the White House announced the use of reserve supplies (January 4) and the government committed funds for vaccine distribution (January 6).

Unique to vaccination Twitter conversations compared to this study's mask-wearing findings, discussions regarding perceived benefits were always below perceived barrier trends, especially later in the pandemic. Perceived barriers reflected difficulties, challenges, conspiracies, negative effects, dangers, and perceived ineffectiveness associated with vaccinations. One potential explanation is that vaccine-related conversations became more complex because of these factors as the pandemic emerged. This complexity likely involved factors outside our study, such as political distrust, contradictory messages, and others. The Twitter conversations seemed to focus on the

XSL•FO
RenderX

likelihood of the disease relative to face masks and the benefits of the behavior, but the vaccination barriers complicated the benefits of vaccination.

The number of vaccines administered peaked as perceived seriousness and perceived susceptibility discussions peaked (April 4). These findings are consistent with the HBM, which suggests behavior will change when the threat of disease increases one's perception of seriousness and susceptibility increases. Interestingly, although the volume of discussion for perceived benefits and perceived barriers related to vaccination nearly crossed on March 4, benefit discussions remained higher than barrier discussions. Altogether, however, it appears that various cues to action beginning March 1 influenced health beliefs and ultimately vaccine behavior. This helps demonstrate that understanding the individual beliefs regarding COVID-19–preventive behaviors in response to various cues to action from these high-credibility sources is crucial toward helping manage an infodemic. Moving forward, public health officials may better manage information and positively influence health beliefs and vaccine behaviors by using traditional risk communication approaches [43]. For example, focusing on building trust through announcing early findings, being transparent with what is known and unknown, respecting public concerns, and planning in advance may serve as important lessons learned from the COVID-19 pandemic, particularly in response to vaccinations [44].

Limitations

There are several limitations of this study. First, as noted in the study purpose justification, the results of the study are primarily exploratory in nature and should be interpreted in this context. Additional research is needed to further confirm the influence of high-profile cues to action and HBM constructs, ideally with some form of case-control or experimental design. Second, although publicly available Twitter data were used in this analysis, these data were collected from a subset of tweets returned from the Twitter Application Programming Interface (API). Because there was no way of knowing the size of the subset in relation to the whole of tweets, or the sampling method used by Twitter to create the subset, there was a potential for bias. Third, because the COVID-19–related tweets were filtered for vaccines or masks, the perceived susceptibility and perceived severity constructs may have been minimized. It may also be described by the HBM construct definitions in Tables 1 and 2, which emphasize COVID-19 itself. Regardless, filtering tweets for mask wearing and vaccines may have inadvertently diminished the susceptibility and seriousness of COVID-19 conversations with the HBM because most of these COVID-19 conversations tended to focus on the perceived benefits and perceived barriers. Fourth, future studies could use topic modeling techniques suitable for short-text documents, such as tweets, to better understand the topics surrounding masks, vaccines, and COVID-19 related to the HBM. In addition, future research might explore specifically which social media messages, communications channels, and voices are most influential on COVID-19–related prevention health beliefs. Fifth, although there are several explanations of the findings in this study, we did not attempt to establish causal relationships. Future longitudinal studies can explore questions of causation.

Conclusion

Throughout the pandemic, experts have provided aggressive recommendations for COVID-19 prevention [21,22,24,45]. During both the COVID-19 pandemic and the infodemic, this study used a machine learning approach to explore health beliefs related to mask-wearing and vaccination recommendations and the possible influence of high-profile cues to action. Findings suggest that although certain health beliefs on Twitter appear to respond to various high-profile cues to action, health belief trends differ between mask wearing and vaccination.

Acknowledgments

The authors would like to acknowledge the work of Kiersten DeCook, Sarah Callaway, McKaylee Smith, Zane Maguet, Si Yang Ke, Emily Freeman, and Anna Leman in hand-labeling the 2 training sets of 3500 tweets.

Authors' Contributions

All authors discussed and collaborated in the conception of the ideas presented in the paper. ESN-T gathered and filtered the data. QS developed the classification models and produced the resulting classification data. SYK developed the data-counting code and produced the graphs. MB and CLH developed the theory and analysis of the results. All authors collaborated on the writing and presentation of the research and contributed to the final manuscript.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Jernigan DB, CDC COVID-19 Response Team. Update: public health response to the coronavirus disease 2019 outbreak - United States, February 24, 2020. MMWR Morb Mortal Wkly Rep 2020 Feb 28;69(8):216-219. [doi: 10.15585/mmwr.mm6908e1] [Medline: 32106216]
2. World Health Organization. Infodemic. URL: https://www.who.int/health-topics/infodemic%23tab=tab_1 [accessed 2021-08-02]

XSL•FO
RenderX

3. Yang K, Pierri F, Hui P, Axelrod D, Torres-Lugo C, Bryden J, et al. The COVID-19 infodemic: Twitter versus Facebook. Big Data Soc 2021 May 05;8(1):205395172110138. [doi: 10.1177/20539517211013861]
4. 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]
5. 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]
6. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb;395(10225):676. [doi: 10.1016/s0140-6736(20)30461-x]
7. 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]
8. Fung IC, Tse ZTH, Fu K. The use of social media in public health surveillance. Western Pac Surveill Response J 2015;6(2):3-6 [FREE Full text] [doi: 10.5365/WPSAR.2015.6.1.019] [Medline: 26306208]
9. 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]
10. 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]
11. Guntuku SC, Buttenheim AM, Sherman G, Merchant RM. Twitter discourse reveals geographical and temporal variation in concerns about COVID-19 vaccines in the United States. Vaccine 2021 Jul 05;39(30):4034-4038 [FREE Full text] [doi: 10.1016/j.vaccine.2021.06.014] [Medline: 34140171]
12. He L, He C, Reynolds T, Bai Q, Huang Y, Li C, et al. Why do people oppose mask wearing? A comprehensive analysis of U.S. tweets during the COVID-19 pandemic. J Am Med Inform Assoc 2021 Jul 14;28(7):1564-1573 [FREE Full text] [doi: 10.1093/jamia/ocab047] [Medline: 33690794]
13. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public discourse against masks in the COVID-19 era: infodemiology study of Twitter data. JMIR Public Health Surveill 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: 10.2196/26780] [Medline: 33720841]
14. 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]
15. Rosenstock IM. Why people use health services. Milbank Mem Fund Q 1966 Jul;44(3):94. [doi: 10.2307/3348967]
16. 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]
17. Guidry JPD, Carlyle KE, LaRose JG, Perrin P, Messner M, Ryan M. Using the Health Belief Model to analyze Instagram posts about Zika for public health communications. Emerg Infect Dis 2019 Jan;25(1):179-180 [FREE Full text] [doi: 10.3201/eid2501.180824] [Medline: 30561302]
18. Du J, Cunningham RM, Xiang Y, Li F, Jia Y, Boom JA, et al. Leveraging deep learning to understand health beliefs about the human papillomavirus vaccine from social media. NPJ Digit Med 2019 Apr 15;2(1):27 [FREE Full text] [doi: 10.1038/s41746-019-0102-4] [Medline: 31304374]
19. Wang H, Li Y, Hutch M, Naidech A, Luo Y. Using tweets to understand how COVID-19-related health beliefs are affected in the age of social media: Twitter data analysis study. J Med Internet Res 2021 Feb 22;23(2):e26302 [FREE Full text] [doi: 10.2196/26302] [Medline: 33529155]
20. Al-Metwali BZ, Al-Jumaili AA, Al-Alag ZA, Sorofman B. Exploring the acceptance of COVID-19 vaccine among healthcare workers and general population using health belief model. J Eval Clin Pract 2021 Oct 07;27(5):1112-1122 [FREE Full text] [doi: 10.1111/jep.13581] [Medline: 33960582]
21. World Health Organization. Advice on the Use of Masks in the Community, During Home Care and in Health Care Settings in the Context of the Novel Coronavirus (2019-nCoV) Outbreak: Interim Guidance, 29 January 2020. Technical Documents. 2019. URL: https://apps.who.int/iris/bitstream/handle/10665/330987/WHO-nCov-IPC_Masks-2020.1-eng.pdf?sequence=1&isAllowed=y [accessed 2021-08-02]
22. World Health Organization. Advice on the Use of Masks in the Context of COVID-19: Interim Guidance, 5 June 2020. Technical Documents. URL: https://apps.who.int/iris/bitstream/handle/10665/332293/WHO-2019-nCov-IPC_Masks-2020.4-eng.pdf?sequence=1&isAllowed=y [accessed 2021-08-02]
23. Berg M, Lin L. Prevalence and predictors of early COVID-19 behavioral intentions in the United States. Transl Behav Med 2020 Oct 08;10(4):843-849 [FREE Full text] [doi: 10.1093/tbm/ibaa085] [Medline: 32893867]
24. Sohrabi C, Alsafi Z, O'Neill N, Khan M, Kerwan A, Al-Jabir A, et al. World Health Organization declares global emergency: a review of the 2019 novel coronavirus (COVID-19). Int J Surg 2020 Apr 26;76:71-76 [FREE Full text] [doi: 10.1016/j.ijsu.2020.02.034] [Medline: 32112977]
25. Kaiser Family Foundation. The US Government and the World Health Organization. URL: https://www.kff.org/coronavirus-covid-19/fact-sheet/the-u-s-government-and-the-world-health-organization/ [accessed 2021-08-02]

XSL•FO
RenderX

26. 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]
27. Chen E. Covid-19-Tweet IDs. URL: https://github.com/echen102/COVID-19-TweetIDs [accessed 2021-12-11]
28. Vaswani A, Shazeer N, Parmar N. Attention is all you need. 2017 Presented at: NIPS'17: 31st International Conference on Neural Information Processing Systems; December 4-9, 2017; Long Beach, CA URL: https://dl.acm.org/doi/10.5555/3295222.3295349
29. Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv 2018:4805. [doi: 10.18653/v1/N19-1423]
30. Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov RR, Le QV. XLNet: generalized autoregressive pretraining for language understanding. 2019 Presented at: NIPS'19: 33rd International Conference on Neural Information Processing Systems; December 8-14, 2019; Vancouver, Canada.
31. Cho K, van Merrienboer M, Bahdanau D, Bengio Y. On the properties of neural machine translation: encoder-decoder approaches. arXiv 2014 [FREE Full text] [doi: 10.3115/v1/w14-4012]
32. Gwet KL. Calculating Chance-Corrected Agreement Coefficients (CAC). URL: https://cran.r-project.org/web/packages/irrCAC/vignettes/overview.html [accessed 2022-09-21]
33. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013 Apr 29;13(1):61 [FREE Full text] [doi: 10.1186/1471-2288-13-61] [Medline: 23627889]
34. nytimes / covid-19-data. URL: https://github.com/nytimes/covid-19-data/blob/master/us.csv [accessed 2022-01-26]
35. AJMC Staff. A Timeline of COVID-19 Vaccine Developments in 2020. URL: https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020 [accessed 2021-02-05]
36. AJMC Staff. A Timeline of COVID-19 Vaccine Developments in 2021. URL: https://www.ajmc.com/view/a-timeline-of-covid-19-vaccine-developments-in-2021 [accessed 2022-02-02]
37. BloombergGraphics / covid-vaccine-tracker-data. URL: https://github.com/BloombergGraphics/covid-vaccine-tracker-data/blob/master/data/historical-global-doses-administered.csv [accessed 2022-02-02]
38. Christensen SR, Pilling EB, Eyring JB, Dickerson G, Sloan CD, Magnusson BM. Political and personal reactions to COVID-19 during initial weeks of social distancing in the United States. PLoS One 2020 Sep 24;15(9):e0239693 [FREE Full text] [doi: 10.1371/journal.pone.0239693] [Medline: 32970761]
39. Li M, Lv G, Hsieh S, Shao R, Yuan J. Pandemic worry and preventive health behaviors during the COVID-19 outbreak. Front Med (Lausanne) 2021 Jun 17;8:700072 [FREE Full text] [doi: 10.3389/fmed.2021.700072] [Medline: 34222295]
40. Kwan RYC, Lee PH, Cheung DSK, Lam SC. Face mask wearing behaviors, depressive symptoms, and health beliefs among older people during the COVID-19 pandemic. Front Med (Lausanne) 2021 Feb 5;8:590936 [FREE Full text] [doi: 10.3389/fmed.2021.590936] [Medline: 33614680]
41. Sim S, Moey K, Tan N. The use of facemasks to prevent respiratory infection: a literature review in the context of the Health Belief Model. Singapore Med J 2014 Mar;55(3):160-167 [FREE Full text] [doi: 10.11622/smedj.2014037] [Medline: 24664384]
42. Cherry TL, James AG, Murphy J. The impact of public health messaging and personal experience on the acceptance of mask wearing during the COVID-19 pandemic. J Econ Behav Organ 2021 Jul;187:415-430 [FREE Full text] [doi: 10.1016/j.jebo.2021.04.006] [Medline: 33994606]
43. Abraham T. Risk and outbreak communication: lessons from alternative paradigms. Bull World Health Organ 2009 Aug 1;87(8):604-607 [FREE Full text] [doi: 10.2471/blt.08.058149] [Medline: 19705010]
44. Latkin CA, Dayton L, Miller JR, Yi G, Jaleel A, Nwosu CC, et al. Behavioral and attitudinal correlates of trusted sources of COVID-19 vaccine information in the US. Behav Sci (Basel) 2021 Apr 20;11(4):56 [FREE Full text] [doi: 10.3390/bs11040056] [Medline: 33924118]
45. World Health Organization. COVID-19 Advice for the Public: Getting Vaccinated. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/covid-19-vaccines/advice%20 [accessed 2021-08-02]

Abbreviations

AUROC: area under the receiver operating characteristic
BERT: bidirectional encoder representations from transformers
BiGRU: bidirectional gated recurrent unit
CDC: Centers for Disease Control and Prevention
DistilBERT: distilled version of bidirectional encoder representations from transformers
FDA: Food and Drug Administration
HBM: Health Belief Model
HHS: Department of Health and Human Services
J&J: Johnson & Johnson

XSL•FO
RenderX

WHO: World Health Organization

Edited by T Mackey; submitted 09.03.22; peer-reviewed by SF Tsao, M Memon; comments to author 01.07.22; revised version received 18.08.22; accepted 10.09.22; published 31.10.22.

Please cite as:
Ke SY, Neeley-Tass ES, Barnes M, Hanson CL, Giraud-Carrier C, Snell Q
COVID-19 Health Beliefs Regarding Mask Wearing and Vaccinations on Twitter: Deep Learning Approach
JMIR Infodemiology 2022;2(2):e37861
URL: https://infodemiology.jmir.org/2022/2/e37861
doi:10.2196/37861
PMID:36348979

©Si Yang Ke, E Shannon Neeley-Tass, Michael Barnes, Carl L Hanson, Christophe Giraud-Carrier, Quinn Snell. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 31.10.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Codeveloping and Evaluating a Campaign to Reduce Dementia Misconceptions on Twitter: Machine Learning Study

Sinan Erturk[1,2], BSc, MSc; Georgie Hudson[1,2], BSc; Sonja M Jansli[1,2], BSc, MSc; Daniel Morris[1,2], BSc; Clarissa M Odoi[1,2], BSc, MSc; Emma Wilson[1,2], BSc, MSc; Angela Clayton-Turner[1], BA, GDip; Vanessa Bray[1], BSc, MA; Gill Yourston[1], PG; Andrew Cornwall[1], BSc (Hons); Nicholas Cummins[1], PhD; Til Wykes[1,2], PhD; Sagar Jilka[1,2,3], PhD

[1]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[2]South London and Maudsley NHS Foundation Trust, London, United Kingdom

[3]Warwick Medical School, University of Warwick, Coventry, United Kingdom

Corresponding Author:
Sagar Jilka, PhD
Warwick Medical School
University of Warwick
B0.20 Medical School Building
Gibbet Hill Road
Coventry, CV4 7HL
United Kingdom
Phone: 44 07708715627
Email: sagar.jilka@warwick.ac.uk

Abstract

Background: Dementia misconceptions on Twitter can have detrimental or harmful effects. Machine learning (ML) models codeveloped with carers provide a method to identify these and help in evaluating awareness campaigns.

Objective: This study aimed to develop an ML model to distinguish between misconceptions and neutral tweets and to develop, deploy, and evaluate an awareness campaign to tackle dementia misconceptions.

Methods: Taking 1414 tweets rated by carers from our previous work, we built 4 ML models. Using a 5-fold cross-validation, we evaluated them and performed a further blind validation with carers for the best 2 ML models; from this blind validation, we selected the best model overall. We codeveloped an awareness campaign and collected pre-post campaign tweets (N=4880), classifying them with our model as misconceptions or not. We analyzed dementia tweets from the United Kingdom across the campaign period (N=7124) to investigate how current events influenced misconception prevalence during this time.

Results: A random forest model best identified misconceptions with an accuracy of 82% from blind validation and found that 37% of the UK tweets (N=7124) about dementia across the campaign period were misconceptions. From this, we could track how the prevalence of misconceptions changed in response to top news stories in the United Kingdom. Misconceptions significantly rose around political topics and were highest (22/28, 79% of the dementia tweets) when there was controversy over the UK government allowing to continue hunting during the COVID-19 pandemic. After our campaign, there was no significant change in the prevalence of misconceptions.

Conclusions: Through codevelopment with carers, we developed an accurate ML model to predict misconceptions in dementia tweets. Our awareness campaign was ineffective, but similar campaigns could be enhanced through ML to respond to current events that affect misconceptions in real time.

(JMIR Infodemiology 2022;2(2):e36871) doi:10.2196/36871

KEYWORDS

machine learning; patient and public involvement; codevelopment; misconceptions; stigma; Twitter; social media

XSL•FO
RenderX

Introduction

Overview

Negative language and opinions concerning dementia are common on social media platforms [1]. On Twitter, dementia is ridiculed, and stigma surrounding the condition is perpetuated [2]. Stigma toward dementia has been attributed to many different factors, including the loss of independence and functioning the condition can cause [3]. An important factor identified from a systematic review is the myths and misconceptions surrounding dementia; this lack of education around the truth of this condition leads to people forming negative, incorrect beliefs about the condition that are represented by stigma [4]. These misconceptions have also been said to directly influence how communities and families respond to people with dementia [5]. In our past work investigating negative language around dementia on Twitter alongside carers, we found that these misconceptions underlie the negative comments found on the platform and concluded that addressing these directly would help promote awareness and education for dementia rather than simply correcting negative language represented by stigma [6]. With a daily average of 500 million tweets [7], identifying misconceptions quickly can only be carried out with machine learning (ML) [8]. Studies have found that ML models are as accurate as humans in recognizing stigma toward bipolar disorder and general mental health issues in social media posts [9,10] and stigma toward dementia in tweets [2]. However, although Oscar et al [2] attempted to sort tweets into a wide range of categories, we focus solely on the identification of misconceptions, with the full involvement of care partners for those living with dementia to maximize involvement of the community [11,12] and to minimize bias in supervised ML [13]. As supervised models are trained on a given set of classifications, we argue that these classifications should be curated with the community.

Only identifying misconceptions will not change public perceptions; however, by identifying people who are posting them on Twitter, these people can be targeted by an educational awareness campaign. To our knowledge, this form of campaign has not yet been undertaken on a social media platform. Similar campaigns around dementia have either focused on awareness around risks for developing dementia [14] or have been delivered through other mediums [15]. An awareness campaign on social media for dementia would therefore be more comparable with campaigns that reduce stigma and misconceptions of mental health problems such as the *Time to Change* antistigma campaigns [16]. These have been effective at increasing positive attitudes toward mental health and reducing discrimination by 11.5% [17,18]. This has in turn helped those with mental health difficulties feel more able to approach mental health services [19]. Although dementia is a neurodegenerative disease, rather than a mental health condition, these findings suggest that an awareness campaign run on Twitter could be effective in reducing misconceptions around dementia. However, these campaigns are run over long periods, and other factors, such as global events, may also mediate or affect this change (eg, Budenz et al [20] showed that mental health stigma significantly increases after a shooting). We therefore sought to track global events in relation to our campaign to observe how our campaign works in a real-world environment where news stories can shape discussions on social media.

We aimed to codevelop a supervised ML model that can detect dementia misconceptions on Twitter with dementia care partners central to the analytical pipeline and to co-design and then deploy an awareness campaign on Twitter to address these misconceptions. Furthermore, we aimed to use the ML model to evaluate the effectiveness of our campaign in reducing misconceptions and track global events that affected misconceptions during the campaign period.

Background

This study is built upon our previously published work that qualitatively examined conversations about dementia on Twitter, working with carers of people living with dementia [6]. We held 3 focus groups with them, across which, they defined search terms for finding both negative and neutral tweets about dementia and developed a framework of 6 categories to classify tweets about dementia with 3 misconception categories and 3 neutral categories. A set of 1500 tweets was rated by care partners into these categories, 6 of them, each categorizing 250. Our previous study [6] covered a thematic analysis of the tweets rated as misconception categories by our group of care partners. For this study, we carried forward the obtained search terms to collect further tweets about dementia and used the set of 1500 categorized tweets by carers to develop an ML model based on their choice of categorization.

Methods

Design

This was a mixed methods study using participatory methods across 2 stages.

Stage 1 involved developing an ML model to distinguish between misconceptions and neutral tweets; stage 2 involved developing, deploying, and evaluating an awareness campaign to tackle dementia misconceptions.

We sought carer opinions across both stages to ensure our methodology was grounded in their perspective [11,12].

Ethics Approval

The study was granted ethics approval from the King's College London Psychiatry, Nursing, and Midwifery Research Ethics Committee (reference HR-19/20-14,565).

Participants

Participants (care partners for people living with dementia) were recruited if they had unpaid experience of caring for someone close to them with a diagnosis of dementia from the Maudsley Biomedical Research Centre's dementia research advisory group (MALADY) [21] and Join Dementia Research, a United Kingdom–wide web-based platform hosted by the National Institute of Health Research. Participants were excluded before data collection if they were unable to give consent or were aged <18 years. Participants were asked for their demographic

XSL•FO
RenderX

information so we could provide characteristics of the carers codeveloping our model and campaign (Table 1).

Table 1. Participant characteristics (N=7).

Characteristics	Participants
Sex (female), n (%)	5 (71)
Age (years), mean (SD)[a]	63.33 (11.79)
Ethnicity, n (%)	
White British	6 (86)
Black or Black British	1 (14)
Employment status, n (%)[a]	
Employed (part time)	1 (14)
Self-employed	1 (14)
Retired	3 (43)
Employment and Support Allowance	1 (14)
Years being a carer, mean (SD)[a]	8.83 (6.59)

[a]n=1; missing data.

Tweet Extraction (Data Collection)

In our previous work, we held a focus group with carers to generate a list of dementia-related keywords, both negative and neutral, for tweet extraction [6]. For this focus group, participants first discussed their experiences with dementia being mentioned on Twitter in a 45-minute discourse facilitated by a research assistant. Afterward, participants were each given an iPad that they used to browse Twitter and were told to input search terms they thought might bring up either negative or neutral tweets about dementia. They then examined the tweets that came up from their search and noted how relevant they felt each search term they used was. Participants could freely discuss this task with each other as they completed it. They collectively agreed upon the most useful keywords that brought up both negative and neutral tweets about dementia. These were used to extract tweets using Twitter's streaming application programming interface [22] with "Tweepy" [23] over 2 extraction phases. The final keywords are indicated in Table S1 in Multimedia Appendix 1 [2,24-29].

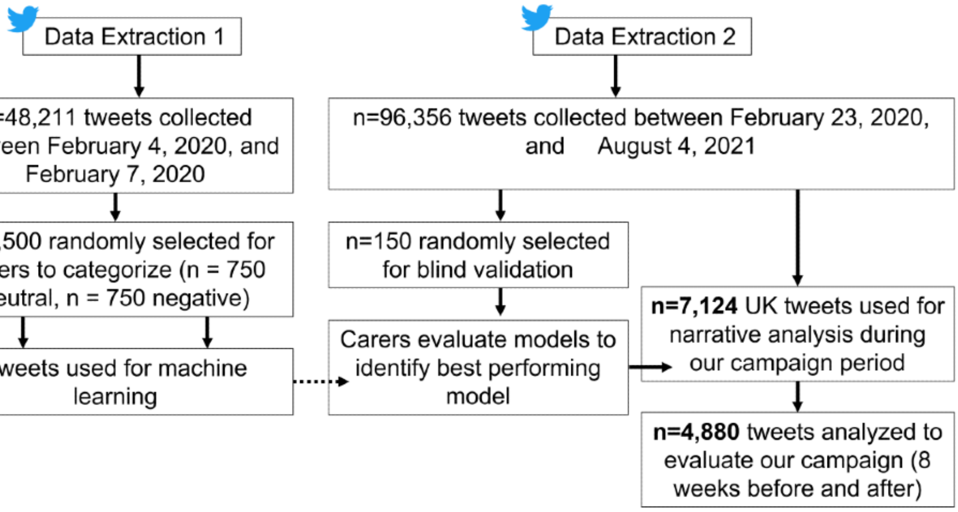
The first extraction was performed in our previous study and provided tweets for our carers to rate and were extracted over a 3-day period (February 4, 2020, to February 7, 2020) to ensure that the tweets were not overly affected by a particular daily event. We collected 48,211 tweets relating to dementia: 35,704 (74.06%) using neutral terms and 12,507 (25.94%) using negative terms. A random sample of 750 tweets with negative keywords and 750 with neutral keywords was extracted for the development of our ML models (N=1500).

The second extraction consisted of 96,356 tweets using the same keywords. We used subsets of this sample to (1) validate our best performing models, (2) explore differences in features for our best performing model, (3) evaluate the effectiveness of our campaign, and (4) understand how global events affected misconceptions during our campaign. We used a much larger date range for the second extraction (February 23, 2020, to April 8, 2021) to ensure that enough tweets from the United Kingdom were extracted, particularly for data collection before (February 23, 2020, to December 2, 2020) and after (January 28, 2021, to April 8, 2021) our campaign. Figure 1 provides an overview of our data sets and procedure.

All tweets collected were original, not retweets. In line with community principles on ethical data practices guidance, all tweets viewed by participants were anonymized to avoid identifying specific individuals [30]. Anonymization included manually removing screen names; specific individuals being mentioned in tweets were censored, unless they were a particular public figure (eg, Donald Trump). It was decided that a mentioned individual was a "public figure" through looking up their username, and if their account was "verified" (representing the account being both genuine and notable), they were deemed to be a public figure.

XSL•FO
RenderX

Figure 1. Tweet extraction diagram.



Procedure

Stage 1: ML

Participatory Involvement

Our model developments had oversight from 7 care partners who attended focus groups and categorized 1500 tweets into neutral and misconception themes. These categories were taken from data used in our previous study [6], where participants built and refined a framework of categories across 2 focus groups, which they then used to rate a set of 250 tweets that they were given (6/7, 86% of the original carers did this process). Half of the categories (3/6, 50%) were agreed upon as different forms of misconceptions and the other half (3/6, 50%) were agreed as different forms of neutral tweets (further detail reported in the study by Hudson and Jansli [6]). We therefore could take each half of the categories and use this set of data as tweets rated as misconceptions, neutral, or neither for the purpose of this study. The number of tweets rated by each participant was collectively decided by the carers as a sample size they could comfortably categorize manually. This size was also manageable from the standpoint of fact-checking tweets, which we left up to the carers' discretion. Carers identified "features" (words or characteristics of tweets) that indicated whether a tweet was a misconception or a neutral tweet. They were told about the nature of features in ML, as we did this to ensure that their involvement at this stage was informed, and they fully understood their contribution to the ML model through this task. They were shown a set of tweets rated by another participant and asked to identify any features that they felt indicated whether a tweet was a misconception and how reliable the feature was. The features considered the most reliable were then taken forward by us to be used in the ML model (Table S2 in Multimedia Appendix 1). Carers also evaluated and selected the best model through a blind validation exercise. Finally, they emphasized accuracy and the number of false negatives to be the key parameters for comparing the performance of different models. Through this, we ensured that our models could be held up to the carers' standards and would therefore be developed according to what they felt was most important.

Importantly, the carers are included as authors on this paper and, as such, have read through and have been able to make comments throughout the writing of this manuscript. Through this, we ensure that we have successfully codeveloped our ML model, awareness campaign, and the paper itself, with the carers.

Preprocessing of Tweets

In accordance with the literature, tweets were preprocessed, which involved them being lemmatized first [31]. This ensured the words in the tweets were in their stem form (eg, "depression," "depressed," and "depressing," would all be converted into "depress"); this removed typos and focused on the meaning of words. We then removed "stop words" to reduce noise as done in previous work [31] and tagged each tweet with the appearance of carer-identified features and extracted 10 additional features based on the literature. This included sentiment (positive or negative tone) and subjectivity (factual to subjective) scores via Python's "TextBlob" library [32]. The other 8 features were tweet descriptive; for example, the length of the tweet [33] and average word length [34] (Multimedia Appendix 1).

Natural language processing methods converted the tweets into their numerical form [35] and we used term frequency–inverse document frequency [24] to vectorize our training set with the default settings in the "Scikit-Learn" library in Python [36]. This generated a data set of features to identify words within the training set that were related to carer-rated misconceptions (Multimedia Appendix 1).

Development of the Supervised ML Models

Given the novelty of this work, we compared the ability of 4 classifiers previously used in health data [37] to test their ability to predict misconceptions. The classifiers used were random forest [38], gradient boosted decision tree [39], support vector machine (SVM) with radial basis function, and SVM with linear kernel [40-42]. Each classifier was created with a 5-fold cross-validation. Hyperparameter optimization was performed for each model, prioritizing accuracy and false negatives while also considering recall and precision. For the random forest and gradient boost, the parameters optimized were the maximum depth and the number of estimators. For the SVMs, the cost

XSL•FO
RenderX

function was optimized. The algorithms for our ML were trained and tested using Scikit-learn (version 0.24; Python Software Foundation) in the programming language Python (version 3.9.0) [43].

ML Blind Validation

As models can perform at similar levels of accuracy during testing [44], we tested levels of agreement between our model and carers by implementing a validation phase for the top 2 models. This serves as further testing for our ML model on a set of tweets independent from those used for training, which is commonly used in validating ML models [45,46]. Our past research has shown that this additional step can help clarify a small difference in accuracies and demonstrate a clearer difference in performance between top-performing models, confirming that this is an important step [31]. We randomly selected 150 tweets from our second sample of 96,356 tweets and split them into 3 batches of 50. A total of 5 care partners then categorized these tweets as misconceptions or neutral tweets. Carers were not shown the model's predicted category (ie, blind validation). When 2 carers agreed on a tweet's category, we took this as the final agreed classification. Tweets without an agreement on category were rated by another carer who decided the final classification. Final carer classifications were compared with our model classifications to investigate agreement.

Stage 2: Campaign to Reduce Misconceptions on Twitter

Participatory Involvement

Carers codeveloped a campaign to combat misconceptions. This was done in two stages:

1. Participants were shown previous dementia awareness campaigns and reflected on what was good and bad about each of them. They then suggested several different focus areas for a campaign, detailing what it should include and how it would address the issue of misconceptions.
2. We combined the suggested focus areas for a campaign, looking for overlaps between suggestions and, from this, developed 3 campaign concepts, focusing on the way language around dementia needs to change, dispelling specific myths or telling the stories of people behind the diagnosis. Each carer assessed the campaign concepts and made suggestions about them including specific quotes to use. We then created infographics for the campaign concept that most carers thought was the best, incorporating selected quotes that were suggested.

Campaign Deployment

We deployed our campaign infographics via our Twitter account "@DementiaReality" for a period of 8 weeks, from December 3, 2020, to January 27, 2021. The campaign targeted UK-based individuals who had previously posted tweets with negative dementia keywords (Table S1 in Multimedia Appendix 1). Our campaign was followed by a poll which asked, "How has a recent dementia tweet made you think differently about dementia?" and Twitter users responded through four choices ("more positively," "more negatively," "no difference," or "didn't see it"). We opted to ask about "any dementia tweet" to ensure that we did not prompt them to remember the original tweet.

Campaign Evaluation

We evaluated our campaign by applying our carer validated ML model through UK-based tweets, posted from 8 weeks before to 8 weeks after our campaign, from our second tweet data set (7124 of 96,356 tweets) to compare the prevalence of dementia misconceptions on Twitter before and after our campaign. Tweets were identified as being from the United Kingdom through the use of geographic longitude and latitude co-ordinates of a reference point (an address specified for all people living in a particular area) associated with the user who posted the tweet.

Data Analysis

Stage 1

Manual Coding of Tweets

We performed independent sample 2-tailed *t* tests and chi-square tests to investigate which features (both carer-identified and literature-defined features) significantly differed between misconceptions and neutral tweets in 1414 tweets. In addition, we also made use of the sklearn library's feature selection with family-wise error in Python to compare this algorithm with our manual tests and confirm their validity. Only statistically significant features were used in our ML model to improve accuracy and reduce noise [47].

Evaluation of the ML Model

We evaluated our models based on accuracy and false negatives and standard ML metrics [48]. Accuracy answers the question, "Overall, how often is the model correct?" and the number of false negatives highlights cases where the model incorrectly classified a tweet as neutral.

ML Blind Validation

To assess the levels of agreement between carers and our 2 best ML models, we performed cross-tabulations, calculated a Cohen κ statistic and a 2×2 chi-square to assess the difference between the models' accuracies by examining the proportion of correct ratings.

Stage 2

We investigated the effect of our campaign on (1) the prevalence of misconceptions among UK Twitter users who discuss dementia and (2) sentiment. We tested whether these outcomes differed in UK-based tweets 8 weeks before and after the campaign, using chi-square tests or 2-tailed *t* tests where appropriate.

Twitter does not allow us to view the users who have been shown our campaign, so it is not possible to directly assess the level of misconceptions of those who had been shown the campaign. To address this, we examined the frequency of misconceptions tweeted by a given user within our second set of extracted tweets. To do this, we classified tweets in this data set using our ML model and examined all identified misconceptions. We separated the tweets by username and identified the average number of days between each user's

XSL•FO
RenderX

misconception tweets. This way, we could demonstrate whether people consistently tweet misconceptions and, therefore, that they would likely to be targeted by the campaign and appear in our evaluation.

We computed the rolling 3-day average of the prevalence of misconceptions and sentiment to investigate changes over our 6-month study period (8 weeks before, 8 weeks during, and 8 weeks after the campaign) and used a time series trend to identify any external influences [49]. To understand how current affairs affected misconceptions, we calculated the mean and SD for sentiment and prevalence at each day across this 6-month period and investigated time points where sentiment and prevalence were –2 to +2 SD from the mean; that is, statistical outliers at 95% probability [50].

Results

Stage 1

Feature Extraction

Carers identified 18 features, 13 (72%) for misconceptions and 5 (28%) for neutral tweets. Carers associated the words "demented" and "senile" as belonging to misconception tweets, as well as tweets where "Donald Trump" and "Nancy Pelosi" are mentioned. Tweets with a URL or those with the words "research" or "memory" were associated with neutral tweets.

Feature Analysis

Carer-Identified Features

Of the 18 features carers identified, 9 (50%) significantly distinguished misconceptions from neutral tweets. These included the mention of Donald Trump (11.97% in misconceptions vs 0% in neutral tweets; χ^2_1=81.6; P<.001) and the occurrence of the word "demented" (46.98% in misconceptions vs 0.16% in neutral tweets; χ^2_1=399.9; P<.001).

Literature-Identified Features

We found significant differences in 8 of the 10 (80%) features with misconceptions being more negative in sentiment (mean −0.04, SD 0.30 vs mean 0.16, SD 0.28; $t_{1,412}$=12.94; P<.001) and shorter (mean characters 139.31, SD 73.12) than neutral tweets (mean 178.97, SD 63.04; $t_{1,412}$=13.71; P<.001).

A full list of features and their significance tests are provided in Tables S3 and S4 in Multimedia Appendix 1. The significance of features from our test run by the "sklearn" algorithm was compared with the manual tests in Table S5 in Multimedia Appendix 1; the difference was minimal, so we proceeded with our manual tests in mind.

Manual Coding

Of the 1500 tweets presented to carers, 86 (5.73%) could not be categorized because the carers felt they could not be sure whether the tweet was a misconception, leaving 1414 for ML: 637 (45.04%) neutral and 777 (54.95%) misconceptions.

ML Model Evaluation

We evaluated our ML models based on various parameters (Table 2). The SVM with a linear kernel and the random forest performed equally well in terms of accuracy (96% each), but the random forest had 7 false negatives, which was slightly less than the SVM with a linear kernel which had 10. Hyperparameter optimization led to our SVM with linear kernel having a cost function of 0.1 and our random forest having a maximum depth of 25 and 500 estimators.

Table 2. Machine learning model comparison.

Parameter	RF[a]	GB[b]	SVM[c]: RBF[d]	SVM linear
Accuracy[e]	0.96	0.95	0.96	0.96
Misclassification rate	0.04	0.05	0.04	0.04
Sensitivity	0.96	0.94	0.93	0.94
Specificity	0.97	0.96	0.99	0.99
False positive rate	0.03	0.04	0.01	0.01
Precision	0.97	0.97	0.99	0.99
False negative rate	0.04	0.06	0.07	0.06
False negatives[e]	7	10	11	10
False positives	3	5	2	1
Area under the receiver operating characteristic curve	0.97	0.95	0.96	0.96

[a]RF: random forest.

[b]GB: gradient boost.

[c]SVM: support vector machine.

[d]RBF: radial basis function.

[e]These parameters were the primary ones used for assessing model performance.

XSL•FO
RenderX

ML Blind Validation

Carers considered 72% (108/150) of the tweets to be misconceptions in the validation data set. We then applied our top two models (SVM with a linear kernel and random forest) to these 150 tweets to select the best performing model:

- SVM with a linear kernel and carers: there was moderate agreement (Cohen κ= 0.538, 95% CI 0.403-0.673; $P<.001$) with agreement on 79% of ratings; there were 26 false negatives. The SVM predicted that 58.6% (88/150) of the tweets were misconceptions.
- Random forest and carers: there was a moderate agreement (Cohen κ=0.581, 95% CI 0.442-0.720; $P<.001$), with agreement on 82% of the ratings and 18 false negatives. The random forest predicted that 72% (108/150) of the tweets were misconceptions.

The random forest was significantly more accurate than the SVM with a linear kernel (n=150; $\chi^2_1=79.9$; $P<.001$).

Stage 2

Campaign Deployment

Our campaign addressed common dementia misconceptions and outlined facts (Figure 2). The graphics were implemented by a graphic designer with quotes suggested by the carers and had a link to more information about dementia on the Alzheimer's Research UK website.

These campaign posters were delivered to 239,360 UK Twitter users who saw at least one of them, and 2.12% (5071/239,360) of the users responded to our evaluation question. Furthermore, 8.05% (408/5071) of the users reported that the campaign had a positive impact, 5.70% (289/5071) reported a negative impact, 10.89% (552/5071) reported no impact, and most (3822/5071, 75.37%) users did not remember seeing it.

Figure 2. Our campaign posters advertised on Twitter.

Campaign Evaluation

We classified UK tweets spanning 8 weeks before the start of our campaign and 8 weeks after the end of our campaign (a total of 16 weeks, N=4880 tweets; Table 3). A chi-square test of independence between the 8-week periods before and after our campaign found no significant difference in prevalence of misconceptions (N=4880; $\chi^2_1=0.8$; $P=.36$). There was also no statistically significant difference in sentiment before and after the campaign ($t_{4878}=1.219$; $P=.22$).

Table 3. Differences in outcome measures 8 weeks before and after our campaign.

Period	Before the campaign (October 8, 2020, to December 2, 2020)	After the campaign (January 28, 2021, to March 24, 2021)
Total number of tweets (n)	2877	2003
Number of misconceptions, n (%)	1035 (35.97)	746 (37.24)
Sentiment, mean (SD)	0.09 (0.30)	0.08 (0.30)

UK-Based Dementia Tweets

We found the prevalence of misconceptions in our set of 7124 UK tweets to be 37%.

Carer-Identified Features

The word "senile" appeared in 13.6% of the misconceptions compared with 0% in neutral tweets (N=7124; $\chi^2_1=640.5$; $P<.001$). Tweets with the appearance of the word "caregiver" did not significantly differ between misconceptions and neutral tweets (N=7124; $\chi^2_1=3.6$; $P=.06$).

Literature-Defined Features

Sentiment was significantly higher in neutral tweets (mean 0.14, SD 0.29) compared with misconceptions (mean −0.03, SD 0.28; $t_{7,122}=5.72$; $P<.001$) and word count was significantly shorter in misconceptions (mean 21.90, SD 14.50) compared with neutral tweets (mean 33.33, SD 12.85; $t_{7,122}=33.56$; $P<.001$).

A full list of feature significances is provided in Tables S6 and S7 in Multimedia Appendix 1.

XSL•FO
RenderX

Frequency of Misconceptions

Our model identified 45,865 tweets as misconceptions within our second set of 96,356 tweets. Table 4 details how often users usually tweeted about misconceptions.

The vast majority of users (45,011/45,865, 98.14%) only tweeted misconceptions as a one-off event, multiple times within a day, or at most within a month. Most users did not continue to tweet misconceptions long after they had first done so.

Table 4. The frequency of users posting misconceptions.

Frequency	Tweets, n (%)
One-off misconception tweet	39,837 (86.85)
Multiple within 1 day	757 (1.65)
Within 1 month	4417 (9.63)
Over a month	854 (1.86)

How Current Affairs Affected Misconceptions and Sentiment Across Our Campaign Period

We identified dates where prevalence was 2 SDs above or below the mean daily prevalence of misconceptions (mean 38%, SD 11%); that is, <17% or >59%. In total, 8 time points fulfilled these criteria; 7 above and 1 below. We also identified dates where sentiment was 2 SDs away from the mean (0.08, SD 0.06); that is, <−0.04 or >0.20. We identified 9 time points: 3 above and 6 below. These points are indicated on Figure 3.

Misconceptions in UK tweets were high, and sentiment was low on the day former president Donald Trump announced that he would rather leave the United States than admit defeat to President Biden [51] (October 18, 2020), with 72% of the tweets being misconceptions and the average sentiment being −0.07. Misconceptions were also high and sentiment low during COVID-19–related events. These included the controversy over the COVID-19 pandemic restriction exemption for hunting in the United Kingdom [51] (December 26, 2020), with 79% of the tweets containing misconceptions, the average sentiment being −0.11 and when reports of a nurse breaking down over empty supermarket shelves went viral [51] (March 21, 2021), with 64% of the tweets containing misconceptions, the average sentiment being −0.05.

Figure 3. Graph of the prevalence of misconceptions and average sentiment on each day for 8 weeks before, during and after our campaign.

Discussion

Principal Findings

We codeveloped and tested an ML model to automatically classify dementia misconceptions with 96% accuracy and a campaign to dispel dementia myths and educate Twitter users on stigmatizing language. We ensured that carers were at the core of our analyses through participatory methods throughout the study. We also show how misconceptions peak and trough as global events shape the Twitter conversations.

Training a model from carer opinions and involving them throughout the study (also including them as authors) has yielded a unique perspective on misconceptions about dementia and how they impact those affected by dementia. This approach differs from previous ML approaches that only used researcher-defined themes [2]. Many features in our study are well established as stigmatizing words or phrases in the literature, such as calling those with the condition "demented," "senile," or diminishing them as "not being all there" [52,53]. However, we also show that an indicator of misconceptions was weaponizing the disease to insult older public figures, most notably politicians, such as Donald Trump; this is in line with the findings of our previous study [6].

XSL•FO
RenderX

Our work improves on previous modeling that detected ridicule as a form of negativity in dementia tweets at 90% accuracy [2]. This demonstrates the value in using larger data sets for training models (eg, Oscar et al [2] used only 331 tweets), as larger training data sets expose a model to a heterogeneous range of language. We also deployed a validation stage that is not commonly used, as noted in a systematic review by Wongkoblap et al [54], so this extra step has no context of comparison within the literature. Our model performed well and was firmly established in the opinions of carers with 96% accuracy, highlighting the effectiveness of community involvement in the ML pipeline.

The campaign we developed did not yield similar benefits from carer involvement and showed clear signs of not being effective. From just our initial polls, we could see the campaign had not left much of an impact, with the vast majority of people not remembering seeing it. Similar campaigns on social media usually assess general awareness of campaigns, without knowing whether that person has seen it before, and so this lack of awareness is uniquely poor [14]. This may be because of the nature of advertisements on Twitter, which are a natural part of a person's feed and thus can easily be scrolled past. In combination with the fact that our funding only allowed for our campaign to be shown once to most people, our campaign was likely not able to have much impact. As such, our finding of little reduction in the prevalence of misconceptions is not surprising, showing that our campaign was ineffective.

Our finding that levels of misconceptions change in response to news events also shows how external factors should ideally be taken into account when running a campaign. By using ML to categorize large amounts of tweets in a short time, notable changes can be tracked, and the news stories associated can be identified, allowing for real-time responses in the campaign, potentially enhancing its effectiveness.

Strengths and Limitations

This study is built on the firm foundation and involvement of those caring for people living with dementia. The opinions of carers were used to fully develop our ML model and our campaign. This perspective is key to classifying dementia misconceptions, as carers are greatly affected by them, and so can provide a unique perspective in identifying tweets that would be the most harmful. None of our participants had a diagnosis of dementia, and this would be an important perspective to incorporate into future work where appropriate. In addition, ML models such as the ones used in this study benefit from larger training sets; given the number of carers and tweets that could reasonably be rated, it is possible that our sample resulted in overfitting. Future research should incorporate larger samples. It is difficult to fully account for spelling mistakes and their frequency within tweets. Although lemmatization accounts for a great deal of these, some spelling mistakes would make it more difficult for our model to correctly use these words. Furthermore, in future research, different approaches to preprocessing and lemmatization could be used, such as the Python library Bidirectional Encoder Representations from Transformers, which has specific uses appropriate for tweets.

Twitter campaigns must competitively bid for "ad space" to show advertisements to users. This may mean that the target audience only has a campaign advertisement appear approximately once on their feed and may explain why 75% of the users did not remember seeing our campaign. Twitter does not provide the names of those targeted by our campaign, so we could not examine the tweets of specific people. Despite this, by examining the general discourse around dementia from tweets posted by people in the United Kingdom, we could indirectly assess how our campaign affected the prevalence of misconceptions: this indirect assessment of the audience being a usual way of assessing web-based campaign effectiveness [14-16]. As we found that the vast majority of users did not continue to tweet misconceptions, long after they had first done so, our study is limited by its inability to directly assess those who viewed our campaign. However, our method of extraction did not provide an exhaustive list of tweets from each user, and as such, this does not necessarily assess all tweets of every user; it is therefore possible that users did indeed tweet misconceptions over time. In the future, it would be important to consider directly assessing users and ensuring that they tweet misconceptions over a long period. Future work must also ensure a competitive campaign budget so that advertisements are shown to users multiple times, as sheer repetition may then have an effect. It is not possible to distinguish world events from the effect of our campaign through this study. Our examination of news stories suggests that they can have an impact on the use of language related to dementia in discussions on Twitter.

Conclusions

This study showed how accurate ML models can be developed alongside carers of people with dementia, highlighting the effectiveness of codevelopment alongside individuals with relevant personal experience. Unfortunately, our campaign seemed unimpactful and ineffective in practice, but from this, we can see the potential in using ML models to assess campaigns. Such assessment could be done in real time, combined with tracking news stories that affect levels of misconceptions, which could be used to tailor the campaign to relative news stories.

Acknowledgments

The authors would like to thank the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre's MALADY advisory group and dementia theme, Zunera Khan, and Miguel Vasconcelos Da Silva for their support on this work. This research was reviewed by a team with experience of mental health problems and their care partners who have been specially trained to advise on research proposals and documentation through the Feasibility and Acceptability Support Team for Researchers: a free, confidential service in England provided by the NIHR Maudsley Biomedical Research Centre via King's College London and South London and Maudsley National Health Service Foundation Trust.

XSL•FO
RenderX

This work was supported by the NIHR Biomedical Research Centre at South London and Maudsley National Health Service Foundation Trust and King's College London (IS-BRC-1215-20018) and Alzheimer's Research UK's Inspire Fund to the corresponding author.

Data Availability

Data are available upon request from the corresponding author. The scripts used are available on GitHub [55].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details about feature extraction and preprocessing of tweets, as well as detailed results about significant features.
[DOCX File , 69 KB - infodemiology_v2i2e36871_app1.docx]

References

1. Robinson P, Turk D, Jilka S, Cella M. Measuring attitudes towards mental health using social media: investigating stigma and trivialisation. Soc Psychiatry Psychiatr Epidemiol 2018 Aug 1;54(1):51-58 [FREE Full text] [doi: 10.1007/s00127-018-1571-5]
2. Oscar N, Fox PA, Croucher R, Wernick R, Keune J, Hooker K. Machine learning, sentiment analysis, and tweets: an examination of Alzheimer's disease stigma on Twitter. J Gerontol B Psychol Sci Soc Sci 2017 Sep 01;72(5):742-751. [doi: 10.1093/geronb/gbx014] [Medline: 28329835]
3. Cipriani G, Borin G. Understanding dementia in the sociocultural context: a review. Int J Soc Psychiatry 2015 Mar 27;61(2):198-204. [doi: 10.1177/0020764014560357] [Medline: 25431401]
4. Cahill S, Pierce M, Werner P, Darley A, Bobersky A. A systematic review of the public's knowledge and understanding of Alzheimer's disease and dementia. Alzheimer Dis Assoc Disord 2015;29(3):255-275. [doi: 10.1097/WAD.0000000000000102] [Medline: 26207322]
5. Morgan DG, Semchuk KM, Stewart NJ, D'Arcy C. Rural families caring for a relative with dementia: barriers to use of formal services. Social Sci Med 2002 Oct;55(7):1129-1142. [doi: 10.1016/s0277-9536(01)00255-6]
6. Hudson G, Jansli SM, Erturk S, Morris D, Odoi CM, Clayton-Turner A, et al. Investigation of carers' perspectives of dementia misconceptions on Twitter: focus group study. JMIR Aging 2022 Jan 24;5(1):e30388 [FREE Full text] [doi: 10.2196/30388] [Medline: 35072637]
7. Twitter usage statistics. Internet Live Stats. URL: https://www.internetlivestats.com/twitter-statistics/ [accessed 2021-04-28]
8. Li A, Jiao D, Zhu T. Detecting depression stigma on social media: a linguistic analysis. J Affect Disord 2018 May;232:358-362. [doi: 10.1016/j.jad.2018.02.087] [Medline: 29510353]
9. Tokmic F, Hadzikadic M, Cook JR, Tcheremissine OV. Development of a behavioral health stigma measure and application of machine learning for classification. Innov Clin Neurosci 2018 Jun 01;15(5-6):34-42 [FREE Full text] [Medline: 30013818]
10. Budenz A, Klassen A, Purtle J, Yom Tov E, Yudell M, Massey P. Mental illness and bipolar disorder on Twitter: implications for stigma and social support. J Ment Health 2020 Apr;29(2):191-199. [doi: 10.1080/09638237.2019.1677878] [Medline: 31694433]
11. Wykes T. Great expectations for participatory research: what have we achieved in the last ten years? World Psychiatry 2014 Feb 04;13(1):24-27 [FREE Full text] [doi: 10.1002/wps.20086] [Medline: 24497239]
12. Rose D. Participatory research: real or imagined. Soc Psychiatry Psychiatr Epidemiol 2018 Aug 21;53(8):765-771 [FREE Full text] [doi: 10.1007/s00127-018-1549-3] [Medline: 29931442]
13. Machine Learning: the Power and Promise of Computers That Learn by Example. London, United Kingdom: The Royal Society; Apr 2017.
14. Van Asbroeck S, van Boxtel MP, Steyaert J, Köhler S, Heger I, de Vugt M, et al. Increasing knowledge on dementia risk reduction in the general population: results of a public awareness campaign. Prev Med 2021 Jun;147:106522 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106522] [Medline: 33744328]
15. Phillipson L, Hall D, Cridland E, Fleming R, Brennan-Horley C, Guggisberg N, et al. Involvement of people with dementia in raising awareness and changing attitudes in a dementia friendly community pilot project. Dementia (London) 2019 Jan 24;18(7-8):2679-2694. [doi: 10.1177/1471301218754455] [Medline: 29363336]
16. Evans-Lacko S, Corker E, Williams P, Henderson C, Thornicroft G. Effect of the Time to Change anti-stigma campaign on trends in mental-illness-related public stigma among the English population in 2003-13: an analysis of survey data. Lancet Psychiatry 2014 Jul;1(2):121-128 [FREE Full text] [doi: 10.1016/S2215-0366(14)70243-3] [Medline: 26360575]
17. Henderson C, Thornicroft G. Stigma and discrimination in mental illness: time to change. Lancet 2009 Jun;373(9679):1928-1930. [doi: 10.1016/S0140-6736(09)61046-1]

XSL•FO
RenderX

18. Sampogna G, Bakolis I, Evans-Lacko S, Robinson E, Thornicroft G, Henderson C. The impact of social marketing campaigns on reducing mental health stigma: results from the 2009-2014 Time to Change programme. Eur Psychiatry 2017 Feb 23;40:116-122 [FREE Full text] [doi: 10.1016/j.eurpsy.2016.08.008] [Medline: 27997875]
19. Corker E, Hamilton S, Henderson C, Weeks C, Pinfold V, Rose D, et al. Experiences of discrimination among people using mental health services in England 2008-2011. Br J Psychiatry Suppl 2013 Apr 02;55(s55):s58-s63. [doi: 10.1192/bjp.bp.112.112912] [Medline: 23553696]
20. Budenz A, Purtle J, Klassen A, Yom-Tov E, Yudell M, Massey P. The case of a mass shooting and violence-related mental illness stigma on Twitter. Stigma Health 2019 Nov;4(4):411-420. [doi: 10.1037/sah0000155]
21. Support for Researchers. NIHR Maudsley Biomedical Research Centre. URL: https://www.maudsleybrc.nihr.ac.uk/patients-public/support-for-researchers/ [accessed 2021-09-16]
22. Docs. Developer Platform. URL: https://dev.twitter.com/streaming/overview [accessed 2021-02-22]
23. API. Tweepy. URL: https://docs.tweepy.org/en/latest/api.html [accessed 2021-02-22]
24. Salton G, McGill MJ. Introduction to Modem Information Retrieval. New York, United States: McGraw-Hill; 1983.
25. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. Ann Emerg Med 2019 Apr;73(4):334-344. [doi: 10.1016/j.annemergmed.2018.11.036] [Medline: 30661855]
26. Neppalli VK, Medeiros M, Caragea C, Caragea D, Tapia A, Halse S. Retweetability analysis and prediction during Hurricane Sandy. In: Proceedings of the ISCRAM 2016 Conference. 2016 Presented at: The 13th International Conference on Information Systems for Crisis Response and Management; May 22-25, 2016; Rio de Janeiro, Brazil.
27. Gretry A, Davis SW, Horvath C, Belei N. How tweet readability and brand hedonism affect consumer engagement. Advances Consum Res 2017;45:629-633 [FREE Full text]
28. Farías D, Patti V, Rosso P. Irony Detection in Twitter. ACM Trans Internet Technol 2016 Aug 22;16(3):1-24. [doi: 10.1145/2930663]
29. Giving you more characters to express yourself. Twitter. URL: https://blog.twitter.com/official/en_us/topics/product/2017/Giving-you-more-characters-to-express-yourself.html [accessed 2022-11-09]
30. Green L. Community principles on ethical data sharing presented at Bloomberg's data for good exchange. DOMINO. 2018 Feb 20. URL: https://www.dominodatalab.com/blog/community-principles-ethical-data-sharing-presented-bloombergs-data-good-exchange/ [accessed 2022-01-10]
31. Jilka S, Odoi CM, van Bilsen J, Morris D, Erturk S, Cummins N, et al. Identifying schizophrenia stigma on Twitter: a proof of principle model using service user supervised machine learning. Schizophrenia (Heidelb) 2022 Feb 07;8(1):1 [FREE Full text] [doi: 10.1038/s41537-021-00197-6] [Medline: 35132080]
32. Du J, Xu J, Song H, Liu X, Tao C. Optimization on machine learning based approaches for sentiment analysis on HPV vaccines related tweets. J Biomed Semantics 2017 Mar 03;8(1):9 [FREE Full text] [doi: 10.1186/s13326-017-0120-6] [Medline: 28253919]
33. Gulliver A, Griffiths KM, Christensen H, Mackinnon A, Calear AL, Parsons A, et al. Internet-based interventions to promote mental health help-seeking in elite athletes: an exploratory randomized controlled trial. J Med Internet Res 2012 Jun 29;14(3):e69 [FREE Full text] [doi: 10.2196/jmir.1864] [Medline: 22743352]
34. Crawford M, Khoshgoftaar TM, Prusa JD, Richter AN, Al Najada H. Survey of review spam detection using machine learning techniques. J Big Data 2015 Oct 5;2(1):23. [doi: 10.1186/s40537-015-0029-9]
35. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. New York: Cham; 2013.
36. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:2825-2830. [doi: 10.3389/fninf.2014.00014]
37. Anjaria M, Guddeti R. Influence factor based opinion mining of Twitter data using supervised learning. In: Proceedings of the 2014 Sixth International Conference on Communication Systems and Networks (COMSNETS). 2014 Presented at: 2014 Sixth International Conference on Communication Systems and Networks (COMSNETS); Jan 06-10, 2014; Bangalore, India. [doi: 10.1109/comsnets.2014.6734907]
38. Breiman L. Random forests. Mach Learn 2001;45:5-32. [doi: 10.1023/A:1010933404324]
39. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001 Oct 1;29(5):1189-1232. [doi: 10.1214/aos/1013203451]
40. Vapnik V. The Nature of Statistical Learning Theory. New York: Springer; 1995.
41. Vapnik V. Statistical Learning Theory. Hoboken, NJ: Wiley; 1998.
42. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995 Sep;20(3):273-297. [doi: 10.1007/BF00994018]
43. Python homepage. Python. URL: http://www.python.org [accessed 2022-01-13]
44. Wong NC, Lam C, Patterson L, Shayegan B. Use of machine learning to predict early biochemical recurrence after robot-assisted prostatectomy. BJU Int 2019 Jan 05;123(1):51-57. [doi: 10.1111/bju.14477] [Medline: 29969172]
45. Ginart A, Das S, Harris J, Wong R, Yan H, Krauss M, et al. Drugs or dancing? Using real-time machine learning to classify streamed "dabbing" homograph Tweets. In: Proceedings of the 2016 IEEE International Conference on Healthcare Informatics

XSL•FO
RenderX

(ICHI). 2016 Presented at: 2016 IEEE International Conference on Healthcare Informatics (ICHI); Oct 04-07, 2016; Chicago, IL, USA. [doi: 10.1109/ichi.2016.97]
46. Shah T. About train, validation and test sets in machine learning. Towards Data Science. 2017 Dec 6. URL: https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7 [accessed 2022-01-13]
47. Paul S. Python feature selection tutorial: a beginner's guide. DataCamp. 2020 Jan. URL: https://www.datacamp.com/community/tutorials/feature-selection-python [accessed 2021-09-16]
48. Thieme A, Belgrave D, Doherty G. Machine learning in mental health. ACM Trans Comput Human Interact 2020 Oct 31;27(5):1-53. [doi: 10.1145/3398069]
49. Bryhn AC, Dimberg PH. An operational definition of a statistically meaningful trend. PLoS One 2011 Apr 28;6(4):e19241 [FREE Full text] [doi: 10.1371/journal.pone.0019241] [Medline: 21552532]
50. Dunn P. Scientific Research and Methodology: An Introduction to Quantitative Research in Science and Health. Australia: Sippy Downs; 2021.
51. Wincalendar homepage. Wincalendar. URL: https://www.wincalendar.com/Calendar-UK/Today [accessed 2021-08-11]
52. Milne A. The 'D' word: reflections on the relationship between stigma, discrimination and dementia. J Ment Health 2010 Jun 04;19(3):227-233. [doi: 10.3109/09638231003728166] [Medline: 20441486]
53. Swaffer K. Dementia: stigma, language, and dementia-friendly. Dementia (London) 2014 Nov 17;13(6):709-716. [doi: 10.1177/1471301214548143] [Medline: 25326220]
54. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J Med Internet Res 2017 Jun 29;19(6):e228 [FREE Full text] [doi: 10.2196/jmir.7215] [Medline: 28663166]
55. Zenedo. URL: https://zenodo.org/record/6956557#.Y2tx9XZBzIU [accessed 2022-11-09]

Abbreviations

ML: machine learning
NIHR: National Institute for Health Research
SVM: support vector machine

Edited by T Mackey; submitted 28.01.22; peer-reviewed by JD Bacsu, P Giabbanelli, S Nagavally; comments to author 29.04.22; revised version received 23.06.22; accepted 15.08.22; published 22.11.22.

Please cite as:
Erturk S, Hudson G, Jansli SM, Morris D, Odoi CM, Wilson E, Clayton-Turner A, Bray V, Yourston G, Cornwall A, Cummins N, Wykes T, Jilka S
Codeveloping and Evaluating a Campaign to Reduce Dementia Misconceptions on Twitter: Machine Learning Study
JMIR Infodemiology 2022;2(2):e36871
URL: https://infodemiology.jmir.org/2022/2/e36871
doi:10.2196/36871
PMID:37113444

©Sinan Erturk, Georgie Hudson, Sonja M Jansli, Daniel Morris, Clarissa M Odoi, Emma Wilson, Angela Clayton-Turner, Vanessa Bray, Gill Yourston, Andrew Cornwall, Nicholas Cummins, Til Wykes, Sagar Jilka. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 22.11.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Profiles and Symptoms of Patients With Long COVID in France: Data Mining Infodemiology Study Based on Social Media

Amélia Déguilhem[1*], MSc; Joelle Malaab[1*], MSc, MPH; Manissa Talmatkadi[1], MSc; Simon Renner[1], PharmD; Pierre Foulquié[1], MSc; Guy Fagherazzi[2], PhD; Paul Loussikian[1], MSc; Tom Marty[1], PharmD; Adel Mebarki[1], MSc; Nathalie Texier[1], PharmD; Stephane Schuck[1], MD

[1]Kap Code, Paris, France

[2]Deep Digital Phenotyping Research Unit, Department of Precision Health, Luxembourg Institute of Health, Strassen, Luxembourg

[*]these authors contributed equally

Corresponding Author:
Joelle Malaab, MSc, MPH
Kap Code
146 Rue Montmartre
Paris, 75002
France
Phone: 33 09 72 60 57 55
Email: malaabjoelle@gmail.com

Abstract

Background: Long COVID—a condition with persistent symptoms post COVID-19 infection—is the first illness arising from social media. In France, the French hashtag #ApresJ20 described symptoms persisting longer than 20 days after contracting COVID-19. Faced with a lack of recognition from medical and official entities, patients formed communities on social media and described their symptoms as long-lasting, fluctuating, and multisystemic. While many studies on long COVID relied on traditional research methods with lengthy processes, social media offers a foundation for large-scale studies with a fast-flowing outburst of data.

Objective: We aimed to identify and analyze Long Haulers' main reported symptoms, symptom co-occurrences, topics of discussion, difficulties encountered, and patient profiles.

Methods: Data were extracted based on a list of pertinent keywords from public sites (eg, Twitter) and health-related forums (eg, Doctissimo). Reported symptoms were identified via the MedDRA dictionary, displayed per the volume of posts mentioning them, and aggregated at the user level. Associations were assessed by computing co-occurrences in users' messages, as pairs of preferred terms. Discussion topics were analyzed using the Biterm Topic Modeling; difficulties and unmet needs were explored manually. To identify patient profiles in relation to their symptoms, each preferred term's total was used to create user-level hierarchal clusters.

Results: Between January 1, 2020, and August 10, 2021, overall, 15,364 messages were identified as originating from 6494 patients of long COVID or their caregivers. Our analyses revealed 3 major symptom co-occurrences: asthenia-dyspnea (102/289, 35.3%), asthenia-anxiety (65/289, 22.5%), and asthenia-headaches (50/289, 17.3%). The main reported difficulties were symptom management (150/424, 35.4% of messages), psychological impact (64/424,15.1%), significant pain (51/424, 12.0%), deterioration in general well-being (52/424, 12.3%), and impact on daily and professional life (40/424, 9.4% and 34/424, 8.0% of messages, respectively). We identified 3 profiles of patients in relation to their symptoms: profile A (n=406 patients) reported exclusively an asthenia symptom; profile B (n=129) expressed anxiety (n=129, 100%), asthenia (n=28, 21.7%), dyspnea (n=15, 11.6%), and ageusia (n=3, 2.3%); and profile C (n=141) described dyspnea (n=141, 100%), and asthenia (n=45, 31.9%). Approximately 49.1% of users (79/161) continued expressing symptoms after more than 3 months post infection, and 20.5% (33/161) after 1 year.

Conclusions: Long COVID is a lingering condition that affects people worldwide, physically and psychologically. It impacts Long Haulers' quality of life, everyday tasks, and professional activities. Social media played an undeniable role in raising and delivering Long Haulers' voices and can potentially rapidly provide large volumes of valuable patient-reported information. Since long COVID was a self-titled condition by patients themselves via social media, it is imperative to continuously include their

XSL•FO
RenderX

perspectives in related research. Our results can help design patient-centric instruments to be further used in clinical practice to better capture meaningful dimensions of long COVID.

(JMIR Infodemiology 2022;2(2):e39849) doi:10.2196/39849

KEYWORDS

long COVID; social media; Long Haulers; difficulties encountered; symptoms; infodemiology study; infodemiology; symptoms; COVID-19; patient-reported outcomes; persistent; condition; topics; discussion; social media; content

Introduction

Long COVID, also known as postacute sequelae of COVID-19, is one of the many repercussions of the COVID-19 pandemic. Patients who once had COVID-19 and experienced lasting symptoms referred to their condition as "long COVID" and themselves as "Long Haulers" [1]. Long COVID is defined as a persistence of symptoms for several weeks after the onset of COVID-19, with over 20% of those afflicted with it describing them after at least 4 weeks, and over 10% of patients after 3 months [2]. Early in the course of the health crisis, scientists focused on studying the novel SARS-CoV-2 and officials rushed to contain the spread of contamination, paying less attention to long-term effects. While infections were once thought of as short-term, in many cases, they became a lingering illness. The exact prevalence of long COVID is yet to be determined. A meta-analysis by Chen et al [3] estimated that 43% of COVID-19–positive individuals have had long COVID, and an even higher proportion for those who were hospitalized during the acute phase of infection [4]. Patients described their symptoms as long-lasting, fluctuating, and multisystemic, most frequently reporting generalized fatigue [5], respiratory ailments [6], neurological and cardiothoracic disorders, and a partial or complete loss of smell and taste [7].

Long COVID is the first illness arising from social media: the original long COVID hashtag (#LongCovid) appeared on Twitter in May 2020 to illustrate a lengthier and more complex course of the disease than described in the early reports from Wuhan, China [1]. The French hashtag #ApresJ20, which translates to "after day 20," described symptoms persisting longer than 20 days after contracting the infection. Patient-led groups on social media swiftly assembled, growing into a hub for information-sharing, support, encouragement, and communication among Long Haulers. In just a few months, discussions about long COVID moved from patients, via various media, to formal clinical and policy entities [1]. This highlighted the role of social media in drawing attention to a condition originally deemed invisible or nonexistent.

Indeed, social media has become an integral part of people's lives over the years. In 2021, the International Telecommunication Union estimated that 4.9 billion people were using the internet [8]. During the COVID-19 pandemic, social media turned into the main source of communication during lockdown [9], witnessing a 17% increase in internet users [8]. Social media has also demonstrated an increasing presence in health care: a rising number of patients have turned to the internet for health-related reasons [10]. Recently, the rise of social media prompted the emergence of infodemiological studies. Studies using data from web-based platforms have proven effective in research, notably in studying influenza-related topics [11-15], HIV/AIDS [16,17], and measles [18]. These studies use data obtained directly from the patient, avoiding lengthy traditional research methods and clinical studies. These especially proved to be useful during the pandemic, as scientists rapidly needed information about the novel coronavirus, and lockdown and social distancing measures disrupted the world. As long COVID rapidly gained awareness owing to social media, one cannot deny the substantial volume of data and respondents on web-based platforms and their impact to ultimately influence public policies. Indeed, there exists a window of opportunity regarding social media as tools for health research [19-22].

While research on long COVID has expanded, many studies relied on web-based surveys or clinical settings [7,23-26]. A limitation of these methods is the lengthy process required to launch the studies, obtain results, and finally to publish them. Social listening, however, offers a foundation for large-scale studies with a fast-flowing outburst of data and the opportunity to listen to patients in real time. The duration of our research spanned 587 days; this will allow us to fill any knowledge gaps that exist beyond the 1-year postinfection mark, ultimately preventing a shortfall in health care's potential during this crisis. Furthermore, this study encompassed multiple social media sources (ie, Twitter, Reddit, Doctissimo, Facebook, and other forums), thus increasing its exhaustivity and possibly inclusivity and representativity.

In this study, we aimed to examine patient feedback on social media regarding their experience with long COVID, using data mining methods. Our aim was to examine the impact of long COVID on Long Haulers by analyzing their main topics of discussion, the difficulties they encountered, and their most reported symptoms.

Methods

Study Design and Population

This observational, retrospective, real-world study encompasses data retrieved from social media posts of individuals with long COVID symptoms and their caregivers. The duration of the study spanned from January 1, 2020, to August 10, 2021.

Data Extraction

Messages written in French, geolocated in France, and posted between January 1, 2020, and August 10, 2021, were included. The data ultimately retrieved are composed of messages from public websites (eg, Twitter) and health-related forums (eg, Doctissimo). Owing to restricted data access and closed groups, only 2 Facebook pages "AprèsJ20" and "Collectif de Malades

XSL•FO
RenderX

Covid 19 au Long Cours" were analyzed, while Instagram and WhatsApp were excluded from this study. Keywords associated with long COVID were identified (eg, "long covid," "chronic covid," "persistent covid," "long term covid," "covid" + "months," and "covid" + "brain fog") and subsequently inserted in the extraction query.

Data extraction was performed by the Brandwatch extractor (Cision Ltd). First, we collected publicly available posts found on Twitter and forums featuring one of the relevant keywords. In parallel, we performed web crawling—or data collection—on the previously selected, publicly accessed Facebook pages. Posts were retrieved along with their associated metadata (eg, author or publication date). In this study, there was no distinction in the treatment of posts from the different platforms.

Preprocessing consisted of selecting only relevant messages based on several exclusion criteria: posts of 5 words or less and those exceeding 10,000 characters were excluded, as they are typically found to be irrelevant. Duplicates, posts not written in French, and sources deemed unsafe or inapplicable to our study (eg, advertising websites and forums related to cars, pets, or animals) were also excluded.

To further advance the filtering process, a supervised machine learning algorithm was applied to identify posts associated with patients' or caregivers' experiences. This algorithm was previously developed using a training set of 12,330 messages related to different health domains (dermatology, tobacco use, and oncology, among others). The method consists of a pipeline featuring 2 XGBoost [27] classifiers (one for caregivers' experiences and one for patients' experiences) applied successively. This method allowed us to identify if a post belonged to a patient, a caregiver, or neither. Both classifiers are based on features combining pronouns and lexical fields describing relatives and pathologies (eg, "my [pronoun] father [relative] has cancer [pathology]"). We trained the algorithm by first identifying the caregivers; this was carried out on the whole data set. To determine patients' messages, we then reapplied the algorithm on the rest of the data set (excluding the already identified caregiver messages). Evaluation of performances yielded F1-scores (a measure of accuracy combining precision and recall) of 88.0% and 87.0% for the caregiver and patient classifier, respectively.

In this study, to assess its performance on COVID-19–related data, the algorithm's performances were measured on a random sample of 700 messages classified as patients' or caregivers' messages. The pipeline predictions were then used to filter out posts that do not describe personal experiences with long COVID.

In this study, only posts from patients and caregivers were considered for analysis.

Ethical Considerations

This study included data from publicly available sources; private groups or web pages were thus excluded from our data extraction process. We did not seek approval as users automatically grant their consent for the reuse of their data when they post on public platforms. Furthermore, the results of this study do not contain any identifiable information and are presented in aggregate. Information such as the name, username or handle, geographic locations, or any other sensitive data were not included.

Data Analysis

Topics of Discussion

Main discussion themes were identified through the examination of all 15,364 posts from patients and caregivers regarding long COVID. This was performed using Biterm Topic Modeling (BTM) with the BTM R package [28]. BTM is a natural language processing–based text mining approach, which clusters similar texts on the basis of common discussion topics and provides lists of words to be interpreted for cluster labeling [29]. Topic modeling considers documents (messages and posts) as a mixture of topics that are a probability distribution over the words of the data set. A post can then be associated to its most prominent topic. BTM provides, for each topic, a list of the 20 highest-probability words and all the posts associated with the topic. Through human interpretation, these lists of words were then used to label the topics, and the associated posts were thoroughly scanned to ensure correct interpretation.

Unmet Needs and Difficulties Encountered

A manual review of posts identified the unmet needs described by patients with long COVID and their caregivers. A total of 450 messages were randomly selected from 3 types of sources: Twitter, forums, and selected public Facebook pages (n=150 messages from each type of source). We considered that this sample was sufficient and reasonable to have an overall view of the different types of unmet needs. To identify the unmet needs and difficulties of patients and caregivers, independent evaluators reviewed this sample via qualitative analysis: difficulties were coded in accordance with a previously set grid of categories to guarantee standardized data labeling and depending on whether the difficulty pertained to a patient, caregiver, or both. The categorization process was not mutually exclusive: the same message could contain multiple difficulties.

Reported Symptoms

To identify patients' and caregivers' reported symptoms, the data set resulting from preprocessing cleaning was analyzed using the MedDRA dictionary. The MedDRA dictionary is governed by a 5-level structure of hierarchy: a system organ class (SOC) is the highest or most general level, which is further subdivided into high-level group terms, high-level terms, preferred terms (PTs), and the most detailed lowest-level terms (LLTs). This last level was used for the detection of reported symptoms to achieve maximal thoroughness [30-32]. Since the MedDRA dictionary lacked terms related to long COVID at the time of the study, we manually added to it a list of symptoms identified through literature review [5-7]. All these terms were then searched in the messages and sorted in accordance with their frequency of occurrence, which allowed us to create a list of PTs of interest by selecting the most recurrent and relevant PT. The last step consisted of manual cleaning of the LLTs associated with the list of PTs, and pooling similar PTs (eg, fatigue and asthenia). Results were then sorted in accordance with the frequency at which they were reported with the top 15 PTs selected for this study. A manual review was then performed

XSL•FO
RenderX

to assess whether the medical concepts of those PTs were indeed long COVID symptoms. Once achieved, the set of LLTs associated with the selected PTs was used for detecting symptoms. Hereinafter, we shall refer to the top 15 PTs as "symptoms' PTs."

Co-occurrences and Standard Profile of a Patient With Long COVID in Terms of Symptoms

Associations were assessed by computing co-occurrences in users' posts as pairs of PTs. Counts of each PT in the total of their posted content was used to cluster users through hierarchical clustering. Users who mentioned at least 2 different symptoms' PTs in their messages were considered (n=289). A heat map was created to clearly display the significant co-occurrences.

Regarding the standard profile of a patient with long COVID, symptoms were displayed in accordance with the volume of posts mentioning them and then aggregated at the user level.

Chronological Monitoring of Symptoms

The evolution of symptoms was monitored through time: we collected messages pertaining to users who had written a minimum of 5 posts and featuring the selected symptoms' PTs within 18 months of their infection dates. We selected a threshold of a minimum of 5 messages from the same user to have enough data to follow their symptoms over time.

For the selected users, using regular expressions of duration and dates (eg, "It has been six months" and "since April") helped determine the estimated date of the COVID-19 infection. Following this information, a manual review of the messages allowed the analysis of symptom evolution per quarter year.

Results

Population and Posts

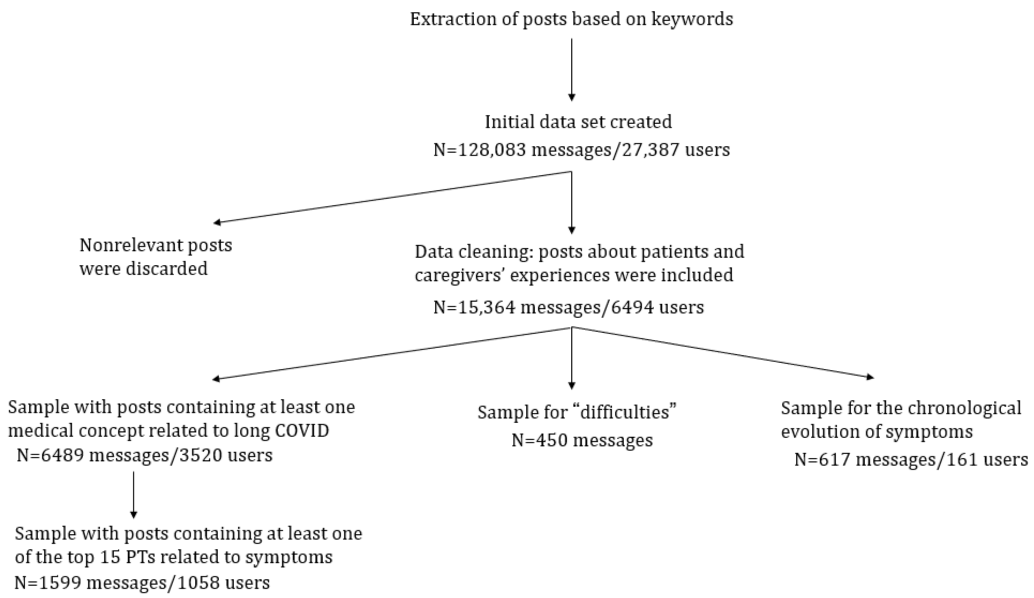
Between January 1, 2020, and August 10, 2021, a total of 128,083 messages were retrieved, which were written by 27,387 French speakers in France, ranging from journalists, politicians, citizens, or individuals recently infected with COVID-19 fearing a long-term progression of symptoms. A total of 21 sources were included in this study; however, the majority of the retrieved data (121,560/128,083, 94.9%) were found on Twitter. Subsequent analyses were not segmented among sources. As previously mentioned, a machine learning algorithm was applied to identify posts associated with patients or caregivers' experiences. As a result, of the 128,083 retrieved messages, 15,364 messages were identified as having originated from 6494 patients with long COVID or their caregivers (Figure 1).

The patient-caregiver algorithm was evaluated on long COVID data through a manual review of a random sample of 700 messages. Comparing pipeline predictions to manual coding yielded the following performance results: an accuracy of 87.5%, F1-score of 89.7%, sensibility of 96.3%, specificity of 75.8%, and a precision of 84.0%. Twitter remained the main source of expression with 93.8% (14,410/15,364) of messages.

Our analysis revealed that the first mention of "covid chronique" (which translates to "chronic COVID-19") appeared on social media on March 16, 2020. Less than a month later on April 12, 2020, the hashtag #apresJ20 was first mentioned on Twitter in the following message (translated from French): "Lack of information for people who continue to have symptoms after D20. It would be nice to share our feelings and feel less alone so I'm opening this poll for those who are still struggling after D20 #COVID19 #afterD20."

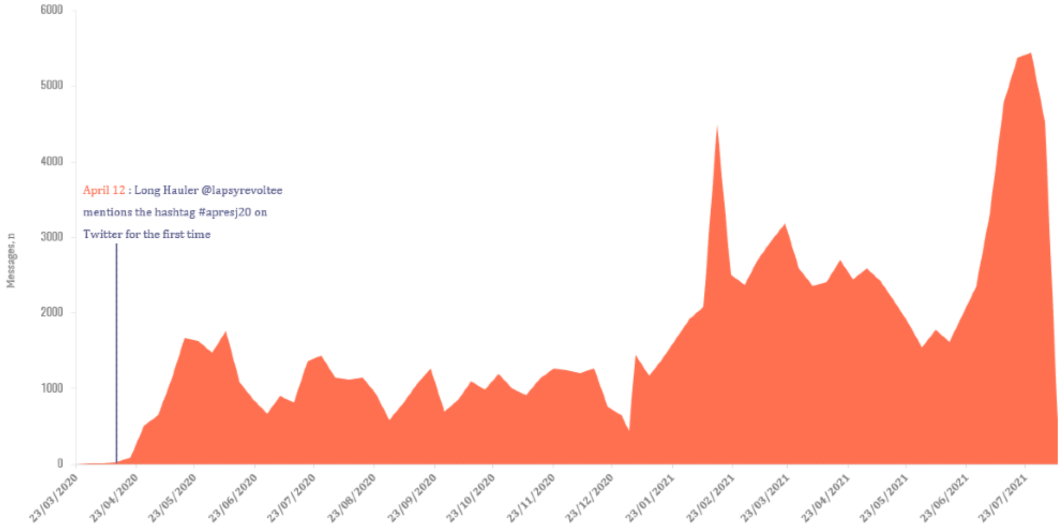
Following the introduction of the hashtag #ApresJ20, it rapidly went viral in France; Twitter and Facebook witnessed a rise in the number of users sharing their experiences with long COVID, particularly after the launch of pages and groups dedicated to this subject (Figure 2).

Figure 1. Flow chart of the data cleaning and sample selection processes. PT: preferred term.



XSL•FO
RenderX

Figure 2. Evolution of the percentages of number of messages per main topic between March 2020 and July 2021.



Data Analysis

Topics of Discussion

Following the application of the BTM on the data set including all the analyzed forums (Multimedia Appendix 1), various discussion topics were identified through human interpretation of each topic's most associated terms. For instance, "vaccine," "protect," and "long" were among the top terms of the topic "Vaccination and Long Covid" after translation. The main revealed topics are featured in Figure 3.

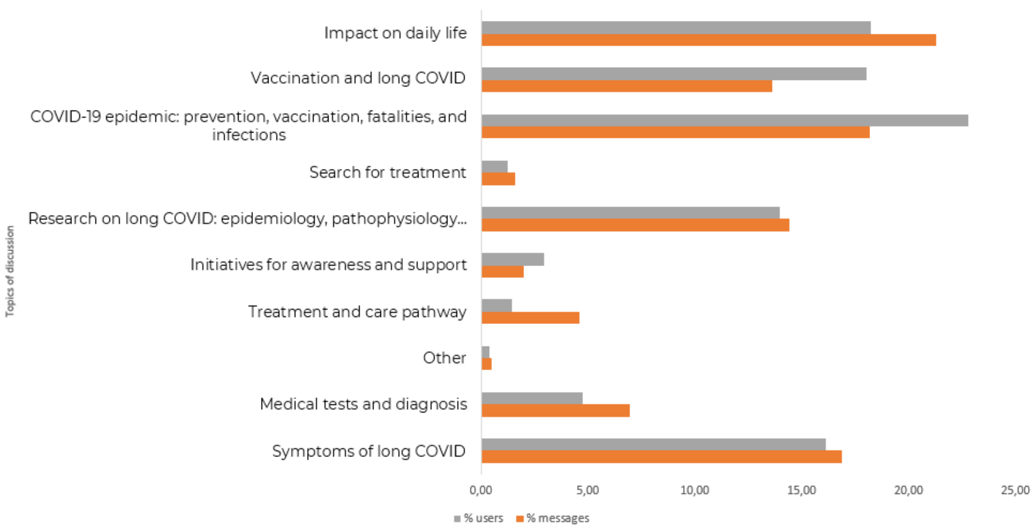
The 5 primary topics of discussion centered around the COVID-19 pandemic in general (2793/15,364, 18.2% of messages) in addition to issues related to long COVID were as follows: impact on daily life (3269/15,364, 21.3%), reported symptoms (2592/15,364, 16.9%), vaccination (2090/15,364, 13.6%), and research (2212/15,364, 14.4%).

The topic "Covid-19 pandemic" was discussed by the highest number of users (1480/6494, 22.8%), while the topic "impact on daily life" received the largest volume of posts (3269/15,364, 21.3%).

The monitoring of these 5 topics through time revealed a peak in the number of posts around the subject of "symptoms" in the first half of 2020 (5/11, 45.5% of the messages posted in March 2020 about long COVID; Figure 4).

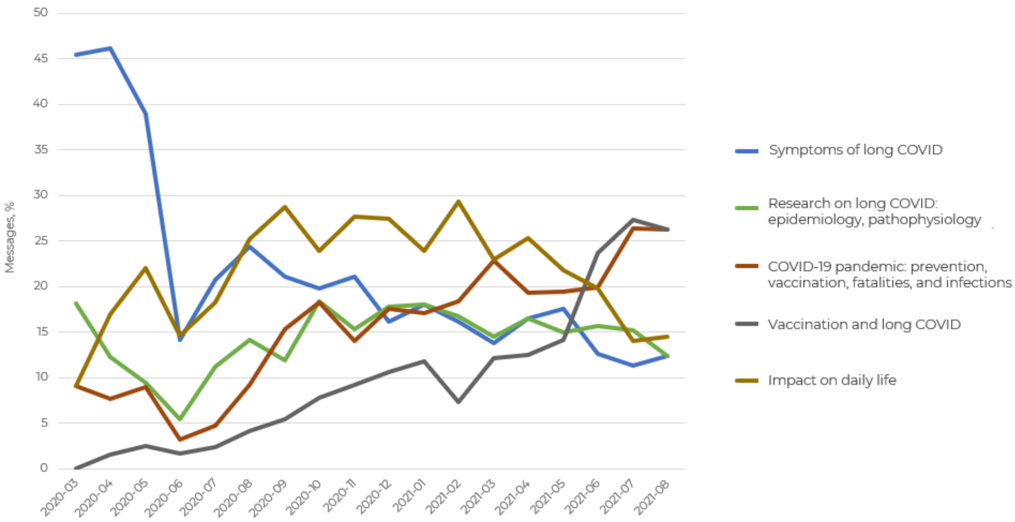
Furthermore, the topic "COVID-19 pandemic" progressively gained momentum and was increasingly discussed over time after the second half of 2020 (31/982, 3.2% of the messages posted in June 2020, ending at 204/775, 26.3% of the messages posted in August 2021; Figure 4). In contrast, discussions around the impact on daily life gradually decreased. On completion of this study, the topics of "vaccination" (204/775, 26.3%) and "COVID-19 pandemic" (204/775, 26.3%) were the most frequently discussed among users (Figure 4).

Figure 3. Proportions of messages and internet users with the main discussion topics.



XSL•FO
RenderX

Figure 4. Chronological evolution of the main discussion topics by the proportion of messages.



Unmet Needs and Difficulties Encountered

Of the 450 messages analyzed, 424 included at least 1 difficulty reported by a patient or caregiver. These messages contained a total of 709 difficulties, which were sorted into 34 categories overall, with a single message possibly containing more than 1 category of difficulties. The top 20 categories of the main reported difficulties are featured in Table 1.

The main difficulties reported by patients in relation to long COVID were the management of their symptoms (150/424, 35.4% of messages), which were described as diverse, lasting several weeks or months, fluctuating over time, and involving relapses. Patients also reported a psychological impact characterized by a fear of the unknown (64/424, 15.1%). Additionally, messages mentioned a feeling of pain (51/424, 12.0%) and a deterioration in general well-being (52/424, 12.3%), particularly owing to intense and chronic fatigue. The impact on daily and professional life, mentioned in 9.4% (40/424) and 8.0% (34/424) of messages, respectively, was described by patients and caregivers as considerably reducing their quality of life. Furthermore, 13.7% (58/424) of the reported difficulties were centered around the lack of recognition of long COVID by health care providers, public and health authorities, or even patients' close circles. In fact, several patients reported that doctors were questioning the clinical validity of their symptoms, and, in some cases, even suggesting that the problems were simply due to stress. A patient detailed her experience with a health care professional in the following message (translated from French):

> *Doctors have a hard time diagnosing long Covid... Some have told me that I was in denial about being pregnant (I haven't had intercourse in several months), or that I was starting menopause (not true after blood work), that it was depression or that it was all in my head! A shame that this disease is still not recognized nor treated*

XSL•FO
RenderX

Table 1. Proportions of messages featuring the main reported difficulties.

Difficulties	Messages, n (%)
Concern with and management of symptoms of long COVID	150 (35.4)
Psychological impact of long COVID on patients and the stress of uncertainty	64 (15.1)
Lack of recognition of long COVID	58 (13.7)
Alteration of the general state of health with chronic fatigue, loss of capacity, brain fog, etc	52 (12.3)
Management of pain	51 (12.0)
Impact on daily activities	40 (9.4)
Professional impact for the patient: part-time work, absences, etc	34 (8.0)
Patient dissatisfaction with the provision of care	34 (8.0)
Fears and management of aggravations and relapses of long COVID	33 (7.9)
Difficulty in accessing care: long waiting time, difficulty taking an appointment, lack of experts on the subject, etc	29 (6.8)
Issues related to the lack of training of health care personnel on long COVID	17 (4.0)
Sequelae of COVID-19	17 (4.0)
Sharing, experiences, and support: discussion groups, social networks, etc	16 (3.8)
Worried or concerned about the future, life expectancy, or difficulty planning ahead	16 (3.8)
Communication and relationship problems: lack of empathy, conveyance of information, medical jargon, etc	11 (2.6)
Disagreement in health management: heterogeneity of medical decisions and opinions, disagreement between the patient and medical team, etc	10 (2.4)
Lack of general knowledge or scientific information about long COVID	9 (2.1)
Multiple treatment failure or ineffective treatments	9 (2.1)
Financial impact of health care for patients	9 (2.1)
Impact of long COVID on the management of comorbidities	7 (1.7)

Reported Symptoms

Overall, 6489 messages posted by 3520 users in the different forums had expressed at least 1 medical concept related to long COVID. The most reported symptoms were revealed on the basis of pooled PTs, with 1599 messages written by 1058 users having expressed at least one of the top 15 PTs. An evaluation of noise (ie, random data errors) on a random sample of 400 messages revealed an 86.0% correct classification, meaning that 86.0% of PTs actually corresponded to symptoms of long COVID. Additionally, an assessment was performed on 10 random messages for each PT, yielding a correct classification of 90.2% on average.

Medical concepts were also categorized in accordance with the MedDRA dictionary's SOC based on 7 organ categories: systemic, respiratory, nervous, psychiatric, musculoskeletal, cardiac, and gastrointestinal. The majority of messages (995/1599, 62.2%) pertained to the "systemic" category (Multimedia Appendix 2).

In addition, health ailments were related to patients' respiratory system (267/1599, 16.7%), nervous system (264/1599, 16.5%), and psychiatric system (252/1599, 15.8%; Multimedia Appendix 2).

The pooling of PTs revealed a range of symptoms related to long COVID with the top 3 most reported ones in patients' and caregivers' messages being asthenia (835/1599, 52.2% of messages), dyspnea (267/1599, 16.7%), and anxiety (242/1599, 15.1%; Multimedia Appendix 3).

Indeed, patients reported a feeling of chronic fatigue and weakness, as described in the following message (translated from French): "I had so much vital energy that I tired everyone around me! Ever since my long Covid, I remain confined, low blood pressure, insane fatigue, seizures of all kinds, yes, I have no more energy!"

Co-occurrences and Standard Profile of a Patient With Long COVID in Terms of Symptoms

Among the patients who described a symptom from among the top 15 symptoms (1584 patients), 41.2% (n=652) of them experienced asthenia, 14.7% (n=233) reported dyspnea, and 12.6% (n=200) experienced anxiety (Figure 5).

Several symptoms were also frequently and simultaneously reported by patients or caregivers. Users who mentioned at least 2 different symptoms' PTs in their messages were considered (n=289). A heat map featuring the co-occurrences that appeared at least 5 times is displayed as a logarithmic scale in Figure 6. The highest proportion of patients (102/289, 35.3%) reported asthenia paired with dyspnea, followed by 22.5% (65/289) of patients who experienced asthenia along with anxiety, while 17.3% (50/289) of patients experienced asthenia in combination with headaches (Figure 5).

XSL•FO
RenderX

Furthermore, the “clustering” method allowed the identification of the 3 main standard profiles of patients in relation to their symptoms: profile A (n=406 patients) reported exclusively 1 symptom of asthenia; profile B (n=129) expressed anxiety (n=129, 100%), asthenia (n=28, 21.7%), dyspnea (n=15, 11.6%), ageusia or loss of a sense of taste (n=3, 2.3%); and finally, profile C (n=141) described dyspnea (n=141, 100%) and asthenia (n=45, 31.9%).

Figure 5. Distribution of reported symptoms related to an individual.

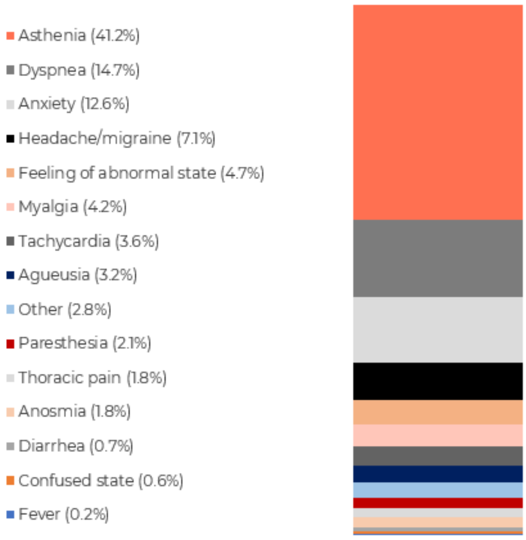
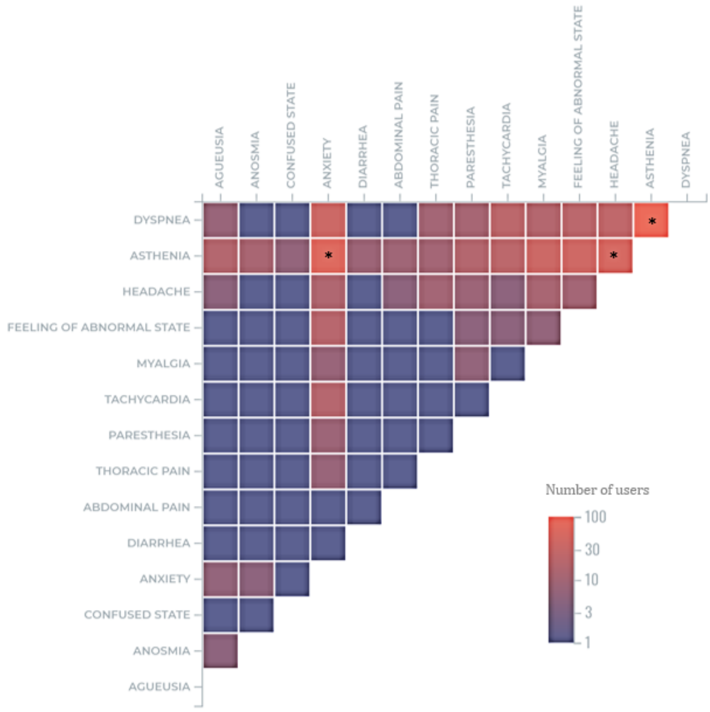


Figure 6. Heat map of the co-occurrences that appeared at least 5 times.



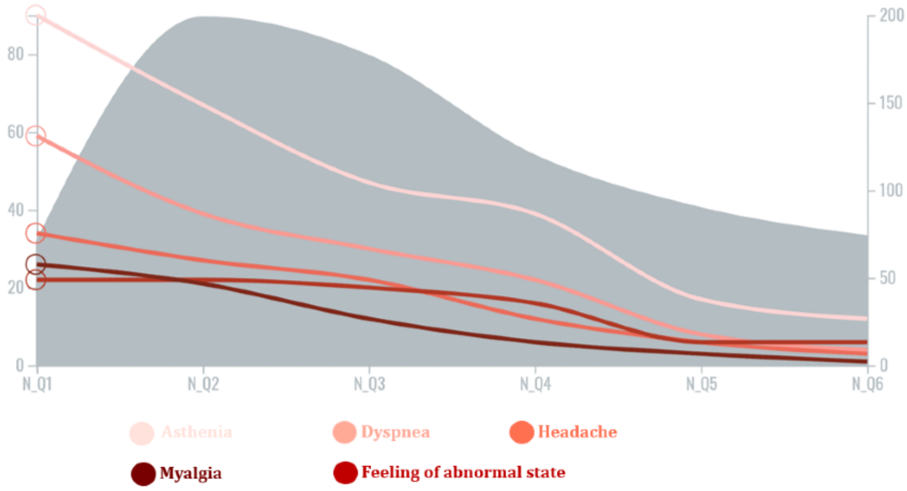
Chronological Monitoring of Symptoms

To monitor the evolution of symptoms through time, the initial data set was filtered to retain users with enough content to be followed: of the 15,364 messages in the initial data set, 3062 posts by 1493 users included a regular expression of duration and dates. Among those users, 330 had posted at least 5 messages; their posts amounted to a total of 1765, including 712 posts (from 217 users) with at least 1 mention of a PT. Finally, we retained 617 posts by 161 users featuring PTs related to symptoms within 18 months of their infection dates. This final data set of messages revealed the following reported symptoms: asthenia, dyspnea, headache, a feeling of abnormal state, and myalgia.

Our analysis showed a peak in the number of messages in the second trimester, followed by a gradual decrease over time (Figure 7). Furthermore, 79 of 161 (49.1%) users continued expressing symptoms between 3 months and 6 months post infection, 32.9% (53/161) between 9 and 12 months, and 20.5% (33/161) between 15 and 18 months.

XSL•FO
RenderX

Figure 7. Evolution of the number of occurrences per symptom and the number of messages with a mention of a symptom per quarter year.



Discussion

Principal Findings

Background

This study revealed—through the lens of Long Haulers—the multifaceted challenges and repercussions associated with long COVID. The main topics of discussion on social media centered around the impact on daily life, reported symptoms, and vaccination. Patients expressed difficulties related to the management of their symptoms, the impact on their mental health, and the impact on their daily and professional lives. Our analyses revealed 3 major symptom co-occurrences: asthenia-dyspnea, asthenia-anxiety, and asthenia-headaches. We identified 3 profiles of patients in relation to their symptoms: profiles A, B, and C, which mainly reported asthenia, anxiety, and dyspnea, respectively. Approximately 49.1% of users (79/161) continued expressing symptoms after more than 3 months post infection, and 20.5% (33/161) after 1 year.

Role of Social Media

The COVID-19 pandemic gained momentum on social media. Met with skepticism from medical professionals, coined as "medical gaslighting" [33,34], Long Haulers turned to web-based platforms, mainly Twitter (93.8% of total messages), to share their experiences. In fact, many infodemiological studies have documented the popularity of Twitter among internet users and used it as the main source of data [20,35-38].

Our study showed multiple peaks in communication coinciding with notable events. The #ApresJ20 launched a nationwide discussion on long COVID, with a peak in the number of messages soon after the hashtag was first mentioned in April 2020. Other peaks later followed in the course of our study in response to various events: in October 2020, after the Apresj20-Association Covid Long France launched along with its website, Facebook, and Twitter pages, offering a support group, patient experiences, resources, and information regarding long COVID [39]; in February 2021, after the association proposed resolutions to the National Assembly for the recognition of individuals with long COVID; and in July 2021, after President Emmanuel Macron extended the health pass and announced mandatory vaccination for certain professions [40]. This further demonstrated the role that social media played in propelling long COVID and echoing the voices of Long Haulers during the health crisis; it also showed that communication on social media mirrored notable events related to long COVID.

Topics of Discussion and Difficulties Encountered

Not only was social media effective for collecting substantial volumes of data, but also it communicated patients' sentiments, perceptions, and pain points. Patients reported physical and psychological sequelae that affected their day-to-day lives, complained of the heavy toll they were experiencing post COVID-19 infection, and criticized the quality of care afforded to them. Health providers' lack of knowledge of long COVID have led to serial misdiagnoses, and patients felt "invisible," uncertain if they will ever be cured of their physical pain, deteriorating well-being, and chronic fatigue.

Our analyses revealed that posts about "symptoms" initially dominated discussion topics, which is consistent with the findings of other infodemiological studies [21,41]. As the condition was relatively unknown in the first semester of 2020, social media served as a medium for patients and caregivers to relate to people with similar symptoms, thus creating a community for support and communication. A lack of information, recognition, and acknowledgment were the main catalysts behind the rallying forces of advocacy groups on social media; this prompted official entities, in the latter half of the year, to officially recognize long COVID and offer information and support for those afflicted with it [42]. Main topics of discussion also revolved around its impact on quality of life. Patients shared their experiences with long COVID and its impact on their everyday activities: mobility, housework, sports, etc. According to a study by Shah et al [43], survivors of COVID-19 reported a considerable impact of long COVID on their quality of life with problems ranging from physical (eg, limited mobility, disrupted usual activities, pain, and discomfort) to psychological (eg, anxiety, stress, and depression). Interestingly, discussions around the impact of long COVID on daily life gradually decreased, suggesting a lighter burden on some patients as their symptoms improved or as they came to accept their condition as their new health baseline.

XSL•FO
RenderX

Long COVID spread to many aspects of patients' lives; it affected their professional activities as they felt incapable of resuming work owing to their fragile state of health. Our findings align with those of Davis et al [7], where patients experienced difficulties going back to their work; among those who did, many reported experiencing relapse and could no longer continue their work activities. Faghy et al [44] corroborated these results, with patients reporting reduced health and capacity to participate in daily and work activities. This highlights the importance of officials' recognition of long COVID as a debilitating condition, hence offering those afflicted with it a proper recovery time before resuming their job. In that case, patients should be able to have access to financial government assistance, flexible work hours, or the possibility of teleworking. A holistic approach to restore their pre–COVID-19 health and quality of life, tackling the numerous and multifaceted challenges that patients have highlighted, is also of utmost importance.

An increasingly mentioned topic throughout our study involved "vaccination"; it reached its peak toward the latter half of 2021, coinciding with the implementation of compulsory vaccination to certain professions. This topic triggered debate among Long Haulers: some of them were apprehensive, while others reported the effectiveness of vaccines in alleviating or curing their condition. While the exact pathophysiology of Long COVID is still unknown, evidence shows that getting vaccinated might attenuate symptoms [45].

Symptoms and Co-occurrences

Long Haulers have reported symptoms affecting various body organs. The SOC revealed that symptoms mainly pertained to the "systemic" category, followed by the respiratory, nervous, and psychiatric systems. According to Nalbandian et al [46], long COVID also affects the excretory, circulatory, integumentary, and endocrine systems. The main reported symptoms were asthenia and dyspnea, which is consistent with the findings of other studies [5,6]. Additionally, patients described experiencing anxiety due to either the effect of long COVID on brain health or to "medical gaslighting." A study by Taquet et al [47] on neurological and psychiatric sequelae in survivors of COVID-19 revealed similar results, as patients reported experiencing anxiety even 6 months post COVID-19 infection.

The map of co-occurrences revealed the most commonly reported symptoms that collocated, and solidified the results obtained in the SOC. It revealed 3 major co-occurrences: asthenia (systemic)-dyspnea (respiratory), asthenia-anxiety (psychiatric), asthenia-headaches (nervous). The clustering method further corroborated these findings, as asthenia, anxiety, and dyspnea were found at the top of the 3 main standard profiles of Long Haulers. This raises concern regarding the damage that long COVID may have on organ systems, and highlights the need for a thorough examination of its repercussions. According to Graham et al [48], the main reported symptoms including "brain fog," persistent fatigue, and depression or anxiety affected Long Haulers' cognition and quality of life [48].

This study showed a lingering effect of long COVID, with patients still reporting symptoms after 6 months and even after 1 year, albeit to a lesser extent. These findings entail a lasting impact on patients in various aspects of their lives. According to the French e-cohort study ComPaRe [49], among patients who were symptomatic after 2 months, 85% of them reported persisting symptoms 1 year after symptom onset. Another study [50] revealed a prevalent feeling of severe fatigue among patients in long COVID online support groups. Whether they are physical, psychological, social, professional, or financial, the debilitating sequelae must be further explored to achieve better management.

Limitations

We recognize several limitations related to our study.

First, our research was limited to French-language social media posts and to individuals of certain socioeconomic demographics and literacy capacity, who have access to the internet and are capable and knowledgeable enough to post messages on social media. However, considering that 93% of the French population comprises internet users [51], we may safely assume that our study is adequately representative of the French Long Haulers.

Another limitation is that the majority of our data were obtained from Twitter. However, given that Twitter is one of the most visited websites in France with more than 16 million monthly users, it is no surprise that most of the data used is this study originated from it [52]. This source has specific features, notably a limit in post length and a high reactivity to events. Other sources such as health discussion forums tend to have less but more complete content on a patient's experience. The disparities among the different sources might have had an impact on some of our analyses such as topic modeling and symptom identification, which are based on content analysis. Although Twitter might have created a bias since it constituted the majority of our data set, Twitter data were essential to our research since the term "long COVID" originated from a tweet [1].

The annotation included a small pool of messages; it might also be prone to the subjective bias of the annotators who performed it. As such, our findings might not be accurately representative of the global population. However, several annotators were involved to limit this bias.

Owing to ethical reasons, our study included only openly accessible web-based networks and, as a result, lacked other platforms with restricted access, such as WhatsApp and Instagram.

Conclusions

Long COVID is a lingering condition that affects the lives of people worldwide, physically and psychologically. It impacts Long Haulers' quality of life, everyday tasks, and professional activities. The role of social media in raising and delivering Long Haulers' voices is undeniable: patients turned to social media to document their negative experiences post COVID-19 infection, search for information regarding their condition, exchange experiences and resources, gain recognition via advocacy groups, and find support in times of uncertainty. It

XSL•FO
RenderX

also has the potential to rapidly provide large volume of valuable patient-reported information. Considering the fact that long COVID was a self-titled condition by the patients themselves, it is imperative to continuously include their perspectives in research on long COVID; ignoring this aspect would simply lead to ignoring the key element of how this condition initially emerged. This study provides a good understanding of patients' perceptions, physical and psychological symptoms, and the difficulties they encountered during their illness. The data presented here can help design patient-centric instruments to be used in clinical practice to better capture meaningful dimensions of long COVID. Further research is imperative to bridge the knowledge gap about long COVID and improve the management of the condition by the health care system.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of sources.
[PNG File , 36 KB - infodemiology_v2i2e39849_app1.png]

Multimedia Appendix 2
Proportions of messages with at least 1 occurrence of the system organ class categorized in accordance with the different organ systems.
[PNG File , 76 KB - infodemiology_v2i2e39849_app2.png]

Multimedia Appendix 3
Proportions of messages featuring at least one occurrence of a PT related to symptoms.
[PNG File , 21 KB - infodemiology_v2i2e39849_app3.png]

References

1. Callard F, Perego E. How and why patients made long Covid. Soc Sci Med 2021 Jan;268:113426 [FREE Full text] [doi: 10.1016/j.socscimed.2020.113426] [Medline: 33199035]
2. Symptômes prolongés suite à une Covid-19 de l'adulte - Diagnostic et prise en charge. Haute Autorité de Santé. 2021. URL: https://www.has-sante.fr/jcms/p_3237041/fr/symptomes-prolonges-suite-a-une-covid-19-de-l-adulte-diagnostic-et-prise-en-charge [accessed 2022-08-04]
3. Chen C, Haupert S, Zimmermann L, Shi X, Fritsche L, Mukherjee B. Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long COVID: a meta-analysis and systematic review. J Infect Dis 2022 Nov 01;226(9):1593-1607 [FREE Full text] [doi: 10.1093/infdis/jiac136] [Medline: 35429399]
4. Ghosn J, Piroth L, Epaulard O, Le Turnier P, Mentré F, Bachelet D, et al. Persistent COVID-19 symptoms are highly prevalent 6 months after hospitalization: results from a large prospective cohort. Clin Microbiol Infect 2021 Jul;27(7):1041.e1-1041.e4 [FREE Full text] [doi: 10.1016/j.cmi.2021.03.012] [Medline: 34125067]
5. Shah W, Hillman T, Playford ED, Hishmeh L. Managing the long term effects of covid-19: summary of NICE, SIGN, and RCGP rapid guideline. BMJ 2021 Jan 22;372:n136. [doi: 10.1136/bmj.n136] [Medline: 33483331]
6. Cares-Marambio K, Montenegro-Jiménez Y, Torres-Castro R, Vera-Uribe R, Torralba Y, Alsina-Restoy X, et al. Prevalence of potential respiratory symptoms in survivors of hospital admission after coronavirus disease 2019 (COVID-19): A systematic review and meta-analysis. Chron Respir Dis 2021;18:14799731211002240 [FREE Full text] [doi: 10.1177/14799731211002240] [Medline: 33729021]
7. Davis HE, Assaf GS, McCorkell L, Wei H, Low RJ, Re'em Y, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine 2021 Aug;38:101019. [doi: 10.1016/j.eclinm.2021.101019] [Medline: 34308300]
8. Statistics. ITUPP Bucharest. 2022. URL: https://www.itu.int:443/en/ITU-D/Statistics/Pages/stat/default.aspx [accessed 2022-08-04]
9. Pérez-Escoda A, Jiménez-Narros C, Perlado-Lamo-de-Espinosa M, Pedrero-Esteban LM. Social networks' engagement during the COVID-19 pandemic in Spain: health media vs. healthcare professionals. Int J Environ Res Public Health 2020 Jul 21;17(14):5261 [FREE Full text] [doi: 10.3390/ijerph17145261] [Medline: 32708231]
10. 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]
11. Bour C, Ahne A, Schmitz S, Perchoux C, Dessenne C, Fagherazzi G. The use of social media for health research purposes: scoping review. J Med Internet Res 2021 May 27;23(5):e25736 [FREE Full text] [doi: 10.2196/25736] [Medline: 34042593]

XSL•FO
RenderX

12. Yom-Tov E, Gabrilovich E. Postmarket drug surveillance without trial costs: discovery of adverse drug reactions through large-scale analysis of web search queries. J Med Internet Res 2013 Jun 18;15(6):e124 [FREE Full text] [doi: 10.2196/jmir.2614] [Medline: 23778053]
13. Seo D, Jo M, Sohn CH, Shin S, Lee J, Yu M, et al. Cumulative query method for influenza surveillance using search engine data. J Med Internet Res 2014 Dec 16;16(12):e289 [FREE Full text] [doi: 10.2196/jmir.3680] [Medline: 25517353]
14. Sharpe JD, Hopkins RS, Cook RL, Striley CW. Evaluating Google, Twitter, and Wikipedia as tools for influenza surveillance using Bayesian change point analysis: a comparative analysis. JMIR Public Health Surveill 2016 Oct 20;2(2):e161 [FREE Full text] [doi: 10.2196/publichealth.5901] [Medline: 27765731]
15. Santos JC, Matos S. Analysing Twitter and web queries for flu trend prediction. Theor Biol Med Model 2014 May 07;11 Suppl 1:S6 [FREE Full text] [doi: 10.1186/1742-4682-11-S1-S6] [Medline: 25077431]
16. Adrover C, Bodnar T, Huang Z, Telenti A, Salathé M. Identifying adverse effects of HIV drug treatment and associated sentiments using Twitter. JMIR Public Health Surveill 2015 Jul 27;1(2):e7 [FREE Full text] [doi: 10.2196/publichealth.4488] [Medline: 27227141]
17. Mavragani A, Ochoa G. Forecasting AIDS prevalence in the United States using online search traffic data. J Big Data 2018 May 19;5(1). [doi: 10.1186/s40537-018-0126-7]
18. Mollema L, Harmsen IA, Broekhuizen E, Clijnk R, De Melker H, Paulussen T, et al. Disease detection or public opinion reflection? Content analysis of tweets, other social media, and online newspapers during the measles outbreak in The Netherlands in 2013. J Med Internet Res 2015 May 26;17(5):e128 [FREE Full text] [doi: 10.2196/jmir.3863] [Medline: 26013683]
19. 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]
20. Odlum M, Yoon S, Broadwell P, Brewer R, Kuang D. How Twitter can support the HIV/AIDS response to achieve the 2030 eradication goal: in-depth thematic analysis of world AIDS day tweets. JMIR Public Health Surveill 2018 Nov 22;4(4):e10262 [FREE Full text] [doi: 10.2196/10262] [Medline: 30467102]
21. Santarossa S, Rapp A, Sardinas S, Hussein J, Ramirez A, Cassidy-Bushrow AE, et al. Understanding the #longCOVID and #longhaulers Conversation on Twitter: multimethod study. JMIR Infodemiology 2022;2(1):e31259 [FREE Full text] [doi: 10.2196/31259] [Medline: 35229074]
22. Jordan S, Hovet S, Fung I, Liang H, Fu K, Tse Z. Using Twitter for public health surveillance from monitoring and prediction to public response. Data 2018 Dec 29;4(1):6. [doi: 10.3390/data4010006]
23. Nikravesh I. Long COVID can negatively impact physical and cognitive function, employment, and quality of life for at least one year. EurekAlert!. 2021. URL: https://www.eurekalert.org/news-releases/932580 [accessed 2022-08-04]
24. Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: A 6-month retrospective cohort study of 273,618 survivors of COVID-19. PLoS Med 2021 Sep;18(9):e1003773 [FREE Full text] [doi: 10.1371/journal.pmed.1003773] [Medline: 34582441]
25. Cox T. How many people get 'long COVID?' More than half, researchers find. The Pennsylvania State University. 2021. URL: https://www.psu.edu/news/research/story/how-many-people-get-long-covid-more-half-researchers-find/ [accessed 2022-08-04]
26. Yong SJ. Long COVID or post-COVID-19 syndrome: putative pathophysiology, risk factors, and treatments. Infect Dis (Lond) 2021 Oct 22;53(10):737-754 [FREE Full text] [doi: 10.1080/23744235.2021.1924397] [Medline: 34024217]
27. Chen T, He T, Benesty M. Xgboostxtreme gradient boosting. R package version 04-2. 2015. URL: https://cran.microsoft.com/snapshot/2017-12-11/web/packages/xgboost/vignettes/xgboost.pdf [accessed 2022-11-17]
28. Yan X. Shortext.Github.Io. Github. URL: https://github.com/xiaohuiyan/xiaohuiyan.github.io/blob/61d74f1b74a26ba136d65c7b3420d1e54440fe44/paper/BTM-WWW13.pdf [accessed 2022-08-11]
29. Yan X, Guo J, Lan Y, Cheng X. A biterm topic model for short texts. 2013 Presented at: WWW '13: 22nd International World Wide Web Conference; May 13-17, 2013; Rio de Janeiro. [doi: 10.1145/2488388.2488514]
30. 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 Jun 22;3(2):e36 [FREE Full text] [doi: 10.2196/publichealth.6577] [Medline: 28642212]
31. Schäfer F, Faviez C, Voillot P, Foulquié P, Najm M, Jeanne J, et al. Mapping and modeling of discussions related to gastrointestinal discomfort in French-speaking online forums: results of a 15-year retrospective infodemiology study. J Med Internet Res 2020 Nov 03;22(11):e17247 [FREE Full text] [doi: 10.2196/17247] [Medline: 33141087]
32. Schück S, Roustamal A, Gedik A, Voillot P, Foulquié P, Penfornis C, et al. Assessing patient perceptions and experiences of paracetamol in France: infodemiology study using social media data mining. J Med Internet Res 2021 Jul 12;23(7):e25049 [FREE Full text] [doi: 10.2196/25049] [Medline: 34255645]
33. Living with Covid19. National Institute for Health and Care Research. 2020. URL: https://evidence.nihr.ac.uk/themedreview/living-with-covid19/ [accessed 2022-11-09]
34. Rubin R. As their numbers grow, COVID-19 "Long Haulers" stump experts. JAMA 2020 Oct 13;324(14):1381-1383. [doi: 10.1001/jama.2020.17709] [Medline: 32965460]

XSL•FO
RenderX

35. Zeraatkar K, Ahmadi M. Trends of infodemiology studies: a scoping review. Health Info Libr J 2018 Jun 04;35(2):91-120 [FREE Full text] [doi: 10.1111/hir.12216] [Medline: 29729073]
36. Bian J, Zhao Y, Salloum RG, Guo Y, Wang M, Prosperi M, et al. Using social media data to understand the impact of promotional information on laypeople's discussions: a case study of Lynch syndrome. J Med Internet Res 2017 Dec 13;19(12):e414 [FREE Full text] [doi: 10.2196/jmir.9266] [Medline: 29237586]
37. Chen T, Dredze M. Vaccine images on Twitter: analysis of what images are shared. J Med Internet Res 2018 Apr 03;20(4):e130 [FREE Full text] [doi: 10.2196/jmir.8221] [Medline: 29615386]
38. Hswen Y, Naslund JA, Brownstein JS, Hawkins JB. Monitoring online discussions about suicide among Twitter users with schizophrenia: exploratory study. JMIR Ment Health 2018 Dec 13;5(4):e11483 [FREE Full text] [doi: 10.2196/11483] [Medline: 30545811]
39. Long COVID France. Long-Covid Europe. URL: https://longcovideurope.org/?p=373 [accessed 2022-08-04]
40. Salvi E. Macron trumpets own record as he announces mandatory vaccines for health staff and Covid 'passports'. Mediapart. 2021. URL: https://www.mediapart.fr/en/journal/france/130721/macron-trumpets-own-record-he-announces-mandatory-vaccines-health-staff-and-covid-passports [accessed 2022-08-04]
41. Miyake E, Martin S. Long Covid: online patient narratives, public health communication and vaccine hesitancy. Digit Health 2021 Nov 29;7:20552076211059649 [FREE Full text] [doi: 10.1177/20552076211059649] [Medline: 34868622]
42. WHO Director-General's opening remarks at the media briefing on COVID-19 - 21 August 2020. World Health Organization. 2020. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---21-august-2020 [accessed 2022-08-04]
43. Shah R, Ali FM, Nixon SJ, Ingram JR, Salek SM, Finlay AY. Measuring the impact of COVID-19 on the quality of life of the survivors, partners and family members: a cross-sectional international online survey. BMJ Open 2021 May 25;11(5):e047680 [FREE Full text] [doi: 10.1136/bmjopen-2020-047680] [Medline: 34035105]
44. Faghy MA, Maden-Wilkinson T, Arena R, Copeland RJ, Owen R, Hodgkins H, et al. COVID-19 patients require multi-disciplinary rehabilitation approaches to address persisting symptom profiles and restore pre-COVID quality of life. Expert Rev Respir Med 2022 May;16(5):595-600. [doi: 10.1080/17476348.2022.2063843] [Medline: 35385677]
45. Li X, Raventós B, Roel E, Pistillo A, Martinez-Hernandez E, Delmestri A, et al. Association between covid-19 vaccination, SARS-CoV-2 infection, and risk of immune mediated neurological events: population based cohort and self-controlled case series analysis. BMJ 2022 Mar 16;376:e068373 [FREE Full text] [doi: 10.1136/bmj-2021-068373] [Medline: 35296468]
46. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr;27(4):601-615 [FREE Full text] [doi: 10.1038/s41591-021-01283-z] [Medline: 33753937]
47. Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. Lancet Psychiat 2021 May;8(5):416-427. [doi: 10.1016/s2215-0366(21)00084-5]
48. Graham EL, Clark JR, Orban ZS, Lim PH, Szymanski AL, Taylor C, et al. Persistent neurologic symptoms and cognitive dysfunction in non-hospitalized Covid-19 "long haulers". Ann Clin Transl Neurol 2021 May 30;8(5):1073-1085 [FREE Full text] [doi: 10.1002/acn3.51350] [Medline: 33755344]
49. Tran V, Porcher R, Pane I, Ravaud P. Course of post COVID-19 disease symptoms over time in the ComPaRe long COVID prospective e-cohort. Nat Commun 2022 Apr 05;13(1):1812 [FREE Full text] [doi: 10.1038/s41467-022-29513-z] [Medline: 35383197]
50. Van Herck M, Goërtz YMJ, Houben-Wilke S, Machado FVC, Meys R, Delbressine JM, et al. Severe fatigue in long COVID: web-based quantitative follow-up study in members of online long COVID support groups. J Med Internet Res 2021 Sep 21;23(9):e30274 [FREE Full text] [doi: 10.2196/30274] [Medline: 34494964]
51. Kemp S. Digital 2022: France. Kepios. 2022. URL: https://datareportal.com/reports/digital-2022-france [accessed 2022-08-04]
52. Audience Internet Global en France en décembre 2021. Médiamétrie. 2022. URL: https://www.mediametrie.fr/fr/audience-internet-global-en-france-en-decembre-2021-0 [accessed 2022-11-09]

Abbreviations

BTM: Biterm Topic Modeling
LLT: lowest-level term
PT: preferred term
SOC: system organ class

XSL•FO
RenderX

Edited by T Mackey; submitted 25.05.22; peer-reviewed by L Patel, A Shaik, W Ceron, SF Tsao; comments to author 01.07.22; revised version received 19.09.22; accepted 01.10.22; published 22.11.22.

Please cite as:
Déguilhem A, Malaab J, Talmatkadi M, Renner S, Foulquié P, Fagherazzi G, Loussikian P, Marty T, Mebarki A, Texier N, Schuck S
Identifying Profiles and Symptoms of Patients With Long COVID in France: Data Mining Infodemiology Study Based on Social Media
JMIR Infodemiology 2022;2(2):e39849
URL: https://infodemiology.jmir.org/2022/2/e39849
doi:10.2196/39849
PMID:36447795

©Amélia Déguilhem, Joelle Malaab, Manissa Talmatkadi, Simon Renner, Pierre Foulquié, Guy Fagherazzi, Paul Loussikian, Tom Marty, Adel Mebarki, Nathalie Texier, Stephane Schuck. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 22.11.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Infodemic Management Using Digital Information and Knowledge Cocreation to Address COVID-19 Vaccine Hesitancy: Case Study From Ghana

Anna-Leena Lohiniva[1*], MA, MSc; Anastasiya Nurzhynska[1*], MA, PhD; Al-hassan Hudi[1], MA, MPhil; Bridget Anim[2], MSc; Da Costa Aboagye[2], MSc

[1]UNICEF Ghana Country Office, Accra, Ghana

[2]Health Promotion Division, Ghana Health Services, Accra, Ghana

[*]these authors contributed equally

Corresponding Author:
Anna-Leena Lohiniva, MA, MSc
UNICEF Ghana Country Office
House No 4-6, Rangoon Ward No 24
Accra
Ghana
Phone: 233 549761859
Fax: 233 000
Email: Lohinivaa@gmail.com

Abstract

Background: Infodemic management is an integral part of pandemic management. Ghana Health Services (GHS) together with the UNICEF (United Nations International Children's Emergency Fund) Country Office have developed a systematic process that effectively identifies, analyzes, and responds to COVID-19 and vaccine-related misinformation in Ghana.

Objective: This paper describes an infodemic management system workflow based on digital data collection, qualitative methodology, and human-centered systems to support the COVID-19 vaccine rollout in Ghana with examples of system implementation.

Methods: The infodemic management system was developed by the Health Promotion Division of the GHS and the UNICEF Country Office. It uses Talkwalker, a social listening software platform, to collect misinformation on the web. The methodology relies on qualitative data analysis and interpretation as well as knowledge cocreation to verify the findings.

Results: A multi-sectoral National Misinformation Task Force was established to implement and oversee the misinformation management system. Two members of the task force were responsible for carrying out the analysis. They used Talkwalker to find posts that include the keywords related to COVID-19 vaccine–related discussions. They then assessed the significance of the posts on the basis of the engagement rate and potential reach of the posts, negative sentiments, and contextual factors. The process continues by identifying misinformation within the posts, rating the risk of identified misinformation posts, and developing proposed responses to address them. The results of the analysis are shared weekly with the Misinformation Task Force for their review and verification to ensure that the risk assessment and responses are feasible, practical, and acceptable in the context of Ghana.

Conclusions: The paper describes an infodemic management system workflow in Ghana based on qualitative data synthesis that can be used to manage real-time infodemic responses.

(JMIR Infodemiology 2022;2(2):e37134) doi:10.2196/37134

KEYWORDS

COVID-19; infodemic management; misinformation; disinformation; social listening; pandemic preparedness; infodemiology; social media; Ghana; vaccination; qualitative methods

XSL•FO
RenderX

Introduction

The COVID-19 pandemic has led to an unprecedented global “infodemic,” which refers to an abundance of rapidly spreading fake news, misinformation, disinformation, and conspiracy theories related to the pandemic. In the ever-expanding digital world, the infodemic has become increasingly problematic as misinformation spreads rapidly through social media channels [1]. A number of recent studies highlight the negative effects of the infodemic on public perceptions of the COVID-19 pandemic [2-4] and reluctance to comply with public health guidance, including willingness to accept a COVID-19 vaccine [5-7].

Infodemic management has been acknowledged by many public health organizations as an important emerging scientific field and critical area of practice during epidemics [8]. It includes the systematic use of risk- and evidence-based analysis and approaches to manage the abundance of information and mitigate misinformation to reduce its impact on health behaviors during health emergencies. The World Health Organization (WHO) has identified a framework to manage infodemics, which includes listening to community concerns and questions, delivering high-quality health information and programming, building resilience to misinformation, and engaging and empowering communities to take positive action [9]. The WHO is encouraging countries to study and pilot strategies to combat the infodemic surge. As the nature of an infodemic is specific to place and time, it is important to establish a process that identifies context-specific solutions [9].

A growing body of literature on social media platforms has been used to address the infodemic [10-14]. Social media data derived from Facebook, Instagram, Twitter, YouTube, blogs, news sites, and messaging platforms provides useful information to pinpoint context-specific issues in real time to allow for the quick identification of public attitudes on issues of public health importance [10-14]. Gathering social media posts on the basis of a set of keywords, used by digital platforms such as Talkwalker, have become popular with organizations as a means to identify relevant misinformation and rumors [14]. Talkwalker is a dashboard tool that collects, processes, and categorizes information around keywords from social media handles. The UNICEF (United Nations International Children's Emergency Fund) is using this platform to identify misinformation and rumors in several countries [15].

Infodemic management benefits from human-centered approaches that encourage knowledge sharing and knowledge cocreation. While definitions vary widely, knowledge cocreation is essentially the bidirectional, interactive development of new knowledge created with input and perspectives from diverse stakeholders including experts and the public. It allows for the development of acceptable and practical interventions that can be better sustained than those that are developed by public health experts alone [16].

In March 2021, Ghana was the first country worldwide to receive COVID-19 vaccines from the COVAX facility. However, by the beginning of 2022, less than half of the target population of 20 million people had received at least one vaccine dose and only about 13% were fully vaccinated. To increase vaccination rates, the GHS instituted a national COVID-19 vaccination day in February 2022 and inaugurated a second campaign coinciding with Africa immunization week in March of the same year [17]. Surveys during the pandemic indicate that the hesitancy is fueled by different factors that are changing over time, such as the fear of side effects and the lack of trust in the vaccines [18,19]. Similar to many other countries, Ghana has witnessed the widespread transmission of misinformation during the pandemic on the web and offline, including the period during promotion of COVID-19 vaccines [20]. For example, early in the pandemic, COVID-19 misinformation included myths that Black people had some immunity against COVID-19, that the hot climate in Africa reduced the replication of the virus, that COVID-19 was only life-threatening in older people, that drinking “akpetashi”—a locally prepared alcoholic drink—cures COVID-19, and that COVID-19 was a biological weapon to target developed economies; all of which had the potential to reduce risk perception among Ghanaians and contribute to lack of compliance with pandemic measures. There have also been various COVID-19 conspiracy theories identified across Africa on various social media platforms, including those in Ghana, ranging from SARS-CoV-2 having been created as a biological weapon to disrupt the economic power of China against other economically prosperous nations including the United States, to the use of local herbs or products being able to cure the disease [21]. In addition, misinformation has fueled mistrust toward the government, particularly in closed social media platforms such as WhatsApp, which has made risk communication efforts challenging for health authorities during the pandemic [22,23].

Social listening to web-based sources is important in Ghana as the number of social media users has increased significantly in recent years. Currently, over 50% of the population has access to the internet and 140% of the population has a mobile phone connection. In early 2022, there were approximately 8.80 million social media users in Ghana, which is approximately 27.4% of the total 32 million total population. WhatsApp is used by almost 90% and Facebook by over 70% of social media users followed by Instagram by approximately 60% of the users and Twitter and Snapchat are used by approximately 45% of the users [24]. Twitter is known to be used by those who want to generate political discussion in Ghana. The African media agency reports that almost as many women as men use the internet in Ghana, and men are 6% more likely to have a presence on the web than women [25]. However, there are likely to be disparities in the usage of social media between urban and rural populations in the country [24].

The Health Promotion Division of Ghana Health Services (GHS) together with the UNICEF Country Office has established an infodemic management system that combines information identification through the Talkwalker platform and knowledge cocreation among a National Misinformation Taskforce to verify potential misinformation and respond to it appropriately. The system was created to strengthen COVID-19 vaccine programming and to combat vaccine hesitancy, which is defined by WHO as a “delay in acceptance or refusal of vaccines despite availability of vaccination services” [26]. This paper describes

XSL•FO
RenderX

the methodology of the infodemic management system in Ghana, which combines digital and human-centered approaches. Some concrete examples are also given to demonstrate how infodemic management operates. The findings of the study can be used to apply the system in other countries that plan to conduct social listening.

Methods

The infodemic management system was developed by the Health Promotion Division of the GHS and UNICEF Ghana Country Office to effectively identify, analyze, and cocreate content to respond to misinformation during the pandemic. The objectives were to improve compliance with public health safety measures, support COVID-19 vaccine programming, and identify factors that may increase vaccine hesitancy and lead to vaccine refusal.

The data collection system was based on Talkwalker, a commercial social listening software platform. It uses machine learning and artificial intelligence to consolidate publicly visible occurrences of given keywords on the internet. Talkwalker functions like a search engine and provides the ability to filter, contextualize, export, and analyze large data sets. It gathers Ghana-specific COVID-19–related posts from open Twitter, YouTube, and other websites by monitoring keywords, phrases, and hashtags. It categorizes relevant posts by sentiment: neural, positive, or negative. Negative posts are of particular interest as they may contain rumors, misinformation, or disinformation. In addition, the platform includes a feature to categorize data as misinformation; another point of interest to infodemic management. It deepens the understanding of the circulating narratives by aggregating numbers related to the total reach, engagement, and demographic information about those who are engaged in these discussions. The limitation of the tool is that it cannot access conversations on Facebook, Instagram, and WhatsApp owing to privacy restrictions. Approximately 70% of the posts retrieved by Talkwalker in relation to COVID-19–specific information in Ghana are published by men, almost half of which were published by adults aged 25-34 years.

The analysis utilizes qualitative methods to classify a post as misinformation and to assess the risk level of the post. It uses Talkwalker algorithms to identify posts, but the final assessment is based on assessing the post given the local context. In particular, the risk level of particular misinformation requires a qualitative assessment of the situation using applied content analysis [27].

The methodology also relies on knowledge cocreation, which is referred to as collaborative knowledge generation by various stakeholders. Knowledge cocreation is a participatory approach to enhance the value and reliability of outcomes and ensure that they benefit all parties [28,29]. The Misinformation Management Task Force cocreates by assessing the risk level and the proposed responses to address misinformation to ensure they are feasible, practical, and acceptable in the context of Ghana.

Results

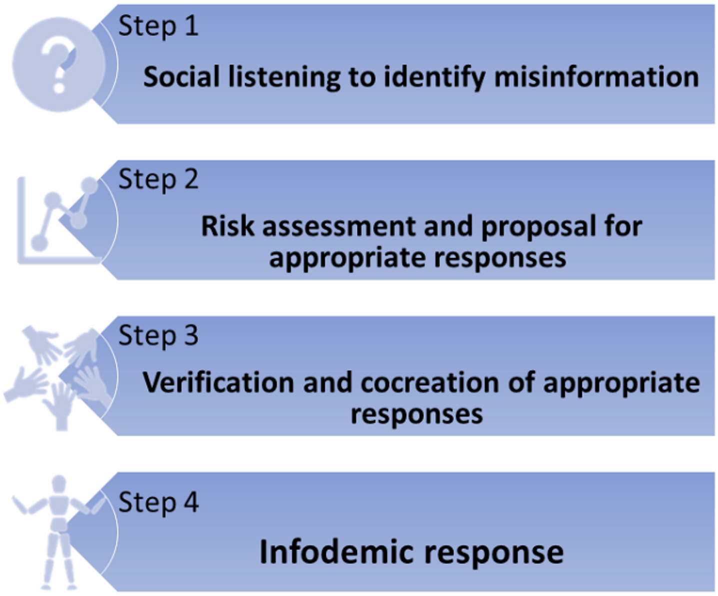
Working Modalities of the Infodemic Management System

A multi-sectoral National Misinformation Task Force was established to implement and oversee the process developed by the GHS, which also appointed members for the task force. The task force was established on the basis of the membership of an existing task force of Risk Communication and Social Mobilization experts and expanded to include other public health experts, media, development partners, and an organization of Ghana fact-checkers, UNICEF, and other critical partners. Since the beginning of the pandemic, the task force has held biweekly web-based meetings with approximately 20 experts. The head of the Health Promotion Division of GHS is the chairperson of the group. UNICEF has provided technical assistance to the group, including capacity-building training on how to use Talkwalker to identify misinformation, how to assess the risk level of misinformation, and how to respond to misinformation.

The infodemic management system includes 4 interlinked steps that are carried out by selected members of the task force on a biweekly basis, including social listening to identify misinformation, risk assessment, and proposal for appropriate information; verification; cocreation of appropriate responses; and infodemic response. Figure 1 shows the workflow of the infodemic management system.

XSL•FO
RenderX

Figure 1. Infodemic management system.



Step 1: Social Listening to Identify Misinformation

The first step is to identify and analyze misinformation through the Talkwalker social media and web-based monitoring platform. Two members of the task force are responsible for carrying out the analysis. They used Talkwalker to determine the number of results (posts that include the keywords for COVID-19 vaccine–related discussions by Talkwalker) during a specific period of time, which is usually a week. The analysts gained an overall understanding of the results by looking into the demographics of those who have generated the results (gender and age) and creating a word cloud to see how the results are thematized. Then, the analysts assessed the engagement rate and potential reach of the results, followed by a review of the results that convey negative sentiments to determine the significance of each individual result. The analyst read each headline of the posts (results) or the entire tweet to decide if it contains misinformation to be included in the analysis. If so, the entire post is extracted from Talkwalker and pasted into a document for further risk analysis and response. Then, the analysts reviewed the rest of the results because even if the reach or the engagement is not high, a result may be potentially risky in the context of Ghana. For example, a post may relate to a historical or political event that is significant in the context of Ghana. At the end of the analysis, the analysts had a list of posts extracted from Talkwalker, which requires verification from public health experts and fact-checkers. If confirmed as misinformation, they are included in the list of misinformation.

Step 2: Risk Assessment and Proposal for Appropriate Responses

The second step includes assessing the risk level of all the posts that were classified as misinformation based on the UNICEF risk assessment matrix that classifies misinformation into low, medium, or high risk levels based on 5 criteria [30]. See Table 1 for the UNICEF risk assessment matrix. If the analysts are able to link a post to more than 2 criteria in a particular risk level, it is categorized as such. If the analysts are able to relate a post to several levels, the post is classified on the basis of the expert opinion of the analyst based on their broad qualitative analysis of contextual and cultural factors surrounding the post. Once the post has a defined risk level, a set of responses are proposed on the basis of common risk communication and misinformation management best practices [29,31-34].

XSL•FO
RenderX

Table 1. UNICEF (United Nations International Children's Emergency Fund) misinformation management matrix (adapted from UNICEF's Vaccine Misinformation Management Guide [30], which is published under Creative Commons Attribution 4.0 International License [35]).

Indicator	Low risk	Medium risk	High risk
Risk to vaccine hesitancy and demand	Low risk to vaccine demand	Potential to trigger vaccine hesitancy	Potential to lead to vaccine refusal
Reach and scope of misinformation	Limited potential to reach scope	Moderate potential to reach scope	Wide cross-country reach or scope
Likelihood of issue spread or escalation	Unlikely to spread on the web or in the community	Spreading on the web or in communities	Spreading rapidly on the web and in community
Capacity to respond	Strong messaging and capacity in place	Some existing messaging and resources to manage crisis	Limited existing messaging and resources to manage crisis
General public trust	Remaining trust in government health services and vaccines	Reduces trust in government health services and vaccines	Increasing mistrust toward government health services and vaccines
Response	Monitor closely and consider prebunking	Debunk and raise trusted voices	Debunk and raise trusted voices

Step 3: Verification and Cocreation of Appropriate Responses

The analysts presented the analysis in a PowerPoint presentation during the task force meetings. The presentation is discussed jointly with the original post, analysis of the risk level, and proposed response. The discussions are the core of the knowledge cocreation during which the task force members view the contents of the posts and the related risk assessment and proposed interventions and actions to ensure that they are appropriate, culturally acceptable, and practical in the context of Ghana [36]. Decisions are based on consensus among the members of the task force [37].

Step 4: Infodemic Response

The evidence and the systematic process of verification, risk assessment, and response proposal are provided to the management for approval. There is also a Message Box containing prepared responses to frequently asked questions including rumors, misinformation, and disinformation. Responses may include press releases, social media posts, and direct communication, among others. The GHS is responsible for implementing the response.

Examples of Implementing the Misinformation Management Workflow

Example 1: Negative Attitude of Health Care Workers

Through Talkwalker, the analysts identified a post that included complaints about negative attitudes of health care staff toward patients in one particular health center. The owner of the post was a young social media influencer with over 10,000 followers, many of whom also actively retweeted the post within their own networks. Analysts considered that a risk, though the allegations of the post itself were not considered particularly threatening as it related to one particular health center. The analysts suggested taking localized action to address the issues with that particular health center. During Misinformation Task Force cocreation, the members rated the post as medium risk in consensus and agreed to take targeted action by training all staff members of that particular health care center in service delivery and customer care.

Example 2: Misinformation About the Side Effects of COVID-19 Vaccines

Through Talkwalker, analysts identified a post on popular news sites about an interview with a local premier league football team coach in which he claimed that the team had lost a game because of their weak physical status after they received the COVID-19 vaccine. Analysts noted that the engagement rates were high and this spread rapidly on social media platforms, particularly on Twitter. The analysts rated it as high risk because it was from a national web-based news outlet, because the interview was with a local celebrity, and because football is a popular sport in Ghana and many fans may be influenced by the post. The analysts suggested taking action on the web in the same news outlet where the post was published. The task force agreed with the high risk level but instead of responding on the web, they decided to contact the football coach directly to gain clarification on his statement, to educate him about the side effects of the COVID-19 vaccine, the Adverse Effects Following Immunization protocol, and, most importantly, to recruit him as a vaccine supporter and encourage him to speak publicly to deliver pro–COVID-19 vaccine messages.

Example 3: Disinformation About the Alleged Lethal Nature of the COVID-19 Vaccine

Using Talkwalker, the analysts identified a retweet about a post that claimed to be published by an award-winning doctor who warned that those who take the COVID-19 vaccine will die within 2 years. The analysts checked with Ghana fact-checkers about the post and learned that it had been viral worldwide for some time already and it had been fact-checked as false. The analysis rated the post as high risk owing to the global spread and because it referenced death. They proposed to the task force that action should be taken to clarify that this was fake news. The task force agreed that the post should be rated as high risk because it described a severe adverse effect, which is known to promote vaccine hesitancy, and because the disinformation claims to have originated by a doctor—a highly respected profession in Ghana—which could contribute to the spread of this misinformation in Ghana. The task force decided to circulate the post with a “Fake News” stamp across various GHS social media channels. At the same time, they posted factual

XSL•FO
RenderX

information about the COVID-19 vaccine and had it circulated across social media channels.

Example 4: Mistrust Toward the COVID-19 Vaccination Program and Health Authorities

Through Talkwalker, the analysts identified a Tweet that accused the GHS of not sharing information about COVID-19–related mortality on its website in a timely manner. It was created by a political activist, with over 20,000 followers, who is known to initiate discussion against the government. The analysts assessed the risk as medium as they did not find any tangible accusations in the post and suggested that no action be taken at this time as a response would only bring more attention to the post. The task force assessed the risk level as high because the post could encourage more politically driven rumors to further spread mistrust against the government. The task force response included issuing an official press release that clarified the data verification process of any statistics displayed by the GHS on its website and highlighting the importance of publishing accurate information. A summary of the examples is provided in Table 2.

Table 2. Examples of infodemic management systems in Ghana.

Talkwalker	Risk assessment	Action	Expected outcome
The attitude of nursing staff in xx health unit is not appropriate. People are not willing to get their COVID-19 vaccines at that location.	Medium risk: potential to increase COVID-19 vaccine hesitancy as people are unwilling to get the vaccine in the health unit. It also impacts the uptake of any services at that particular health unit.	Training of the health care unit staff members in customer service.	Improved health systems through more service-oriented staff.
Misinformation was spread by a national football coach who claimed having lost a game because all players were vaccinated and it made them weak.	High risk: football is a popular sport in Ghana and the coach is seen as a local celebrity. Accordingly, the misinformation can spread rapidly among football fans.	Personal contact with the football coach to understand his claims, provide information about the COVID-19 vaccine, and encourage him to publicly advocate for the vaccine.	Gaining the football club as a vaccine supporter that can disseminate positive COVID-19 vaccine messages as needed.
Disinformation by an alleged doctor that all who have taken the COVID-19 vaccine will die in 2 years.	High risk: potential to increase vaccine hesitancy and contribute to refusal to take the vaccine because it was from an alleged doctor and relates to the severe adverse effect of the vaccine. The disinformation was also circulating widely and rapidly.	Fact-check and, once verified fake, circulate the news with a fake news stamp. Simultaneously run factual information about the COVID-19 vaccine across different social media platforms.	Stopped circulation of the fake news.
Rumors that GHS is faking COVID-19 death statistics as the numbers on the website do not correspond with numbers available on social media.	High risk: potential to decrease trust toward the COVID-19 vaccination program and Ghana health services.	Issue a press release and explain that sometimes there is a lag in GHS numbers owing to the verification process to ensure that the numbers are correct and highlight how important it is for GHD to verify any information before publishing it on the website.	Improved trust towards GHS reporting procedures.

Discussion

Principal Findings

This paper described an infodemic management system developed and implemented in Ghana. The system relies on data collection through a digital platform and on human-centered approaches to verify the findings with appropriate response mechanisms. The system has been used to identify COVID-19 misinformation, disinformation, and rumors, which were addressed in a timely manner.

The implementation of the infodemic management system in Ghana highlights the critical role of qualitative inquiry in social listening as it allows for a greater understanding of the positions, perceptions, and potential misinformation and disinformation among population groups in order to assess the potential risk and take appropriate action in a timely and targeted manner. For example, a post from a football coach may not be significant in a country where football is less popular, but in the context of Ghana, it was assessed as a high risk that has the potential to spread fast and raise high emotions. Talkwalker cannot carry out such an interpretation, which aligns with a number of studies that have pointed out the limitations of machines. Although machine learning methods have been developed to solve real-world problems, they are not sufficient by themselves in critical decision-making approaches [38,39]. Digital platforms still have limitations to interpret and contextualize data [40]. In addition, digital platforms cannot commonly identify whether the information is misinformation or disinformation; a critical differentiation essential to infodemic response processes [29]. The use of a digital platform together with a qualitative analysis aligns with a UNICEF MENA (UNICEF in the Middle East and North Africa region) case study that showed the importance of involving human minds in digital data interpretation to create a shared sense of reality that fosters engagement and connections with the communities and facilitates risk communication and community engagement responses [15].

XSL•FO
RenderX

The implementation of the infodemic management system in Ghana has also highlighted that knowledge cocreation can be implemented even in crisis situations. Knowledge cocreation has been identified in a number of studies as an effective approach to discover, share, and blend knowledge for practical use, allowing stakeholders to learn about the applied implications of knowledge use and to collectively create actionable recommendations [41,42]. Cocreation can act as capacity-building for those who participate [43]. Ideally, cocreation will allow the task force members to build their misinformation management skills so that in the future, the system can run without support from external stakeholders such as UNICEF. A systematic approach to detecting, analyzing, and responding to an infodemic also often facilitates official approvals for press releases or other responses [12]. The Misinformation Task Force was developed by merging an existing working group with the task force instead of creating a new structure, which has been a successful model in other countries such as Finland where social listening was built into existing working groups [12].

The infodemic management system in Ghana also has limitations. The Talkwalker posts and interactions are mainly published by men and young adults, excluding the voices of women, youth, and older people. In addition, here are still significant numbers of people, particularly in vulnerable populations such as low-income individuals and those who cannot read and write, who are not reached by digital platforms [44]. Accordingly, there is a need to merge offline listening systems with the infodemic management system, such as the perspectives of community leaders, who are highly respected in Ghana, and women, who play a key role in the vaccination decision-making of their children in Ghana [45,46]. Moreover, Talkwalker does not include Facebook or WhatsApp, which are two of the most popular social media sites in Ghana [21]. The qualitative inquiry of the system has also weaknesses. The process of identifying misinformation relies on the analyst's decision and is based on their own reflectivity including their worldview, beliefs, attitudes, and skills [47], which may present bias as to what type of information is determined as disinformation or misinformation and how significant risk it is perceived. Cocreation methods can mitigate bias [48]. Other strategies to minimize bias should be considered, such as engaging more task force members with different backgrounds in the assessment and analysis process [48]. In the future, studies should be conducted to measure the impact of the system and the various infodemic response strategies implemented by the task force.

Conclusions

The paper has described an infodemic management system workflow based on a mix of digital and human-centered methods, including effective social listening through a social media management platform, qualitative analysis process, and cocreation through a national task force of experts, resulting in context-specific, real-time infodemic responses.

Acknowledgments

The authors wish to thank the Ghana Health Services (Health Promotion Division and Expanded Programme on Immunization), Ministry of Information in Ghana, Food and Drug Authority in Ghana, FactSpace West Africa (GhanaFacts), USAID Breakthrough Project, and the World Health Organization for their support for the project. The authors wish to extend special thanks to Joel Abekuliya.

Conflicts of Interest

None declared.

References

1. De Paor S, Heravi B. Information literacy and fake news: How the field of librarianship can help combat the epidemic of fake news. J Acad Librariansh 2020 Sep;46(5):102218 [FREE Full text] [doi: 10.1016/j.acalib.2020.102218]
2. Geldsetzer P. Knowledge and perceptions of COVID-19 among the general public in the United States and the United Kingdom: a cross-sectional online survey. Ann Intern Med 2020 Jul 21;173(2):157-160. [doi: 10.7326/m20-0912]
3. Islam M, Sarkar T, Khan S, Mostofa Kamal AH, Hasan S, 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. Kim HK, Ahn J, Atkinson L, Kahlor LA. Effects of COVID-19 misinformation on information seeking, avoidance, and processing: a multicountry comparative study. Sci Commun 2020 Sep 13;42(5):586-615. [doi: 10.1177/1075547020959670]
5. 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]
6. 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]
7. 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]

XSL•FO
RenderX

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. An ad hoc WHO technical consultation managing the COVID-19 infodemic: call for action. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240010314 [accessed 2022-01-31]
10. Shah Z, Surian D, Dyda A, Coiera E, Mandl KD, Dunn AG. Automatically appraising the credibility of vaccine-related web pages shared on social media: a twitter surveillance study. J Med Internet Res 2019 Nov 04;21(11):e14007 [FREE Full text] [doi: 10.2196/14007] [Medline: 31682571]
11. 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]
12. 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) [FREE Full text] [doi: 10.2807/1560-7917.ES.2020.25.13.2000317] [Medline: 32265008]
13. Steffens MS, Dunn AG, Wiley KE, Leask J. How organisations promoting vaccination respond to misinformation on social media: a qualitative investigation. BMC Public Health 2019 Oct 23;19(1):1348 [FREE Full text] [doi: 10.1186/s12889-019-7659-3] [Medline: 31640660]
14. Tran HTT, Lu S, Tran HTT, Nguyen BV. Social media insights during the COVID-19 pandemic: infodemiology study using big data. JMIR Med Inform 2021 Jul 16;9(7):e27116 [FREE Full text] [doi: 10.2196/27116] [Medline: 34152994]
15. Tackling the infodemic using conversational insights - Case Study of UNICEF Response in MENA. United Nations Office for the Coordination of Humanitarian Affairs. 2021. URL: https://reliefweb.int/report/world/tackling-infodemic-using-conversational-insights-case-study-unicef-response-mena [accessed 2022-01-31]
16. Abbate T, Codini AP, Aquilani B. Knowledge co-creation in Open Innovation Digital Platforms: processes, tools and services. JBIM 2019 Aug 05;34(7):1434-1447. [doi: 10.1108/jbim-09-2018-0276]
17. Ghana finds success in COVID-19 mass vaccination campaigns. World Health Organization. 2022. URL: https://www.afro.who.int/countries/ghana/news/ghana-finds-success-covid-19-mass-vaccination-campaigns [accessed 2022-05-24]
18. Botwe B, Antwi W, Adusei J, Mayeden R, Akudjedu T, Sule S. COVID-19 vaccine hesitancy concerns: Findings from a Ghana clinical radiography workforce survey. Radiography (Lond) 2022 May;28(2):537-544 [FREE Full text] [doi: 10.1016/j.radi.2021.09.015] [Medline: 34654631]
19. Alhassan RK, Aberese-Ako M, Doegah PT, Immurana M, Dalaba MA, Manyeh AK, et al. COVID-19 vaccine hesitancy among the adult population in Ghana: evidence from a pre-vaccination rollout survey. Trop Med Health 2021 Dec 16;49(1):96 [FREE Full text] [doi: 10.1186/s41182-021-00357-5] [Medline: 34915939]
20. Durizzo K, Asiedu E, Van der Merwe A, Van Niekerk A, Günther I. Managing the COVID-19 pandemic in poor urban neighborhoods: The case of Accra and Johannesburg. World Dev 2021 Jan;137:105175 [FREE Full text] [doi: 10.1016/j.worlddev.2020.105175] [Medline: 32904458]
21. 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]
22. Ogola G. Africa and the Covid-19 iformation framing crisis. MaC 2020 Jun 25;8(2):440-443. [doi: 10.17645/mac.v8i2.3223]
23. Tabong P, Segtub M. Misconceptions, misinformation and politics of COVID-19 on social media: a multi-level analysis in Ghana. Front Commun 2021 May 5;6 [FREE Full text] [doi: 10.3389/fcomm.2021.613794]
24. DIGITAL 2022: GHANA. Kepios. URL: https://datareportal.com/reports/digital-2022-ghana [accessed 2022-05-24]
25. Women's Rights Online: Closing the digital gender gap for a more equal world. World Wide Web Foundation. 2020. URL: http://webfoundation.org/docs/2020/10/Womens-Rights-Online-Report-1.pdf [accessed 2022-06-24]
26. 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]
27. Iosifidis P, Nicoli N. The battle to end fake news: A qualitative content analysis of announcements on how it combats disinformation. Int Commun Gaz 2019 Oct 10;82(1):60-81. [doi: 10.1177/1748048519880729]
28. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: literature review and case study. Milbank Q 2016 Jun 06;94(2):392-429 [FREE Full text] [doi: 10.1111/1468-0009.12197] [Medline: 27265562]
29. Godinho MA, Borda A, Kariotis T, Molnar A, Kostkova P, Liaw S. Knowledge co-creation in participatory policy and practice: Building community through data-driven direct democracy. Big Data & Society 2021 May 21;8(1):205395172110194. [doi: 10.1177/20539517211019430]
30. Vaccine Misinformation Management Guide. UNICEF. URL: https://vaccinemisinformation.guide/ [accessed 2022-05-24]
31. The COVID-19 Vaccine Communication Handbook. European Union. URL: https://hackmd.io/@scibehC19vax/home [accessed 2022-05-24]
32. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]

XSL•FO
RenderX

33. Petit V. The Behavioral Drivers Model. UNICEF. URL: https://www.unicef.org/mena/media/5586/file/The_Behavioural_Drivers_Model_0.pdf%20.pdf [accessed 2022-01-31]
34. Razai MS, Oakeshott P, Esmail A, Wiysonge CS, Viswanath K, Mills MC. COVID-19 vaccine hesitancy: the five Cs to tackle behavioural and sociodemographic factors. J R Soc Med 2021 Jun 02;114(6):295-298 [FREE Full text] [doi: 10.1177/01410768211018951] [Medline: 34077688]
35. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2022-07-05]
36. Xu Y, Lai KK, Leung WKJ. A consensus-based decision model for assessing the health systems. PLoS One 2020 Aug 18;15(8):e0237892 [FREE Full text] [doi: 10.1371/journal.pone.0237892] [Medline: 32810150]
37. Maadi M, Akbarzadeh Khorshidi H, Aickelin U. A Review on human-AI interaction in machine learning and insights for medical applications. Int J Environ Res Public Health 2021 Feb 22;18(4):2121 [FREE Full text] [doi: 10.3390/ijerph18042121] [Medline: 33671609]
38. Zerilli J, Knott A, Maclaurin J, Gavaghan C. Algorithmic decision-making and the control problem. Minds Mach 2019 Dec 11;29(4):555-578. [doi: 10.1007/s11023-019-09513-7]
39. Wu J, Liu TYA, Hsu W, Ho JH, Lee C. Performance and limitation of machine learning algorithms for diabetic retinopathy screening: meta-analysis. J Med Internet Res 2021 Jul 03;23(7):e23863 [FREE Full text] [doi: 10.2196/23863] [Medline: 34407500]
40. UNESCO. URL: https://en.unesco.org/sites/default/files/journalism_fake_news_disinformation_print_friendly_0.pdf [accessed 2022-01-31]
41. Langley J, Wolstenholme D, Cooke J. 'Collective making' as knowledge mobilisation: the contribution of participatory design in the co-creation of knowledge in healthcare. BMC Health Serv Res 2018 Jul 25;18(1):585 [FREE Full text] [doi: 10.1186/s12913-018-3397-y] [Medline: 30045726]
42. Langley J, Wolstenholme D, Cooke J. 'Collective making' as knowledge mobilisation: the contribution of participatory design in the co-creation of knowledge in healthcare. BMC Health Serv Res 2018 Jul 25;18(1):585 [FREE Full text] [doi: 10.1186/s12913-018-3397-y] [Medline: 30045726]
43. Sasu D. Number of active social media users in Ghana 2017-2020. Statista. URL: https://www.statista.com/statistics/1171445/number-of-social-media-users-ghana/ [accessed 2022-01-31]
44. Amo-Adjei J, Nurzhynska A, Essuman R, Lohiniva A. Trust and willingness towards COVID-19 vaccine uptake: a mixed-method study in Ghana, 2021. Arch Public Health 2022 Feb 21;80(1):64 [FREE Full text] [doi: 10.1186/s13690-022-00827-0] [Medline: 35189963]
45. Fuseini K, Kalule-Sabiti I, Lwanga C. Dynamics of women’s autonomy in household decision-making in Ghana. J Comp Fam Stud 2019 Dec;50(4):293-312. [doi: 10.3138/jcfs.50.4.002]
46. Wadams M, Park T. Qualitative research in correctional settings: researcher bias, western ideological influences, and social justice. J Forensic Nurs 2018;14(2):72-79. [doi: 10.1097/JFN.0000000000000199] [Medline: 29781967]
47. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. Plast Reconstr Surg 2010;126(2):619-625. [doi: 10.1097/prs.0b013e3181de24bc]
48. Pearce T, Maple M, Shakeshaft A, Wayland S, McKay K. What is the co-creation of new knowledge? A content analysis and proposed definition for health interventions. Int J Environ Res Public Health 2020 Mar 26;17(7):2229 [FREE Full text] [doi: 10.3390/ijerph17072229] [Medline: 32224998]

Abbreviations

GHS: Ghana Health Services
UNICEF: The United Nations International Children's Emergency Fund
UNICEF MENA: The United Nations International Children's Emergency Fund in the Middle East and North Africa region
WHO: World Health Organization

Edited by T Purnat; submitted 08.02.22; peer-reviewed by M Steffens, AM Auvinen; comments to author 10.05.22; revised version received 24.05.22; accepted 10.06.22; published 12.07.22.

Please cite as:
Lohiniva AL, Nurzhynska A, Hudi AH, Anim B, Aboagye DC
Infodemic Management Using Digital Information and Knowledge Cocreation to Address COVID-19 Vaccine Hesitancy: Case Study From Ghana
JMIR Infodemiology 2022;2(2):e37134
URL: https://infodemiology.jmir.org/2022/2/e37134
doi:10.2196/37134
PMID:35854815

XSL•FO
RenderX

©Anna-Leena Lohiniva, Anastasiya Nurzhynska, Al-hassan Hudi, Bridget Anim, Da Costa Aboagye. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 12.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

COVID-19 Misinformation Detection: Machine-Learned Solutions to the Infodemic

Nikhil Kolluri[1*], BS; Yunong Liu[2*]; Dhiraj Murthy[3], BA, MSc, DPhil

[1]Computational Media Lab, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, United States

[2]School of Engineering, College of Science and Engineering, University of Edinburgh, Edinburgh, United Kingdom

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

[*]these authors contributed equally

Corresponding Author:
Dhiraj Murthy, BA, MSc, DPhil
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 volume of COVID-19–related misinformation has long exceeded the resources available to fact checkers to effectively mitigate its ill effects. Automated and web-based approaches can provide effective deterrents to online misinformation. Machine learning–based methods have achieved robust performance on text classification tasks, including potentially low-quality-news credibility assessment. Despite the progress of initial, rapid interventions, the enormity of COVID-19–related misinformation continues to overwhelm fact checkers. Therefore, improvement in automated and machine-learned methods for an infodemic response is urgently needed.

Objective: The aim of this study was to achieve improvement in automated and machine-learned methods for an infodemic response.

Methods: We evaluated three strategies for training a machine-learning model to determine the highest model performance: (1) COVID-19–related fact-checked data only, (2) general fact-checked data only, and (3) combined COVID-19 and general fact-checked data. We created two COVID-19–related misinformation data sets from fact-checked "false" content combined with programmatically retrieved "true" content. The first set contained ~7000 entries from July to August 2020, and the second contained ~31,000 entries from January 2020 to June 2022. We crowdsourced 31,441 votes to human label the first data set.

Results: The models achieved an accuracy of 96.55% and 94.56% on the first and second external validation data set, respectively. Our best-performing model was developed using COVID-19–specific content. We were able to successfully develop combined models that outperformed human votes of misinformation. Specifically, when we blended our model predictions with human votes, the highest accuracy we achieved on the first external validation data set was 99.1%. When we considered outputs where the machine-learning model agreed with human votes, we achieved accuracies up to 98.59% on the first validation data set. This outperformed human votes alone with an accuracy of only 73%.

Conclusions: External validation accuracies of 96.55% and 94.56% are evidence that machine learning can produce superior results for the difficult task of classifying the veracity of COVID-19 content. Pretrained language models performed best when fine-tuned on a topic-specific data set, while other models achieved their best accuracy when fine-tuned on a combination of topic-specific and general-topic data sets. Crucially, our study found that blended models, trained/fine-tuned on general-topic content with crowdsourced data, improved our models' accuracies up to 99.7%. The successful use of crowdsourced data can increase the accuracy of models in situations when expert-labeled data are scarce. The 98.59% accuracy on a "high-confidence" subsection comprised of machine-learned and human labels suggests that crowdsourced votes can optimize machine-learned labels to improve accuracy above human-only levels. These results support the utility of supervised machine learning to deter and combat future health-related disinformation.

XSL•FO
RenderX

(JMIR Infodemiology 2022;2(2):e38756) doi:10.2196/38756

KEYWORDS

COVID-19; misinformation; machine learning; fact-checking; infodemiology; infodemic management; model performance; model accuracy; content analysis

Introduction

Background

Low information quality has led to adverse health outcomes for individuals during the COVID-19 pandemic [1-3]. Claims were being made on social media of dangerous home remedies and perceived preventative measures (eg, gargling with bleach-infused water) [4]. Low-quality and biased sources of information can be more alluring to some, as they easily capture attention and offer simpler solutions with unambiguous evidence. Due to their persuasive, "simple" messaging [2], these sources can appear more convincing to some because they confirm existing biases or better align with ideological narratives. Information veracity around COVID-19 is fundamentally important to the health outcomes of individuals worldwide [5]. For example, the information that has been circulated in social media spaces that masks do not prevent COVID-19 transmission or that wearing a mask is unhealthy [6] has been a major issue in terms of increased cases in the United States, but also in India, Brazil, and Turkey. Social media represent a key avenue where COVID-19–related disinformation and misinformation have been disseminated [7].

To tackle this misinformation, manual intervention alone is insufficient. Indeed, in the first quarter of 2020 alone, English-language fact checks of COVID-19–related content jumped 900% [8]. Despite checks increasing, there are a limited number of fact checkers. Moreover, they cannot check the high volume of content that needs evaluation [8]. Thus, creating any interventions for providing automated solutions to evaluate the credibility of COVID-19–related content being circulated remains critical.

In this study, we importantly compared COVID-19–related, general, and combined data sets for veracity classification applications, and developed a successful bidirectional long short-term memory (Bi-LSTM) machine-learning model (achieving internal and external validation accuracies of 93% and 75%, respectively). When crowdsourced human labels agreed with machine-learned outputs, the accuracy of 90% exceeded that of either approach alone. Our study provides critical, empirical evidence that small amounts of human labeling and machine learning can be an effective infodemic response to health disinformation.

Misinformation and Disinformation

Misinformation is defined as "incorrect or misleading information" [9]. For example, a family member likely does not have intent to mislead you when they provide misinformation about politics or health, as they believe what they are sharing is actually true. Although misinformation is not inherently intentional, it can also cause real harm, as seen with COVID-19 misinformation being attributed to fatalities [10]. Disinformation refers to intentionally and surreptitiously disseminated false information aiming to obscure the truth [11]. Although both words refer to incorrect or inaccurate information, only disinformation is intentionally incorrect. A well-known example of a disinformation campaign is the 2016 Russian hacking of the Hillary Clinton campaign, and distribution of politically damaging propaganda on Facebook, Twitter, YouTube, and Instagram [12]. Russia's social media disinformation campaign was found to have likely influenced the 2016 US election [13].

COVID-19 and Social Media

Early COVID-19–related research was critical in documenting keywords, topics that were emerging, as well as temporal patterns [14-16]. Some work specifically highlighted instances of rumors [17], racism against individuals of Asian descent, and released data sets [18]. Other studies documented COVID-19–related misinformation and disinformation [8,19]. This work found that misinformation was widely diffused, which included that neem leaves can cure coronavirus [20], certain ethnic and racial groups were immune (particularly if they had darker skin), individuals in warmer countries would not be affected, and the disease was no more harmful than the common flu [21].

Other studies used machine-learned methods to try to classify misinformation and disinformation that was being circulated online [22-24]. By training machine-learned classifiers on labeled misinformation and disinformation data sets, these approaches were able to achieve accuracy ranging from 16.7% to 96% as measured by F1 scores. Early work was mostly focused on deploying rapid results rather than optimizing classifiers for the best accuracy to COVID-19–specific misinformation and disinformation. The presumption was that there would be a reasonable similarity of misinformation detection approaches more broadly with the misinformation being spread during COVID-19. As studies emerged, it became clear that COVID-19–specific data sets and platforms were needed.

COVID-19–Related Misinformation Data Sets, Machine Learning, and Automated Detection

Due to the vast amount of COVID-19–related information circulating in public domains, automatic machine-learned identification and classification remains a critical method for detecting harmful content at scale. Six machine-learning algorithms with ensemble learning were used to study COVID-19–related Twitter data [25]. Combinations of several machine-learning approaches and natural language processing (NLP) are being used to develop large-scale misinformation detection. For example, ReCOVery, a repository for COVID-19 news credibility checking, evaluates various machine-learned methods [26]. One of the key issues hindering machine-learned methods remains the lack of large, verified, and labeled misinformation data sets [27]. A reason for this lack is that

XSL•FO
RenderX

robust labeled data sets require involvement of humans with specific domain knowledge. Moreover, misinformation is a diverse dynamic phenomenon that changes rapidly [28]. Additionally, there remains a dearth of automated solutions that are scalable to incorporate content from multiple platforms. Although global studies indicate a high prevalence of misinformation (which disproportionately impacts low-income countries) [29], currently available data sets may not be large enough to be scalable [30].

To help address this gap, FakeCovid is a database of 5182 fact-checked news articles that uses 40 languages from 105 countries and classifies data using machine learning [31]. COVIDLIES is another database comprising 6761 expert-annotated COVID-19–related tweets [22]. Effective NLP methodology has also been used for detecting COVID-19 misinformation through YouTube videos by studying user comments [23]. More than 100 million Twitter messages have been collected and classified to build the "Infodemic Risk Index" to estimate the magnitude of exposure to misinformation across various regions and countries [2]. A manually labeled data set related to COVID-19 misinformation was released [32]. COVID-19–specific data sets have also been developed with non-English–language content, including Arabic [33], Portuguese [34], Italian [35], Chinese [36], and multiple Indic languages [37]. Machine-learned approaches have also been developed to complement manually labeled data sets related to COVID-19 [35].

Machine-Learning Methods for Text Classification

NLP applications for text classification include news categorization, sentiment analysis, emotion detection, and authorship attribution [38,39]. Most classical machine-learning models in text classification tasks extract features (eg, bag of words) from the documents and then feed them to a classifier to make a prediction [38]. Note that, following prior works [40], we use the word "classical" to describe traditional supervised and unsupervised machine-learning methods.

The classical machine-learning models have some limitations, including tedious feature engineering in the process to extract hand-crafted features and the fact that they are difficult to generalize to new tasks due to their strong reliance on domain knowledge when designing features [38]. Deep-learning models make use of embedding models to map text into a feature vector with lower dimensions, thus limiting the need to rely on hand-crafted features (which often require domain knowledge) [38]. ELMo [41], a 3-layer Bi-LSTM model with 93 million parameters developed in 2017, achieved better performance than the previous most popular word2vec models [42,43] developed by Google in 2013. In 2018, OpenAI developed Generative Pre-trained Transformer (GPT) [42], and Google developed Bidirectional Encoder Representations from Transformers (BERT) [43], which inspired the creation of several different pretrained language models (PLMs) of large size based on transformers [38]. For example, XLNet, a generalized autoregressive pretraining method, allows for the learning of bidirectional contexts, and its autoregressive formulation overcomes some limitations of BERT [44]. Moreover, Facebook developed RoBERTa [45], which is trained on a larger data set than BERT. Large models based on transformers, including BERT, RoBERTa, and XLNet, achieved a high level of success in many NLP tasks [43-45].

Objective

The objective of this study was to ameliorate the impact of online misinformation through automated, machine-learned, and scalable methods. Our study sought to answer the following three core research questions (RQs):

RQ1: Can approaches leveraging automated and scalable strategies such as machine learning, information retrieval, and crowdsourcing help combat misinformation when information growth exceeds fact-checker capabilities?

RQ2: Does training a machine-learning model on only COVID-19–related misinformation data, only on general misinformation data, or on both result in the highest performance on COVID-19–related data?

RQ3: Does combining crowdsourced labels with machine-learning model outputs improve accuracy over either approach individually?

Methods

Machine-Learned Classification

We first developed a classifier using the CoAID data set [46]; specifically, the 05-01-2020 and 07-01-2020 folders of the CoAID data set were used. Since there are more pieces of news deemed to be accurate ("true") than those deemed to be inaccurate ("false"), we included all inaccurate news, but limited the quantity of true news to be equal to the amount of false news to have a balanced data set. For the Bi-LSTM model, we split our input data into a training set (75%) and test set (25%). Pandas [47] and scikit-learn [48] were used in our classifier development and implementation.

We evaluated different architectures, dropouts, activation functions, optimizers, regularizers, and batch sizes. We ultimately chose an embedding layer, Bi-LSTM layer, Dropout layer with a rate of 0.7, and Dense layer with a 1-dimensional output and sigmoid activation function. We used an Adam optimizer with a learning rate of 0.0001, binary cross-entropy loss, and a batch size of 1. The Bi-LSTM model has a kernel regularizer with l_1 and l_2 regularization factors of 1e-5 and 1e-4, respectively. In addition, we employed several state-of-the-art models for text classification, including PLMs such as BERT, RoBERTa, and XLNet. We selected RoBERTa, as it is an optimized BERT approach, and XLNet, as it is an autoregressive BERT-like model. We employed four transformers: BERT-base [43], XLNet [44], and two models fine-tuned on RoBERTa-base [45,49,50] for this specific classification task on the 7 data sets described in Table 1 for 3 epochs with default training arguments in HuggingFace Trainer [51]. Moreover, we trained a convolutional neural network (CNN) model for text classification [52], as this method has been extensively used in text classification [38].

All source code files for our models are publicly available as open source [53].

XSL•FO
RenderX

Table 1. Data set sources and specifications.

Data set	Source	Time range	Size (number of articles)			Type
			Noncredible news	True news	Total	
CoAID[a]	Tweets	Until May 1, 2020	572	1324	1896	COVID-19–specific
FNN[b]	PolitiFact	N/A[c]	472	797	1270	General news
FNN	Gossip Cop	N/A	16,818	5335	22,153	General news
Validation data set 1[d]	Poynter.org (noncredible news); Washington Post, Associated Press, Politico (true news)	July 20, 2020, to August 8, 2020	3874	3177	7051	COVID-19–specific
Validation data set 2[d]	Poynter.org (noncredible news); BBC, AXIOS, CBS News, The Globe and Mail (true news)	January 20, 2020, to June 15, 2022	14,398	16,232	30,630	COVID-19–specific

[a]Only the 05-01-2020 folder of the CoAID data set was used.

[b]FNN: FakeNewsNet.

[c]N/A: not applicable.

[d]Scraped with the query term "COVID-19."

Data Evaluation

To develop our external validation data sets, we used data from Poynter [54], which had several thousand instances of COVID-19–related content with a "false" label. For "true" news, we inherited article accuracy from the credibility of the media source on which the documents were published, following an approach similar to the ReCOVery [26] and CoAID [46] COVID-19–related data sets. We created two external validation data sets with different "true" news sources to test the generalization ability of the models. The first external validation data set consists of ~4000 pieces of false-news content scraped from Poynter and ~3000 pieces of true-news content collected from several news outlets that we deemed to be reliable by inheriting source credibility. We used NewsAPI's application programming interface [55] to retrieve content from the following news outlets: Reuters, BBC, The Wall Street Journal, The Washington Post, Associated Press, and Politico. We searched for articles from July 20, 2020, to August 8, 2020, with the query term "COVID-19." With these parameters, we queried just over 3000 news articles and stored their labels, titles, sources, descriptions, URLs, and publication dates. The second external validation data set consists of ~14,000 pieces of noncredible news scraped from Poynter in the time range from March 20, 2020, to February 23, 2022, and ~16,000 pieces of true news scraped from BBC, AXIOS, CBS News, and The Globe and Mail with the query term "COVID-19" in the time range from January 20, 2020, to June 15, 2022. In total, after removing elements due to nonapplicable Poynter labels, the first data set had 7051 labeled pieces of COVID-19–related content within the time range from July 20, 2020, to August 8, 2020, and the second data set had 30,630 pieces of COVID-19–related content within the time range from January 20, 2020, to June 15, 2022.

We developed methods to evaluate whether training a machine-learning model on only COVID-19–related misinformation data, only on general misinformation data, or on both would result in the highest performance on new, unseen COVID-19 data sets. When evaluating general data sets, FakeNewsNet (FNN) [56,57] provided a data format matching our needs and with a sufficient volume for the scale of our training. For COVID-19–related data, we found that CoAID, a COVID-19 health care misinformation data set, with 1896 news articles, 183,564 related user engagements, 516 social platform posts about COVID-19, and ground truth labels [46], allowed us to achieve high internal validation accuracy in preliminary trials. To be as consistent across the two data sets as possible, we drew from standard benchmarking practices performed on data sets using default machine-learning model implementations. We trained on 7 different combinations of data sources to mimic different situations in the real world: (1) only CoAID, used to mimic the situation when sufficient topic-specific data are available; (2) partial (using only the 05-01-2020 folder of the CoAID data set) CoAID and FNN; (3) partial CoAID and PolitiFact; (4) partial CoAID and the GossipCop content from FNN, used to mimic the situation when we have a limited quantity of topic-specific data; (5) FNN; (6) PolitiFact; and (7) GossipCop, used to mimic the situation when no topic-specific data are available. For three classical models (support vector machine [SVM], logistic regression [LR], and Bernoulli naïve Bayes [BNB]) and six deep-learning models (Bi-LSTM, BERT-based model, two RoBERTa-based models [45,49,50], XLNet [44], and Text-CNN [52]) on all seven data source combinations, we computed precision, recall, and F1-score for both internal validation and the two external validation data sets described above. These were taken as a weighted average of both labels and rounded to the nearest hundredth, as detailed in Multimedia Appendix 1-3, and are available as a CSV file on our data repository [53].

Ethics Considerations

The University of Texas at Austin Institutional Review Board (IRB) approved this study for human subjects research on April 20, 2021 (STUDY00000962). Informed consent from all study participants was obtained.

Crowdsourced Classification

We recruited annotators from the crowdsourcing platform Prolific to vote on pieces of news content from the data set we

XSL•FO
RenderX

created. On Prolific, we set the study distribution to "standard sample," which launched the study to the whole participant pool [58]. In line with the IRB protocol, we limited voting to US residents only. We established approximately 10 rounds of Prolific tasks with each participant being paid varying amounts of ~$8 an hour, which resulted in 31,441 votes from 756 voters.

After completing the crowdsourced voting, we then processed the data both manually and with Python scripts for usability. We removed duplicate votes for the same label (two "true" votes) and votes from Prolific IDs that we could not find in the set of IDs reported to us by Prolific. The processed data set had more than 6800 pieces of content with at least 3 votes for either the "true" or "false" label. We took the initial ground truth labels from Poynter and credible news sources and mapped them to 0 or 1. "True" was coded as 1 and "false" was coded as 0. Additionally, "correct" labels were coded as 1 (2 labels), and all other labels were converted to 0 (690 labels). Mapping our labels to 0 or 1 allowed us to collect certain metrics for our data set. Some examples from the crowdsourced data set are provided in Table 2 (also see Multimedia Appendix 1). Voter soft labels of 0.0 or 1.0 indicate that the vote results are concordant (ie, all votes were for the same label), whereas a voter soft label range of 0.4-0.6 implies that (nearly) half of the voters have different opinions.

We also computed the percentage of agreeing decisions, which we defined as the probability that the label decided on by the crowdsourced votes was the same as the ground truth label. The percentage of agreeing decisions (human voter accuracy) was ~0.73, or 73%. We also calculated interannotator agreements to determine the agreement among voters. As the number of voters varied (from 3 to 7) for each piece of news content, Cohen and Fleiss κ statistics were not suitable for our data set. We therefore computed the percent agreement between users to determine interrater reliability (68.5%) for our data. As percent agreement does not take chance agreement into consideration, we calculated Krippendorff α (0.428). As percent agreement is considered to be acceptable when above 75% [59] and α is "acceptable at 0.667≤α≤0.823 and unacceptable at α<0.667" [60], there was low agreement among all voters in the crowdsourced data. Ultimately, crowdsourced voters had low accuracy (~73%) when identifying COVID-19–related noncredible content, and there was a high level of disagreement among them. Given that this data set was not used as the ground truth, but rather to evaluate whether labeled data from nonexperts could improve model performance, low agreement is not an issue for our use case. Moreover, this low agreement indicates that nonprofessionals respond to misinformation differently rather than consistently.

Given this high level of variability, we next evaluated whether our crowdsourced data could actually improve machine-learning model predictions. With this in mind, we developed and answered the following questions: (1) Which model best predicted crowdsourced labels? (2) Can model performance be improved after being blended with crowdsourced labels? (3) Which model performs best when blended with crowdsourced labels? (4) If we only take the subset of the data set where machine-learning models and human votes have agreeing labels, will the performance of prediction be improved?; if so, which model has the highest performance?

Table 2. Examples from the crowdsourced data set.

News title	Ground truth	Voter soft label[a]	Voter label	Total votes	Results
Concordant human votes					
COVID-19 pandemic derails Germany's push for migrant integration-Reuters	1	1.0	1	3	Correctly classified by humans
Photo shows the last meeting of a Turkish doctor who died due to COVID-19 with his child in Munich	0	1.0	1	4	Misclassified by humans
3M brings on another lobbying firm	1	1.0	1	5	Correctly classified by humans
Video shows that the Italian government/Brisbane police used zombie robots/drones to chase their citizen and make them stay home	0	0.0	0	4	Correctly classified by humans
British vaccine provokes immune response in first human studies	1	0.0	0	3	Misclassified by humans
This video shows a woman eating a bat soup in Wuhan	0	0.0	0	5	Correctly classified by humans
Discordant human votes					
An emergency department closed in a Spanish hospital	0	0.5	1	6	Misclassified by human
Majority of Caledonian hotel jobs under review in Edinburgh	1	0.5	1	4	Correctly classified by humans
England v Ireland: Captain Eoin Morgan relishes 'new journey' in ODI series	1	0.6	1	5	Correctly classified by humans
Panic scene in Germany with people rushing into a supermarket	0	0.4	0	5	Correctly classified by humans

[a]Voter soft label is calculated by the number of true labels/total votes.

XSL•FO
RenderX

Results

Machine-Learned Classification

RQ1 asks whether automated systems can help combat COVID-19–related misinformation. We found that machine learning predicts veracity better than random. We developed a Bi-LSTM model trained on the CoAID data set. Specifically, we used 1257 entries from CoAID for training and tested our model on 419 entries from CoAID. We achieved a weighted average F1-score of 0.93 (with equal precision, recall, and accuracy) across both labels. Using the same model, the external validation results on our data set was an F1-score of 0.75, with equal precision, recall, and accuracy. In addition, we fine-tuned BERT-base, RoBERTa-fake-news, Fake-News-BERT-Detect, XLNet, and trained Text-CNN on 7 data set combinations and tested them on the two external validation data sets. The results are shown in Multimedia Appendix 1-2. We achieved accuracies of up to 91%, 93%, 97%, 94%, and 87% on the first external validation data set from BERT-base, RoBERTa-fake-news, Fake-News-BERT-Detect, XLNet, and trained Text-CNN, respectively. Accuracies of up to 93%, 84%, 93%, 91%, and 85% were achieved on the second external data sets from the same models. Given these results, RQ1 can be answered in the affirmative.

Data Evaluation

RQ2 asks whether training a machine-learning model on only COVID-19–related misinformation data, on only general misinformation data, or on both results in the highest performance on COVID-19–related data. We found that machine-learned models benefit from COVID-19–related data. Specifically, after training on 7 different data sets (see Multimedia Appendix 1-3), RQ2 can be answered as follows: for classical models, the combination of topic-specific and general-topic data results in the best performance; however, pretrained models benefit from purely topic-specific data the most. In this study, we investigated the efficacy of three scenarios: (1) training on COVID-19–related misinformation, (2) training on non-COVID-19–related misinformation, and (3) training on both COVID-19–related misinformation and non-COVID-19–related misinformation. Our results indicate that including COVID-19–related misinformation (in our case CoAID data) helped—or, at least, maintained—model performance.

Examples of classical classification models include LR, SVM, BNB, hidden Markov model, and random forests [39]. In our experiment, classical models used included LR, SVM, and BNB. All three classical models shown in Multimedia Appendix 3 achieved the best accuracy when trained on the combination of CoAID and PolitiFact, whereas for deep-learning pretrained models, which have already "studied" the behavior of the English language, the best model performance was obtained when fine-tuned on CoAID only (see Multimedia Appendix 1-3). In instances where we are lacking additional COVID-19–related misinformation content, our findings suggest that incorporation of prior misinformation data sets in conjunction with COVID-19–specific misinformation data sets could potentially be useful to detect new COVID-19–related misinformation when using classical models. However, using PLMs (eg, BERT), which normally have much better performance on language tasks than classical models, fine-tuning on a topic-specific data set tended to give a better result. By combining COVID-19–related (ie, CoAID) and broad, multitopic misinformation data sets (ie, FNN, GossipCop, and PolitiFact), we evaluated the performance of our machine-learning models. Combining labeled data sets from different sources coupled with various machine-learning models is a novel contribution of our study in terms of producing a scalable and generalizable framework. As detailed in Multimedia Appendix 1-3, we found that the accuracy of models where we used only GossipCop data sets was very low. The lowest BNB accuracy we obtained (0.37) was also obtained for GossipCop, indicating the important role that labeled data sets play in the validity of misinformation detection. As GossipCop is considered a credible source of celebrity news, the labeled data sets of GossipCop are specific and have limited value to COVID-19 misinformation detection on their own. Conversely, combining CoAID and GossipCop as the input data to train our models significantly improved the accuracy (0.64) for the BNB model (Multimedia Appendix 3). As the best result, an accuracy of 96.55% was achieved when we fine-tuned Fake-News-BERT-Detect using only the CoAID data set (Multimedia Appendix 1). With these findings, RQ2 can be answered positively.

Crowdsourced Classification

RQ3 asks whether combining crowdsourced labels with machine-learning model outputs improves accuracy over either approach individually. We found that combining human votes with machine-learned outputs allowed us to create higher performance models. Specifically, deep-learning models are able to predict human votes at an accuracy up to 70%. Combining human votes with machine-learned outputs allowed us to create a model with 99.1% accuracy. We achieved accuracy up to 98.59% when only considering the subset where model and human votes agreed.

We first evaluated how well our models could predict our crowdsourced values or the labels we generated from our Prolific labeling (see Multimedia Appendix 4-9). A label of 0 indicates that most voters voted false, while a label of 1 indicates that greater than or equal to half of the voters voted true. Using the models trained on the 7 data set combinations and testing on our data set of 7051 votes, the success at predicting the crowdsourced values from Prolific had accuracies up to 0.70 (see Multimedia Appendix 7). All values were rounded to the nearest hundredth.

Second, we blended the soft predictions (ie, probabilities) from the models and soft vote (combining the probabilities of each prediction in contrast to hard voting, which chooses the prediction that receives the most votes) results from crowdsourcing data in different proportions to assess both the maximum improvements and highest accuracies that can be achieved after blending. The soft vote results were computed by taking the number of votes for label 1 (credible) and dividing by the number of total votes. The results shown in Table 3

XSL•FO
RenderX

(predictions from blended models) were calculated by the following formula:

a×(soft predictions from model)+(1–a)×(soft vote results from crowdsourcing data)

Table 3 illustrates that models had higher accuracy on average after blending, and the highest accuracy we achieved was 99.1% on the first external validation data set (when blending 10% of user vote results with 90% of the machine-learning model prediction). Therefore, we found that models trained on general news were improved. Those models achieved much higher accuracies (up to 99.7%) after blending with user vote results. This represents a considerable improvement over the human vote accuracy of ~73%. As shown in Table 3, when a=0.9, the performance of Text-CNN trained on GossipCop could be improved from 42.6% to 99.1% after blending with crowdsourced data.

Third, as discussed in the Machine-Learned Classification section above, the machine-learning models had accuracies ranging from 41% to 98% and the human votes had approximately 73% accuracy. Out of the 7051 pieces of content, 39.24%-69.58% (for the best-performing model) showed agreement in both the human votes and the machine-learning model. We were therefore able to make reduced sets of 2766 to 4906 pieces of content. For each piece of content, we assigned its label to whichever value both the machine-learning model and human votes agreed on. Using this approach, our best accuracy was 98.59% (see Multimedia Appendix 10), which was from the Fake-News-BERT-Detect model fine-tuned on the CoAID data set. This is in comparison with an accuracy of 73% for human votes and 96.55% for the entire validation data set. All models achieved the best performance when the models were previously fine-tuned on COVID-19–specific data sets (ie, CoAID).

The performance of models trained/fine-tuned on a general-topic data set could be improved with crowdsourced data (eg, in low-data situations such as pandemics). Specifically, the base model achieved an accuracy of 71.01% on the whole validation data set. For example, for the subset, we achieved an accuracy of 89.96% at best (by BERT-base fine-tuned on PolitiFact). In addition, models trained on the combination general-topic and COVID-19–specific data set were also improved by this approach. Specifically, accuracies of up to 89.93% on the whole data sets (see Multimedia Appendix 1) were improved to up to 96.26% (for the subset). Practically speaking, both credibility tests could be applied to a piece of content and receive a label of "true" or "false" with up to 98.59% accuracy. Combining human votes with machine-learned outputs therefore outperformed models with human votes alone. Our response to RQ3 is that both blending crowdsourced labels with model predictions and reducing the data set to a "high-confidence" data subset increased model performance.

Table 3. Analysis of accuracy for blended models, evaluated on the first external validation data set.

Metric	a=0.9	a=0.7	a=0.5	a=0.3	a=0.1
Average improvement	0.069	0.082	0.084	0.063	0.029
Maximum improvement					
Maximum improvement	0.565	0.562	0.463	0.385	0.415
Model name	Text-CNN trained on GossipCop	Text-CNN trained on GossipCop	Text-CNN trained on GossipCop	Text-CNN trained on GossipCop	Fake-News-BERT-Detect fine-tuned on GossipCop
Model accuracy (before blending)	0.426	0.426	0.426	0.426	0.302
Model accuracy (after blending)	0.991	0.981	0.889	0.804	0.717
Best performance					
Model name	Text-CNN trained on CoAID	Text-CNN trained on CoAID	Text-CNN trained on CoAID and PolitiFact	Text-CNN trained on GossipCop	Text-CNN trained on PolitiFact
Model accuracy (before blending)	0.874	0.874	0.798	0.426	0.499
Model accuracy (after blending)	0.991	0.984	0.891	0.804	0.728

Discussion

Principal Results

Our results indicate that RQ1 (which asks whether automated systems and scalable strategies can help combat misinformation) can be answered in the affirmative. The models we trained showed an accuracy of 98% on our first external validation data set (of ~7000 posts and true news from July 20, 2020, to August 8, 2020) and an accuracy of 93% on our second validation data set (of ~15,000 posts and true news from January 20, 2020, to June 15, 2022). Labeling by fact-checkers can be time-consuming, labor-intensive, and expensive, whereas machine-learning models can be used at will and at scale once trained. These results support our finding that machine learning significantly improves fact checking given the reality that human fact-checkers are overburdened and cannot feasibly keep up with the increasing volume of online misinformation.

Regarding RQ2 (which asks what kind of data set is most helpful to machine learning), we found that training/fine-tuning on pandemic-specific content tends to result in higher accuracy. Specifically, our best-performing models were fine-tuned on COVID-19 topic content only. We evaluated three classical

XSL•FO
RenderX

models and five deep-learning models trained on seven different data sets, including one topic-specific data set (CoAID only), three general-topic data sets (FNN, GossipCop, and PolitiFact), and three combinations of topic-specific and general-topic data sets (CoAID and FNN, GossipCop and CoAID, PolitiFact and CoAID). Classical models achieved the best accuracy when trained on a combination of general-topic and COVID-19–specific data (the combination of CoAID and PolitiFact), while deep-learning PLMs (eg, BERT), which have already been trained on English-language text and therefore could be considered as having "studied" the behavior of the English language, obtained the best model performance when fine-tuned on a COVID-19–specific data set (ie, CoAID).

Regarding RQ3, which asks whether combining crowdsourced labels with models can improve model performance, we found that blending crowdsourced labels with model predictions increased model performance. The blended model (crowdsourced votes mixed with a machine-learning model) was able to achieve an accuracy of 99.1%. Given that the accuracy of crowdsourced votes was 73% and the highest accuracy of our machine-learning models was 96.55%, our results therefore show that crowdsourcing can be used in conjunction with machine learning to boost accuracy. In addition, models trained on general news could be improved to achieve much higher accuracies after blending with user vote results. Specifically, we found improvements of up to 57.1% after blending (see Table 3). That being said, the performance of models trained/fine-tuned on a general-topic data set could only be improved when considering the subset. With neither crowdsourcing nor machine learning requiring time from expert fact-checkers, both are viable options for addressing COVID-19 and other health-related misinformation at scale.

Future Work

Future work can further optimize our machine-learning model and extend and develop our labeled data set. Moreover, we hope that our findings encourage others to develop COVID-19–specific disinformation and misinformation data sets. As the quantity of COVID-19–related labeled data increases, the combination of COVID-19–related labeled data and general misinformation data should be further evaluated and benchmarked by others to enhance machine-learning model accuracy. Our results would therefore benefit from replication in future work with a data set consisting of both COVID-19–related and broad, multitopic content. Since we only crowdsourced votes for the first external validation data set (which spans one month), future work could crowdsource vote results on the second validation data set to strengthen the validity of our conclusions. Furthermore, the size of the crowdsourcing data set is relatively small (31,441 pieces of content and 4.46 average votes each), which could be strengthened with the accumulation of more votes and would increase the generalizability of our results. Thus, future work would benefit from extending our framework to a larger crowdsourced data set. Since collecting crowdsourced data could be time-consuming, using machine-learning models to generate pseudohuman votes can potentially be another way to strengthen the crowdsourced data set. After collecting crowdsourced data for a small news data set, the pseudohuman votes model trained on that data set can be used to predict human labels on a larger data set. This method would be especially useful with unlabeled news data sets, on which we could simulate human votes in the absence of ground truth labels.

Future work could also measure whether there are sufficient advantages of using machine-learning models rather than expert fact-checkers (given that the former method allows for cheaper and quicker large-scale data labeling). There is also the possibility that machine-learning models and professional fact-checkers combined together could deliver better results. For example, fact-checkers could use models to flag news to speed up their work, and the results from fact-checkers could be used to refine models. Human-in-loop models could be developed by using this method. A live news browser displaying news alongside fact-checker results or model predictions (if no fact-checker is available) could help assess credibility even when there is more misinformation than experts can check manually. Lastly, future work could further examine the relationship between crowdsourced outputs and ground truth labels for COVID-19–related data, a line of inquiry we minimally investigated in this study. Specifically, future work could examine when humans are more likely to make misjudgments by exploring the scenarios in which crowdsourced and ground truth labels are most likely to disagree. Research could explore crowdsourced data in different problem domains to identify the misinformation in problem domains that interventions should pay most attention to, using metrics such as the disagreement between human votes and ground truth labels.

Limitations

A limitation of our work is that our study did not rigorously test the ceiling of possible model optimization on all combinations of FNN and CoAID models. Another minor limitation is that we assigned "false" to all labels (except two "correct" labels) in the Poynter data set when evaluating our model, even though a small portion of labels could be interpreted as true (<0.5% with labels such as "half true" and "mostly true"). The crowdsourced data set quality was potentially limited due to the number of votes per item and the time span of the labeled data set. Lastly, we were only able to crowdsource votes for the first external validation data set due to time and funding constraints.

Conclusion

Manual fact checking is unable to cope with the large volumes of COVID-19–related misinformation that now exists [8]. To help address the proliferation of COVID-19–related misinformation, we developed an automated, machine-learned, and scalable approach. Since the best-performing models we evaluated were fine-tuned on COVID-19–specific content only, topic-specific data sets are much more helpful than general-topic data sets or the combination of the two. The 96.55% and 94.6% accuracy on the first and second external validation data set, respectively, suggest that machine learning can be used to achieve significantly better than random results for the difficult task of determining the veracity of COVID-19–related content. Our study also found that in the cases when only considering the reduced set of the content that both human votes and model

XSL•FO
RenderX

outputs agreed on, the models achieved up to 99.1% accuracy. Models trained/fine-tuned on general-topic content can be improved to an acceptable level after combining with human votes, and may be used to supplement limited amounts of topic-specific content in low-data situations (eg, pandemics) to increase accuracy.

Our findings also suggest that machine-learning models can be augmented with the labels of lay, crowdsourced voters to boost accuracy without additional input from expert fact-checkers. Blending human votes with model prediction results achieved an accuracy up to 99.1% (by combining 10% of a human vote label with 90% of a label from the model). We have released our topic-related data set of 7000 ground truth and crowdsourced labels, machine-learning model, and code in open-source form to promote the development by others of automated, scalable solutions to the COVID-19 infodemic.

COVID-19 infodemic responses need to acknowledge that misinformation can be amorphous and highly decentralized. The machine-learned and automated approaches developed in this study rely on text features, making them powerful in that they can be extended (eg, by researchers or technology companies) to study a variety of platforms and contexts (eg, news and social media) in which online misinformation exists. Automation and machine learning offer the ability to exchange a small decrease in accuracy for scalability, which is an important consideration when misinformation growth exceeds fact-checking capabilities as continues to be the case during the COVID-19 pandemic.

Acknowledgments

The authors wish to thank Kami Vinton for her insightful comments and suggestions, as well as for her assistance proofreading the manuscript. This work was supported by Good Systems, a research Grand Challenge at the University of Texas at Austin, and an Undergraduate Research Fellowship at The University of Texas at Austin.

Authors' Contributions

NK and DM jointly architected the study, wrote the first version of the manuscript, and collaborated to obtain crowdsourcing funding. DM obtained further funding for the deep-learning aspects of the project. NK wrote all of the code for the first version of the manuscript, performed the experiments for the first version of the manuscript, collected crowdsourced data, and provided data for Multimedia Appendix 4 and part of Multimedia Appendix 5. YL wrote substantial sections of the manuscript revision, performed the experiments during the revision process, provided the second validation data set, performed the experiment regarding deep-learning models, and provided data for the other tables. All authors collaborated on the revised manuscript. NK and YL contributed equally to the study and should be viewed as joint first authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Model performances on the first external validation data set.
[DOCX File , 15 KB - infodemiology_v2i2e38756_app1.docx]

Multimedia Appendix 2
Model performances on the second external validation data set.
[DOCX File , 15 KB - infodemiology_v2i2e38756_app2.docx]

Multimedia Appendix 3
Benchmarking results using classical models.
[DOCX File , 20 KB - infodemiology_v2i2e38756_app3.docx]

Multimedia Appendix 4
Results for the bidirectional long short-term memory (Bi-LSTM) model trained on CoAID and tested on crowdsourced labels.
[DOCX File , 14 KB - infodemiology_v2i2e38756_app4.docx]

Multimedia Appendix 5
Results for BERT-base tested on crowdsourced labels.
[DOCX File , 14 KB - infodemiology_v2i2e38756_app5.docx]

Multimedia Appendix 6
Results for RoBERTa-Fake-News tested on crowdsourced labels.

XSL•FO
RenderX

[DOCX File , 14 KB - infodemiology_v2i2e38756_app6.docx]

Multimedia Appendix 7
Results for Fake-News-BERT-Detect tested on crowdsourced labels.
[DOCX File , 14 KB - infodemiology_v2i2e38756_app7.docx]

Multimedia Appendix 8
Results for XLNet tested on crowdsourced labels.
[DOCX File , 14 KB - infodemiology_v2i2e38756_app8.docx]

Multimedia Appendix 9
Results for Text-CNN tested on crowdsourced labels.
[DOCX File , 14 KB - infodemiology_v2i2e38756_app9.docx]

Multimedia Appendix 10
Model performances on the reduced set of content when human and machine-learned votes agree.
[DOCX File , 17 KB - infodemiology_v2i2e38756_app10.docx]

References

1. Shim J, Lee Y, Ahn H. A link2vec-based fake news detection model using web search results. Expert Syst Appl 2021 Dec;184:115491. [doi: 10.1016/j.eswa.2021.115491]
2. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat Hum Behav 2020 Dec;4(12):1285-1293. [doi: 10.1038/s41562-020-00994-6] [Medline: 33122812]
3. 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. [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
4. Litman L, Rosen Z, Rosenzweig C, Weinberger-Litman SL, Moss AJ, Robinson J. Did people really drink bleach to prevent COVID-19? A tale of problematic respondents and a guide for measuring rare events in survey data. MedRxiv. Preprint published on January 2, 2021. URL: https://www.medrxiv.org/content/10.1101/2020.12.11.20246694v3 [accessed 2022-08-12]
5. An ad hoc WHO technical consultation managing the COVID-19 infodemic: call for action, 7-8 April 2020. World Health Organization, Institutional Repository for Information Sharing. 2020. URL: https://apps.who.int/iris/handle/10665/334287 [accessed 2020-12-03]
6. Khazan O. How a bizarre claim about masks has lived on for months. The Atlantic. 2020 Oct 09. URL: https://www.theatlantic.com/politics/archive/2020/10/can-masks-make-you-sicker/616641/ [accessed 2022-04-11]
7. Bridgman A, Merkley E, Loewen P, Owen T, Ruths D, Teichmann L, et al. The causes and consequences of COVID-19 misperceptions: understanding the role of news and social media. HKS Misinfo Review 2020 Jun 18;19:1-18 [FREE Full text] [doi: 10.37016/mr-2020-028]
8. Simon F, Howard PN, Nielsen RK. Types, sources, and claims of COVID-19 misinformation. Reuters Institute. 2020 Apr 07. URL: https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation [accessed 2022-08-12]
9. Misinformation. Merriam-Webster Dictionary. URL: https://www.merriam-webster.com/dictionary/misinformation [accessed 2022-07-05]
10. Coleman A. 'Hundreds dead' because of Covid-19 misinformation. BBC. 2020 Aug 12. URL: https://www.bbc.co.uk/news/world-53755067 [accessed 2022-07-17]
11. Disinformation. Merriam-Webster Dictionary. URL: https://www.merriam-webster.com/dictionary/disinformation [accessed 2022-07-05]
12. MacLellan K, Kerry F. Britain says Russian troll factory is spreading disinformation on social media. Reuters. URL: https://www.reuters.com/world/europe/britain-says-russian-troll-factory-is-spreading-disinformation-social-media-2022-04-30/ [accessed 2022-07-17]
13. Jamieson KH. Cyberwar: how Russian hackers and trolls helped elect a president: what we don't, can't, and do know. Oxford, UK: Oxford University Press; 2020.
14. Banda JM, Tekumalla R, Wang G, Yu J, Liu T, Ding Y, et al. A large-scale COVID-19 Twitter chatter dataset for open scientific Research—an international collaboration. Epidemiologia 2021 Aug 05;2(3):315-324. [doi: 10.3390/epidemiologia2030024]
15. 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]
16. Gao Z, Yada S, Wakamiya S. NAIST COVID: Multilingual COVID-19 Twitter and Weibo dataset. arXiv. Preprint published on April 17, 2020. URL: https://arxiv.org/abs/2004.08145 [accessed 2022-08-12]

XSL•FO
RenderX

17. Haouari F, Hasanain M, Suwaileh R. ArCOV19-Rumors: Arabic COVID-19 Twitter dataset for misinformation detection. arXiv. Preprint published on March 13, 2021. URL: https://arxiv.org/abs/2010.08768 [accessed 2022-08-12]
18. He B, Ziems C, Soni S, Ramakrishan N, Yang D, Kumar S. Racism is a virus: anti-Asian hate and counterspeech in social media during the COVID-19 crisis. arXiv. Preprint published on November 10, 2021. URL: https://arxiv.org/abs/2005.12423?context=cs.CL [accessed 2022-08-12]
19. 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]
20. FALSE: A claim that neem leaves can cure the novel coronavirus and relieve its symptoms has been shared thousands of times in multiple Facebook posts. Poynter. 2020 Mar 22. URL: https://tinyurl.com/yc22cz8z [accessed 2022-07-17]
21. 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. Preprint published on March 31, 2020. URL: https://arxiv.org/abs/2003.13907 [accessed 2022-08-12]
22. Hossain T, Logan RI, Ugarte A. COVIDLies: detecting COVID-19 misinformation on social media. 2020 Presented at: 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020; November 20, 2020; virtual. [doi: 10.18653/v1/2020.nlpcovid19-2.11]
23. Serrano JCM, Papakyriakopoulos O, Hegelich S. NLP-based feature extraction for the detection of COVID-19 misinformation videos on YouTube. 2020 Presented at: 1st Workshop on NLP for COVID-19 at ACL 2020; July 2020; virtual.
24. Dharawat A, Lourentzou I, Morales A, Zai CX. Drink bleach or do what now? Covid-HeRA: A dataset for risk-informed health decision making in the presence of COVID19 misinformation. arXiv. Preprint published on October 17, 2020. URL: https://arxiv.org/abs/2010.08743 [accessed 2022-08-12]
25. 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]
26. Zhou X, Mulay A, Ferrara E, Zafarani R. ReCOVery: a multimodal repository for COVID-19 news credibility research. arXiv. Preprint published on June 9, 2020. URL: https://arxiv.org/abs/2006.05557 [accessed 2022-08-12]
27. 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]
28. Zhang X, Ghorbani AA. An overview of online fake news: characterization, detection, and discussion. Inf Process Manag 2020 Mar;57(2):102025. [doi: 10.1016/j.ipm.2019.03.004]
29. Cha M, Cha C, Singh K, Lima G, Ahn Y, Kulshrestha J, et al. Prevalence of misinformation and factchecks on the COVID-19 pandemic in 35 countries: observational infodemiology study. JMIR Hum Factors 2021 Feb 13;8(1):e23279 [FREE Full text] [doi: 10.2196/23279] [Medline: 33395395]
30. Asr FT, Taboada M. Big Data and quality data for fake news and misinformation detection. Big Data & Society 2019 Jan;6(1). [doi: 10.1177/2053951719843310]
31. Shahi GK, Nandini D. FakeCovid--a multilingual cross-domain fact check news dataset for COVID-19. arXiv. Preprint published on June 19, 2020. URL: https://arxiv.org/abs/2006.11343 [accessed 2022-08-12]
32. Patwa P, Sharma S, Pykl S, Guptha V, Kumari G, Akhtar M, et al. Fighting an infodemic: COVID-19 fake news dataset. In: Chakraborty T, Shu K, Bernard HR, Liu H, Akhtar MS, editors. Combating Online Hostile Posts in Regional Languages during Emergency Situation. CONSTRAINT 2021. Communications in Computer and Information Science, vol 1402. Cham: Springer; 2021.
33. Addawood A. Coronavirus: Public Arabic Twitter Data Set. OpenReview. 2020 Aug 12. URL: https://openreview.net/forum?id=ZxjFAfD0pSy [accessed 2022-08-12]
34. Melo T, Figueiredo CMS. A first public dataset from Brazilian twitter and news on COVID-19 in Portuguese. Data in Brief 2020;32:106179. [doi: 10.1016/j.dib.2020.106179]
35. Rovetta A, Bhagavathula AS. COVID-19-related web search behaviors and infodemic attitudes in Italy: infodemiological Study. JMIR Public Health Surveill 2020 May 05;6(2):e19374 [FREE Full text] [doi: 10.2196/19374] [Medline: 32338613]
36. Yang C, Zhou X, Zafarani R. CHECKED: Chinese COVID-19 fake news dataset. Soc Netw Anal Min 2021;11(1):58 [FREE Full text] [doi: 10.1007/s13278-021-00766-8] [Medline: 34178179]
37. Kar D, Bhardwaj M, Samanta S, Azad AP. No rumours please! A multi-indic-lingual approach for COVID fake-tweet detection. arXiv. Preprint published on October 14, 2020. URL: https://arxiv.org/abs/2010.06906 [accessed 2022-08-12]
38. Minaee S, Kalchbrenner N, Cambria E, Nikzad N, Chenaghlu M, Gao J. Deep learning–based text classification. ACM Comput Surv 2022 Apr 30;54(3):1-40. [doi: 10.1145/3439726]
39. Alam T, Khan A, Alam F. Bangla text classification using transformers. arXiv. Preprint published on November 9, 2020. URL: https://arxiv.org/abs/2011.04446 [accessed 2022-08-12]
40. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S. Quantum machine learning. Nature 2017 Sep 13;549(7671):195-202. [doi: 10.1038/nature23474] [Medline: 28905917]
41. Peters ME, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, et al. Deep contextualized word representations. arXiv. Preprint published on February 15, 2018. URL: https://arxiv.org/abs/1802.05365 [accessed 2022-08-12]

XSL•FO
RenderX

42. Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language understanding by generative pre-training. Amazon Simple Storage System (S3). 2018. URL: https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf [accessed 2022-08-12]
43. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv. Preprint published on October 11, 2018. URL: https://arxiv.org/abs/1810.04805 [accessed 2022-08-12]
44. Yang Z, Dai Z, Yang Y, Carbonell J, Salakhutdinov R, Le QV. XLNet: generalized autoregressive pretraining for language understanding. 2019 Presented at: 33rd Conference on Neural Information Processing Systems (NeurIPS 2019); December 8-14, 2019; Vancouver, BC.
45. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized bert pretraining approach. arXiv. Preprint publihsed on Jul, 26 2019. URL: https://arxiv.org/abs/1907.11692 [accessed 2022-08-12]
46. Cui L, Lee D. CoAID: COVID-19 healthcare misinformation dataset. arXiv. Preprint published on May 22, 2020. URL: https://arxiv.org/abs/2006.00885 [accessed 2022-08-12]
47. The Pandas Development Team. Zenodo. 2020 Dec 07. URL: https://zenodo.org/record/4309786#.YvfOIC9E3mp [accessed 2022-08-12]
48. Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:14. [doi: 10.3389/fninf.2014.00014] [Medline: 24600388]
49. Liu J. Transformer-fake-news-detection. GitHub. 2022. URL: https://github.com/Jiayif/Transformer-Fake-News-Detection [accessed 2022-08-12]
50. Tatti GV. roberta-fake-news. Hugging Face. 2021. URL: https://huggingface.co/ghanashyamvtatti/roberta-fake-news [accessed 2022-08-12]
51. Thomas W, Lysandre D, Victor S, Julien C, Clement D, Anthony M, et al. Transformers: State-of-the-Art Natural Language Processing. 2020 Presented at: 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations; October 2020; online. [doi: 10.18653/v1/2020.emnlp-demos.6]
52. Gong M, Shou L, Lin W, Sang Z, Yan Q, Yang Z, et al. NeuronBlocks: building your NLP DNN models like playing Lego. arXiv. Preprint published on April 21, 2019. URL: https://arxiv.org/abs/1904.09535 [accessed 2022-08-12]
53. COVID-19 misinformation detection: machine-learned solutions to the infodemic. GitHub. URL: https://github.com/yunongLiu1/COVID-19-Misinformation-Detection--Machine-Learned-Solutions-to-the-Infodemic [accessed 2022-08-13]
54. Poynter. 2020. URL: https://www.poynter.org/ [accessed 2022-04-11]
55. News API. URL: https://newsapi.org/ [accessed 2020-12-06]
56. Shu K, Mahudeswaran D, Wang S, Lee D, Liu H. FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media. Big Data 2020;8(3):171-188. [doi: 10.1089/big.2020.0062]
57. Shu K, Sliva A, Wang S, Tang J, Liu H. Fake news detection on social media: a data mining perspective. arXiv. Preprint published on August 7, 2017. URL: https://arxiv.org/abs/1708.01967 [accessed 2022-08-12]
58. Setting up a study on Prolific. Prolific. URL: https://researcher-help.prolific.co/hc/en-gb/articles/4407449546002-Audience [accessed 2022-07-17]
59. Inter-rater reliability IRR: definition, calculation. Statistics How To. URL: https://www.statisticshowto.com/inter-rater-reliability/ [accessed 2022-07-17]
60. Shabankhani B, Charati JY, Shabankhani K, Cherati SK. Survey of agreement between raters for nominal data using Krippendorff's alpha. Arch Pharma Pract 2020;10(S1):160-164.

Abbreviations

BERT: Bidirectional Encoder Representations from Transformers
Bi-LSTM: bidirectional long short-term memory
BNB: Bernoulli naïve Bayes
CNN: convolutional neural network
FNN: FakeNewsNet
GPT: Generative Pre-trained Transformer
IRB: institutional review board
LR: logistic regression
NLP: natural language processing
PLM: pretrained language model
RQ: research question
SVM: support vector machine

XSL•FO
RenderX

Edited by W Ahmed; submitted 14.04.22; peer-reviewed by J Banda, N Döring; comments to author 08.06.22; revised version received 20.07.22; accepted 08.08.22; published 25.08.22.

Please cite as:
Kolluri N, Liu Y, Murthy D
COVID-19 Misinformation Detection: Machine-Learned Solutions to the Infodemic
JMIR Infodemiology 2022;2(2):e38756
URL: https://infodemiology.jmir.org/2022/2/e38756
doi:10.2196/38756
PMID:37113446

©Nikhil Kolluri, Yunong Liu, Dhiraj Murthy. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 25.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Emotions and Incivility in Vaccine Mandate Discourse: Natural Language Processing Insights

Hannah Stevens[1], BA; Muhammad Ehab Rasul[1], MA; Yoo Jung Oh[1], MA

University of California, Davis, Davis, CA, United States

Corresponding Author:
Hannah Stevens, BA
University of California, Davis
1 Shields Ave
Davis, CA, 95616
United States
Phone: 1 530 752 0966
Email: hrstevens@ucdavis.edu

Abstract

Background: Despite vaccine availability, vaccine hesitancy has inhibited public health officials' efforts to mitigate the COVID-19 pandemic in the United States. Although some US elected officials have responded by issuing vaccine mandates, others have amplified vaccine hesitancy by broadcasting messages that minimize vaccine efficacy. The politically polarized nature of COVID-19 information on social media has given rise to incivility, wherein health attitudes often hinge more on political ideology than science.

Objective: To the best of our knowledge, incivility has not been studied in the context of discourse regarding COVID-19 vaccines and mandates. Specifically, there is little focus on the psychological processes that elicit uncivil vaccine discourse and behaviors. Thus, we investigated 3 psychological processes theorized to predict discourse incivility—namely, anxiety, anger, and sadness.

Methods: We used 2 different natural language processing approaches: (1) the Linguistic Inquiry and Word Count computational tool and (2) the Google Perspective application programming interface (API) to analyze a data set of 8014 tweets containing terms related to COVID-19 vaccine mandates from September 14, 2021, to October 1, 2021. To collect the tweets, we used the Twitter API Tweet Downloader Tool (version 2). Subsequently, we filtered through a data set of 375,000 vaccine-related tweets using keywords to extract tweets explicitly focused on vaccine mandates. We relied on the Linguistic Inquiry and Word Count computational tool to measure the valence of linguistic anger, sadness, and anxiety in the tweets. To measure dimensions of post incivility, we used the Google Perspective API.

Results: This study resolved discrepant operationalizations of incivility by introducing incivility as a multifaceted construct and explored the distinct emotional processes underlying 5 dimensions of discourse incivility. The findings revealed that 3 types of emotions—anxiety, anger, and sadness—were uniquely associated with dimensions of incivility (eg, toxicity, severe toxicity, insult, profanity, threat, and identity attacks). Specifically, the results showed that anger was significantly positively associated with all dimensions of incivility (all $P<.001$), whereas sadness was significantly positively related to threat ($P=.04$). Conversely, anxiety was significantly negatively associated with identity attack ($P=.03$) and profanity ($P=.02$).

Conclusions: The results suggest that our multidimensional approach to incivility is a promising alternative to understanding and intervening in the psychological processes underlying uncivil vaccine discourse. Understanding specific emotions that can increase or decrease incivility such as anxiety, anger, and sadness can enable researchers and public health professionals to develop effective interventions against uncivil vaccine discourse. Given the need for real-time monitoring and automated responses to the spread of health information and misinformation on the web, social media platforms can harness the Google Perspective API to offer users immediate, automated feedback when it detects that a comment is uncivil.

(JMIR Infodemiology 2022;2(2):e37635) doi:10.2196/37635

KEYWORDS

vaccine hesitancy; COVID-19; vaccine mandates; natural language processing; incivility; LIWC; Linguistic Inquiry and Word Count; Twitter

XSL•FO
RenderX

Introduction

Background

The emergence of the novel coronavirus (COVID-19) has caused 5,878,328 confirmed deaths worldwide as of February 2022, along with 423,437,674 confirmed infections [1]. Despite vaccine availability, vaccine hesitancy has inhibited public health officials' efforts to mitigate the COVID-19 pandemic, especially in the United States [2]. Although some US elected officials have responded by issuing vaccine mandates, others have amplified vaccine hesitancy by broadcasting messages that minimize vaccine efficacy [3,4].

With 68% of American adults reporting social media as a source of their news diet [5], social media platforms such as Twitter have become important communication channels for US politicians to share their agendas [6]. As a result, social media have become a prominent source of political information and misinformation, including information surrounding COVID-19 vaccines [7-11]. The politically polarized nature of COVID-19 information on social media has given rise to an infodemic, wherein health attitudes often hinge more on political ideology than science [12-15]. As a result, political affiliation influences negative sentiment toward the vaccine [16]. Such negative sentiment may foster uncivil discourse toward the vaccines and mandates [17,18].

Incivility on social media platforms has been widely studied and discussed in both political and health contexts, among others [19-25]. However, to the best of our knowledge, incivility has not been studied in the context of discourse regarding COVID-19 vaccines and mandates. Specifically, there is little focus on the psychological processes that elicit uncivil vaccine discourse. We aimed to bridge this gap by conducting a computational analysis of tweets. In this study, we investigated the role of negative emotion in predicting uncivil posts about COVID-19 vaccine mandates on Twitter. Ultimately, we argue that a more nuanced understanding of the psychological processes underlying uncivil vaccine discourse has practical implications for public health interventions.

The Role of Negative Emotion in Vaccine Mandate Incivility

Incivility has become a salient point of discussion in social media research. However, scholars across fields have found it difficult to conceptualize incivility. Incivility has been defined in a variety of ways, including impoliteness, profanity, and specific discriminatory acts (eg, former US president Trump caught on a hot mic in 2016 praising nonconsensual sexual encounters with women) [26-29]. Papacharissi [29] supplements this definition by including threat—in this case to democracy—as uncivil. Other scholars have operationalized incivility as including the use of all capital letters, accusations of lying, pejorative speech, ideologically extreme language, exaggerated argument, and misinformation [26,30-33]. Despite these inconsistent operationalizations, incivility is a concept that is nuanced and varies across individuals, perhaps because it is bound by cultural perceptions and understandings of what uncivil discourse is [16,18]. Informed by the operational inconsistency of incivility outlined in the literature, we conceptualize incivility as a multifaceted construct encompassing a diversity of uncivil behaviors, including toxicity, severe toxicity, profanity, threats, insults, and identity attacks in discourse. Recent studies have argued that uncivil behaviors are related to toxicity on social media platforms [34]. Tromble [28] asserts that profanity and insulting language constitute key indicators of uncivil behaviors. Likewise, scholars have argued that identity attacks and threatening language that aims to morally attack individuals or groups are also aspects of incivility and uncivil discourse [35]. We now shift our attention to explaining what causes incivility.

Incivility does not have a single cause; instead, varying forms of uncivil behaviors are a result of diverse psychological processes. For example, a user may post profane content because they are anxious, whereas a user might make an insulting comment because they are angry. However, scholars often obscure these distinct underlying psychological mechanisms by conceptualizing incivility as a one-dimensional process with a unitary explanation [19,21]. In the context of COVID-19 vaccines and mandates, emotional responses such as anger and anxiety among other negative emotions are salient in the discourse about the pandemic [36,37]. In fact, studies have found negative emotions such as anger and anxiety to play a role in driving vaccine hesitancy [38]. We investigated 3 psychological processes that are likely to predict discourse incivility—namely, anxiety, anger, and sadness.

Anxiety and Incivility

Anxiety about the safety of the COVID-19 vaccine, paired with dismissive attitudes toward COVID-19's threat, has a sizable segment of the United States indicating their unwillingness to get vaccinated [38-40]. In line with extant theory asserting that fear-based aggression is the most prevalent when a perceived threat is inescapable [41-43], a fear of harm from the vaccine, as perpetuated by elected officials and media alike, is often followed by avoidance strategies (eg, refusing the vaccine) [9-11,44]. Accordingly, policies that mandate the hesitant to get vaccinated inhibit the ability to escape the threat, and as a result, individuals may react with incivility. Indeed, stress and anxiety have been demonstrated to predict a wealth of uncivil behaviors, including cyber aggression and bullying during COVID-19 [45-47]. Thus, we posit the following.

> Hypothesis (H) 1: Anxiety will positively predict post incivility.

Anger and Incivility

COVID-19 vaccine mandates have drawn the ire of segments of the United States, including political elites and media outlets who have fueled public outrage about the threat to personal freedoms that vaccine mandates impose [48,49]. Simultaneously, the lack of confidence in vaccine safety and efficacy has segments of the population feeling threatened by the health risks they perceive to be associated with the vaccine. Anger can be understood as an adaptive response to a threat [44]; indeed, a study by Featherstone and Zhang [44] found vaccine misinformation to negatively impact attitudes toward vaccines through anger. Although anger has the functional value of

XSL•FO
RenderX

suppressing fear and potentiating a sense of personal control in the face of threat, it can also propel uncivil behavior, including acts of aggression and dismissiveness directed toward those with opposing views [50-52]. Thus, we can expect anger to foster incivility in COVID-19 vaccine mandate discourse.

H2: Anger will positively predict incivility.

Sadness and Incivility

Feelings of sadness have been linked with uncivil behavior, including acts of cyber aggression [47,53]. The freedom to travel, remain employed, socialize in groups, eat in restaurants, go to the gym, and more is increasingly determined by one's vaccination status [54,55]. Thus, mandates that prohibit the unvaccinated from participating in the relationships and activities available to those who are vaccinated may exacerbate existing sadness and depression induced by preexisting COVID-19 lifestyle disruptors [56,57]. Furthermore, social exclusion can elicit sadness and feelings that a group (ie, the unvaccinated) has experienced wrongs that must be righted—a mindset political scientists have coined "victimhood" [58]. Victimhood mentality may prompt individuals to retaliate against vaccine mandates and manifest as uncivil behaviors. Accordingly, we predict the following.

H3: Sadness will positively predict incivility.

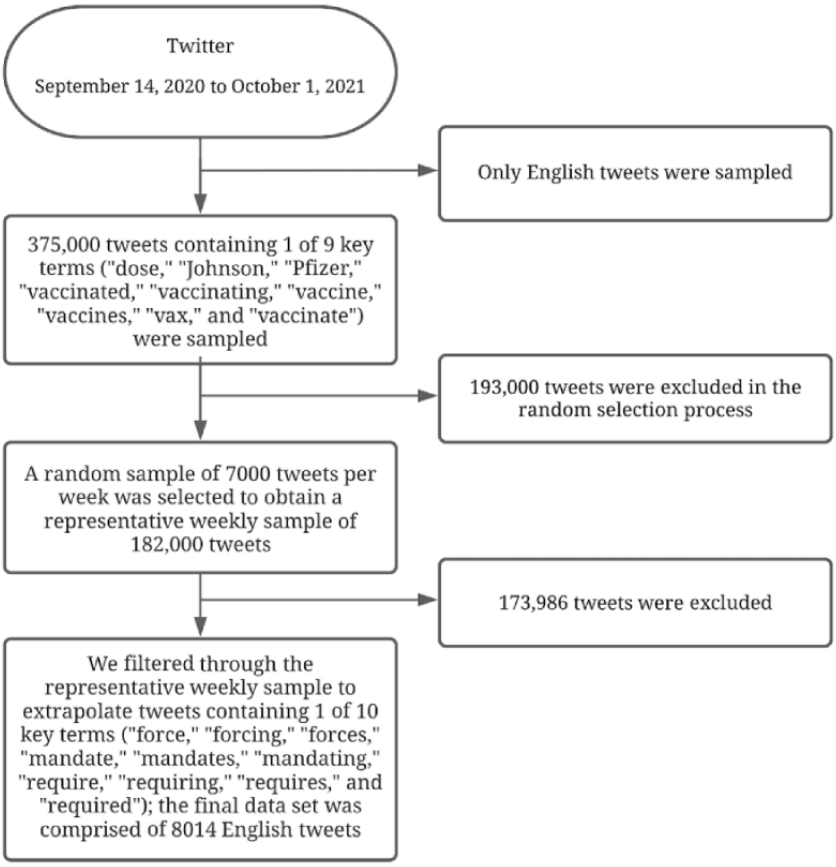
Methods

Data Collection

The sample comprised posts shared to Twitter, a popular platform for seeking and sharing health information on the web, including (mis)information about vaccination and vaccines [7-11]. We opted to curate a list of vaccine-related words and scraped tweets containing those words. We curated a list of words that we believed would collect tweets related to the vaccine, without introducing bias into the data set. For example, "shot" was not included, because we noticed that it scraped tweets about gunshots, which are unrelated to the COVID-19 vaccine. The text of the 8014 tweets contained terms related to COVID-19 vaccine mandates (eg, "Moderna," "required," and "mandating") from September 14, 2020, to October 1, 2021. See Figure 1 for a flowchart of the data collection process.

Twitter's code-free application programming interface (API) Tweet Downloader Tool (version 2) was used to extract posts about COVID-19 vaccine mandates. We were interested in words that would identify tweets about COVID-19 vaccine mandates rather than the COVID-19 vaccine generally. Thus, we filtered through a data set of 375,000 vaccine-related tweets posted from September 14, 2020, to October 1, 2021, to extrapolate tweets specifically related to vaccine mandates (eg, "forcing," "required," and "mandating") from September 14, 2020, to October 1, 2021; the final sample contained 8014 tweets.

Figure 1. Flowchart of the data collection process.



XSL•FO
RenderX

Natural Language Processing Procedures

The data were analyzed using 2 different natural language processing approaches: (1) the Linguistic Inquiry and Word Count (LIWC) computational tool [59] and (2) the Google Perspective API [60].

LIWC Sentiment Analysis

LIWC is a natural language processing tool that measures psychological processes in texts by counting the percentage of words in a given tweet that fall into prespecified categories. It has been validated and used in investigations of mental health during the COVID-19 pandemic (eg, LGBTQ+ youth mental health) [12,61]. In contrast to other sentiment analysis lexicons that generate the valence of emotion (eg, Afinn and Bing, which assign texts a score from negative to positive) without extrapolating discrete emotions and sentiment analysis lexicons that produce binary outcomes (eg, NRC), we wanted a continuous measure of the extent to which texts had a particular sentiment [62]. Although there are multiple tools that continuously capture sentiment and emotions using natural language processing methods (eg, IBM Watson) [63], we specifically used the LIWC dictionary for emotion classification, because compared to the aforementioned natural language processing tools, the LIWC dictionary has been validated in multiple studies, and thus, we considered that it would present a more accurate estimate of the level of emotions reflected in the textual data. We leveraged LIWC to measure the valence of linguistic anger (eg, "frustrated," and "annoyed"), sadness (eg, "hopeless," and "miserable"), and anxiety (eg, "afraid," and "stressed") in texts [59]. Tweets had an average anxiety score of 0.79 (SD 1.67), an average anger score of 0.11 (SD 0.75) and an average sadness score of 0.09 (SD 0.52).

Google Perspective API Machine Learning Analysis

To measure dimensions of post incivility, we used the Google Perspective API to measure levels of toxicity, severe toxicity, insult, profanity, threat, and identity attacks in tweets related to vaccine mandates (see Table 1) [60]. The Google Perspective API is a tool designed by Google's Counter-Abuse Technology Team that measures incivility in web-based posts.

The Google Perspective API model is trained by human coders on a data set of millions of comments from a variety of web-based sources, including forums (eg, Wikipedia). The model is robust and has been used in a variety of contexts, from political incivility to rape culture to COVID-19 vaccine information [21,64,65]. For example, Hopp et al [64] asked respondents to self-report the degree to which they engage in uncivil communication on the web and then correlated that with trace data of participants' social media content. The results indicated that those who self-disclose engaging in uncivil social media behavior also tend to generate uncivil content on social media, measured via the Google Perspective API. These dimensions of incivility have been tested across multiple domains and trained on substantial amounts of human-annotated comments [60].

Table 1. Incivility variable attributes.

Attribute name	Perspective API[a] description [60]	Example post[b]
Severe toxicity	"A very hateful, aggressive, disrespectful comment or otherwise very likely to make a user leave a discussion or give up on sharing their perspective."	"F*ck the vaccine and f*ck COVID, this should not be required period!!!"
Identity attack	"Negative or hateful comments targeting someone because of their identity."	"DO NOT COMPLY. Screw liberals and their idiotic vaccine mandate."
Insult	"Insulting, inflammatory, or negative comment towards a person or a group of people."	"Bank accounts are frozen for protesting mandates. How many more vaccines will you take before you wisen up? Wake up you stupid little sheep."
Profanity	"Swear words, curse words, or other obscene or profane language."	"It must be hard to be a victim of the vaccine mandate. A**holes on the internet FROTH at the F*CKING mouth to dismiss your experience."
Threat	"Describes an intention to inflict pain, injury, or violence against an individual or group."	"I'll put a bullet in someone who tries to force my kid to get the vaccine."

[a]API: application programming interface.

[b]Curse words have been censored to make the table suitable for publication.

Ethical Considerations

No personally identifiable information was included in this study. The institutional review board recognizes that the analysis of publicly available data does not constitute human subjects research. This study only used information in the public domain; thus, ethical review and approval was not required.

Results

Factor Analysis of Dimensions of Uncivil Discourse

Prior to hypothesis testing, we conducted a repeated measures ANOVA to assess whether to model dimensions of incivility together or separately. The main effect for the within-subjects factor was significant ($F_{4,32052}$=930.44; P<.001), indicating significant differences among identity attack, insult, profanity, threat, and severe toxicity (see Table 2).

XSL•FO
RenderX

Tukey comparisons were used to test marginal mean differences in each combination of incivility dimensions. There were significant differences between each combination, except identity attack and profanity (see Table 3). Thus, we concluded that the 5 dimensions of incivility should be assessed separately in the main analysis.

Table 2. Means table for within-subject variables (N=8014).

Incivility dimension	Mean (SD)
Severe toxicity	0.10 (0.14)
Identity attack	0.12 (0.12)
Insult	0.18 (0.20)
Profanity	0.12 (0.18)
Threat	0.17 (0.15)

Table 3. The marginal means contrasts for each combination of within-subject variables for the repeated measures ANOVA.

Contrast	Difference	SE	*t* test (df)	*P* value
Severe toxicity – identity attack	–0.02	0.001	–15.11 (8013)	<.001
Severe toxicity – insult	–0.08	0.001	–66.07 (8013)	<.001
Severe toxicity – profanity	–0.02	0.0008	–25.79 (8013)	<.001
Severe toxicity – threat	–0.06	0.001	–43.18 (8013)	<.001
Identity attack – insult	–0.06	0.002	–36.78 (8013)	<.001
Identity attack – profanity	–0.004	0.002	–2.39 (8013)	.12
Identity attack – threat	–0.05	0.002	–30.34 (8013)	<.001
Insult – profanity	0.06	0.001	43.06 (8013)	<.001
Insult – threat	0.01	0.002	6.30 (8013)	<.001
Profanity – threat	–0.04	0.002	–21.48 (8013)	<.001

Logistic Regression Analyses

Dichotomizing the Data

The skewed distribution of the data necessitated that we dichotomize the incivility dimensions for regression. The Google Perspective API recommends flagging a comment as having an attribute if it scores a 0.7 or higher—thus, this value was used to dichotomize the data for logistic regression [60]. Of the 8014 tweets, 53 (0.66%) contained identity attacks, 405 (5.05%) contained insults, 317 (3.96%) contained profanity, 137 (1.71%) contained threats, and 91 (1.14%) contained severe toxicity.

For hypothesis testing, we conducted 5 logistic regression analyses to assess whether anger, anxiety, and sadness in posts predicted uncivil tweets (see Table 4 and Figure 2). Variance inflation factors for anxiety, sadness, and anger on all dimensions of incivility were less than 1.5, indicating there was not any multicollinearity between our independent variables.

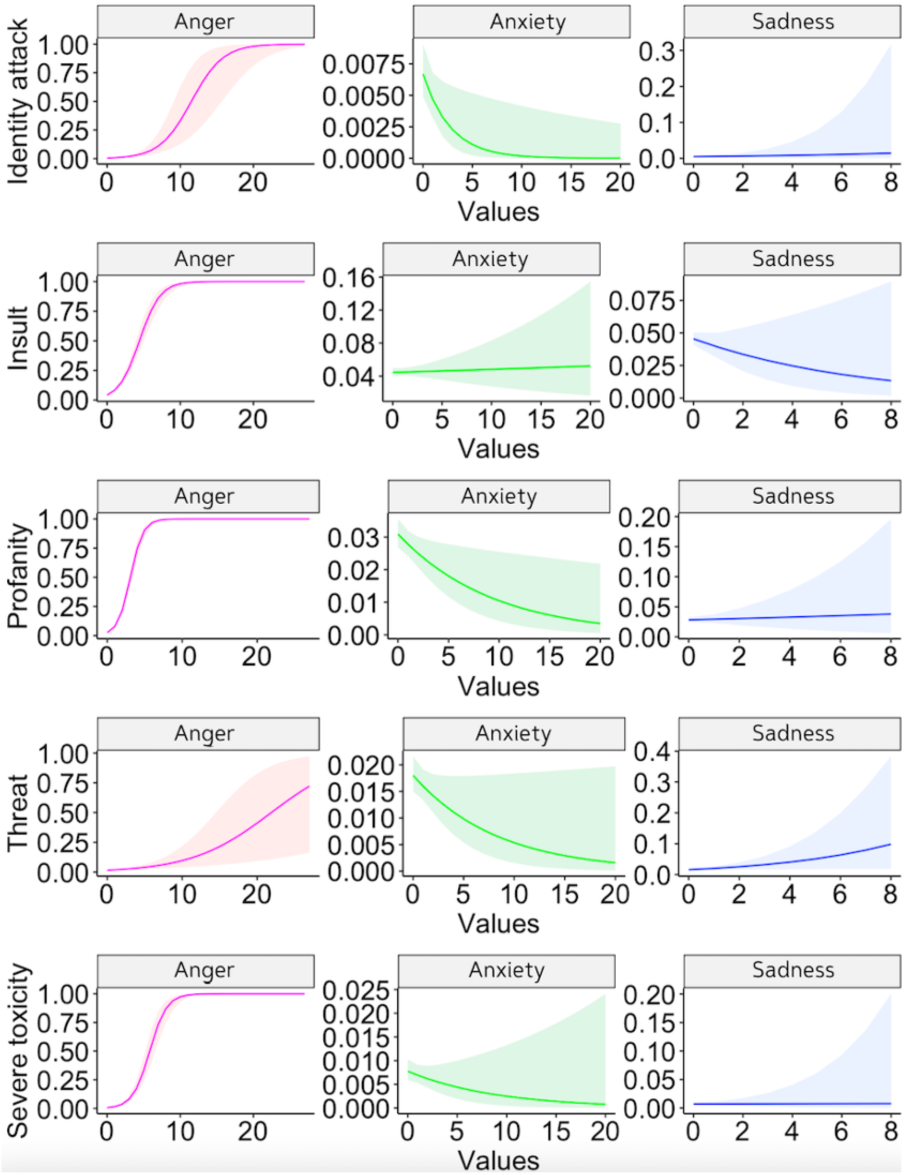
XSL•FO
RenderX

Table 4. Binary logistic regression results with anxiety, anger, and sadness predicting dimensions of incivility. McFadden R2 was used to calculate model fit.

Variable	Odds ratio (95% CI)	*B*	*P* value	R^2	χ^2_3
Threat				.01	18.78
(Intercept)		–4.04	<.001		
Anxiety	0.88 (0.78-1.01)	–.12	.06		
Sadness	1.27 (1.02-1.58)	.24	.04		
Anger	1.21 (1.10-1.33)	.19	<.001		
Identity attack				.09	58.64
(Intercept)		–5.06	<.001		
Anxiety	0.70 (0.50-0.96)	–.36	.03		
Sadness	1.15 (0.74-1.77)	.14	.54		
Anger	1.59 (1.40-1.80)	.46	<.001		
Profanity				.22	567.15
(Intercept)		–3.58	<.001		
Anxiety	0.90 (0.81-0.98)	–.11	.02		
Sadness	1.04 (0.83-1.31)	.04	.75		
Anger	3.27 (2.93-3.67)	1.19	<.001		
Insult				.08	258.25
(Intercept)		–3.13	<.001		
Anxiety	1.01 (0.95-1.07)	.008	.79		
Sadness	0.85 (0.67-1.10)	–.16	.22		
Anger	2.03 (1.85-2.23)	.71	<.001		
Severe toxicity				.24	239.27
(Intercept)		–.45	<.001		
Anxiety	0.89 (0.75-1.06)	–.11	.20		
Sadness	1.01 (0.65-1.57)	.01	.96		
Anger	2.37 (2.12-2.66)	.86	<.001		

XSL•FO
RenderX

Figure 2. Negative emotion predicting the odds of severe toxicity, threat, profanity, insult, and identity attack. Scores for anger, anxiety, and sadness were computed using the Linguistic Inquiry and Word Count computerized coding tool that measures psychological processes in texts by counting the percentage of words in a given tweet that fall into prespecified categories.



Anxiety

We found that the effect of anxiety on identity attack (*B*=–.36; odds ratio [OR] 0.70; *P*=.03) and profanity (*B*=–.11; OR 0.90; *P*=.02) were significant. However, contrary to our prediction that linguistic anxiety would increase incivility (H1), the results indicated that anxiety decreased the odds of identity attacks and profanity by approximately 30.48% and 10.43%, respectively. The results also reflected a stronger relationship between anxiety and identity attack than profanity. No other significant differences were found.

Anger

Consistent with our hypothesis (H2), the effect of anger on all 5 dimensions of incivility was significant (all *P*<.001). The results revealed that anger predicted the odds of profanity, insult, and severe toxicity to a greater extent than identity attacks and threats. The effect of the anger on threat (*B*=.19; OR 1.21; *P*<.001) and identity attack (*B*=.46; OR 1.59; *P*<.001) indicated that a 1-unit increase in anger increased the odds of threats by approximately 20.67% and identity attacks by approximately 58.9%. The effect of anger on insult (*B*=.71; OR 2.03; *P*<.001) and severe toxicity (*B*=.86; OR 2.37; *P*<.001) indicated that an increase in anger increased the odds of insults by approximately 103.15% and severe toxicity by approximately 137.29%. The results indicated that anger increased the odds of profanity the most (approximately 227.49%; *B*=1.19; OR 3.27; *P*<.001) when compared to the other 4 dimensions.

Sadness

H3 predicted that sadness will be positively associated with the level of incivility expressed in tweets. Our results showed that the effect of sadness on threat was significant (*B*=.24; OR 1.27; *P*=.04), indicating that a 1-unit increase in sadness increased the odds of threats by approximately 26.86%. Sadness did not have a significant effect on any other dimension of incivility.

XSL•FO
RenderX

Discussion

Principal Findings

Incivility has been understood as a multifaceted construct, encompassing the breadth of conceptual and operational definitions offered in the literature. This study resolved discrepant operationalizations of incivility by introducing incivility as a multifaceted construct and explored the distinct emotional processes underlying 5 dimensions of discourse incivility. The findings reveal that 3 types of emotions—anxiety, anger, and sadness—were significantly associated with dimensions of incivility. With regard to the relationship between anxiety and incivility, we found that the anxiety was negatively associated with identity attacks and profanity expressed in Twitter posts. Individuals who expressed higher levels of anger were more likely to engage in all 5 dimensions of incivility, including profanity, insults, severe toxicity, identity attacks, and threats. Lastly, our findings revealed that sadness was positively associated with uncivil behavior, especially threats.

Comparison With Prior Work

Individuals who expressed higher anxiety were less likely to engage in uncivil behaviors such as posting hateful comments targeting individuals with a specific identity or using profane language in their posts. We suspect that individuals' anxiety may have decreased the level of uncivil expressions about vaccine mandate policy, because individuals who are anxious about COVID-19 and its health consequences are more likely to seek ways to contain the threat (ie, spread of COVID-19) and exhibit positive attitudes and behaviors toward policies related to restricting the spread of COVID-19. Namely, when novel threatening stimuli are encountered and feelings of anxiety are induced, people may be motivated to attend to the issue at hand [66]. In line with this idea, previous studies suggest that anxiety can be an indicator of a "functional fear" that predicts individuals' positive attitudes and behaviors (eg, compliance) toward COVID-19–related measures and policies [67]. For instance, an extant work shows that COVID-19–related anxiety and health-related fears were associated with more protective health behaviors and higher vaccine acceptance [68,69].

It is noteworthy that anger, unlike anxiety or sadness, predicted all dimensions of incivility, demonstrating that this emotion is the strongest predictor of incivility.

Evidence from previous studies has shown that prolonged risk and uncertainty about the level of risk can elicit anger and conflict within the community [70]. People have experienced increased levels of anger during the pandemic [71], and those who express anger have also exhibited disbelief toward COVID-19 vaccines [72]. Moreover, it has been shown that political polarization regarding the issues of vaccination and vaccine mandates has further fueled public outrage among groups with conflicting political views [51,52]. Thus, the strong association between anger and uncivil behaviors can be due to both social disruptions caused by the wide spread of COVID-19 and political conflicts partly induced by media outlets.

Lastly, as the level of sadness increased, individuals were more likely to exhibit verbal intentions to inflict pain and hurt other individuals or groups. Such aggression toward other people, especially exhibiting intentions to hurt others, may be explained by depression and victimhood. Approximately over 2 years of the COVID-19 pandemic, individuals worldwide have experienced prolonged social isolation and lifestyle disruptions, which have led them to be depressed [56,57]. Furthermore, the direct health impacts of the spread of COVID-19 have caused many individuals to become the victims of multiple losses such as a loss of financial security, loss of family members, and loss of physical/mental health and general safety [73,74]. However, sadness may have been strongly associated with viewing themselves as victims of COVID-19, which could have led them to issue threats to others who were favorable toward vaccine mandates. Additionally, this victimhood mentality [58] among the unvaccinated may have also been high because they are prohibited from participating in relationships and activities available to those who are vaccinated. This prohibition may have led them to feel socially excluded and in turn prompt threats toward the outgroup members—proponents of vaccine mandates.

Limitations

Although the findings shed light on the psychological processes underlying vaccine mandate incivility, this study is not without limitations. The LIWC computational tool does not measure the nuances afforded by human coders. Although we endeavored to minimize this limitation by using well-validated measures [59], future work might employ human coders to analyze the specific topics related to uncivil discourse. Additionally, we focused on posts shared to Twitter and therefore cannot generalize our findings about incivility to other social media platforms. Given the role of platform community norms in predicting incivility, future work should investigate how incivility manifests itself on different platforms. Likewise, Twitter users are wealthier, younger, and more liberal than the wider population of Americans [75], and the sample was limited to English-speaking Twitter users, which makes it difficult to generalize our findings to the entire US population. Additionally, we acknowledge that social media posting data could have been biased based on individuals' geographical area (eg, city and state), whether they were local residents or visitors in the area at the time of the post, as well as the types of activities completed during the course of a day [76,77]. These factors may have contributed to our study findings. Lastly, we did not measure how many different users were included in each stage in the data collection process. Future work should elucidate the extent to which a small number of active users produce uncivil vaccine mandate content.

Conclusions

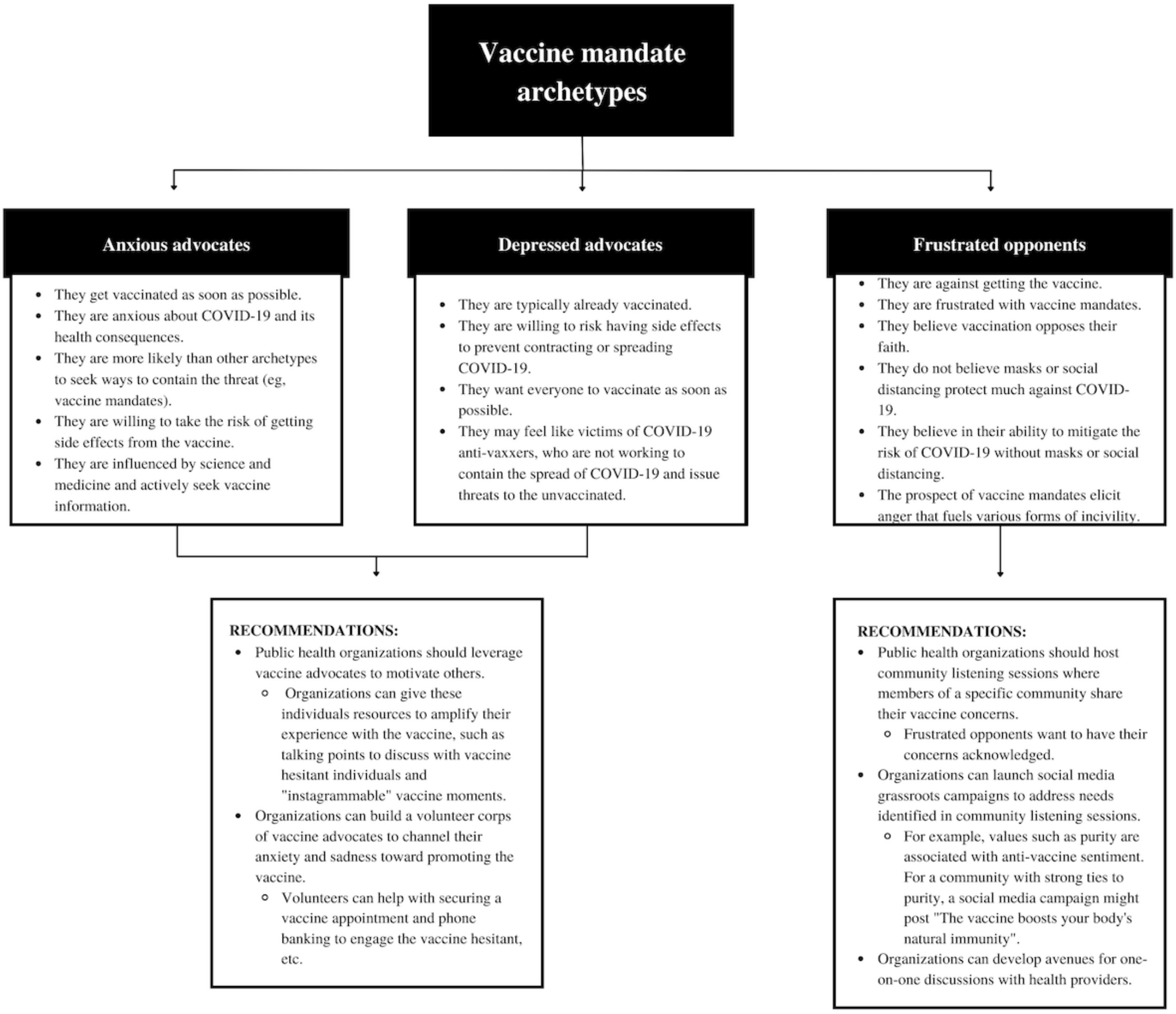
The results suggest that our multidimensional approach to incivility is a promising alternative to understanding and intervening in the psychological processes underlying uncivil vaccine discourse. Given the need for real-time monitoring and automated responses to the spread of health information and misinformation on the web, social media platforms can harness the Google Perspective API to offer users immediate, automated feedback when it detects that a comment is uncivil [78]. Furthermore, the Perspective API is available in 17

XSL•FO
RenderX

languages—from Arabic to Korean, enabling the study of uncivil health discourse in non-English posts. Future work should explore cross-cultural differences in uncivil health discourse.

Vaccine hesitancy still remains a threat to global health, and this work demonstrates that distinct emotional processes underlie distinct attitudes toward vaccines and vaccine-related policies. It is important for health practitioners and policy makers to first acknowledge negative emotions associated with vaccines and vaccine mandates while emphasizing the safety of COVID-19 vaccines in health campaigns, which would provide aid in reducing vaccine hesitancy. One avenue public health officials can take to combat vaccine hesitancy while simultaneously affirming discrete negative emotions toward the vaccine is by holding COVID-19 community listening sessions, where officials can hear directly from communities about COVID-19 concerns, including vaccination (see Figure 3 for an overview) [79]. After officials have a better understanding of the specific emotional processes underlying a communities' vaccine hesitancy, public health campaigns can tailor messages to address these concerns (see Figure 3) [80,81].

Figure 3. Concrete recommendations for promoting vaccine uptake based on underlying emotions.



Conflicts of Interest

None declared.

References

1. WHO coronavirus (COVID-19) dashboard. World Health Organization. URL: https://covid19.who.int/ [accessed 2022-08-21]
2. Chang K. Fauci and other top U.S. health officials stress the urgency of vaccinations and their safety. The New York Times. 2021 Apr 25. URL: https://www.nytimes.com/2021/04/25/health/fauci-vaccines-safety.html [accessed 2022-08-21]
3. See where 12 million U.S. employees are affected by government vaccine mandates. The New York Times. 2021 Dec 18. URL: https://www.nytimes.com/interactive/2021/12/18/us/vaccine-mandate-states.html [accessed 2022-08-21]

XSL•FO
RenderX

4. Sabahelzain MM, Hartigan-Go K, Larson HJ. The politics of COVID-19 vaccine confidence. Curr Opin Immunol 2021 Aug;71:92-96 [FREE Full text] [doi: 10.1016/j.coi.2021.06.007] [Medline: 34237648]
5. Hitlin P, Olmstead K. The science people see on social media. Pew Research Center. 2018 Mar 21. URL: https://www.pewresearch.org/science/2018/03/21/the-science-people-see-on-social-media/ [accessed 2022-09-06]
6. Buccoliero L, Bellio E, Crestini G, Arkoudas A. Twitter and politics: evidence from the US presidential elections 2016. Journal of Marketing Communications 2018 Aug 16;26(1):88-114 [FREE Full text] [doi: 10.1080/13527266.2018.1504228]
7. Benis A, Chatsubi A, Levner E, Ashkenazi S. Change in threads on Twitter regarding influenza, vaccines, and vaccination during the COVID-19 pandemic: artificial intelligence-based infodemiology study. JMIR Infodemiology 2021 Oct 14;1(1):e31983 [FREE Full text] [doi: 10.2196/31983] [Medline: 34693212]
8. De Choudhury M, Morris MR, White RW. Seeking and sharing health information online: comparing search engines and social media. Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI'14 New York, New York, USA: ACM Press; 2014 Apr 26 Presented at: CHI'14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 24 to May 1, 2014; Toronto, ON p. 1365-1376. [doi: 10.1145/2556288.2557214]
9. Featherstone JD, Bell RA, Ruiz JB. Relationship of people's sources of health information and political ideology with acceptance of conspiratorial beliefs about vaccines. Vaccine 2019 May 21;37(23):2993-2997. [doi: 10.1016/j.vaccine.2019.04.063] [Medline: 31031028]
10. Ruiz JB, Bell RA. Predictors of intention to vaccinate against COVID-19: results of a nationwide survey. Vaccine 2021 Feb 12;39(7):1080-1086 [FREE Full text] [doi: 10.1016/j.vaccine.2021.01.010] [Medline: 33461833]
11. Ruiz J, Featherstone JD, Barnett GA. Identifying vaccine hesitant communities on twitter and their geolocations: a network approach. 2021 Jan 05 Presented at: Proceedings of the 54th Hawaii International Conference on System Sciences; January 5, 2021; Kauai, HI p. 3964-3970. [doi: 10.24251/hicss.2021.480]
12. Stevens H, Palomares NA. Constituents' inferences of local governments' goals and the relationship between political party and belief in COVID-19 misinformation: cross-sectional survey of Twitter followers of state public health departments. JMIR Infodemiology 2022 Feb 10;2(1):e29246. [doi: 10.2196/29246]
13. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194 [FREE Full text] [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
14. Muhamad JW, Merle P. Identity and information overload: examining the impact of health messaging in times of crisis. In: O'Hair HD, O'Hair MJ, editors. Communicating Science in Times of Crisis. Hoboken, NJ: Wiley; 2021:110-127.
15. Scoville C, McCumber A, Amironesei R, Jeon J. Mask refusal backlash: the politicization of face masks in the American public sphere during the early stages of the COVID-19 pandemic. Socius 2022 Apr 25;8:237802312210931. [doi: 10.1177/23780231221093158]
16. Kennedy J. Populist politics and vaccine hesitancy in Western Europe: an analysis of national-level data. Eur J Public Health 2019 Jun 01;29(3):512-516. [doi: 10.1093/eurpub/ckz004] [Medline: 30801109]
17. Landler M. Vaccine mandates rekindle fierce debate over civil liberties. The New York Times. 2021 Dec 10. URL: https://www.nytimes.com/2021/12/10/world/europe/vaccine-mandates-civil-liberties.html [accessed 2022-09-06]
18. Largent EA, Persad G, Sangenito S, Glickman A, Boyle C, Emanuel EJ. US public attitudes toward COVID-19 vaccine mandates. JAMA Netw Open 2020 Dec 01;3(12):e2033324 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.33324] [Medline: 33337490]
19. Masullo Chen G, Muddiman A, Wilner T, Pariser E, Stroud NJ. We should not get rid of incivility online. Soc Media Soc 2019 Jul 16;5(3):205630511986264. [doi: 10.1177/2056305119862641]
20. Muddiman A, Stroud NJ. News values, cognitive biases, and partisan incivility in comment sections: uncivil comments. J Commun 2017 Jul 18;67(4):586-609. [doi: 10.1111/jcom.12312]
21. Stevens H, Acic I, Taylor LD. Uncivil reactions to sexual assault online: Linguistic features of news reports predict discourse incivility. Cyberpsychol Behav Soc Netw 2021 Dec;24(12):815-821. [doi: 10.1089/cyber.2021.0075] [Medline: 34515520]
22. Sun Q, Wojcieszak M, Davidson S. Over-time trends in incivility on social media: evidence from political, non-political, and mixed sub-reddits over eleven years. Front Polit Sci 2021 Nov 2;3:741605. [doi: 10.3389/fpos.2021.741605]
23. Theocharis Y, Barberá P, Fazekas Z, Popa SA. The dynamics of political incivility on Twitter. SAGE Open 2020 May 13;10(2):215824402091944. [doi: 10.1177/2158244020919447]
24. Wojcieszak M, Azrout R, Vreese D. Waving the red cloth: media coverage of a contentious issue triggers polarization. Public Opin Q 2018 Dec 09;82(1):87-109. [doi: 10.1093/poq/nfx040]
25. Dubey AD. The resurgence of cyber racism during the COVID-19 pandemic and its aftereffects: analysis of sentiments and emotions in tweets. JMIR Public Health Surveill 2020 Oct 15;6(4):e19833 [FREE Full text] [doi: 10.2196/19833] [Medline: 32936772]
26. Muddiman A. Personal and public levels of political incivility. Int J Commun 2017 Aug;11:3182-3202.
27. Muddiman A, Kearney M. What counts as incivility in the Trump era? an inductive approach to political incivility. 2018 Presented at: 2018 Annual Meeting of the Political Communication Division of the American Society of Political Scientists; August 30 to September 2, 2018; Boston, MA.

XSL•FO
RenderX

28. Tromble R. Twitter "Healthy Conversations". Rebekah Tromble. URL: https://www.rebekahtromble.net/political-communication [accessed 2022-08-21]
29. Papacharissi Z. Democracy online: civility, politeness, and the democratic potential of online political discussion groups. New Media & Society 2016 Jun 30;6(2):259-283. [doi: 10.1177/1461444804041444]
30. Sobieraj S, Berry JM. From incivility to outrage: political discourse in blogs, talk radio, and cable news. Polit Commun 2011 Feb 09;28(1):19-41. [doi: 10.1080/10584609.2010.542360]
31. Coe K, Kenski K, Rains SA. Online and uncivil? patterns and determinants of incivility in newspaper website comments. J Commun 2014 Jun 16;64(4):658-679. [doi: 10.1111/jcom.12104]
32. Gervais BT. Incivility online: affective and behavioral reactions to uncivil political posts in a web-based experiment. J Inf Technol Politics 2015 Jan 14;12(2):167-185. [doi: 10.1080/19331681.2014.997416]
33. Stroud NJ, Scacco JM, Muddiman A, Curry AL. Changing deliberative norms on news organizations' Facebook sites. J Comput-Mediat Comm 2014 Oct 16;20(2):188-203. [doi: 10.1111/jcc4.12104]
34. Kim JW, Guess A, Nyhan B, Reifler J. The distorting prism of social media: how self-selection and exposure to incivility fuel online comment toxicity. J Commun 2021 Sep 03;71(6):922-946. [doi: 10.1093/joc/jqab034]
35. Rossini P. Disentangling uncivil and intolerant discourse in online political talk. In: A Crisis of Civility. New York, NY: Routledge; 2019:142-157.
36. Lwin MO, Sheldenkar A, Lu J, Schulz PJ, Shin W, Panchapakesan C, et al. The evolution of public sentiments during the COVID-19 pandemic: case comparisons of India, Singapore, South Korea, the United Kingdom, and the United States. JMIR Infodemiology 2022 Feb 10;2(1):e31473. [doi: 10.2196/31473]
37. 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]
38. Chou WS, Budenz A. Considering emotion in COVID-19 vaccine communication: addressing vaccine hesitancy and fostering vaccine confidence. Health Commun 2020 Oct 30;35(14):1718-1722. [doi: 10.1080/10410236.2020.1838096] [Medline: 33124475]
39. 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 Jul 22;11(8):1119-1128. [doi: 10.1177/1948550620940539]
40. Wiysonge CS, Ndwandwe D, Ryan J, Jaca A, Batouré O, Anya BM, et al. Vaccine hesitancy in the era of COVID-19: could lessons from the past help in divining the future? Hum Vaccin Immunother 2022 Dec 31;18(1):1-3 [FREE Full text] [doi: 10.1080/21645515.2021.1893062] [Medline: 33684019]
41. Blanchard DC. Translating dynamic defense patterns from rodents to people. Neurosci Biobehav Rev 2017 May;76(Pt A):22-28. [doi: 10.1016/j.neubiorev.2016.11.001] [Medline: 28434585]
42. Carver CS, Harmon-Jones E. Anger is an approach-related affect: evidence and implications. Psychol Bull 2009 Mar;135(2):183-204. [doi: 10.1037/a0013965] [Medline: 19254075]
43. Kessels LTE, Ruiter RAC, Wouters L, Jansma BM. Neuroscientific evidence for defensive avoidance of fear appeals. Int J Psychol 2014 Apr;49(2):80-88 [FREE Full text] [doi: 10.1002/ijop.12036] [Medline: 24811878]
44. Featherstone JD, Zhang J. Feeling angry: the effects of vaccine misinformation and refutational messages on negative emotions and vaccination attitude. J Health Commun 2020 Sep 01;25(9):692-702. [doi: 10.1080/10810730.2020.1838671] [Medline: 33103600]
45. Rief W. Fear of adverse effects and COVID-19 vaccine hesitancy: recommendations of the treatment expectation expert group. JAMA Health Forum 2021 Apr 16;2(4):e210804. [doi: 10.1001/jamahealthforum.2021.0804]
46. Barlett CP, Rinker A, Roth B. Cyberbullying perpetration in the COVID-19 era: an application of general strain theory. J Soc Psychol 2021 Jul 04;161(4):466-476. [doi: 10.1080/00224545.2021.1883503] [Medline: 33543673]
47. Wang Q, Luo X, Tu R, Xiao T, Hu W. COVID-19 information overload and cyber aggression during the pandemic lockdown: the mediating role of depression/anxiety and the moderating role of Confucian responsibility thinking. Int J Environ Res Public Health 2022 Jan 29;19(3):1540 [FREE Full text] [doi: 10.3390/ijerph19031540] [Medline: 35162578]
48. Brnovich M. Why states are fighting Biden's vaccine mandates Internet. Fox News. 2022 Jan 24. URL: https://www.foxnews.com/opinion/why-states-fighting-bidens-vaccine-mandates-arizona-ag-mark-brnovich [accessed 2022-08-21]
49. Krugman P. What to do with our pandemic anger. The New York Times. 2022 Feb 08. URL: https://www.nytimes.com/2022/02/07/opinion/covid-unvaccinated-anger.html [accessed 2022-08-21]
50. Novaco RW. Anger. In: Zeigler-Hill V, Shackelford TK, editors. Encyclopedia of Personality and Individual Differences. Cham, Switzerland: Springer; Apr 22, 2020:166-171.
51. Kerr J, Panagopoulos C, van der Linden S. Political polarization on COVID-19 pandemic response in the United States. Pers Individ Dif 2021 Sep;179:110892 [FREE Full text] [doi: 10.1016/j.paid.2021.110892] [Medline: 34866723]
52. Pennycook G, Rand DG. The psychology of fake news. Trends Cogn Sci 2021 May;25(5):388-402 [FREE Full text] [doi: 10.1016/j.tics.2021.02.007] [Medline: 33736957]

XSL•FO
RenderX

53. Savage J, Verhulst B, Copeland W, Althoff RR, Lichtenstein P, Roberson-Nay R. A genetically informed study of the longitudinal relation between irritability and anxious/depressed symptoms. J Am Acad Child Adolesc Psychiatry 2015 May;54(5):377-384 [FREE Full text] [doi: 10.1016/j.jaac.2015.02.010] [Medline: 25901774]
54. Balch B. The cost of being unvaccinated is rising — will people be willing to pay the price? Association of American Medical Colleges. 2021 Oct 26. URL: https://www.aamc.org/news-insights/cost-being-unvaccinated-rising-will-people-be-willing-pay-price [accessed 2022-08-21]
55. Money L, Lin RGII, Breijo S. L.A. County plans to require proof of vaccination at indoor bars, nightclubs, breweries, wineries. Los Angeles Times. 2021 Sep 15. URL: https://www.latimes.com/california/story/2021-09-15/la-to-require-covid-19-vaccines-at-bars-nightclubs-breweries-wineries [accessed 2022-08-21]
56. Bignardi G, Dalmaijer ES, Anwyl-Irvine AL, Smith TA, Siugzdaite R, Uh S, et al. Longitudinal increases in childhood depression symptoms during the COVID-19 lockdown. Arch Dis Child 2020 Dec 09:791-797 [FREE Full text] [doi: 10.1136/archdischild-2020-320372] [Medline: 33298552]
57. Palgi Y, Shrira A, Ring L, Bodner E, Avidor S, Bergman Y, et al. The loneliness pandemic: loneliness and other concomitants of depression, anxiety and their comorbidity during the COVID-19 outbreak. J Affect Disord 2020 Oct 01;275:109-111 [FREE Full text] [doi: 10.1016/j.jad.2020.06.036] [Medline: 32658811]
58. Horwitz RB. Politics as victimhood, victimhood as politics. J Policy Hist 2018 Jun 21;30(3):552-574. [doi: 10.1017/s0898030618000209]
59. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. The Universityt of Texas at Austin. 2015. URL: https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf [accessed 2022-08-21]
60. Perspective API. URL: https://developers.perspectiveapi.com/s/about-the-api-attributes-and-languages [accessed 2022-08-21]
61. Stevens HR, Acic I, Rhea S. Natural language processing insight into LGBTQ+ youth mental health during the COVID-19 pandemic: longitudinal content analysis of anxiety-provoking topics and trends in emotion in LGBTeens microcommunity subreddit. JMIR Public Health Surveill 2021 Aug 17;7(8):e29029 [FREE Full text] [doi: 10.2196/29029] [Medline: 34402803]
62. Miazga J, Hachaj T. Evaluation of most popular sentiment lexicons coverage on various datasets. 2019 Oct 08 Presented at: SSIP 2019: Proceedings of the 2019 2nd International Conference on Sensors, Signal and Image Processing; October 8-10, 2019; Prague, Czech Republic p. 86-90. [doi: 10.1145/3365245.3365251]
63. High R. The era of cognitive systems: an inside look at IBM Watson and how it works. Redbooks. 1994. URL: https://www.redbooks.ibm.com/redpapers/pdfs/redp4955.pdf [accessed 2022-09-06]
64. Hopp T, Vargo CJ, Dixon L, Thain N. Correlating self-report and trace data measures of incivility: a proof of concept. Soc Sci Comput Rev 2018 Dec 17;38(5):584-599. [doi: 10.1177/0894439318814241]
65. Pascual-Ferrá P, Alperstein N, Barnett DJ. A multi-platform approach to monitoring negative dominance for COVID-19 vaccine-related information online. Disaster Med Public Health Prep 2021 May 03:1-24. [doi: 10.1017/dmp.2021.136] [Medline: 33938423]
66. Marcus GE, Neuman WR, Mackuen M. Affective Intelligence and Political Judgment. Chicago, IL: The University of Chicago Press; 2000.
67. Harper CA, Satchell LP, Fido D, Latzman RD. Functional fear predicts public health compliance in the COVID-19 pandemic. Int J Ment Health Addict 2021 Apr 27;19(5):1875-1888 [FREE Full text] [doi: 10.1007/s11469-020-00281-5] [Medline: 32346359]
68. Stevens HR, Oh YJ, Taylor LD. Desensitization to fear-inducing COVID-19 health news on Twitter: observational study. JMIR Infodemiology 2021 Jul 16;1(1):e26876 [FREE Full text] [doi: 10.2196/26876] [Medline: 34447923]
69. 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]
70. Smith LE, Duffy B, Moxham-Hall V, Strang L, Wessely S, Rubin GJ. Anger and confrontation during the COVID-19 pandemic: a national cross-sectional survey in the UK. J R Soc Med 2021 Feb;114(2):77-90 [FREE Full text] [doi: 10.1177/0141076820962068] [Medline: 33115327]
71. Shanahan L, Steinhoff A, Bechtiger L, Murray AL, Nivette A, Hepp U, et al. Emotional distress in young adults during the COVID-19 pandemic: evidence of risk and resilience from a longitudinal cohort study. Psychol Med 2022 Apr;52(5):824-833 [FREE Full text] [doi: 10.1017/S003329172000241X] [Medline: 32571438]
72. Hurley D. An upshot of COVID-19—neurologists face instances of increased hostility and strategies to deal with it. Neurology Today 2022 Feb 03;22(3):8-31. [doi: 10.1097/01.nt.0000821684.69901.4f]
73. Wallace CL, Wladkowski SP, Gibson A, White P. Grief during the COVID-19 pandemic: considerations for palliative care providers. J Pain Symptom Manage 2020 Jul;60(1):e70-e76 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.04.012] [Medline: 32298748]
74. Jacobson NC, Lekkas D, Price G, Heinz MV, Song M, O'Malley AJ, et al. Flattening the mental health curve: COVID-19 stay-at-home orders are associated with alterations in mental health search behavior in the United States. JMIR Ment Health 2020 Jun 01;7(6):e19347 [FREE Full text] [doi: 10.2196/19347] [Medline: 32459186]
75. Hitlin P, Olmstead K. The science people see on social media. Pew Research Center. 2018 Mar 21. URL: https://www.pewresearch.org/science/2018/03/21/the-science-people-see-on-social-media/ [accessed 2022-09-06]

XSL•FO
RenderX

76. Padilla JJ, Kavak H, Lynch CJ, Gore RJ, Diallo SY. Temporal and spatiotemporal investigation of tourist attraction visit sentiment on Twitter. PLoS One 2018 Jun 14;13(6):e0198857 [FREE Full text] [doi: 10.1371/journal.pone.0198857] [Medline: 29902270]
77. Frank MR, Mitchell L, Dodds PS, Danforth CM. Happiness and the patterns of life: a study of geolocated tweets. Sci Rep 2013 Sep 12;3:2625 [FREE Full text] [doi: 10.1038/srep02625] [Medline: 24026340]
78. Mackey T, Baur C, Eysenbach G. Advancing infodemiology in a digital intensive era. JMIR Infodemiology 2022 Feb 14;2(1):e37115. [doi: 10.2196/37115]
79. Bhat A, Browning-McNee LA, Ghauri K, Winckler S. COVID-19 vaccine confidence project. J Am Pharm Assoc (2003) 2022 Jan;62(1):288-295.e2 [FREE Full text] [doi: 10.1016/j.japh.2021.06.006] [Medline: 34246577]
80. Chou WY, Burgdorf CE, Gaysynsky A, Hunter CM. COVID-19 vaccination* communication: applying behavioral and social science to address vaccine hesitancy and foster vaccine confidence. National Institutes of Health. 2020 Dec. URL: https://www.obssr.od.nih.gov/sites/obssr/files/inline-files/OBSSR_VaccineWhitePaper_FINAL_508.pdf [accessed 2021-08-21]
81. How to tailor COVID-19 vaccine information to your specific audience. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/vaccines/covid-19/hcp/tailoring-information.html [accessed 2022-08-21]

Abbreviations

API: application programming interface
H: hypothesis
LIWC: Linguistic Inquiry and Word Count
OR: odds ratio

Edited by A Mavragani; submitted 28.02.22; peer-reviewed by R Gore, Y Li; comments to author 21.07.22; revised version received 21.08.22; accepted 30.08.22; published 13.09.22.

Please cite as:
Stevens H, Rasul ME, Oh YJ
Emotions and Incivility in Vaccine Mandate Discourse: Natural Language Processing Insights
JMIR Infodemiology 2022;2(2):e37635
URL: https://infodemiology.jmir.org/2022/2/e37635
doi:10.2196/37635
PMID:36188420

©Hannah Stevens, Muhammad Ehab Rasul, Yoo Jung Oh. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 13.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Information Sharing Behaviors of Dietitians and Twitter Users in the Nutrition and COVID-19 Infodemic: Content Analysis Study of Tweets

Esther Charbonneau[1,2], MSc, RD; Sehl Mellouli[1,3], PhD; Arbi Chouikh[3], MSc; Laurie-Jane Couture[2]; Sophie Desroches[1,2], PhD, RD

[1]Centre Nutrition, Santé et Société, Institute of Nutrition and Functional Foods, Université Laval, Quebec City, QC, Canada

[2]School of Nutrition, Université Laval, Quebec City, QC, Canada

[3]Faculty of Business Administration, Université Laval, Quebec City, QC, Canada

Corresponding Author:
Sophie Desroches, PhD, RD
Centre Nutrition, Santé et Société
Institute of Nutrition and Functional Foods
Université Laval
Pavillon des services
2440 Hochelaga Blvd
Quebec City, QC, G1V 0A6
Canada
Phone: 1 418 656 2131 ext 405564
Email: sophie.desroches@fsaa.ulaval.ca

Abstract

Background: The COVID-19 pandemic has generated an infodemic, an overabundance of online and offline information. In this context, accurate information as well as misinformation and disinformation about the links between nutrition and COVID-19 have circulated on Twitter since the onset of the pandemic.

Objective: The purpose of this study was to compare tweets on nutrition in times of COVID-19 published by 2 groups, namely, a preidentified group of dietitians and a group of general users of Twitter, in terms of themes, content accuracy, use of behavior change factors, and user engagement, in order to contrast their information sharing behaviors during the pandemic.

Methods: Public English-language tweets published between December 31, 2019, and December 31, 2020, by 625 dietitians from Canada and the United States, and Twitter users were collected using hashtags and keywords related to nutrition and COVID-19. After filtration, tweets were coded against an original codebook of themes and the Theoretical Domains Framework (TDF) for identifying behavior change factors, and were compared to reliable nutritional recommendations pertaining to COVID-19. The numbers of likes, replies, and retweets per tweet were also collected to determine user engagement.

Results: In total, 2886 tweets (dietitians, n=1417; public, n=1469) were included in the analyses. Differences in frequency between groups were found in 11 out of 15 themes. Grocery (271/1417, 19.1%), and diets and dietary patterns (n=507, 34.5%) were the most frequently addressed themes by dietitians and the public, respectively. For 9 out of 14 TDF domains, there were differences in the frequency of usage between groups. "Skills" was the most used domain by both groups, although they used it in different proportions (dietitians: 612/1417, 43.2% vs public: 529/1469, 36.0%; $P<.001$). A higher proportion of dietitians' tweets were accurate compared with the public's tweets (532/575, 92.5% vs 250/382, 65.5%; $P<.001$). The results for user engagement were mixed. While engagement by likes varied between groups according to the theme, engagement by replies and retweets was similar across themes but varied according to the group.

Conclusions: Differences in tweets between groups, notably ones related to content accuracy, themes, and engagement in the form of likes, shed light on potentially useful and relevant elements to include in timely social media interventions aiming at fighting the COVID-19–related infodemic or future infodemics.

(JMIR Infodemiology 2022;2(2):e38573) doi:10.2196/38573

XSL•FO
RenderX

KEYWORDS

nutrition; COVID-19; dietitians; Twitter; public; themes; behavior; content accuracy; user engagement; content analysis; misinformation; disinformation; infodemic

Introduction

Background

On January 7, 2020, Chinese health authorities officially announced the emergence of the disease caused by the 2019 novel coronavirus [1] or SARS-CoV-2, a new strain of coronavirus [2]. COVID-19 was then declared a pandemic by the World Health Organization on March 11, 2020 [3], and as of March 24, 2022, the infection resulted in 470,223,960 confirmed cases and 6,094,326 deaths worldwide [4]. COVID-19 is characterized by symptoms ranging from cough and fever [5] to severe pneumonia and central nervous system damage [6]. Besides potential long-term health consequences with long COVID-19 [7], the infection led to serious social and economic repercussions [8]. To date, COVID-19 vaccines are the only man-made product (as opposed to infection-induced immunity) able to build one's immunity against SARS-CoV-2 [9].

From its onset, the pandemic has triggered multiple studies as clinical data were rapidly needed to face and fight the infection [10]. One area of study that retained researcher attention was related to the link between COVID-19 and nutrition. Indeed, concerns have been raised about certain nutrition-related health conditions, namely, diabetes, obesity, and cardiovascular diseases, as these could potentially elevate one's risk of experiencing severe COVID-19 [11]. Moreover, the roles played by nutrients, foods, and other types of supplements in immunity and inflammation have been studied extensively. For instance, Iddir et al [12] studied the role of certain nutrients and phytochemicals in reducing oxidative stress and inflammation, and underlined the importance of an optimal nutritional status in immunity. Furthermore, the pandemic has given rise to food- and nutrition-related changes in individuals, including those pertaining to food security [13], weight [14], and food habits [15]. These new data have led health organizations to develop recommendations and guidelines with regard to the appropriate food habits and nutritional care to follow during the COVID-19 pandemic [16,17].

In parallel, social media are equally being used as sources of health information and as platforms to disseminate health-related recommendations [18]. More specifically, Twitter, a microblogging site that permits real-time communication of 280-character tweets with followers [19], is considered a useful public health tool to share health-related information and engage with the public. As a matter of fact, it has been used by health professionals to provide information, educate people, share updates, disseminate new research, and raise public awareness of health matters like nutrition, infectious diseases, and sanitary emergencies [20,21]. However, concerns have been raised regarding the reliability and accuracy of the information found on social media such as Twitter [18].

Indeed, recently, the COVID-19 pandemic has played a major role in demonstrating how social media can be helpful as well as detrimental. The pandemic has led to what the World Health Organization calls an "infodemic," an overabundance of information online and offline, which may be true or false. Although an infodemic is not solely characterized by false information, it certainly contributes to its propagation. This situation can result in different repercussions, including damage to physical and mental health, increased stigma and conflict, and a lack of compliance with public health measures [22]. Moreover, at the beginning of the pandemic, Twitter was criticized, as most of the false information circulating on the platform was not verified [23]. However, efforts have been made by the microblogging service to counter misleading information [24]. Before going further, the types of false information should be distinguished. False information includes both misinformation and disinformation. Although the former is unintentional, the latter is done deliberately, in order to cause harm [25]. Both terms will be used jointly in this paper, as it can be hypothesized that both take place, but it is not part of the objectives of this study to determine the intent behind false information sharing.

Nutrition has received interest from researchers, official health organizations, and the general population since the beginning of the pandemic. In parallel, social media posts to this effect have also risen, and it is possible that misinformation and disinformation have also reached some of these communication platforms. Knowing this, some sources of information, including lay people, can be unreliable and could contribute to the proliferation and dissemination of misinformation and disinformation on nutrition-related topics. Conversely, dietitians are recognized as nutrition experts and should be prioritized when seeking information on food and nutrition [16]. A comparison between dietitians and general Twitter users relative to nutrition-related tweets has the potential to support the need for exercising caution when using Twitter, given the infodemic and the presence of unverified information on the microblogging site at the start of the pandemic [22,23], as well as for emphasizing the important role of dietitians on social media. Additionally, to our knowledge, only a few studies have documented misinformation or disinformation related to nutrition and COVID-19 altogether. These studies were however focused on specific aspects of nutrition such as immunity boosting claims [9,26].

Influence of Social Media on Behavior

Researchers have started exploring how social media publications regarding COVID-19 could influence intention, behavior, and protection against the virus [27-29]. More specifically, Al-Dmour et al showed that the use of social media platforms, including Facebook, Instagram, and Twitter, results in public protection from the infection, through the mediating effects of public health awareness and public health behavioral changes [29]. These results support the potential influence of social media publications over users. Nonetheless, food- and nutrition-related behaviors have not been investigated in that sense. Moreover, given that misinformation and disinformation

XSL•FO
RenderX

can be found in tweets, it is important to explore the factors used, intentionally or not, in publications, as they could potentially influence behavior. To this end, behavior change theoretical models can be useful to highlight such factors, as well as to understand behavior change. Such models have also been used to build social media interventions aiming to modify health-related behaviors like vaccination [30]. One of these models, the Theoretical Domains Framework (TDF), was initially developed to resolve the issue of having an overabundance of theoretical models and constructs aiming to explain behavior change [31], and is most often being used in implementation research in an array of settings, including health care, namely to identify the facilitators and barriers to implementing evidence-based behaviors or to design interventions [32]. Recently, the TDF has also been applied to content analysis of social media publications to determine the factors explaining COVID-19 vaccine hesitancy [33], thus providing a strategy to explore the behavior change potential of social media.

Objectives and Research Questions

The aim of this study was to compare the information sharing behaviors of registered dietitians (RDs) and Twitter users during the infodemic by analyzing their tweets related to nutrition in times of COVID-19. To do so, we compared the tweets of the 2 groups in terms of their themes, the user engagement they generated, content accuracy, and whether tweets included behavior change factors. To this end, we elaborated some research questions to be answered. Research questions are normally inquisitive in nature and better suited for exploratory studies where too little data are available to develop hypotheses [34], as in this study. The research questions were as follows:

1. What are the differences between dietitians' tweets and the public's tweets in terms of the themes they discuss?

2. What are the differences between dietitians' tweets and the public's tweets in terms of the engagement they receive from users?

3. What is the difference in content accuracy between dietitians' tweets and the public's tweets?

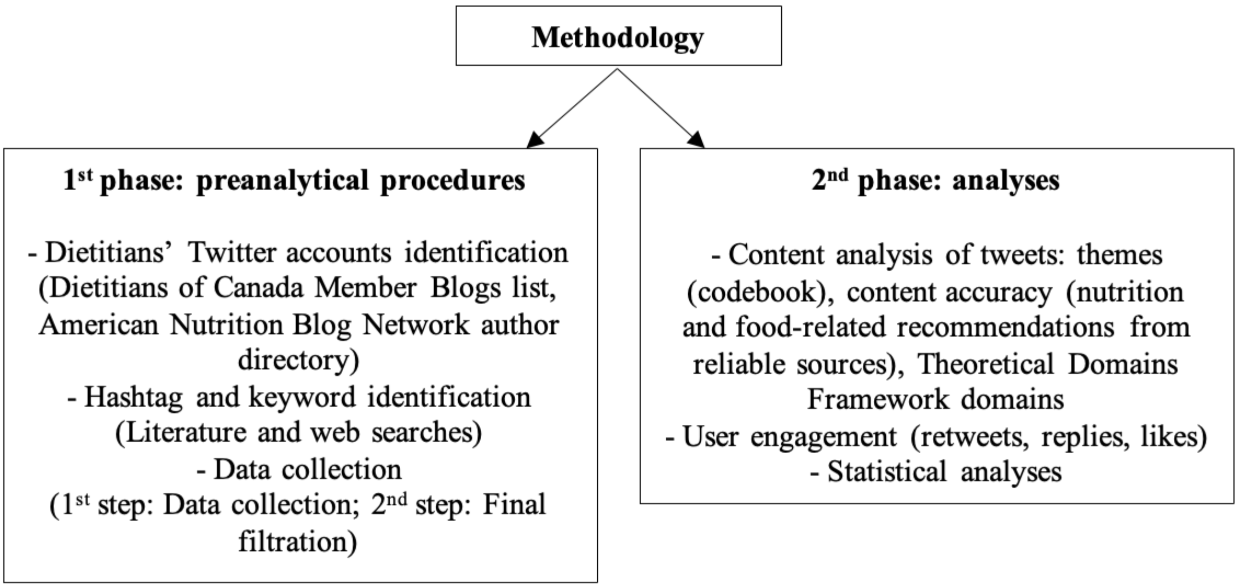
4. What are the differences between dietitians' tweets and the public's tweets in terms of the TDF domains they use, and could their tweets influence behavior?

Methods

Overview

This study's methods can be divided in 2 phases, namely, preanalytical procedures and analyses, as represented in Figure 1.

Figure 1. Study's synthesized methodology.

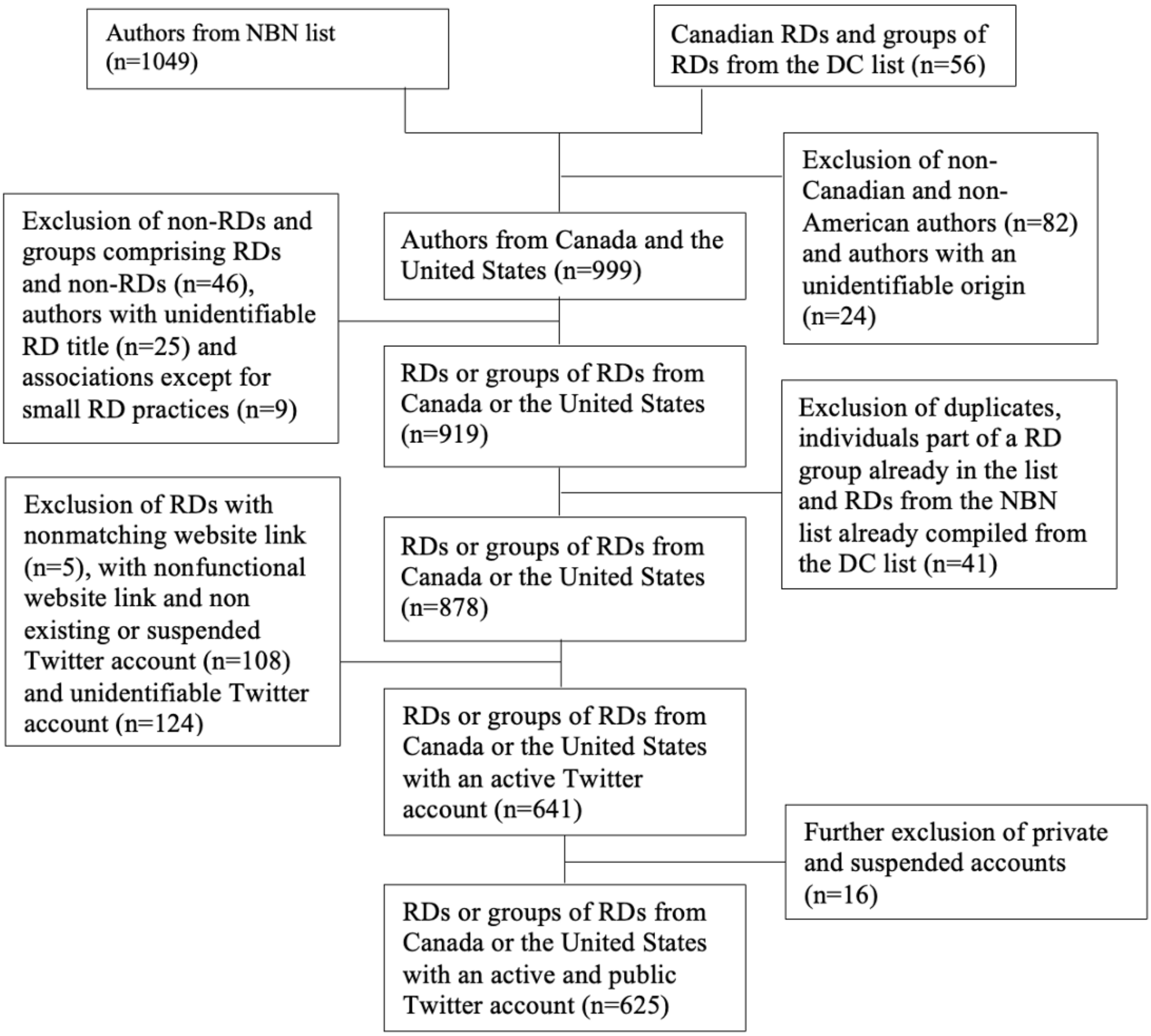


Dietitians' Twitter Account Identification

In order to identify our sample of RDs from Canada and the United States with Twitter accounts, the Dietitians of Canada Member Blogs list [35] (n=56 as of October 2020) and the American Nutrition Blog Network author directory [36] (n=1049 as of October 2020) were used. Both directories were reviewed to create a list of RDs (n=641), which included their name, website title, and Twitter handle. From this list, 16 RDs were excluded owing to suspended or private accounts. The final list thus comprised a total of 625 Twitter accounts. The steps are detailed in Figure 2.

XSL•FO
RenderX

Figure 2. Flow chart detailing the steps for creating the registered dietitians (RDs) list using the Nutrition Blog Network (NBN) author directory and the Dietitians of Canada (DC) Member Blogs list.



Hashtag and Keyword Identification

A predetermined list of 2561 hashtags and keywords related to COVID-19, 41 hashtags related to nutrition, and 16 hashtags related to both was used to filter tweets from the public and RDs (eg, "coronavirus," "#immunity," "#coviddiet," "#health," and "#nutrition"). The method for identifying hashtags and keywords was inspired by previous studies [37-39]. First, the list was built based upon searches on Facebook, Instagram, and Twitter of hashtags and keywords relevant to COVID-19, nutrition, or both. Second, it was enriched through literature [39-45] and web searches [46-49]. Two websites, Tagdef [46] and besthashtags [47], are generally used to find currently trending hashtags. The terms "COVID and nutrition," "COVID-19 and nutrition," "coronavirus and nutrition," and "corona and nutrition" were used to obtain hashtags related to these topics. Moreover, the literature was searched with terms related to nutrition, COVID-19, and Twitter to find studies containing relevant hashtags. Finally, we verified each keyword and hashtag to ensure its relevance.

Data Collection

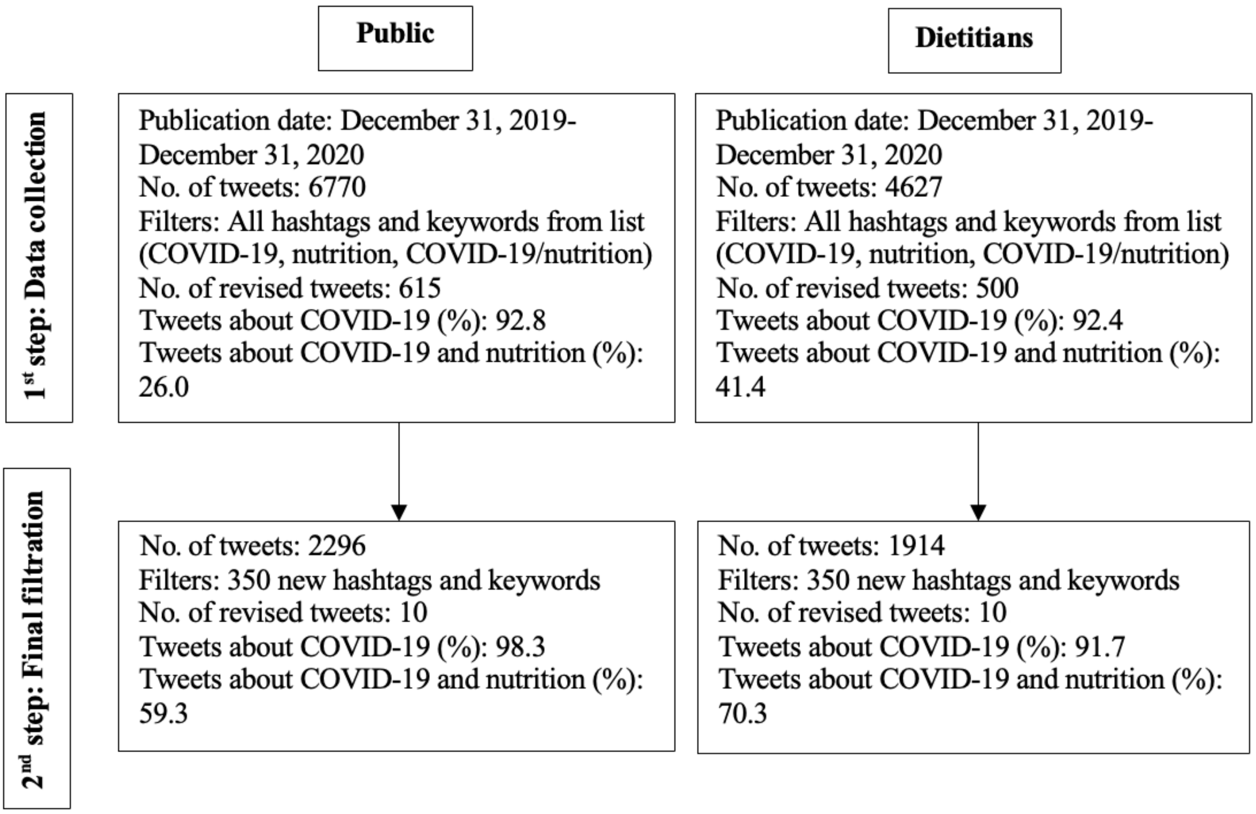
To be considered for the analysis, tweets had to be written in English, discuss at least one aspect of nutrition in times of COVID-19, and be published between December 31, 2019, and December 31, 2020. December 31, 2019, marks the date when cases of an unknown acute respiratory disease in Wuhan were first reported by Chinese health authorities [1]. Conversely, publications containing no written content or link to supplementary information were excluded. Tweets were collected in 2 steps using the Twitter Premium Application Programming Interface (API), which permits access to the Twitter archive. The publication date, author name, description, and country of origin (when available), as well as the numbers of likes, replies, and retweets were collected. Moreover, tweets from Twitter users were filtered to avoid having RDs from our list in that subsample.

Thus, the first step consisted of collecting the data using a predetermined list of hashtags and keywords, which resulted in 6670 tweets for the public group and 4627 tweets for the

XSL•FO
RenderX

dietitian group. After revising a subsample of each group, we observed that only 26.0% and 41.4% of the public's tweets and dietitians' tweets, respectively, were about both nutrition and COVID-19. The predetermined list of hashtags and keywords was thus enriched to render our data more specific to COVID-19 and nutrition. First, using tweets pertaining to COVID-19/nutrition from our 2 revised subsamples (see step 1 in Figure 3), 2 coders noted all the hashtags and keywords about nutrition and COVID-19/nutrition that were not already in our predetermined list (eg, #weightloss and #COVIDbaking). Second, these were compiled in a new list of 332 hashtags and keywords referring to nutrition and another 18 referring to COVID-19 and nutrition. Then, in the second step, the public and dietitian samples were submitted to a final filtration using this list of 350 hashtags and keywords. This process allowed the generation of 2 samples more specific to COVID-19 and nutrition. The steps are detailed in Figure 3.

Figure 3. Steps detailing tweet collection resulting in the final samples for analysis.



Considering the difficulties associated with this type of data collection [50] and that it was important not to reduce data representativeness, it was agreed not to further iterate the data sets. The final sample thus included a total of 4210 tweets (RDs: n=1914; public: n=2296). These 2296 tweets from the public were published by 1043 users. During coding, tweets that were still not related to COVID-19 and nutrition altogether were documented but not analyzed. Often, this happened when tweets contained hashtags or keywords about COVID-19 but not about nutrition specifically as in the following case: "[…] #needsofchildren #artathome #StatHomeStayCreative #coronavirustips." Hence, out of the 1914 tweets in the dietitian group, 1417 were included in the final analyses described below. As for the public group, 1469 out of 2296 tweets were analyzed. Thus, there were a total of 2886 tweets in both groups. When associated content was available through links in the tweet, it was also coded as part of the tweet.

Analyses

Research Question #1: Themes

The infodemic has generated multiple discussions on social media, which can reduce access to reliable information [51]. Defining the themes discussed about nutrition and COVID-19 on Twitter helps in determining which of them need to be more or less addressed by reliable sources of information on nutrition. Themes are patterns of information that represent categories to be analyzed [52]. This analysis was conducted to determine what subjects (RDs and the public) discuss with regard to the nutrition and COVID-19–related infodemic. Coders followed an iterative process based on the methodologies of similar studies to inductively create a codebook of themes [53,54]. First, 2 team members (EC and LJC) each elaborated a list of themes based on a review of the same 100 tweets published by Twitter users from the public. Second, common themes were put together to create an initial codebook. Third, since there were discrepancies between coders, each reviewed the same 50 tweets published by RDs, which led to the improvement of the initial codebook. Fourth, themes for which there was still no agreement

XSL•FO
RenderX

were settled by SD. Then, a codebook comprising 16 themes was established. Fifth, after the first round of reliability coding, the theme "Stress and Anxiety" was eliminated and added to the theme changed from "Physical Activity" to "Other Lifestyle Habits." It was thought relevant to address COVID-19–related lifestyle habits, but it was decided to regroup them into a single theme as they were not specific to nutrition. A final codebook including 15 nonmutually exclusive themes was then established and used by the same 2 investigators to categorize the 2886 tweets. Saturation, which was determined by identifying the point where all themes had been addressed at least once, was reached after 105 tweets for the dietitian group and 71 tweets for the public group. The theme frequency was compared between groups. Based on the tweet publication date, the frequency was also determined in the first 2 waves of the pandemic (first: December 31, 2019, to July 31, 2020; second: August 1, 2020, to December 31, 2020) and compared between waves. To render a more precise description and comparison of themes, statistics were used to compare theme frequencies between groups.

Research Question #2: User Engagement

Members of the public are not necessarily reliable sources of information on nutrition, while dietitians are considered reliable sources. This can become problematic when members of the public generate more engagement in their posts than their expert counterparts. In order to find out whether certain themes were more popular than others from a reader's perspective, the user engagement generated by themes was evaluated based on the numbers of likes, replies, and retweets associated with tweets. More specifically, for both subsamples separately, the mean numbers of likes, replies, and retweets for a single tweet were calculated for each theme. The means were then compared between groups to determine if certain themes were more popular in one group than the other. Additionally, the proportion of dietitians' tweets related to COVID-19 and nutrition out of their total yearly publications was calculated to evaluate their own engagement in this conversation on Twitter.

Research Question #3: Content Accuracy

To determine tweets' content accuracy and thus reveal the presence of misinformation, 2 team members (EC and LJC) compared the 2886 tweets against evidence-based nutrition and food-related recommendations regarding COVID-19. First, a database of recommendations from reliable and expert sources that covered COVID-19 and nutrition-related themes, such as Dietitians of Canada, Health Canada, and the Academy of Nutrition and Dietetics, was elaborated through web searches. However, when a tweet's content was too specific to be compared to the aforementioned recommendations, it became necessary to use more specialized sources of information (eg, PubMed and Mayo Clinic). For instance, the following tweet's content could not be found in our database of recommendations: "If your body happens to change during the pandemic, it could be because of stress […]." Second, during coding, coders read the tweet and verified its information using one or many reliable recommendations pertaining to the specific content of that tweet. If its content was in line with the recommendation, it was deemed accurate. If the content differed from the recommendation in any way, it was deemed inaccurate. Thus, tweets were categorized as accurate, inaccurate, or not applicable. The "not applicable" category was used when it was impossible to determine the tweet's accuracy for one or more of the following reasons: (1) the tweet is sharing a recipe or meal idea, (2) it is formulated as a question, (3) it reports on study results, and (4) it is considered as a nonscientific declaration or an opinion. For this study, it was decided that although study results pertaining to COVID-19 and nutrition could be compared to other studies, which are part of a body of evidence still in development, they include emerging data and not suggestions or advice to be followed. Moreover, they are too preliminary and specific to their study's methodology and population to be compared against nutritional recommendations about COVID-19. Moreover, although opinions or nonscientific declarations can be based on unsupported claims, for this study, it was decided that they could not be evaluated for accuracy. Indeed, this category could include tweets related to, for instance, what the users ate that day, a new nutrition-related habit they developed during the pandemic, or words of encouragement for workers in the food industry. As the evaluation went on from April through July 2021 and was then based on the current and available recommendations at that time, it is possible that the categorization would be different at the time when this paper has been written or published. Nevertheless, we made sure to use the most up to date information by regularly verifying updates in recommendations and available documentation. Saturation, which was determined by identifying the point where the 3 possible categorizations had been coded at least once, was reached after 25 tweets for the dietitian group and 13 tweets for the public group. Finally, the frequencies of accurate and inaccurate tweets were compared between groups. The frequencies of the nonapplicable categorization and of the 4 reasons why a tweet's accuracy could not be evaluated were also compared between groups. Moreover, further analyses were performed to compare the numbers of accurate and inaccurate mentions for each theme, so as to bring out those more frequently inaccurate than accurate.

Research Question #4: TDF Domains

Acting upon misinformation and disinformation can have detrimental effects. Therefore, to verify if tweets could potentially influence readers' behaviors, the 2886 tweets were deductively coded by 2 team members (EC and LJC) using the second version of the TDF [32]. The TDF does not serve as the theoretical lens for the whole study but solely to conduct an analysis aiming to determine whether tweets carry factors that could influence individual behavior. To our knowledge, the TDF has only been applied once before to tweets [33] and is thus a new application to be explored. Table 1 presents the 14 domains reflecting the cognitive, affective, social, and environmental factors influencing behaviors and their descriptions. This behavior change framework was chosen because it facilitates categorization during coding, as distinctive domains can be identified within tweets. Moreover, the TDF is highly comprehensive as it is based upon 33 theories and 128 theoretical constructs related to behavior change [55]. Thus, this model is useful to analyze a wide range of behaviors, which is the case in this study.

XSL•FO
RenderX

Table 1. Description of the Theoretical Domains Framework domains.

Domain	Description [32]
Knowledge	Awareness of something
Skills	Ability or competence developed through practice
Social and professional role and identity	Individual behaviors and qualities displayed in a social or work setting
Beliefs about capabilities	Recognition of one’s competences and abilities that can be put to constructive use
Optimism	Confidence that goals and desires will be reached
Beliefs about consequences	Expectancies about outcomes of a behavior in a situation
Reinforcement	Increasing the probability of a behavior with a stimulus
Intentions	Decision to accomplish a behavior or to act in a certain way
Goals	Mental representations of outcomes one wants to attain
Memory, attention, and decision processes	Ability to remember information, focus, and choose between different alternatives
Environmental context and resources	Situational or environmental aspect of one’s life that encourages or discourages the adoption of an adaptive behavior, skill, or competence
Social influences	Interpersonal processes that lead one to modify their thoughts, feelings, or behaviors
Emotion	Complex reaction by which one attempts to manage a personally significant matter or event
Behavioral regulation	Something done to manage or change one’s actions

The 14 domains were not mutually exclusive. Saturation, which was determined by identifying the point where all domains had been addressed at least once, was reached after 54 tweets for the dietitian group and 13 tweets for the public group. The frequency of each domain was compared between groups. Exploratory analyses were also conducted to reveal the most and least frequent domains for each theme.

Lastly, intercoder agreement, which measures the degree of similarity in codes assigned to a data set by different coders, was determined so as to preserve the consistency of results during individual coding [56]. Thus, the first round of reliability coding was performed where the 2 coders (EC and LJC) analyzed 100 tweets from each group according to the 3 content analyses described above, after which coders met to establish consensus. As scores for some themes and domains were too low, a second round of reliability coding was completed where both coders each analyzed 50 tweets from each group and met again to establish consensus. As scores obtained for themes and domains were satisfying, it was agreed that coding could be initiated. The kappa scores are presented in Table 2. Kappa scores ranging from 0.61 to 0.80 demonstrate substantial agreement between coders, while scores ranging from 0.81 to 1.00 are interpreted as almost perfect agreement [57]. For the rest of the sample, both team members coded 850 and 914 tweets from the dietitian group, respectively, which included 1 round of reliability coding of 100 tweets. They also coded 1000 and 1146 tweets from the general public group, respectively, including 2 rounds of reliability coding of 100 tweets each.

Table 2. Kappa scores obtained after 2 rounds of reliability coding.

Group	COVID-19/nutrition or not (1st round)	Content accuracy (1st round)	Themes (1st and 2nd rounds)	Domains (1st and 2nd rounds)
Public	0.78	0.67	0.54 and 0.65	0.42 and 0.63
Dietitian	0.95	0.78	0.51 and 0.79	0.66 and 0.75

Statistical Analyses

Statistical analyses were performed in SAS OnDemand for Academics (SAS Institute Inc). A *P* value ≤.05 (2 sided) was considered significant. This level of significance is often chosen in research [58]. The *P* value is the probability that measures the likeliness of a difference between groups being due to chance [59]. Chi-square tests were used to compare theme frequencies between groups and between the 2 waves of the pandemic. Chi-square tests were also used to compare the frequencies of the TDF domains, accurate/inaccurate categorization, and reasons for nonapplicability between groups. Comparisons of the frequencies of inaccurate and accurate mentions for each theme were also conducted using the chi-square test. The Fisher exact test was used instead of the chi-square test when at least one cell contained less than 5 data points [60]. Differences in means of likes, replies, and retweets per tweet between dietitians and the public were assessed by the nonparametric version of the *t* test for continuous data, that is, the Wilcoxon rank-sum test, as the data were not normally distributed and samples were independent [61,62].

Ethical Considerations

The Université Laval Research Ethics Board exempted this project from ethical review as analyses were completed with publicly available content. However, complete examples of tweets have not been presented in order to preserve the anonymity of the Twitter users.

XSL•FO
RenderX

Results

Research Question #1: Themes

The number of themes about nutrition and COVID-19 found in this study supports the fact that the infodemic has also reached this thematic. Table 3 shows the number of times each theme was addressed by both groups. In our sample, grocery, and diets and dietary patterns were the most frequently discussed themes by dietitians (271/1417, 19.1%) and the public (507/1469, 34.5%), respectively. Furthermore, many differences were found between the groups. For instance, weight loss was a more frequently discussed theme among the public than among dietitians (106/1469, 7.2% vs 24/1417, 1.7%; *P*<.001). Conversely, immune health was more frequently addressed by dietitians than by the public (177/1417, 12.5% vs 87/1469, 5.9%; *P*<.001).

Table 3. Comparison of theme frequencies between groups.

Theme	Description	Dietitian group (N=1417), n (%)	Public group (N=1469), n (%)	*P* value
Weight loss	Tips, mention, desire, and promotion. Not necessarily due to the pandemic.	24 (1.7)	106 (7.2)	<.001
Cooking and recipes	Sharing of recipes or meal/snack ideas. Mentions of what the next meal will be.	215 (15.2)	214 (14.6)	.65
Immune health	Linking nutrients, supplements, and foods, as well as physical activity, healthy eating, and hydration with immunity.	177 (12.5)	87 (5.9)	<.001
Food support and food system	Food support programs, food services/systems, buying local, gardening, and food insecurity.	206 (14.5)	59 (4.0)	<.001
Specific foods	Mention, consumption, or promotion of foods of various nutritional values.	178 (12.6)	487 (33.2)	<.001
Alcohol consumption	Reference to alcohol or mention of consumption.	19 (1.3)	86 (5.9)	<.001
Nutrients and supplements	Mention or promotion of a nutrient or supplement, regardless of immunity.	80 (5.7)	81 (5.5)	.88
Overeating	Mention of eating a large quantity of food in one sitting.	18 (1.3)	65 (4.4)	<.001
Food tips and recommendations	Hydration, suggestion of certain foods or practices, healthy restaurant food choices, and sanitary measures in restaurants.	253 (17.9)	108 (7.4)	<.001
Food changes	Modification of food choices, habits, and offers due to the pandemic, except for diets.	173 (12.2)	149 (10.1)	.08
Body appearance	References to physical appearance regardless of weight loss; includes weight gain.	86 (6.1)	67 (4.6)	.07
Diets and dietary patterns	Mention or promotion of diets, dietary patterns, and related practices.	26 (1.8)	507 (34.5)	<.001
Other lifestyle habits	References to physical activity (without mention of weight loss), stress/anxiety, sleep, tobacco, and cannabis.	259 (18.3)	453 (30.8)	<.001
Grocery	Food safety, in-store sanitary measures, healthy food choices at the store, ways to reduce grocery bills, and increased/decreased availability of products.	271 (19.1)	68 (4.6)	<.001
Health care system	Changes in dietetics practice, underlying health conditions, and nutrition of infected patients.	209 (14.8)	23 (1.6)	<.001

Comparison of themes between the first 2 waves of the pandemic revealed that none of the themes were more frequently addressed in the second wave than in the first by either of the groups. Indeed, 83.0% of dietitians' tweets were published during the first wave. Weight loss (*P*=.03), cooking and recipes (*P*<.001), specific foods (*P*=.03), food tips and recommendations (*P*=.003), grocery (*P*<.001), and health care system (*P*<.001) were more frequently addressed in the first wave than in the second. As for the public, they published 93.7% of their tweets during the first wave. Food tips and recommendations (*P*<.001), physical appearance (*P*=.02), and diets and dietary patterns (*P*=.03) were more frequent in the first wave than in the second wave. These results indicate that the first wave generally led to more discussions than the second wave.

Research Question #2: Social Media Engagement

Tables 4-6 show the comparisons of the mean numbers of retweets, replies, and likes per tweet for each theme between groups. The results revealed that dietitians constantly received a higher number of retweets per tweet than the public. Conversely, the public had more replies per tweet than dietitians. However, the public rarely had more than one reply per tweet, indicating that replies were seldom used by readers to manifest their engagement in both groups. Furthermore, while

XSL•FO
RenderX

engagement by replies and retweets depended on the group rather than the theme, engagement by likes varied between groups according to the theme. Indeed, weight loss, immune health, food support and food system, nutrients and supplements, and food tips and recommendations were more popular when addressed by dietitians, as other lifestyle habits generated more interest in the public's tweets. Moreover, it was observed that out of 73,323 English-language tweets published by dietitians during the 1-year period, only 1417 (1.9%) pertained to COVID-19 and nutrition. Lastly, there was no difference in the number of followers between groups. In the dietitian group, retweet and follower counts were not associated (r=0.04; P=.16), while there was an association between like and follower counts (r=0.12; P<.001).

Table 4. Comparison of the mean number of retweets per tweet between groups.

Theme	Dietitian group		Public group		*P* value
	Number of tweets	Number of retweets per tweet, mean (SD)	Number of tweets	Number of retweets per tweet, mean (SD)	
Weight loss	24	23.96 (77.49)	106	0.066 (0.42)	<.001
Cooking and recipes	215	149.91 (2181.49)	214	0.028 (0.17)	<.001
Immune health	177	11.99 (65.36)	87	0.092 (0.33)	<.001
Food support and food system	206	569.18 (5040.55)	59	0.017 (0.13)	<.001
Specific foods	178	182.53 (23.97)	487	0.037 (0.22)	<.001
Alcohol consumption	19	76.16 (327.37)	86	0.047 (0.21)	<.001
Nutrients and supplements	80	9.61 (43.43)	81	0.11 (0.39)	<.001
Food overconsumption	18	7.78 (23.37)	65	0.015 (0.12)	<.001
Food tips and recommendations	253	45.71 (677.18)	108	0.14 (0.50)	<.001
Food changes	173	1197.90 (15693.72)	149	0.013 (0.12)	<.001
Body appearance	86	242.36 (1424.06)	67	0.045 (0.21)	<.001
Diets and dietary patterns	26	5.31 (15.71)	507	0.018 (0.15)	<.001
Other lifestyle habits	259	22.37 (141.21)	453	0.15 (0.67)	<.001
Grocery	271	1176.28 (9564.04)	68	0.074 (0.31)	<.001
Health care system	209	65.30 (624.76)	23	0 (0)	<.001

Table 5. Comparison of the mean number of replies per tweet between groups.

Theme	Dietitian group		Public group		*P* value
	Number of tweets	Number of replies per tweet, mean (SD)	Number of tweets	Number of replies per tweet, mean (SD)	
Weight loss	24	0 (0)	106	0.75 (5.073)	.02
Cooking and recipes	215	0 (0)	214	0.32 (0.99)	<.001
Immune health	177	0 (0)	87	0.023 (0.15)	.04
Food support and food system	206	0 (0)	59	0.068 (0.31)	.001
Specific foods	178	0 (0)	487	0.44 (1.15)	<.001
Alcohol consumption	19	0 (0)	86	0.52 (1.49)	.01
Nutrients and supplements	80	0 (0)	81	0.21 (1.03)	.003
Food overconsumption	18	0 (0)	65	0.74 (1.57)	.01
Food tips and recommendations	253	0 (0)	108	0.13 (0.91)	.002
Food changes	173	0 (0)	149	0.66 (1.43)	<.001
Body appearance	86	0 (0)	67	1.06 (6.37)	<.001
Diets and dietary patterns	26	0 (0)	507	0.55 (2.57)	.005
Other lifestyle habits	259	0 (0)	453	0.080 (0.34)	<.001
Grocery	271	0 (0)	68	0.13 (0.39)	<.001
Health care system	209	0 (0)	23	0.044 (0.21)	.003

XSL•FO
RenderX

Table 6. Comparison of the mean number of likes per tweet between groups.

Theme	Dietitian group		Public group		*P* value
	Number of tweets	Number of likes per tweet, mean (SD)	Number of tweets	Number of likes per tweet, mean (SD)	
Weight loss	24	14.58 (32.60)	106	1.44 (3.73)	.03
Cooking and recipes	215	2.44 (5.88)	214	2.02 (5.20)	.18
Immune health	177	5.23 (30.95)	87	0.29 (0.59)	<.001
Food support and food system	206	2.67 (6.67)	59	0.41 (1.04)	.007
Specific foods	178	2.76 (7.66)	487	2.04 (4.72)	.65
Alcohol consumption	19	4.84 (13.87)	86	2.21 (5.37)	.66
Nutrients and supplements	80	2.00 (5.53)	81	1.02 (4.83)	.003
Food overconsumption	18	1.61 (2.48)	65	3.03 (6.22)	.51
Food tips and recommendations	253	1.66 (5.51)	108	0.69 (2.98)	<.001
Food changes	173	1.67 (2.90)	149	2.40 (5.30)	.46
Body appearance	86	5.70 (18.49)	67	1.78 (4.70)	.81
Diets and dietary patterns	26	13.85 (66.33)	507	2.11 (4.85)	.52
Other lifestyle habits	259	1.57 (4.23)	453	1.63 (11.74)	.04
Grocery	271	1.89 (6.06)	68	0.92 (1.66)	.34
Health care system	209	2.13 (8.89)	23	0.22 (0.42)	.06

Research Question #3: Content Accuracy

Content accuracy analyses revealed the presence of misinformation, but mostly in the public's tweets. In fact, a higher proportion of dietitians' tweets were accurate compared with the public's tweets (*P*<.001). For dietitians, out of a total of 575 tweets for which accuracy could be evaluated, 532 (92.5%) were accurate. As for the public, out of 382 tweets, 250 (65.5%) were accurate. Table 7 shows the comparison of the number of accurate and inaccurate tweets per theme. Weight loss was considered problematic as it had more inaccurate than accurate tweets. All other differences were in favor of accuracy.

Table 7. Content accuracy of individual themes.

Theme	Accurate (N=782), n (%)	Inaccurate (N=175), n (%)	*P* value
Weight loss	11 (1.4)	30 (17.1)	<.001
Cooking and recipes	45 (5.8)	15 (8.6)	.16
Immune health	128 (16.4)	77 (44.0)	<.001
Food support and food system	91 (11.6)	1 (0.6)	<.001
Specific foods	105 (13.4)	28 (16.0)	.37
Alcohol consumption	20 (2.6)	7 (4.0)	.30
Nutrients and supplements	78 (10.0)	26 (14.9)	.06
Food overconsumption	12 (1.5)	4 (2.3)	.51
Food tips and recommendations	224 (28.6)	17 (9.7)	<.001
Food changes	57 (7.3)	4 (2.3)	.02
Body appearance	14 (1.8)	3 (1.7)	<.99
Diets and dietary patterns	35 (4.5)	22 (12.6)	<.001
Other lifestyle habits	219 (28.0)	34 (19.4)	.02
Grocery	215 (27.5)	14 (8.0)	<.001
Health care system	74 (9.5)	5 (2.9)	.004

Furthermore, 842 (59.4%) of the dietitians' tweets and 1087 (74.0%) of the public's tweets were deemed not applicable for accuracy evaluation. More specifically, there were differences between groups for 3 reasons out of 4. First, a recipe or meal idea was shared more often in the public's tweets than in dietitians' tweets (332/1087, 30.5% vs 205/842, 24.4%; *P*=.003).

XSL•FO
RenderX

Second, no difference was found between groups when tweets were formulated as a question. Third, study results were more frequently reported in dietitians' tweets than in the public's tweets (118/842, 14.0% vs 8/1087, 0.7%; *P*<.001). Fourth, opinions or nonscientific declarations were more frequently shared in the public's tweets than in dietitians' tweets (806/1087, 74.2% vs 551/842, 65.4%; *P*<.001).

Research Question #4: TDF Domains

Table 8 shows the number of times the groups used each TDF domain in their tweets. In both cases, the TDF domain *skills* was the most used, although it appeared to be more frequently used by dietitians than by the public (612/1417, 43.2% vs 529/1469, 36.0%; *P*<.001). Other differences were also revealed between groups. Additionally, in both groups, it was found that the environmental context and resources, and more specifically, the pandemic, acted as important factors in the adoption of specific behaviors such as exercising at home or modifying a diet.

Table 9 depicts the most and least referenced domains per theme. This puts into light the TDF domains mostly associated with each theme or thematic category and could potentially be used to encourage behaviors related to the said themes or categories. Generally, themes related to weight management (weight loss, body appearance, diets, and dietary patterns) were associated with goals, and environmental context and resources. Food- and supplement-related themes (cooking and recipes, immune health, specific foods, alcohol consumption, nutrients, and supplements) were mostly associated with knowledge, skills, and environmental context and resources. Furthermore, these same 3 TDF domains (knowledge, skills, and environmental context and resources) were equally associated with themes about the food and health care systems (food support and food system, grocery, and health care system). Finally, lifestyle habit–related themes (food overconsumption, food tips and recommendations, food changes, and other lifestyle habits) were more commonly paired with skills, environmental context and resources, and behavioral regulation. Thus, for instance, goal setting could be considered when trying to lose weight or skills development could be implemented to encourage cooking.

Table 8. Comparison of the frequency of Theoretical Domains Framework domains between groups.

Domain	Dietitian group (N=1417), n (%)	Public group (N=1469), n (%)	*P* value
Knowledge	576 (40.7)	265 (18.0)	<.001
Skills	612 (43.2)	529 (36.0)	<.001
Social and professional role and identity	123 (8.7)	17 (1.2)	<.001
Beliefs about capabilities	100 (7.1)	114 (7.8)	.47
Optimism	121 (8.5)	106 (7.2)	.19
Beliefs about consequences	354 (25.0)	306 (20.6)	.008
Reinforcement	303 (21.4)	375 (25.5)	.009
Intentions	43 (3.0)	64 (4.4)	.06
Goals	61 (4.3)	290 (19.7)	<.001
Memory, attention, and decision processes	105 (7.4)	49 (3.3)	<.001
Environmental context and resources	471 (33.2)	482 (32.8)	.81
Social influences	50 (3.5)	41 (2.8)	.26
Emotion	130 (9.2)	61 (4.2)	<.001
Behavioral regulation	246 (17.4)	465 (31.7)	<.001

XSL•FO
RenderX

Table 9. The frequency of Theoretical Domains Framework domains for individual themes.

Theme	Most frequent domain	Frequency, n (%)	Least frequent domain	Frequency, n (%)
Weight loss (N=130)	Goals	59 (45.4)	Memory, attention and decision processes, and emotion	3 (2.3)
Cooking and recipes (N=429)	Skills	343 (80.0)	Social and professional role and identity	5 (1.2)
Immune health (N=264)	Knowledge	200 (75.8)	Intentions	2 (0.8)
Food support and food system (N=265)	Environmental context and resources	153 (57.7)	Social influences	5 (1.9)
Specific foods (N=665)	Environmental context and resources	273 (41.1)	Social and professional role and identity, and emotion	6 (0.9)
Alcohol consumption (N=105)	Environmental context and resources	60 (57.1)	Optimism and social influences	2 (1.9)
Nutrients and supplements (N=161)	Knowledge	105 (65.2)	Social and professional role and identity, and emotion	2 (1.2)
Food overconsumption (N=83)	Environmental context and resources	55 (66.3)	Social and professional role and identity	1 (1.2)
Food tips and recommendations (N=361)	Skills	258 (71.5)	Intentions	8 (2.2)
Food changes (N=322)	Environmental context and resources	232 (72.1)	Social and professional role and identity	9 (2.8)
Body appearance (N=153)	Environmental context and resources	69 (45.1)	Social and professional role and identity	3 (2.0)
Diets and dietary patterns (N=533)	Environmental context and resources	326 (61.2)	Social and professional role and identity, and memory, attention, and decision processes	4 (0.8)
Other lifestyle habits (N=712)	Behavioral regulation	389 (54.6)	Social and professional role and identity	15 (2.1)
Grocery (N=339)	Skills	205 (60.5)	Social influences	7 (2.1)
Health care system (N=232)	Knowledge	101 (43.5)	Social influences	4 (1.7)

Discussion

Principal Findings

This study found differences between dietitians' tweets and the public's tweets about the themes they discuss, the engagement they received from users, the TDF domains they used, and their content accuracy.

Differences about more frequently discussed themes were found between groups. Grocery was the most addressed theme by dietitians. Immune health, food support and food system, food tips and recommendations, grocery, and health care system were also more frequent in this group than in the public group. Conversely, the public group was mostly interested in discussing diets and dietary patterns, while weight loss, specific foods, alcohol consumption, food overconsumption, diets and dietary patterns, and other lifestyle habits emerged as more salient themes in this group than in the dietitian group.

Indeed, concerns have been raised by the population over grocery store safety practices, grocery bills, and an altered food supply [63], with the latter even leading to food shortages and elevated prices [64]. Furthermore, nutrition-induced immunity has been extensively addressed in the literature since the onset of the pandemic. However, online and social media posts on "immunity boosting" have contributed to the spread of misinformation and disinformation [9,26]. It is therefore possible that these were considered by dietitians as 2 areas of concern needing to be addressed by health professionals. Furthermore, results from Twitter users do not come as a surprise as these themes have been subjects of concern in the population during the pandemic. For instance, a survey conducted among adults from the Canadian province of Quebec revealed that weight-related concerns increased in 43% of participants [15]. Moreover, changes in dietary patterns and choices as well as alcohol consumption during the pandemic have been reported in different studies [13,65], just like modifications in weight and physical activity [66,67].

Moreover, as could be expected, thematic analyses between waves demonstrated that most of the discussions on nutrition and COVID-19 took place during the first wave, but more so in the case of dietitians. These results are supported by other studies. For instance, between January and October 2020, Google Search trends about COVID-19 and wine, ginger, 5G network spread, and the sun generally peaked in March and April 2020 [68]. Similarly, Chinese social media posts on COVID-19 misinformation peaked in February and March 2020 before slowly decreasing through May 2020 [69]. The disease novelty, concerns, sudden interest, anxiety, need for information,

XSL•FO
RenderX

and necessity to adapt to an out-of-ordinary situation possibly drove the conversation.

In addition, contrary to our expectations, no general thematic popularity was revealed across the 3 types of user engagement reactions, as only the number of likes differed between groups according to the theme. As opposed to the study by Hand et al [37], where individual RDs did not receive retweets of their heart failure–related tweets, the retweet count for dietitians was fairly elevated in this study. Dietitians constantly received more retweets and the public received more replies. Retweet behavior could partly be explained by the dietitians' authoritativeness, associated with their accurate knowledge of food and nutrition [70]. Moreover, although follower count could potentially influence dietitians' higher retweet values and high variability [70], no difference in follower count was found between groups. Moreover, in the dietitian group, there was no association between retweet and follower values and only a weak association between like and follower values. Similarly, Harris et al [38] showed that the number of followers was not associated with retweets or likes in their study. Discrepancies in replies in their study could not be justified by any of the predictors analyzed. While the results are mixed, they are still promising considering that dietitians received more retweets, that retweet dissemination was exponential, and that where differences were found, dietitians received more likes than the public in all cases but one (ie, other lifestyle habits). Studies on the factors of engagement in nutrition-related tweets and differences in the types of reactions are warranted to optimize interest in dietitians' tweets.

Contrary to other studies that have used the TDF to analyze specific aspects of nutrition or COVID-19, the model served a different purpose in this paper, as multiple nutrition and COVID-19–related behaviors were evaluated in tweets. Hence, all domains were addressed, suggesting that tweets could potentially contribute to behavior change. Additionally, differences were found between groups. However, in general, literature on the TDF mostly addressed the facilitators and barriers to the implementation of various behaviors by specific groups, which differs from how it was used in this study and renders the group comparison difficult. For instance, research on COVID-19 vaccine uptake has shown that themes related to the TDF domains of knowledge, beliefs about consequences, environmental context and resources, social influence, and emotion explain hesitancy [33], while facilitators have been found in beliefs about consequences [71]. Furthermore, another study found that 13 out of the 14 TDF domains explained nurses' physical activity and eating behavior [72]. These factors compare to those in this study, which further implies that tweets could partly influence behavior. Nonetheless, barriers and facilitators are group and behavior dependent and might not apply in this context. Thus, studies on the barriers and facilitators to the implementation of specific nutrition-related behaviors (eg, grocery shopping habits) in times of a pandemic are warranted to determine how tweets should be phrased to influence behaviors.

Furthermore, a high proportion of tweets were considered not applicable for accuracy evaluation, which could be explained by the fact that Twitter is a means "to share quickly where one is, and what one is doing, thinking, or feeling" [73]. Therefore, especially in the public's case, it still might not spontaneously be used to share verifiable facts and guidelines. This brings up the question as to whether Twitter represents the most useful or detrimental platform to seek health, nutrition, and pandemic-related information. However, for those tweets that were evaluated, as expected, a larger proportion of dietitians' tweets about nutrition and COVID-19 were accurate compared with the public's tweets. The higher quality and accuracy of dietitians' blog posts compared with those of nondietitians has been shown before [74], although studies making these comparisons on social media are lacking. This study is one of the first to cast light on the difference in social media post accuracy between dietitians and the public.

Practical Implications

Content accuracy results support the dietitians' role in sharing reliable information on nutrition during a pandemic. Health and governmental agencies should make use of their valuable expertise during health crises, namely by identifying and allying with dietitians who are present and active on social media. This collaboration could also result in more sustained engagement not only in the COVID-19 and nutrition discourse on Twitter but also in other nutrition-related situations and conditions on the part of dietitians.

Moreover, differences in themes addressed by groups, engagement in the form of likes, and theme inaccuracy shed light on the themes that should be prioritized, further discussed, and made more engaging by dietitians to counter the potentially inaccurate tweets of the public. For instance, other lifestyle habits were more interesting to readers when addressed by the public, while weight loss had more inaccurate than accurate tweets. Characterizing the conversation on nutrition and COVID-19 is equally necessary to bring other health professionals to help dietitians in their work toward reducing misinformation and disinformation on Twitter.

Likewise, knowing the behavior change factors employed by each group helps in orienting social media interventions aiming at the adoption of favorable pandemic-related practices. It does so by prioritizing behavior change techniques associated with the most popular determinants (eg, skills), by further integrating ones that tend to be less used or ones recognized as facilitators and barriers of similar behaviors, and by considering the fact that a pandemic acts as a socioenvironmental factor that largely influences behavior.

Lastly, comparison of the frequency of tweets between waves demonstrated that most of the conversation on COVID-19 and nutrition happened during the first few months of the pandemic. Thus, efforts should be made early to counter misinformation and disinformation. Without giving support to a piece of false information, it becomes important to correct it as soon as it starts to spread widely [75,76]. This underlines the importance of being prepared by building timely social media interventions that will not overload readers with information and the importance of encouraging platforms, such as Twitter, to be ready to put in place countermeasures early during a crisis.

XSL•FO
RenderX

Limitations

This study is not without limitations. First, although the methodology used to collect and validate tweets was rigorous, some of the keywords and hashtags were not specific to COVID-19 or nutrition, but were only related to it (eg, mask, disinfectant, and health). This resulted in a data collection that was possibly very sensitive but not specific enough. However, during coding, tweets were manually filtered to only keep those pertaining to the research theme. Hence, a lesser number of COVID-19 and nutrition-specific words should have been used to collect tweets. A keywords list should indeed be reviewed iteratively before initiating data collection [50]. Second, our use of the TDF differs from its prior use in research. Therefore, no similar methodology was available to inform our coding with the model, which could possibly be improved upon given the low initial kappa scores. For example, Griffith et al categorized tweets in a few themes before mapping these onto the TDF [33]. Third, the number of themes in the codebook and assigned to a given tweet should be limited to reduce the variability between coders. Fourth, the RD sample was potentially not representative of groups of dietitians outside of Canada and the United States. Similarly, although an efficient strategy was adopted to identify RDs, the use of the Dietitians of Canada Member Blogs list and the Nutrition Blog Network author directory potentially excluded a relatively high number of dietitians active on Twitter. Finally, it is possible that health professionals, including dietitians, were part of the public sample, which could have potentially influenced accuracy results. Nevertheless, we ensured that no dietitian from our sample was present in the public group.

Conclusion

This study sheds light on the information sharing behaviors of RDs from Canada and the United States, and Twitter users in the COVID-19 and nutrition infodemic on Twitter. Differences were found in discussed themes, use of TDF domains, content accuracy, and generated user engagement. Studies and results like these are needed to support the role of practical, timely, and theory-informed social media interventions led by dietitians, as well as other health professionals specialized in their respective fields, for encouraging sound and evidence-based pandemic-related practices and behaviors.

Acknowledgments

Virginie Drolet-Labelle, RD, who is a candidate of the Master of Nutrition at the School of Nutrition, Université Laval, helped with dietitian Twitter account identification. Alexandra Bédard, PhD, RD, who is a research professional at the Institute of Nutrition and Functional Foods, and Centre Nutrition, santé et société, Université Laval, assisted with statistical analyses. This project was supported financially by a grant and a scholarship from the Centre Nutrition, santé et société (NUTRISS), as well as a scholarship from the Fonds de nutrition publique de l'Université Laval.

Conflicts of Interest

None declared.

References

1. Patel A, Jernigan DB. Initial public health response and interim clinical guidance for the 2019 novel coronavirus outbreak — United States, December 31, 2019–February 4, 2020. Centers for Disease Control and Prevention. URL: https://stacks.cdc.gov/view/cdc/84807 [accessed 2021-09-05]
2. Questions and answers on COVID-19: Basic facts. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/covid-19/questions-answers/questions-answers-basic-facts [accessed 2021-09-05]
3. Timeline: WHO's COVID-19 response. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline#! [accessed 2021-09-05]
4. COVID-19 situation update worldwide. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases [accessed 2022-03-27]
5. COVID-19: Symptoms, treatment, what to do if you feel sick. Government of Canada. URL: https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/symptoms.html [accessed 2021-09-05]
6. Samudrala PK, Kumar P, Choudhary K, Thakur N, Wadekar GS, Dayaramani R, et al. Virology, pathogenesis, diagnosis and in-line treatment of COVID-19. Eur J Pharmacol 2020 Sep 15;883:173375 [FREE Full text] [doi: 10.1016/j.ejphar.2020.173375] [Medline: 32682788]
7. Long COVID or Post-COVID Conditions. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/long-term-effects/index.html [accessed 2021-09-05]
8. Impact of COVID-19 on people's livelihoods, their health and our food systems. World Health Organization. URL: https://www.who.int/news/item/13-10-2020-impact-of-covid-19-on-people%27s-livelihoods-their-health-and-our-food-systems [accessed 2021-09-05]
9. Rachul C, Marcon AR, Collins B, Caulfield T. COVID-19 and 'immune boosting' on the internet: a content analysis of Google search results. BMJ Open 2020 Oct 26;10(10):e040989 [FREE Full text] [doi: 10.1136/bmjopen-2020-040989] [Medline: 33109677]

XSL•FO
RenderX

10. Abbott R, Bethel A, Rogers M, Whear R, Orr N, Shaw L, et al. Characteristics, quality and volume of the first 5 months of the COVID-19 evidence synthesis infodemic: a meta-research study. BMJ Evid Based Med 2022 Jun;27(3):169-177 [FREE Full text] [doi: 10.1136/bmjebm-2021-111710] [Medline: 34083212]
11. People with Certain Medical Conditions. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html [accessed 2021-09-06]
12. Iddir M, Brito A, Dingeo G, Fernandez Del Campo SS, Samouda H, La Frano MR, et al. Strengthening the Immune System and Reducing Inflammation and Oxidative Stress through Diet and Nutrition: Considerations during the COVID-19 Crisis. Nutrients 2020 May 27;12(6):1562 [FREE Full text] [doi: 10.3390/nu12061562] [Medline: 32471251]
13. Lamarche B, Brassard D, Lapointe A, Laramée C, Kearney M, Côté M, et al. Changes in diet quality and food security among adults during the COVID-19-related early lockdown: results from NutriQuébec. Am J Clin Nutr 2021 Apr 06;113(4):984-992 [FREE Full text] [doi: 10.1093/ajcn/nqaa363] [Medline: 33398347]
14. Pellegrini M, Ponzo V, Rosato R, Scumaci E, Goitre I, Benso A, et al. Changes in Weight and Nutritional Habits in Adults with Obesity during the "Lockdown" Period Caused by the COVID-19 Virus Emergency. Nutrients 2020 Jul 07;12(7):2016 [FREE Full text] [doi: 10.3390/nu12072016] [Medline: 32645970]
15. Pandémie, habitudes de vie, qualité du sommeil et préoccupation à l'égard du poids - 23 février 2021. Institut national de santé publique du Québec. URL: https://www.inspq.qc.ca/covid-19/sondages-attitudes-comportements-quebecois/habitudes-de-vie-fevrier-2021 [accessed 2021-09-06]
16. Advice for the general public about COVID-19. Dietitians of Canada. URL: https://www.dietitians.ca/News/2020/Advice-for-the-general-public-about-COVID-19?lang=en-CA [accessed 2022-01-17]
17. Handu D, Moloney L, Rozga M, Cheng FW. Malnutrition Care During the COVID-19 Pandemic: Considerations for Registered Dietitian Nutritionists. J Acad Nutr Diet 2021 May;121(5):979-987 [FREE Full text] [doi: 10.1016/j.jand.2020.05.012] [Medline: 32411575]
18. 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]
19. Counting characters when composing Tweets. Twitter Developer Platform. URL: https://developer.twitter.com/en/docs/counting-characters [accessed 2021-09-06]
20. Breland JY, Quintiliani LM, Schneider KL, May CN, Pagoto S. Social Media as a Tool to Increase the Impact of Public Health Research. Am J Public Health 2017 Dec;107(12):1890-1891. [doi: 10.2105/AJPH.2017.304098] [Medline: 29116846]
21. Hart M, Stetten N, Islam S, Pizarro K. Twitter and Public Health (Part 2): Qualitative Analysis of How Individual Health Professionals Outside Organizations Use Microblogging to Promote and Disseminate Health-Related Information. JMIR Public Health Surveill 2017 Oct 04;3(4):e54 [FREE Full text] [doi: 10.2196/publichealth.6796] [Medline: 28978500]
22. Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation. World Health Organization. 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 2021-09-06]
23. Brennen J, Simon F, Howard P, Nielsen R. Types, Sources, and Claims of COVID-19 Misinformation. Reuters Institute. URL: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-04/Brennen%20-%20COVID%2019%20Misinformation%20FINAL%20%283%29.pdf [accessed 2021-09-17]
24. COVID-19 misleading information policy. Twitter Help Center. URL: https://help.twitter.com/en/rules-and-policies/medical-misinformation-policy [accessed 2021-09-17]
25. Wardle C, Derakhshan H. INFORMATION DISORDER: Toward an interdisciplinary framework for research and policy making. Council of Europe. URL: https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c [accessed 2021-09-17]
26. Wagner DN, Marcon AR, Caulfield T. "Immune Boosting" in the time of COVID: selling immunity on Instagram. Allergy Asthma Clin Immunol 2020;16:76 [FREE Full text] [doi: 10.1186/s13223-020-00474-6] [Medline: 32905318]
27. Farooq A, Laato S, Islam AKMN. Impact of Online Information on Self-Isolation Intention During the COVID-19 Pandemic: Cross-Sectional Study. J Med Internet Res 2020 May 06;22(5):e19128 [FREE Full text] [doi: 10.2196/19128] [Medline: 32330115]
28. 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]
29. Al-Dmour H, Masa'deh R, Salman A, Abuhashesh M, Al-Dmour R. Influence of Social Media Platforms on Public Health Protection Against the COVID-19 Pandemic via the Mediating Effects of Public Health Awareness and Behavioral Changes: Integrated Model. J Med Internet Res 2020 Aug 19;22(8):e19996 [FREE Full text] [doi: 10.2196/19996] [Medline: 32750004]
30. Li L, Wood CE, Kostkova P. Vaccine hesitancy and behavior change theory-based social media interventions: a systematic review. Transl Behav Med 2022 Feb 16;12(2):243-272 [FREE Full text] [doi: 10.1093/tbm/ibab148] [Medline: 34850217]

XSL•FO
RenderX

31. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Mar;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
32. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Jun 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
33. 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]
34. Creating Research Questions. Norwegian University of Life Sciences. URL: https://nmbu.instructure.com/courses/2280/pages/creating-research-questions?module_item_id=15795 [accessed 2022-02-01]
35. Member Blogs. Dietitians of Canada. URL: https://www.dietitians.ca/News/Member-Blogs?lang=en-CA [accessed 2020-11-03]
36. Nutrition Blog Network Facebook page. A site featuring 900+ blogs written by registered dietitians (RDs). Search for blogs by topic or na. Nutrition Blog Network. 2010 Apr 22. URL: https://www.facebook.com/nutritionblognetwork
37. Hand RK, Kenne D, Wolfram TM, Abram JK, Fleming M. Assessing the Viability of Social Media for Disseminating Evidence-Based Nutrition Practice Guideline Through Content Analysis of Twitter Messages and Health Professional Interviews: An Observational Study. J Med Internet Res 2016 Nov 15;18(11):e295 [FREE Full text] [doi: 10.2196/jmir.5811] [Medline: 27847349]
38. Harris JK, Duncan A, Men V, Shevick N, Krauss MJ, Cavazos-Rehg PA. Messengers and Messages for Tweets That Used #thinspo and #fitspo Hashtags in 2016. Prev Chronic Dis 2018 Jan 04;15:E01 [FREE Full text] [doi: 10.5888/pcd15.170309] [Medline: 29300696]
39. Turner-McGrievy GM, Beets MW. Tweet for health: using an online social network to examine temporal trends in weight loss-related posts. Transl Behav Med 2015 Jun 29;5(2):160-166 [FREE Full text] [doi: 10.1007/s13142-015-0308-1] [Medline: 26029278]
40. Lucini D, Gandolfi CE, Antonucci C, Cavagna A, Valzano E, Botta E, et al. #StayHomeStayFit: UNIMI's approach to online healthy lifestyle promotion during the COVID-19 pandemic. Acta Biomed 2020 Sep 07;91(3):e2020037 [FREE Full text] [doi: 10.23750/abm.v91i3.10375] [Medline: 32921731]
41. Lange KW. Food science and COVID-19. Food Science and Human Wellness 2021 Jan;10(1):1-5. [doi: 10.1016/j.fshw.2020.08.005]
42. 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. [doi: 10.1109/access.2020.3010033]
43. Rao HR, Vemprala N, Akello P, Valecha R. Retweets of officials' alarming vs reassuring messages during the COVID-19 pandemic: Implications for crisis management. Int J Inf Manage 2020 Dec;55:102187 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102187] [Medline: 32836644]
44. Melotte S, Kejriwal M. A Geo-Tagged COVID-19 Twitter Dataset for 10 North American Metropolitan Areas over a 255-Day Period. Data 2021 Jun 16;6(6):64. [doi: 10.3390/data6060064]
45. 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]
46. Tagdef. URL: https://tagdef.com/fr/ [accessed 2022-08-30]
47. best-hashtags. URL: http://best-hashtags.com/ [accessed 2022-08-30]
48. What's trending during coronavirus pandemic? A definitive guide to the most used hashtags. The National. URL: https://www.thenationalnews.com/arts-culture/what-s-trending-during-coronavirus-pandemic-a-definitive-guide-to-the-most-used-hashtags-1.996208 [accessed 2022-06-16]
49. Seven trending hashtags about COVID-19 on social media. Media Update. URL: https://www.mediaupdate.co.za/social/148423/seven-trending-hashtags-about-covid-19-on-social-media [accessed 2022-06-16]
50. Kim Y, Huang J, Emery S. Garbage in, Garbage Out: Data Collection, Quality Assessment and Reporting Standards for Social Media Data Use in Health Research, Infodemiology and Digital Disease Detection. J Med Internet Res 2016 Mar 26;18(2):e41 [FREE Full text] [doi: 10.2196/jmir.4738] [Medline: 26920122]
51. Gottlieb M, Dyer S. Information and Disinformation: Social Media in the COVID-19 Crisis. Acad Emerg Med 2020 Jul;27(7):640-641 [FREE Full text] [doi: 10.1111/acem.14036] [Medline: 32474977]
52. Fereday J, Muir-Cochrane E. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. International Journal of Qualitative Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
53. Alvarez-Mon MA, Llavero-Valero M, Sánchez-Bayona R, Pereira-Sanchez V, Vallejo-Valdivielso M, Monserrat J, et al. Areas of Interest and Stigmatic Attitudes of the General Public in Five Relevant Medical Conditions: Thematic and Quantitative Analysis Using Twitter. J Med Internet Res 2019 May 28;21(5):e14110 [FREE Full text] [doi: 10.2196/14110] [Medline: 31140438]

XSL•FO
RenderX

54. van Draanen J, Krishna T, Tsang C, Liu S. Keeping up with the times: how national public health and governmental organizations communicate about cannabis on Twitter. Subst Abuse Treat Prev Policy 2019 Sep 12;14(1):38 [FREE Full text] [doi: 10.1186/s13011-019-0224-3] [Medline: 31511026]
55. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
56. Chapter 5. Achieving Reliability. Colorado State University. URL: https://wac.colostate.edu/docs/books/codingstreams/chapter5.pdf [accessed 2022-01-25]
57. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
58. du Prel J, Hommel G, Röhrig B, Blettner M. Confidence interval or p-value?: part 4 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2009 May;106(19):335-339 [FREE Full text] [doi: 10.3238/arztebl.2009.0335] [Medline: 19547734]
59. Whitley E, Ball J. Statistics review 3: hypothesis testing and P values. Crit Care 2002 Jun;6(3):222-225 [FREE Full text] [doi: 10.1186/cc1493] [Medline: 12133182]
60. Bédard A, Corneau L, Dodin S, Lemieux S. Sex Differences in the Effects of Repeated Taste Exposure to the Mediterranean Diet: A 6-month Follow-up Study. Can J Diet Pract Res 2016 Sep;77(3):125-132. [doi: 10.3148/cjdpr-2015-052] [Medline: 26916988]
61. Mann Whitney U Test (Wilcoxon Rank Sum Test). Boston University School of Public Health. URL: https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_nonparametric/bs704_nonparametric4.html [accessed 2022-01-25]
62. McIntosh A, Sharpe M, Lawrie S. 9 - Research methods, statistics and evidence-based practice. In: Johnstone E, Lawrie S, Sharpe M, Owens D, McIntosh A, editors. Companion to Psychiatric Studies (Eighth Edition). London, United Kingdom: Churchill Livingstone; 2010:157-198.
63. COVID-19: May 2020. International Food Information Council. URL: https://foodinsight.org/wp-content/uploads/2020/04/IFIC-COVID-19-May-2020.pdf [accessed 2021-11-07]
64. Will COVID-19 Threaten Availability and Affordability of our Food? U.S. Department of Agriculture. URL: https://www.usda.gov/media/blog/2020/04/16/will-covid-19-threaten-availability-and-affordability-our-food [accessed 2021-11-07]
65. Di Renzo L, Gualtieri P, Pivari F, Soldati L, Attinà A, Cinelli G, et al. Eating habits and lifestyle changes during COVID-19 lockdown: an Italian survey. J Transl Med 2020 Jun 08;18(1):229 [FREE Full text] [doi: 10.1186/s12967-020-02399-5] [Medline: 32513197]
66. Reyes-Olavarría D, Latorre-Román P, Guzmán-Guzmán IP, Jerez-Mayorga D, Caamaño-Navarrete F, Delgado-Floody P. Positive and Negative Changes in Food Habits, Physical Activity Patterns, and Weight Status during COVID-19 Confinement: Associated Factors in the Chilean Population. Int J Environ Res Public Health 2020 Jul 28;17(15):5431 [FREE Full text] [doi: 10.3390/ijerph17155431] [Medline: 32731509]
67. Kriaucioniene V, Bagdonaviciene L, Rodríguez-Pérez C, Petkeviciene J. Associations between Changes in Health Behaviours and Body Weight during the COVID-19 Quarantine in Lithuania: The Lithuanian COVIDiet Study. Nutrients 2020 Oct 13;12(10):3119 [FREE Full text] [doi: 10.3390/nu12103119] [Medline: 33065991]
68. Nsoesie EO, Cesare N, Müller M, Ozonoff A. COVID-19 Misinformation Spread in Eight Countries: Exponential Growth Modeling Study. J Med Internet Res 2020 Dec 15;22(12):e24425 [FREE Full text] [doi: 10.2196/24425] [Medline: 33264102]
69. Zhang S, Pian W, Ma F, Ni Z, Liu Y. Characterizing the COVID-19 Infodemic on Chinese Social Media: Exploratory Study. JMIR Public Health Surveill 2021 Feb 05;7(2):e26090 [FREE Full text] [doi: 10.2196/26090] [Medline: 33460391]
70. Zhou J, Liu F, Zhou H. Understanding health food messages on Twitter for health literacy promotion. Perspect Public Health 2018 May 07;138(3):173-179. [doi: 10.1177/1757913918760359] [Medline: 29513075]
71. Williams L, Gallant AJ, Rasmussen S, Brown Nicholls LA, Cogan N, Deakin K, et al. Towards intervention development to increase the uptake of COVID-19 vaccination among those at high risk: Outlining evidence-based and theoretically informed future intervention content. Br J Health Psychol 2020 Nov;25(4):1039-1054. [doi: 10.1111/bjhp.12468] [Medline: 32889759]
72. Power BT, Kiezebrink K, Allan JL, Campbell MK. Understanding perceived determinants of nurses' eating and physical activity behaviour: a theory-informed qualitative interview study. BMC Obes 2017;4:18 [FREE Full text] [doi: 10.1186/s40608-017-0154-4] [Medline: 28491327]
73. Social Media for Academic Libraries. Western University. URL: https://ir.lib.uwo.ca/cgi/viewcontent.cgi?article=1030&context=fimspub [accessed 2021-11-08]
74. Toth J, O'connor C, Hartman B, Dworatzek P, Horne J. "Detoxify or Die": Qualitative Assessments of Ontario Nutritionists' and Dietitians' Blog Posts Related to Detoxification Diets. Can J Diet Pract Res 2019 Sep 01;80(3):116-121. [doi: 10.3148/cjdpr-2018-047] [Medline: 30724103]
75. 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]
76. What Role Should Newsrooms Play in Debunking COVID-19 Misinformation? Nieman Reports. URL: https://niemanreports.org/articles/what-role-should-newsrooms-play-in-debunking-covid-19-misinformation/ [accessed 2021-10-04]

XSL•FO
RenderX

Abbreviations

RD: registered dietitian
TDF: Theoretical Domains Framework

Edited by T Mackey; submitted 07.04.22; peer-reviewed by D MacKay, A Farooq; comments to author 09.06.22; revised version received 09.07.22; accepted 27.07.22; published 16.09.22.

Please cite as:
Charbonneau E, Mellouli S, Chouikh A, Couture LJ, Desroches S
The Information Sharing Behaviors of Dietitians and Twitter Users in the Nutrition and COVID-19 Infodemic: Content Analysis Study of Tweets
JMIR Infodemiology 2022;2(2):e38573
URL: https://infodemiology.jmir.org/2022/2/e38573
doi:10.2196/38573
PMID:36188421

©Esther Charbonneau, Sehl Mellouli, Arbi Chouikh, Laurie-Jane Couture, Sophie Desroches. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 16.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Data Exploration and Classification of News Article Reliability: Deep Learning Study

Kevin Zhan[1*]; Yutong Li[1*], BSc; Rafay Osmani[2]; Xiaoyu Wang[3]; Bo Cao[1], PhD

[1]Department of Psychiatry, University of Alberta, Edmonton, AB, Canada

[2]Department of Cell Biology, University of Alberta, Edmonton, AB, Canada

[3]Department of Computing Science, University of Alberta, Edmonton, AB, Canada

[*]these authors contributed equally

Corresponding Author:
Yutong Li, BSc
Department of Psychiatry
University of Alberta
4-142 KATZ Group Centre for Pharmacy and Health Research
87 Avenue and 114 Street
Edmonton, AB, T6G 2E1
Canada
Phone: 1 403 926 6628
Email: yutong5@ualberta.ca

Abstract

Background: During the ongoing COVID-19 pandemic, we are being exposed to large amounts of information each day. This "infodemic" is defined by the World Health Organization as the mass spread of misleading or false information during a pandemic. This spread of misinformation during the infodemic ultimately leads to misunderstandings of public health orders or direct opposition against public policies. Although there have been efforts to combat misinformation spread, current manual fact-checking methods are insufficient to combat the infodemic.

Objective: We propose the use of natural language processing (NLP) and machine learning (ML) techniques to build a model that can be used to identify unreliable news articles online.

Methods: First, we preprocessed the ReCOVery data set to obtain 2029 English news articles tagged with COVID-19 keywords from January to May 2020, which are labeled as reliable or unreliable. Data exploration was conducted to determine major differences between reliable and unreliable articles. We built an ensemble deep learning model using the body text, as well as features, such as sentiment, Empath-derived lexical categories, and readability, to classify the reliability.

Results: We found that reliable news articles have a higher proportion of neutral sentiment, while unreliable articles have a higher proportion of negative sentiment. Additionally, our analysis demonstrated that reliable articles are easier to read than unreliable articles, in addition to having different lexical categories and keywords. Our new model was evaluated to achieve the following performance metrics: 0.906 area under the curve (AUC), 0.835 specificity, and 0.945 sensitivity. These values are above the baseline performance of the original ReCOVery model.

Conclusions: This paper identified novel differences between reliable and unreliable news articles; moreover, the model was trained using state-of-the-art deep learning techniques. We aim to be able to use our findings to help researchers and the public audience more easily identify false information and unreliable media in their everyday lives.

(JMIR Infodemiology 2022;2(2):e38839) doi:10.2196/38839

KEYWORDS

COVID-19; deep learning; news article reliability; false information; infodemic; ensemble model

Introduction

The onset of the COVID-19 pandemic has given the world more to battle. The world has faced a barrage of false information during the "infodemic," which is defined as the spread of a large amount of information that includes misleading or false information during a pandemic [1,2]. Due to quarantine and increased restrictions, information is trafficked to the public via

XSL•FO
RenderX

social media and news sources; consequently, false information propagates at a larger scale and faster rate. Despite available public health guidelines, there is still a large presence of false and misleading information online, comprising around 20% of articles on major social media sites, such as Twitter [3]. Although the proportion of shared false information is less than evidence-informed guidelines, false information spreads at a faster rate because it contains inflammatory information [4,5]. Furthermore, infodemic management is an important aspect in maintaining public trust in scientific guidance [1]. Hence, we need to construct methods to deter the spread of false information online and identify potential sources of false news.

The abundance of fake or false news online can be instances of misinformation or disinformation and often lacks the reliability and credibility in content [6-8]. Disinformation is defined as the intentional spread of false information, while misinformation is the negligent sharing of false information [6]. Hereafter, we will not differentiate between disinformation and misinformation, as we will refer to them together as false information. False news can be categorized into 6 groups: propaganda, advertisement, manipulation, satire, parody, and fabrication [6]. Although news organizations and social media companies have implemented measures to flag and delete false news, the rate of manual false news detection is not fast enough to compete with its rapid spread through social media [9,10]. Approximately 62% of US adults obtain news from social media sites; thus, faster fact checking is critical to ensure false information spread is reduced [11]. As such, the spread of false news has resulted in public confusion, potentially associated with the antimask and vaccine rhetoric [10]. Presently, one of the most common methods to detect false news online is through human-curated fact-checking websites, such as Snopes, to flag false information [12]. Although this method may be accurate, it is inefficient due to the large amount of false news generated during the COVID-19 pandemic [10]. Thus, automatic news article reliability detection is needed.

Current false news detection using machine learning (ML) on social media has been researched extensively. Various textual features from news pages are used to predict reliability of the articles. The use of multiple features to predict the presence of false information is a common theme within current false information detection studies. The use of multiple features can improve the performance of an ML model. For example, Reis et al [13] used textual features (eg, semantic and lexical features) and news source features (eg, credibility of the news organization) as inputs for the ML model. Using traditional classifiers, such as random forest and extreme gradient boosting (XGBoost), a performance of 0.85 and 0.86 area under the curve (AUC) was achieved, respectively [13]. Elhadad et al [14] used a voting ensemble method, in addition to feature engineering, for sentiment and part-of-speech tagging. Singhania et al [15] created a 3-level HAN model using input from words, sentences, and the headline level of a news article. Similar studies have proposed that other lexical features, such as n-grams, term frequency–inverse document frequency (TF-IDF), and probabilistic context-free grammar (PCFG) have also been used as features for misinformation prediction using deep learning [16]. Accordingly, feature engineering provides higher performance metrics as well as improved interpretability. These features allow the model to focus on the important elements, which allows for reliability prediction, especially in news articles, despite high heterogeneity and noise between samples. To build on what other false information research has found, as well as to identify important new factors that contribute to false information detection, we created a final ensemble model using the ReCOVery data set [17].

Ensemble methods were implemented to further improve the performance of misinformation detection within news articles. Ensemble model usage can benefit model performance by improving the ability to generalize to data on which the model has not been trained [18]. Kumar et al [19] demonstrated improvement in performance after the use of an ensemble model, where the use of an ensemble deep learning model with a convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) was able to achieve higher performance than a CNN or long short-term memory (LSTM) model alone, with a performance of 88.78% accuracy versus 73.29% and 80.62% for the CNN and LSTM, respectively. Due to the size of news articles, a bidirectional gated recurrent unit (BiGRU) was selected as the first model in the ensemble [20]. This model is a type of recurrent neural network (RNN) that functions well on sequential text data. A BiGRU solves the vanishing gradient problem, where the model trains on long news articles and "forgets" information from the start of the articles. This model is made of many neurons or cells, each with an update gate to control what new information is added at each word and a reset gate to control how much old information is retained. A BiGRU's bidirectional nature allows it to process each sample from the beginning and end of the article. Compared to other state-of-the-art natural language processing (NLP) models, such as LSTM, a gated recurrent unit (GRU) has lower parameters, making it quicker to train and use [21,22]. A quicker model is important as a large number of news articles are released each day; thus, a model for false information detection needs to be both accurate and fast in order to keep up with the number of new articles. XGBoost is another model included within our ensemble model. One strength of XGboost is its exceptional ability at learning from tabular data [23,24]. As a gradient boosted tree model, it is faster than a neural network and works better on the low-dimensionality output from the first model following feature extraction. Furthermore, XGBoost has been shown to outperform deep learning models for tabular data as the hyperparameter search is shorter [24]. Additionally, XGBoost combined with deep learning models in an ensemble model yields better results than an ensemble model with multiple deep learning models or classical ML models [24].

This study aims to provide a potential solution to the multifaceted false information problem through an ensemble deep learning model to classify the reliability of news articles using the ReCOVery data set. We hypothesize that sentiment, readability, lexical categories, and other text characteristics in news articles can be used together as inputs for news reliability classification improvement. We also explore differences in the sentiment or tone of reliable and unreliable information, which can be used to classify the reliability of the text. The outcome

XSL•FO
RenderX

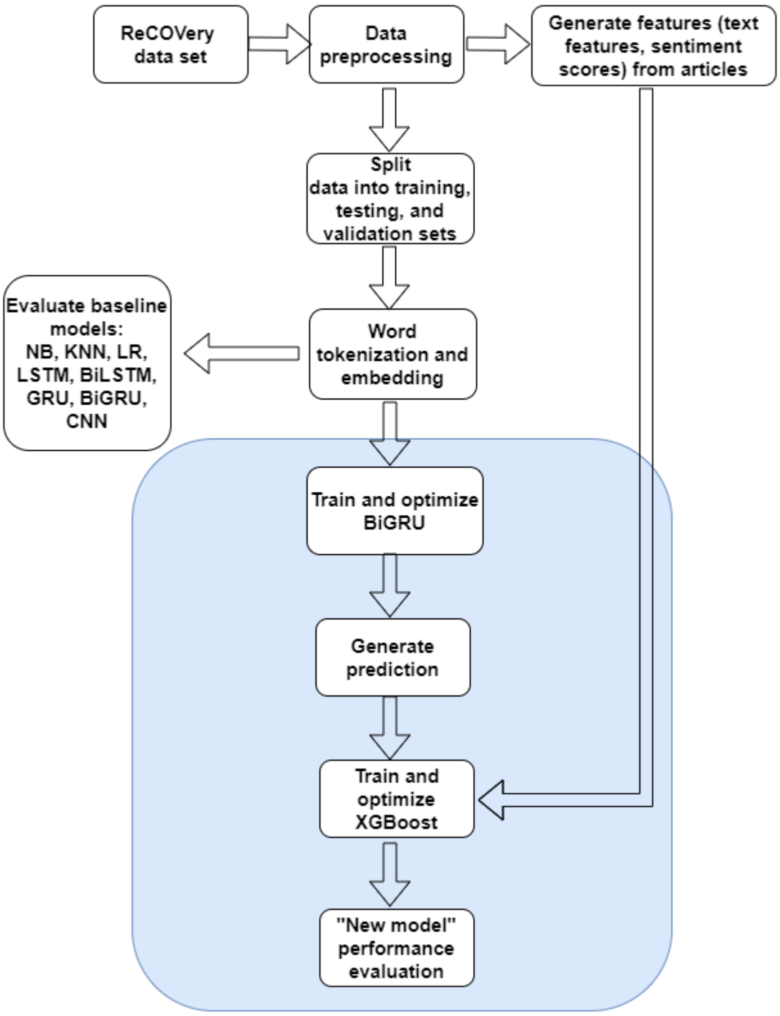
of our study may advance news reliability classification and help researchers and the public identify unreliable news articles in their everyday lives.

Methods

Workflow

First, data preprocessing was completed using the ReCOVery data set, which included removing stop words, links and Universal Resource Locators (URLs), and duplicate articles (Figure 1). Conversion of abbreviations and numbers to words was also completed within the preprocessing step. Following the preprocessing of the data, we performed feature engineering to create readability and sentiment scores, as well as extract lexical categories from the text (Figure 1). The preprocessed data were split into training, validation, and testing sets. Word tokenization and embedding were performed on the training and validation sets. Once tokenization and embedding were completed, 9 different ML models were trained and evaluated on the validation set to determine the best-performing model. We refer to naive Bayes (NB), K-nearest neighbors (KNNs), and logistic regression (LR) as traditional ML models as they are not deep learning models. The best-performing model was the ensemble model containing a bidirectional GRU and XGBoost ensemble "new model," as highlighted in blue in Figure 1.

Figure 1. Details of workflow for data exploration and "new model" construction (highlighted in blue). CNN: convolutional neural network; BiGRU: bidirectional gated recurrent unit; BiLSTM: bidirectional long short-term memory; GRU: gated recurrent unit; KNN: K-nearest neighbor; LR: logistic regression; LSTM: long short-term memory; NB: naive Bayes; XGBoost: extreme gradient boosting.



Data Description

The ReCOVery data set was our main source of data for news articles connected to Twitter posts [17]. It focuses on the reliability of news articles from a wide array of news sources and contains 2029 articles from ~2000 different news outlets from different countries (filtered from January to May 2020) that are related to COVID-19 news [17]. Each article was labeled as either 0 for unreliable or 1 as reliable according to the NewsGuard score [17]. The NewsGuard score was developed by journalists to label the reliability of an online article. Using a scale of 0-100, the NewGuard gives points to articles that accomplish credible and transparent news practices. Online articles with a score above 60 are labeled with a "green" rating as reliable sources, and scores below 60 are labeled with a "red" rating as unreliable sources [17,25]. In addition to the NewsGuard score, ReCOVery uses Media Bias/Fact Check,

XSL•FO
RenderX

which checks the correctness of news sources according to the article subjectivity and ranks articles from "very high" to "very low" in terms of factual reporting [17,26]. Reliable articles have a NewsGuard score higher than 90, with a "very high" or "high" rating on Media Bias/Fact Check. Unreliable articles have a NewsGuard score lower than 30, with a "mixed," "low," or "very low" factual rating on Media Bias/Fact Check [17]. The ReCOVery data set combined the NewsGuard and Media Bias/Fact Check scores to create the final news article reliability score.

Preprocessing

Prior to data analysis, the article text and tweet data were subjected to multiple preprocessing steps. The purpose of preprocessing was to clean the data so that the deep learning model could more efficiently detect patterns in the data. The steps taken to preprocess the article text included the removal of duplicates articles or tweets; common stop words, such as "the" and "a"; and all links and non-English characters. Lemmatization of the article text was also completed, in addition to the conversion of acronyms to full terms.

Preprocessing was conducted using Python libraries, such as Pandas and Natural Language Toolkit [27,28]. A total of 1346 reliable articles and 648 unreliable articles were used for model training. Additionally, 34 articles were removed as they had less than 100 words, which limited the validity of reliability analysis. Following preprocessing, features from the news articles such as text characteristics, readability, and sentiment were extracted for analysis and to be included as input to the deep learning model.

Sentiment Analysis

Sentiment analysis was applied to the body text of reliable and unreliable articles. This was implemented through Valence Aware Dictionary and Sentiment Reasoner (VADER) and TextBlob, which are open source tools for determining predominant sentiment, polarity, and subjectivity [29,30]. The analysis relies on lexicographic analysis to map the text features of each article to different scores with regard to sentiment, polarity, and intensity. In terms of sentiment, the articles have a continuous score between 0 and 1, including both endpoints, with 1 representing that the article contains the specified sentiment as the predominant sentiment. For example, if an article has a positive sentiment of 1, this means the article contains the highest-possible positive sentiment. VADER and TextBlob were imported into Python and applied to the body text of articles within the data set. The total proportion of articles with a positive, negative, and neutral sentiment were determined through library functions within VADER and TextBlob.

Text Analysis

After preprocessing, the body text of articles was analyzed. The most common words from reliable and unreliable articles were determined. They are presented in a frequency bar graph to demonstrate the major differences between unreliable and reliable articles (Figures 2 and 3, respectively). Another feature included within the deep learning model was the text length and readability of the newspaper articles. The length of the articles was assessed using the character length of the article sentences and overall article length. Readability was assessed using 6 different readability metrics from the py-readability-metrics library: the Flesch-Kincaid grade level, Gunning fog index, Coleman-Liau index, Dale-Chall index, automated readability index (ARI), and Linsear Write index [31]. The aforementioned readability metrics are used to determine the grade level necessary to understand a written document based on the sentence length and word length [32].

The Flesch-Kincaid grade level is a scale modified from the Flesch-Kincaid reading ease index that compares the ratio of words per sentence and the ratio of syllables per word [33]. The values for this scale linearly indicate the estimated US grade level of a text. For example, a grade of 10-12 would indicate that the target reader is at the high school level, whereas scores higher than 12 are graduate-level texts [33]. Similarly, the Coleman-Liau index and the ARI both assess character and word frequency to approximate the US grade level required to read a text [34]. The Gunning fog index assesses the frequency of difficult words in a text and is a linear range between 0 and 20: a score of 16-20 is at the graduate level [35]. Similarly, the Dale-Chall index evaluates the frequency of difficult words but is scaled so that a score of 9-10 represents a university graduate–level text [31,36-38]. Lastly, the Linsear Write index was developed to assess the readability of technical texts, and its score represents the years of formal US education needed to understand a text, similar to the previous indices [39].

Topic analysis was performed using Empath, a neural network–based lexicon [40]. Empath is able to determine whether a certain sentence has the lexical categories of politics, religion, contentment, and approximately 200 more categories [40]. By processing the text with Empath, we derived 194 lexical categories that were used as additional features that were concatenated with the previous text, sentiment, and readability features in the final deep learning model. The extracted lexical categories from Empath increased the amount of information the deep learning model trained on for each article and allowed for better interpretability as differences in topic frequencies could also be evaluated. For each of the lexical categories, a mean count for reliable and unreliable articles was derived, along with the *t* test and the *P* value (Table 1).

XSL•FO
RenderX

Figure 2. Number of occurrences for keywords in unreliable news articles (N=298,498 words).

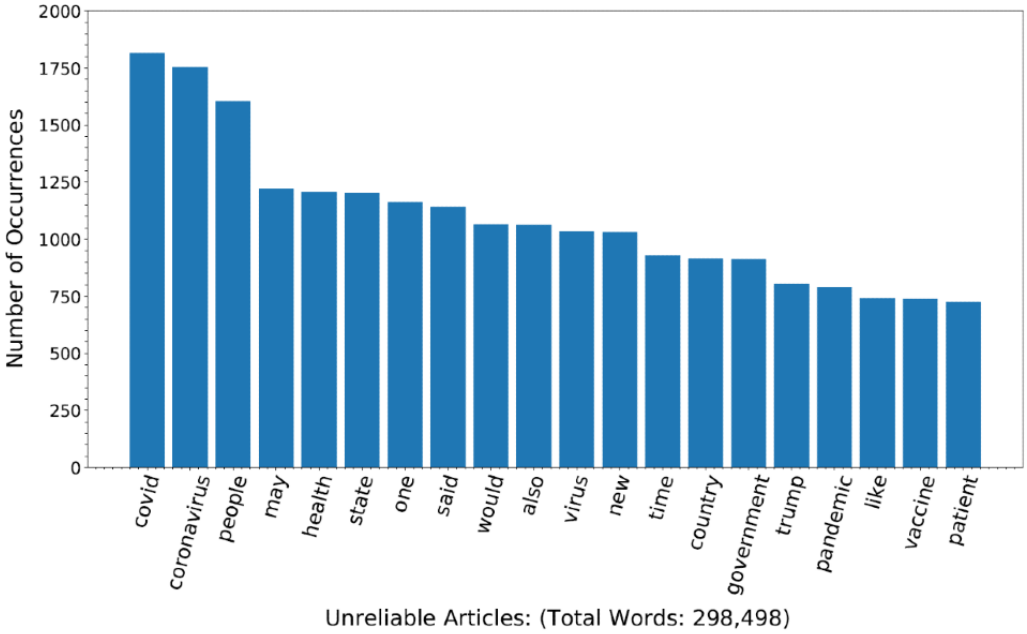
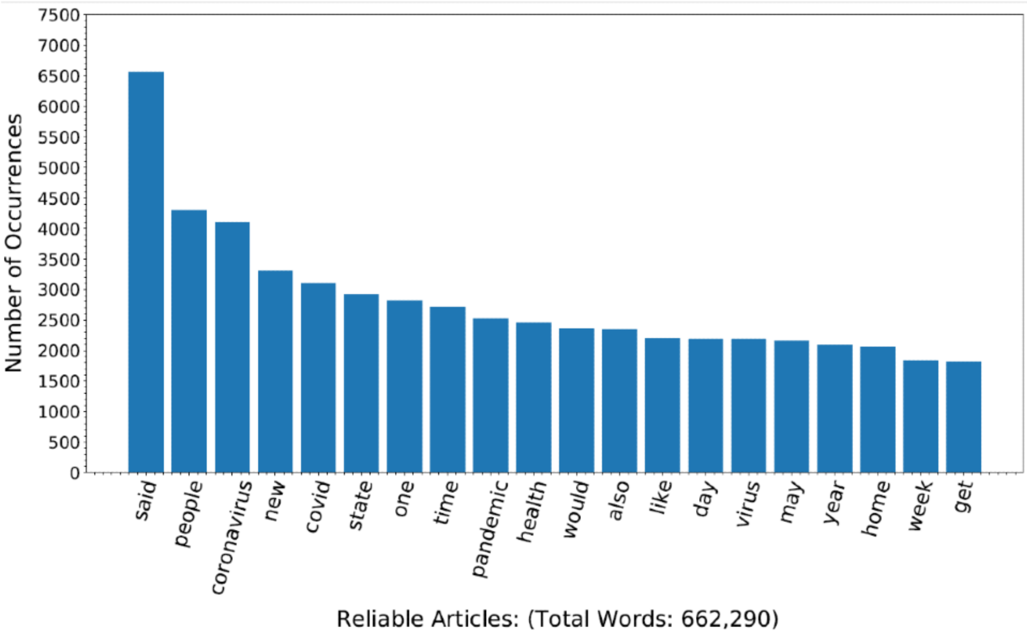


Figure 3. Number of occurrences of keywords in reliable news articles (N=662,290 words).



XSL•FO
RenderX

Table 1. Top 10 lexical categories from Empath (a neural network–based topic analysis tool) in reliable and unreliable news articles selected by Empath. The reliable and unreliable means is the mean counts of each lexical category being classified into reliable and unreliable news articles, respectively.

Lexical category	*t* (*df*)	*P* value	Reliable mean (SD)	Unreliable mean (SD)
magic	–7.91 (1992)	<.001	0.19 (0.60)	0.51 (1.22)
power	–7.16 (1992)	<.001	1.28 (2.20)	2.16 (3.24)
business	7.15 (1992)	<.001	8.58 (10.54)	5.31 (7.10)
work	6.89 (1992)	<.001	5.78 (8.82)	3.28 (3.89)
contentment	6.18 (1992)	<.001	0.70 (1.61)	0.29 (0.72)
office	6.14 (1992)	<.001	3.02 (4.37)	1.88 (2.60)
dispute	–6.11 (1992)	<.001	1.58 (2.48)	2.35 (2.94)
morning	5.87 (1992)	<.001	1.06 (1.87)	0.59 (1.11)
legend	–5.85 (1992)	<.001	0.34 (0.92)	0.64 (1.31)
blue collar job	5.83 (1992)	<.001	0.62 (1.75)	0.21 (0.68)

Tokenization

As ML models only accept numerical inputs, the text data must be tokenized. This process involves a word-index dictionary, where each word in the data set is converted to a numerical value or index, which corresponds to that word in the dictionary. For example, a word such as “coronavirus” might be presented to a ML model as the value 1234. As each unique word creates a unique index number, the “vocabulary” or total number of unique words in the data set can be a problem, especially if the data set is large, since words that appear once or twice in the data set generally do not contribute to the efficacy of the model. We limited the vocabulary size to 20,000 (51.73%) out of a total of 38,663 unique words from the training data. This excluded words that were used only once in the data set, because these words would not be useful to the model—Zipf’s law reaffirms that having larger vocabulary sizes gives diminishing returns as we frequently use a small proportion of their total vocabulary [41,42]. Furthermore, there are various estimates regarding the vocabulary size of an average native English speaker, with around 20,000 being a reasonable estimate for the vocabulary size [43,44]. Articles were also 0-padded to a size of 3500 words, which was the size of the longest article to ensure uniformity of the model input.

Word Embedding

Following tokenization, the data were processed using word embedding, a form of unsupervised ML. Word embedding places the data points of individual words into an embedding space with high dimensionality. Inside this embedding space, each word is represented as a vector with words that are similar to each other being located in close proximity. As such, word embedding allows hidden relationships between similar words to be quantified for ML analysis. Although a new word embedding layer could be trained and fitted on our data set, there exist pretrained word embedding models that are more efficient to use. For the article text data, we leveraged Global Vectors for Word Representation (GloVE), which is a commonly used word embedding model trained on hundreds of thousands of Wikipedia articles, which have an embedding space of 100 dimensions [45].

Machine Learning Classification

The data were randomly split into training, testing, and validation subsets for deep learning. The ratio of these subsets was 8:1:1, respectively. Of the 1994 articles, 1595 (79.99%) were in the training subset, 199 (9.98%) were in the validation subset, and 200 (10.03%) were in the testing subset. The training and validation data were used to build the model to classify between reliable and unreliable articles, while the testing data were used to evaluate the model’s performance. The splitting of the data followed by model training and evaluation were repeated 10-folds so that each article could be included in the training set. An average was taken between the performance metrics obtained from training on each fold. We evaluated the performance of multiple ML models on the data set (NB, KNNs, LR, LSTM, GRU, BiLSTM, BiGRU, and CNN) to determine the best models for reliability detection. The settings or hyperparameters were optimized either experimentally or using Gridsearch, which tests all combinations of hyperparameters for each of the aforementioned ML models.

Finally, we developed an ensemble model using a lightly trained BiGRU to generate an initial reliability prediction, which was then combined with the text features, readability, sentiment, and Empath-classified lexical categories. This was then used to train an XGBoost model with 10-fold cross-validation.

This paper uses several evaluation metrics that rely on the results from the confusion matrix. These metrics were derived from correct predictions by the model, such as true positive (TP) and true negative (TN), as well as incorrect predictions, such as false positive (FP) and false negative (FN). Accuracy is the total proportion of correct predictions, but this evaluation metric is not as effective when there is a class imbalance. Sensitivity refers to the proportion of correctly predicted unreliable articles, while specificity refers to the proportion of correctly predicted reliable articles. The AUC score shows the performance of the model at different TP and FP rates [46].

$$\text{Sensitivity (recall)} = TP/(TP + FN)$$

$$\text{Specificity} = TN/(TN + FP)$$

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

XSL•FO
RenderX

Ethical Considerations

The data used in this paper did not need ethics approval as they were accessed through the open access ReCOVery data set GitHub, as cited in Zhou et al [17].

Results

Data Exploration

Data exploration was performed and features, such as readability, sentiment, and lexical categories, were combined with the full news article text data to train an ensemble model. An ensemble method using BiGRU and XGBoost was created using 1346 reliable articles and 648 unreliable articles.

During data exploration, we found that the average text length in terms of the average word length and sentence length was longer in unreliable articles compared to reliable articles (Table 2). The Flesch-Kincaid grade level, the Dale-Chall index, the ARI, the Coleman-Liau index, the Gunning fog index, and the Linsear Write index indicated that reliable articles are easier to read compared to unreliable articles (Table 2). From the average frequency of 194 Empath-derived lexical categories, 110 (56.7%) were significantly different between reliable and unreliable articles (Multimedia Appendix 1). Most frequent words in unreliable and reliable articles were also visualized (Figures 2 and 3, respectively). Unreliable articles had higher rates of negative sentiment, while reliable articles had higher rates of neutral sentiment (Table 3). Performance metrics of various trained ML models as well as the new ensemble model were determined (Table 3).

Table 2. Text length and readability metrics for reliable (N=1346) and unreliable (N=648) online news articles. The text length was expressed as the average sentence length and word length. Readability was expressed using the Flesch-Kincaid grade level, the Dale-Chall readability index, the ARI[a], the Coleman-Liau index, the Gunning fog index, and the Linsear Write index.

Metrics	Reliable mean (SD)	Unreliable mean (SD)	*t (df)*	*P* value
Average word length (characters)	6.14 (0.27)	6.32 (1.66)	–3.93 (1992)	<.001
Average sentence length (words)	23.67 (5.17)	26.38 (7.06)	–9.70 (1992)	<.001
Flesch-Kincaid grade level	12.68 (2.63)	14.39 (3.37)	–12.38 (1992)	<.001
Gunning fog index	14.87 (2.72)	16.42 (3.33)	–11.00 (1992)	<.001
Coleman-Liau index	10.85 (1.87)	11.82 (2.46)	–9.72 (1992)	<.001
Dale-Chall index	10.21 (0.96)	10.70 (1.02)	–10.53 (1992)	<.001
ARI	13.41 (3.30)	15.43 (4.47)	–11.41 (1992)	<.001
Linsear Write index	16.42 (4.02)	18.73 (5.31)	–10.80 (1992)	<.001

[a]ARI: automated readability index.

Table 3. Comparison of sentiment polarity (0=least expression of sentiment in interest, 1=most expression of sentiment in interest) between reliable (N=1346) and unreliable (N=648) news articles in terms of sentiment of the sentences within news articles. Differences between the frequencies of sentences possessing positive, neutral, or negative sentiment were analyzed with a 2-sample independent *t* test.

Sentiment	Reliable mean (SD)	Unreliable mean (SD)	*t (df)*	*P* value
Negative	0.066 (0.042)	0.076 (0.039)	–5.46 (1992)	<.001
Neutral	0.850 (0.054)	0.840 (0.050)	4.37 (1992)	<.001
Positive	0.084 (0.035)	0.085 (0.035)	–0.095 (1992)	.92

Text Analysis

After removal of stop words, the most frequent words in reliable and unreliable articles were examined. The highest word frequencies for unreliable and reliable articles are illustrated in frequency bar graphs (Figures 2 and 3). Common words between reliable and unreliable news articles were COVID-19–related keywords, such as "coronavirus," "virus," and "pandemic." The differences were related to political undertones, such as "Trump" and "government." Additionally, the Empath lexicon tool was applied to the text to yield lexical categories. The average count for each lexical category was determined for reliable and unreliable text. The differences in means were then compared using *t* tests. There were a total of 194 lexical categories that significantly differed in frequency between reliable and unreliable texts (Multimedia Appendix 1 and Table 1). In Table 1, we display the top 10 lexical categories with the lowest *P* value. Categories included "magic," "power," "business," "work," "contentment," "office," "dispute," "morning," "legend," and "blue collar job." The lexical categories "business," "work," "contentment," "office," "morning," and "blue collar job" had higher mean counts for the reliable articles compared to the unreliable articles. The lexical categories "magic," "power," "legend," and "dispute" had lower mean counts for the reliable articles compared to the unreliable articles. In terms of text characteristics, there was a significant difference in the average sentence length between reliable and unreliable news articles, with reliable articles containing shorter sentences at 23.67 (SD 5.17) words per sentence compared to unreliable articles containing 26.38 (SD 7.06) words per sentence (Table 2). Additionally, the average word lengths were 6.14 (SD 0.27) and 6.32 (SD 1.66) for reliable and unreliable

XSL•FO
RenderX

articles, respectively. In addition to text length, we also analyzed the differences in readability between reliable and unreliable articles. The readability indices used were the Flesch-Kincaid grade level, the Dale-Chall index, the ARI, the Coleman-Liau index, the Gunning fog index, and the Linsear Write index. As shown in Table 2, unreliable articles were less readable, as indicated by all 6 readability indices. Since these text features are important in differentiating between reliable and unreliable news articles, they were input into our final deep learning model.

Sentiment Analysis

Using VADER, the sentences from the articles were classified into positive, neutral, and negative sentiments. The sentiment score ranged from 0 to 1, with 1 denoting strong presentation of the sentiment of interest. For reliable articles, the means for the negative, neutral, and positive sentiments scores were 0.066 (SD 0.042), 0.850 (SD 0.054), and 0.084 (SD 0.035), respectively (Table 3). For unreliable articles, the means for the negative, neutral, and positive sentiment scores were 0.076 (SD 0.039), 0.840 (SD 0.050), and 0.084 (SD 0.035), respectively.

Machine Learning Analysis

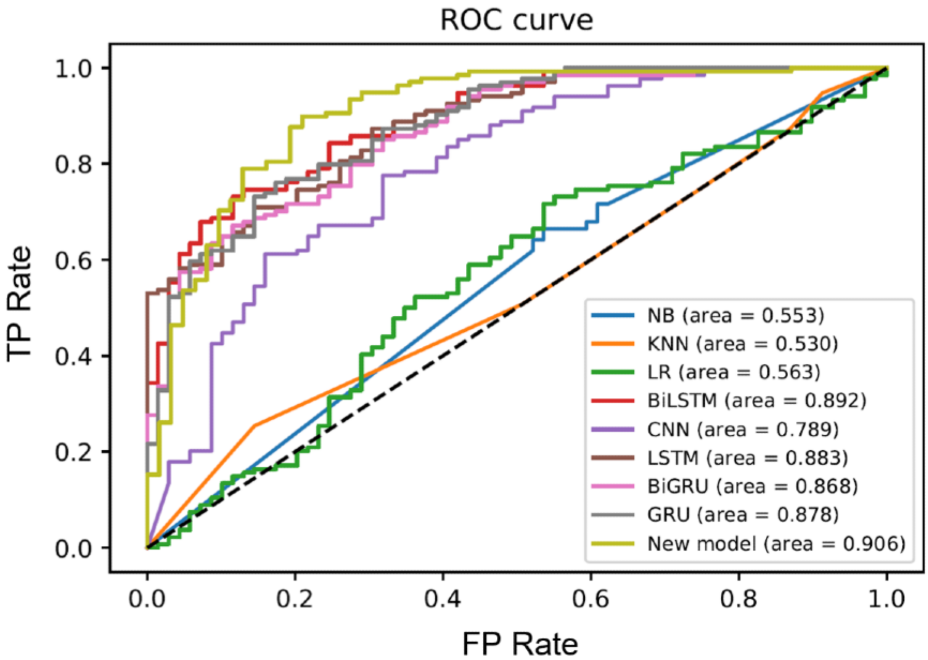
After the newspaper article data were passed through GloVE word embedding, the text data were split 10-folds for cross-validation. The traditional ML models included LR, KNNs, and NB. The AUC values (Figure 4) were generated, in addition to sensitivity and recall values (Table 4).

Next, the deep learning models were fit to the data. Each model included the GloVE word embedding prior to training. Optimization of hyperparameters for the deep learning models was completed using GridSearchCV from the ML Python scikit-learn library. The hyperparameters optimized were batch size, epochs, dropout rate, neuron number, optimizer type, learning rate, and activation function type. Each model had varying hyperparameters that yielded the best results.

The deep learning models that were assessed were LSTM, GRU, BiLSTM, BiGRU, and CNN. Similar to traditional ML models, the AUC, specificity, and recall were determined as performance metrics (Table 4).

Lastly, an ensemble model was developed using the BiGRU and XGBoost. Our new model was first evaluated on the ReCOVery testing subset. A confusion matrix for our new model was generated, as shown in Figure 5. The AUC, specificity, and sensitivity for our new deep learning model were 0.906, 0.835, and 0.945, respectively (Table 4).

Figure 4. Receiver operating characteristic (ROC) curve and AUC scores with the corresponding color for both traditional ML models (KNN, LR,NB) and deep learning models (BiLSTM, CNN, LSTM, BiGRU, GRU, new model). AUC: area under the curve; BiGRU: bidirectional gated recurrent unit; BiLSTM: bidirectional long short-term memory; CNN: convolutional neural network; FP: false positive; GRU: gated recurrent unit; KNN: K-nearest neighbor; LR: logistic regression; LSTM: long short-term memory; ML: machine learning; NB: naive Bayes; TP: true positive.



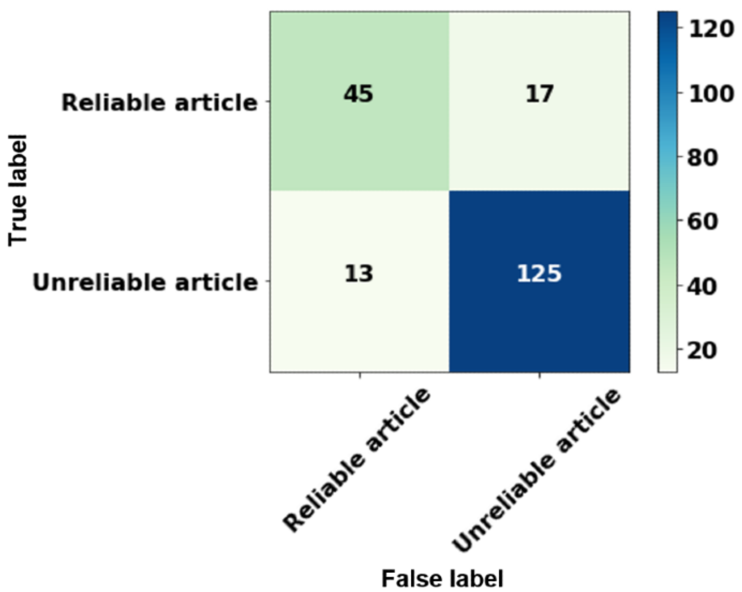
XSL•FO
RenderX

Table 4. Performance metrics for the ReCOVery validation data set for traditional ML[a] models (KNN[b], LR[c], NB[d]), and deep learning models (BiLSTM[e], CNN[f], LSTM[g], BiGRU[h], GRU[i], new model).

Model	Specificity	Sensitivity	AUC[j]
LR	0.720	0.575	0.563
KNN	0.660	0.739	0.530
NB	0.700	0.627	0.553
BiLSTM	0.810	0.925	0.892
CNN	0.792	0.851	0.789
LSTM	0.829	0.903	0.883
BiGRU	0.791	0.963	0.868
GRU	0.804	0.918	0.878
New model	0.835	0.945	0.906

[a]ML: machine learning.
[b]KNN: K-nearest neighbor.
[c]LR: logistic regression.
[d]NB: naive Bayes.
[e]BiLSTM: bidirectional long short-term memory.
[f]CNN: convolutional neural network.
[g]LSTM: long short-term memory.
[h]BiGRU: bidirectional gated recurrent unit.
[i]GRU: gated recurrent unit.
[j]AUC: area under the curve.

Figure 5. Confusion matrix for ReCOVery validation subset on trained new ensemble model with BiGRU and XGBoost. BiGRU: bidirectional gated recurrent unit; XGBoost: extreme gradient boosting.



Discussion

Principal Findings

This study demonstrates an ensemble model with BiGRU and XGBoost for text reliability classification using the ReCOVery data set with a specificity, sensitivity, and AUC of 0.835, 0.945, and 0.906, respectively [17]. Through our data analysis, we demonstrated that unreliable news articles have lower readability and higher sentence length. They also include more negative and less neutral sentiments and contain more polarizing lexical categories in comparison to reliable articles.

Data Usage

With regard to using news articles to build a classification model, an important consideration is the generalizability of the model. To ensure that the model is generalizable, the data used to train the model must be diverse in nature. A shortcoming of

XSL•FO
RenderX

many deep learning misinformation detection studies is the focus on detecting misinformation from a narrow range of news sources, or locations [17,47]. Because of the homogenous nature of the data set used to train these models, many misinformation detection models are potentially less generalizable [47]. An example would be CoAID, a data set constructed from COVID-19–related news articles and social media posts from December 1, 2019, to September 1, 2020. A shortcoming of the CoAID data set would be the lower number of news sources used for the data set as 9 reliable news sources were included during the data collection process [48]. CoVerifi is a study that used the CoAID data set to create a web-based tool to check whether an online news article was credible [49]. Another notable data set is the COVID-19-FAKES data set containing 61,711 tweets with misinformation and 2,985,399 tweets without misinformation [50,51]. Silva et al [51] used the COVID-19-FAKES data set to obtain insights into predictive features for the presence of misinformation in tweets and the differential engagement in tweets with and without misinformation [51]. Hence, we used the ReCOVery data set for the diverse nature of the news articles as they range from ~2000 different news outlets from different countries [17].

Sentiment Analysis

VADER was used to evaluate sentiment at a lexicon-based level due to its high accuracy, with an *F*1 classification accuracy of 0.96 and computational economy [29]. Although VADER has become a staple in NLP for sentiment analysis, 2 key shortcomings to consider are its inability to recognize sarcasm/satire and its reduced accuracy when dealing with 3-class analyses (negative, neutral, and positive) [52].

From the distribution of articles with reliable versus unreliable news articles, it can be observed that reliable articles contain less negative sentiment in comparison to unreliable articles as they had a lower negative sentiment polarity score (Table 3). This is in line with observations of news content in the literature, as Arif et al [53] discussed how individuals searching for negative terms on the internet can lead to more biased articles. To emphasize the importance of sentiment in differentiating fake and real news, Paschen [54] concluded that the titles and body text of fake news articles contain more negative content, such as anger and disgust, compared to real news articles. Fake news is more likely to display negative sentiment to drive a specific narrative for profit, which supports our finding that there are a greater number of negative unreliable sources than neutral or positive unreliable sources.

We observed a difference between the number of neutral reliable and neutral unreliable articles, with more neutral sentiment in reliable articles in comparison to unreliable articles (Table 3). A neutral sentiment scoring for reliable data sources implies impartiality and objectivity when discussing the subject matter [55].

Many ML studies have targeted sentiment as a feature to predict misinformation in a variety of written information online because of the different sentiment valence between reliable and unreliable text due to the aforementioned reasons [56]. Because of the differing nature of sentiment between texts of differing reliability, sentiment analysis was used in the context of filtering out negative messages on social media, spam filtering, among other applications [56]. In agreement with our findings, Ajao et al [57] determined that unreliable tweets often contain more negative sentiment in comparison to reliable tweets due to how authors of unreliable tweets use negative emotions to better propagate their message. They also showed that the use of sentiment can boost support vector machine (SVM) accuracy when the sentiment is considered in addition to textual features [57]. Hence, sentiment was a feature selected for our model.

Text Analysis

The words themselves were observed to be quite similar to one another between the 2 groups because the subject matter of both reliable and unreliable sources is the same: COVID-19. Additionally, many of the most frequently occurring words are mere transitional words that are likely to be found in the majority of English literature.

Interestingly, the most frequently occurring word in reliable sources was “said” (Figure 3). This is likely due to “said” being used to quote political figures and leaders in the scientific field. The reliability of articles in this case is a consequence of the articles citing reliable sources of information. Another observable trend is the increasing number of politically charged words found in unreliable articles. Words such as “country,” government,” and “Trump” were amongst the most frequent words for unreliable sources but not for reliable articles (Figure 2). This communicates a pattern of political commentary occurring in unreliable sources [58]. We can anticipate that articles discussing political content in the context of COVID-19 are likely interested in propagating an agenda—hence, the unreliability. For example, Chen et al [59] found interplay between COVID-19 misinformation propagation and the 2020 US presidential elections with regard to mask use and mail-in ballots. Specifically, health information has been politicized to push political agendas and attack political opponents. In addition to frequently occurring words, lexical categories extracted from Empath and similar models allows us to evaluate the difference in topic frequencies between reliable and unreliable news articles [40]. The use of lexical categories extracted from Empath and similar models can increase model performance compared to using only raw text data [60-63].

Another feature we decided to explore and include in our final deep learning model is the readability and length of the news articles. Readability has been shown to be predictive of misinformation. In the study by Santos et al [64], articles from a frequent source of fake news could be differentiated using only article readability scores with an SVM algorithm with an accuracy of 92% [64]. Similarly, in a study by Zhou et al [65], various metrics were explored based on their ability to classify reliable versus unreliable news articles. It was determined using random forests that readability is among the top 5 in terms of contribution to the model, alongside sentiment [65].

Machine Learning Classification

In the original ReCOVery study, Zhou et al [17] created a baseline prediction performance for news article reliability and found that a precision of 0.721-0.836 and 0.421-0.667 can be obtained for reliable and unreliable news articles, respectively.

XSL•FO
RenderX

A recall of 0.705-0.829 and 0.441-0.667 can be obtained for reliable and unreliable news articles, respectively [17]. The features used in the baseline model ranged from text lexical categories, rhetorical structure, and visual information within news articles. Zhou et al [17] also tested the model on traditional ML models, such as SVMs, or deep learning algorithms, such as CNNs with unimodal and multimodal features. Other studies have also explored the use of the ReCOVery data set for false information classification. One such study is by Raj and Meel [66], where a novel deep learning model, the Allied Recurrent and Convolutional Neural Network (ARCNN), was created using both image and textual features within news articles to detect misinformation. The performance of the ARCNN was tested using 6 COVID-19 fake news data sets, with ReCOVery as 1 of the data sets, achieving an accuracy, precision, recall, and *F*1 score of 80.98%, 53.85%, 58.33%, and 56.00%, respectively [66]. Another study using the ReCOVery data set for model development explored the use of multiple languages for fake news detection to improve model performance [67]. Finally, Wahle et al [68] used the ReCOVery data set as 1 of 6 COVID-19 misinformation data sets to evaluate the performance of 15 transformer-based ML models to determine the generalizability of different transformer models. Differing from the aforementioned studies, we were able to demonstrate that the use of readability, text characteristics, sentiment, and lexical categories can improve upon the original ReCOVery data set baseline models [17]. Hence, we demonstrate the importance of the aforementioned text features to improve upon news article reliability prediction. Furthermore, we show that the combination of multiple inputs and consideration of the chosen model can increase ML model accuracy in the context of NLP.

In our final proposed model, the BiGRU with XGBoost and feature engineering was the best-performing model. A BiGRU is adept at capturing temporal data in long sequences, as bidirectional models can better capture the context of the text [46]. During the experimentation with these models on ReCOVery data, we found that all deep learning models outperformed the traditional ML models because deep learning models are better able to handle more complex data [46,69]. Furthermore, we chose to use the GRU algorithm, which is a variant of the recurrent neural network, in addition to the LSTM algorithm due to the increased performance on longer text compared to LSTM [21]. To further increase the performance of our model, an ensemble model was built, as combining multiple predictions can yield more accurate predictions [70].

Strengths

A strength of our investigation is that it not only had the main goal of creating a deep learning model for reliability prediction but also identified significant trends in text and sentiment for reliable and unreliable news articles. An investigation focused solely on a deep learning model has a "black box" problem where the mechanisms used by the deep learning model are not visible and are contained within its many complex hidden layers [71]. As such, a data exploration approach coupled with the deep learning model is able to better visualize and portray article reliability classification. Furthermore, our paper examined news articles, which had the advantage of being more normalized in text compared to tweets and social media as, each article was written with a professional approach. As such, less data were removed during preprocessing due to grammatical or spelling errors. Using news articles as data also avoided the problem of low hydration that Twitter misinformation data sets suffer from when tweets are removed by Twitter.

Limitations and Future Directions

There are a number of ways our project could be further refined. First, expanding the number of total available data would be valuable as there are nearly twice as much data for reliable sources as unreliable. Furthermore, being able to web-scrape Facebook postings and Reddit threads would allow us to expand our scope of access and evaluate other high-traffic sources of information. Incorporating clustering models would also increase the specificity of our search and create a more accurate model that can consider what aspect of COVID-19 is being discussed when determining reliability. Due to the high accuracy of our model, as shown by the results, our model can be commercialized as a web app that allows users to assess, to a high degree of confidence, the reliability of the article they are reading. Moreover, it can also be used to determine the sentiment scoring of an article to determine whether they want to engage in that specific literature.

Although this model specifically identifies COVID-19–related information, it could also be trained for other types of misinformation. As discussed previously, most current methods to combat misinformation online are through the use of human-moderated fact-checking websites. Examples include Twitter's Birdwatch program, where independent users can flag posts they deem untrustworthy [72]. Other methods used include Facebook's fact-checking service, which manually labels posts or websites containing misinformation as untrustworthy and removes them from public view [73]. Furthermore, warnings are placed below posts containing COVID-19 information to warn readers regarding potential misinformation contained within posts [73]. Even though there are numerous instances of fact checking, the major issue that arises is the inefficiency in manual fact checking [74]. Hence, new fact-checking methods aim toward automating the fact-checking process. The first example of a fact-checking website is the Bot Sentinel automated Twitter fact-checking software, which can be installed by users to monitor spam accounts [75]. Bot Sentinel uses ML technology to classify posts or profiles as reliable or unreliable to an accuracy of 95% [75].

Conclusion

In conclusion, we demonstrated that readability, sentiment, text characteristics, and lexical categories are important in differentiating between reliable and unreliable news articles, as it was shown that unreliable articles are less readable, have more negative sentiment, and have more political lexical categories. The aforementioned features were used to achieve above-the-baseline performance within the original ReCOVery data set, with a specificity, sensitivity, and AUC of 0.835, 0.945, and 0.906, respectively, using our new ensemble deep learning model. Hence, the application of readability, sentiment, and lexical categories using our new model can help determine the dependability of news articles and better improve upon pre-existing models that do not use these features.

XSL•FO
RenderX

COVID-19 has brought to light the importance of developing an automated reliability assessor for news articles, as human-moderated fact-checking methods may be inefficient. Because readability, sentiment, and lexical categories can be used to improve upon pre-existing reliability classification models, we show that automated reliability detection may be an alternate way to determine new article reliability in the future, which will help news readers identify articles containing potentially unreliable information.

Acknowledgments

We would like to thank the science, technology, engineering, and mathematics (STEM) fellowship team for organizing the National Undergraduate Big Data Challenge 2021 and JMIR for sponsoring this publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mean (SDs) scores for Empath categories of reliable and unreliable news articles.

[XLSX File (Microsoft Excel File), 31 KB - infodemiology_v2i2e38839_app1.xlsx]

References

1. 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]
2. World Health Organization. Infodemic. 2022. URL: https://www.who.int/health-topics/infodemic [accessed 2022-06-15]
3. Obiała J, Obiała K, Mańczak M, Owoc J, Olszewski R. COVID-19 misinformation: accuracy of articles about coronavirus prevention mostly shared on social media. Health Policy Technol 2021 Mar 10;10(1):182-186 [FREE Full text] [doi: 10.1016/j.hlpt.2020.10.007] [Medline: 33163352]
4. Figueira Á, Oliveira L. The current state of fake news: challenges and opportunities. Procedia Comput Sci 2017;121:817-825. [doi: 10.1016/j.procs.2017.11.106]
5. 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]
6. Tandoc E, Lim Z, Ling R. Defining "fake news". Digit J 2018;6(2):137-153. [doi: 10.1002/9781119011071.iemp0300]
7. Lazer DMJ, 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]
8. Tandoc EC. The facts of fake news: a research review. Sociol Compass 2019 Jul 25;13(9):e12724. [doi: 10.1111/soc4.12724]
9. de Oliveira NR, Pisa PS, Lopez MA, de Medeiros DSV, Mattos DMF. Identifying fake news on social networks based on natural language processing: trends and challenges. Information 2021 Jan 18;12(1):38. [doi: 10.3390/info12010038]
10. Torabi Asr F, Taboada M. Big data and quality data for fake news and misinformation detection. Big Data Soc 2019 May 23;6(1):205395171984331. [doi: 10.1177/2053951719843310]
11. Clayton K, Blair S, Busam JA, Forstner S, Glance J, Green G, et al. Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. Polit Behav 2019 Feb 11;42(4):1073-1095. [doi: 10.1007/s11109-019-09533-0]
12. Krause NM, Freiling I, Beets B, Brossard D. Fact-checking as risk communication: the multi-layered risk of misinformation in times of COVID-19. J Risk Res 2020 Apr 22;23(7-8):1052-1059. [doi: 10.1080/13669877.2020.1756385]
13. Reis JCS, Correia A, Murai F, Veloso A, Benevenuto F. Supervised learning for fake news detection. IEEE Intell Syst 2019 Mar;34(2):76-81. [doi: 10.1109/mis.2019.2899143]
14. Elhadad M, Li K, Gebali F. An ensemble deep learning technique to detect COVID-19 misleading information. In: Advances in Networked-Based Information Systems. Manhattan, NY: Springer International; 2020 Aug 20 Presented at: International Conference on Network-Based Information Systems; 2020; Victoria, Canada p. 163-175. [doi: 10.1007/978-3-030-57811-4_16]
15. Singhania S, Fernandez N, Rao S. 3HAN: a deep neural network for fake news detection. 2017 Presented at: 24th International Conference on Neural Information Processing (ICONIP 2017); 2017; Guangzhou, China. [doi: 10.1007/978-3-319-70096-0_59]
16. Khan JY, Khondaker MTI, Afroz S, Uddin G, Iqbal A. A benchmark study of machine learning models for online fake news detection. Mach Learn Appl 2021 Jun;4:100032. [doi: 10.1016/j.mlwa.2021.100032]
17. Zhou X, Mulay A, Ferrara E, Zafarani R. ReCOVery: a multimodal repository for COVID-19 news credibility research. 2020 Presented at: CIKM '20: 29th ACM International Conference on Information & Knowledge Management; October 19-23, 2020; Virtual Event Ireland p. 3205-3212. [doi: 10.1145/3340531.3412880]

XSL•FO
RenderX

18. Ganaie M, Hu M, Malik A, Tanveer M, Suganthan P. Ensemble deep learning: a review. arXiv Preprint posted online April 6, 2021 [FREE Full text] [doi: 10.1016/j.engappai.2022.105151]
19. Kumar S, Asthana R, Upadhyay S, Upreti N, Akbar M. Fake news detection using deep learning models: a novel approach. Trans Emerging Tel Tech 2019 Nov 05;31(2):e3767. [doi: 10.1002/ett.3767]
20. Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv Preprint posted online December 11, 2014. [doi: 10.48550/arXiv.1412.3555]
21. Yang S, Yu X, Zhou Y. LSTM and GRU neural network performance comparison study: taking yelp review dataset as an example. 2020 Presented at: 2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI); June 12-14, 2020; Shanghai, China. [doi: 10.1109/iwecai50956.2020.00027]
22. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997 Nov 15;9(8):1735-1780. [doi: 10.1162/neco.1997.9.8.1735] [Medline: 9377276]
23. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. 2016 Presented at: KDD '16: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA. [doi: 10.1145/2939672.2939785]
24. Shwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need. Inf Fusion 2022 May;81:84-90. [doi: 10.1016/j.inffus.2021.11.011]
25. NewsGuard Technologies. Rating Process and Criteria. 2022. URL: https://www.newsguardtech.com/ratings/rating-process-criteria/ [accessed 2022-06-14]
26. Zandt D. About Media Bias / Fact Check. URL: https://mediabiasfactcheck.com/about/ [accessed 2022-06-14]
27. Loper E, Bird S. NLTK: the natural language toolkit. arXiv Preprint posted online May 17, 2002 [FREE Full text] [doi: 10.3115/1118108.1118117]
28. McKinney W. Data structures for statistical computing in Python. 2010 Presented at: 9th Python in Science Conference (SciPy 2010); June 28-July 3, 2010; Austin, TX p. 56-61. [doi: 10.25080/majora-92bf1922-00a]
29. Hutto C, Gilbert E. VADER: a parsimonious rule-based model for sentiment analysis of social media text. 2014 Presented at: 8th International AAAI Conference on Weblogs and Social Media; June 1-4, 2014; Ann Arbor, MI URL: https://ojs.aaai.org/index.php/ICWSM/article/view/14550
30. Loria S. TextBlob: Simplified Text Processing. URL: https://textblob.readthedocs.io/en/dev/ [accessed 2022-09-13]
31. DiMAscio C. py-readability-metrics. URL: https://github.com/cdimascio/py-readability-metrics [accessed 2022-03-01]
32. Si L, Callan J. A statistical model for scientific readability. 2001 Presented at: CIKM01: 10th International Conference on Information and Knowledge Management; October 5-10, 2001; Atlanta, GA p. 574-576. [doi: 10.1145/502585.502695]
33. Hoke B. Comparison of Recreational Reading Books Levels Using the Fry Readability Graph and the Flesch-Kincaid Grade Level. 1999. URL: https://eric.ed.gov/?id=ED428333 [accessed 2022-06-16]
34. Karmakar S, Zhu Y. Visualizing multiple text readability indexes. 2010 Presented at: International Conference on Education and Management Technology (ICEMT 2010); November 2-4, 2010; Cairo, Egypt. [doi: 10.1109/icemt.2010.5657684]
35. Gunning R. The fog index after twenty years. J Bus Commun 2016 Sep 16;6(2):3-13. [doi: 10.1177/002194366900600202]
36. Burke V, Greenberg D. Determining readability: how to select and apply easy-to-use readability formulas to assess the difficulty of adult literacy materials. Adult Basic Educ Lit J 2010;4(1):34-42.
37. Dale E, Chall J. A formula for predicting readability. Educ Res Bull 1948;27(1):11-28.
38. Scott B. The New Dale-Chall Readability Formula: A Vocabulary-Based Readability Formula. 2003. URL: https://www.readabilityformulas.com/new-dale-chall-readability-formula.php [accessed 2022-06-17]
39. Scott B. How to Use the Linsear Write Readability Formula to Grade Your Text. 2003. URL: https://readabilityformulas.com/linsear-write-readability-formula.php [accessed 2022-06-17]
40. Fast E, Chen B, Bernstein M. Empath: understanding topic signals in large-scale text. 2016 Presented at: ACM CHI 2016: 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 4647-4657. [doi: 10.1145/2858036.2858535]
41. Gao L, Zhou G, Luo J, Huang Y. Word embedding with Zipf's context. IEEE Access 2019;7:168934-168943. [doi: 10.1109/access.2019.2954691]
42. Sigurd B, Eeg-Olofsson M, van Weijer J. Word length, sentence length and frequency - Zipf revisited. Studia Linguist 2004 Apr;58(1):37-52. [doi: 10.1111/j.0039-3193.2004.00109.x]
43. Brysbaert M, Stevens M, Mandera P, Keuleers E. How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. Front Psychol 2016;7:1116 [FREE Full text] [doi: 10.3389/fpsyg.2016.01116] [Medline: 27524974]
44. Goulden R, Nation P, Read J. How large can a receptive vocabulary be? Appl Linguist 1990 Dec;11(4):341-363. [doi: 10.26686/wgtn.12560441]
45. Pennington J, Socher R, Manning C. Glove: Global Vectors for Word Representation. 2014 Presented at: 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); October 25-29, 2014; Doha, Qatar p. 1532-1543. [doi: 10.3115/v1/d14-1162]
46. Elfaik H, Nfaoui E. Deep bidirectional LSTM network learning-based sentiment analysis for Arabic text. J Intell Syst 2021;30(1):395-412. [doi: 10.1515/jisys-2020-0021]

XSL•FO
RenderX

47. Suprem A, Pu C. Evaluating generalizability of fine-tuned models for fake news detection. arXiv Preprint posted online May 15, 2022 . [doi: 10.48550/arXiv.2205.07154]
48. Cui L, Lee D. CoAID: COVID-19 healthcare misinformation dataset. arXiv Preprint posted online May 22, 2020 [FREE Full text]
49. Kolluri NL, Murthy D. CoVerifi: a COVID-19 news verification system. Online Soc Netw Media 2021 Mar;22:100123 [FREE Full text] [doi: 10.1016/j.osnem.2021.100123] [Medline: 33521412]
50. Elhadad M, Li K, Gebali F. COVID-19-FAKES: a Twitter (Arabic/English) dataset for detecting misleading information on COVID-19. In: Advances in Intelligent Networking and Collaborative Systems. Manhattan, NY: Springer International; 2021 Presented at: International Conference on Intelligent Networking and Collaborative Systems; 2020; Victoria p. 256-268. [doi: 10.1007/978-3-030-57796-4_25]
51. Silva M, Ceschin F, Shrestha P, Brant C, Fernandes J, Silva CS, et al. Predicting misinformation and engagement in COVID-19 Twitter discourse in the first months of the outbreak. Association for Computing Machinery 2020 Nov;37(4):1-24 [FREE Full text]
52. Maynard D, Greenwood M. Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. 2014 Presented at: Ninth International Conference on Language Resources and Evaluation (LREC'14); May 2014; Reykjavik, Iceland URL: http://www.lrec-conf.org/proceedings/lrec2014/index.html
53. Arif N, Al-Jefri M, Bizzi I, Perano G, Goldman M, Haq I, et al. Fake news or weak science? Visibility and characterization of antivaccine webpages returned by Google in different languages and countries. Front Immunol 2018 Jun 5;9:1215 [FREE Full text] [doi: 10.3389/fimmu.2018.01215]
54. Paschen J. Investigating the emotional appeal of fake news using artificial intelligence and human contributions. J Prod Brand Manag 2019 May 06;29(2):223-233. [doi: 10.1108/jpbm-12-2018-2179]
55. Dey A, Rafi R, Hasan PS, Arko S, Chakrabarty A. Fake news pattern recognition using linguistic analysis. 2018 Presented at: Joint 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2nd International Conference on Imaging, Vision Pattern Recognition (IcIVPR); June 25-29, 2018; Kitakyushu, Japan. [doi: 10.1109/iciev.2018.8641018]
56. Alonso MA, Vilares D, Gómez-Rodríguez C, Vilares J. Sentiment analysis for fake news detection. Electronics 2021 Jun 05;10(11):1348. [doi: 10.3390/electronics10111348]
57. Ajao O, Bhowmik D, Zargari S. Sentiment aware fake news detection on online social networks. 2019 Presented at: ICASSP 2019: 2019 IEEE International Conference on Acoustics, Speech and Signal Processing; May 12-May 17, 2019; Brighton, UK p. 2507-2511. [doi: 10.1109/icassp.2019.8683170]
58. Lee T. The global rise of “fake news” and the threat to democratic elections in the USA. Public Adm Policy 2019 Jul 01;22(1):15-24. [doi: 10.1108/pap-04-2019-0008]
59. Chen E, Chang H, Rao A, Lerman K, Cowan G, Ferrara E. COVID-19 misinformation and the 2020 U.S. presidential election. HKS Misinformation Review. 2021 Mar 3. URL: https://misinforeview.hks.harvard.edu/article/covid-19-misinformation-and-the-2020-u-s-presidential-election/ [accessed 2022-09-13]
60. Agarwal I, Rana D, Panwala K, Shah R, Kathiriya V. Analysis of contextual features’ granularity for fake news detection. Research Square Preprint posted online May 27, 2022 . [doi: 10.21203/rs.3.rs-1676328/v1]
61. Sekulić I, Gjurković M, Šnajder J. Not just depressed: bipolar disorder prediction on reddit. aclanthology Preprint posted online October, 2018 . [doi: 10.18653/v1/w18-6211]
62. Deb K, De H, Chatterjee S, Pal A. Studying borderline personality disorder using machine learning. 2022 Presented at: 16th International Conference on Ubiquitous Information Management and Communication (IMCOM); January 3-5, 2022; Seoul, South Korea. [doi: 10.1109/imcom53663.2022.9721800]
63. Jubair F, Salim N, Al-Karadsheh O, Hassona Y, Saifan R, Abdel-Majeed M. Sentiment analysis for Twitter chatter during the early outbreak period of COVID-19. 2021 Presented at: 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI); December 12-17, 2021; Yogyakarta, Indonesia. [doi: 10.1109/isriti54043.2021.9702837]
64. Santos R, Pedro G, Leal S. Measuring the impact of readability features in fake news detection. 2020 Presented at: 12th Language Resources and Evaluation Conference; May 11-16, 2020; Marseille, France p. 1404-1413 URL: https://aclanthology.org/2020.lrec-1.176
65. Zhou X, Jain A, Phoha VV, Zafarani R. Fake news early detection: a theory-driven model. Digit Threats Res Pract 2020 Jun 30;1(2):1-25. [doi: 10.1145/3377478]
66. Raj C, Meel P. ARCNN framework for multimodal infodemic detection. Neural Netw 2022 Feb;146:36-68. [doi: 10.1016/j.neunet.2021.11.006] [Medline: 34839091]
67. Dementieva D, Panchenko A. Cross-lingual evidence improves monolingual fake news detection. 2021 Presented at: 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop; August 2021; Online p. 310-320. [doi: 10.18653/v1/2021.acl-srw.32]
68. Wahle J, Ashok N, Raus T, Meuschke N, Ghosal T, Gipp B. Testing the generalization of neural language models for COVID-19 misinformation detection. In: Information for a Better World: Shaping the Global Future. iConference 2022. Lecture Notes in Computer Science, Vol 13192. Cham: Springer; 2022 Presented at: Information for a Better World: Shaping the Global Future. iConference 2022; 2022; Virtual p. 381-392. [doi: 10.1007/978-3-030-96957-8_33]

XSL•FO
RenderX

69. Colas F, Brazdil P. Comparison of SVM and some older classification algorithms in text classification tasks. 2006 Presented at: Artificial Intelligence in Theory and Practice, IFIP 19th World Computer Congress, TC 12: IFIP AI 2006 Stream; August 21-24, 2006; Santiago, Chile p. 169-178. [doi: 10.1007/978-0-387-34747-9_18]
70. Sagi O, Rokach L. Ensemble learning: a survey. WIREs Data Mining Knowl Discov 2018 Feb 27;8(4):e1249. [doi: 10.1002/widm.1249]
71. Samek W, Wiegand T, Müller K. Explainable artificial intelligence: understanding, visualizing and interpreting deep learning models. arXiv Preprint posted online August 28, 2017 [FREE Full text]
72. Roth Y, Pickles N. Updating Our Approach to Misleading Information. URL: https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information [accessed 2021-07-03]
73. Meta Business Help Centre. About Fact-Checking on Facebook. URL: https://www.facebook.com/business/help/2593586717571940 [accessed 2022-09-13]
74. Nakov P, Corney D, Hasanain M. Automated fact-checking for assisting human fact-checkers. arXiv Preprint posted online March 13, 2021 [FREE Full text] [doi: 10.24963/ijcai.2021/619]
75. Bot Sentinel. More than just bots. URL: https://botsentinel.com/info/about [accessed 2022-09-13]

Abbreviations

ARCNN: Allied Recurrent and Convolutional Neural Network
ARI: automated readability index
AUC: area under the curve
BiGRU: bidirectional gated recurrent unit
BiLSTM: bidirectional long short-term memory
CNN: convolutional neural network
FN: false negative
FP: false positive
GloVE: Global Vectors for Word Representation
GRU: gated recurrent unit
KNN: K-nearest neighbor
LR: logistic regression
LSTM: long short-term memory
ML: machine learning
NB: naive Bayes
NLP: natural language processing
SVM: support vector machine
TN: true negative
TP: true positive
VADER: Valence Aware Dictionary and sEntiment Reasoner
XGBoost: extreme gradient boosting

Edited by T Mackey; submitted 19.04.22; peer-reviewed by W Ceron, SF Tsao; comments to author 09.06.22; revised version received 25.08.22; accepted 10.09.22; published 22.09.22.

Please cite as:
Zhan K, Li Y, Osmani R, Wang X, Cao B
Data Exploration and Classification of News Article Reliability: Deep Learning Study
JMIR Infodemiology 2022;2(2):e38839
URL: https://infodemiology.jmir.org/2022/2/e38839
doi:10.2196/38839
PMID:36193330

©Kevin Zhan, Yutong Li, Rafay Osmani, Xiaoyu Wang, Bo Cao. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 22.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Role of Information Boxes in Search Engine Results for Symptom Searches: Analysis of Archival Data

Lorien C Abroms[1], PhD; Elad Yom-Tov[2], PhD

[1]Milken Institute School of Public Health, George Washington University, Washington, DC, United States

[2]Microsoft Research Israel, Hertzliya, Israel

Corresponding Author:
Elad Yom-Tov, PhD
Microsoft Research Israel
Alan Turing 3
Hertzliya, 4672415
Israel
Phone: 972 779391359
Email: eladyt@yahoo.com

Abstract

Background: Search engines provide health information boxes as part of search results to address information gaps and misinformation for commonly searched symptoms. Few prior studies have sought to understand how individuals who are seeking information about health symptoms navigate different types of page elements on search engine results pages, including health information boxes.

Objective: Using real-world search engine data, this study sought to investigate how users searching for common health-related symptoms with Bing interacted with health information boxes (info boxes) and other page elements.

Methods: A sample of searches (N=28,552 unique searches) was compiled for the 17 most common medical symptoms queried on Microsoft Bing by users in the United States between September and November 2019. The association between the page elements that users saw, their characteristics, and the time spent on elements or clicks was investigated using linear and logistic regression.

Results: The number of searches ranged by symptom type from 55 searches for cramps to 7459 searches for anxiety. Users searching for common health-related symptoms saw pages with standard web results (n=24,034, 84%), itemized web results (n=23,354, 82%), ads (n=13,171, 46%), and info boxes (n=18,215, 64%). Users spent on average 22 (SD 26) seconds on the search engine results page. Users who saw all page elements spent 25% (7.1 s) of their time on the info box, 23% (6.1 s) on standard web results, 20% (5.7 s) on ads, and 10% (10 s) on itemized web results, with significantly more time on the info box compared to other elements and the least amount of time on itemized web results. Info box characteristics such as reading ease and appearance of related conditions were associated with longer time on the info box. Although none of the info box characteristics were associated with clicks on standard web results, info box characteristics such as reading ease and related searches were negatively correlated with clicks on ads.

Conclusions: Info boxes were attended most by users compared with other page elements, and their characteristics may influence future web searching. Future studies are needed that further explore the utility of info boxes and their influence on real-world health-seeking behaviors.

(JMIR Infodemiology 2022;2(2):e37286) doi:10.2196/37286

KEYWORDS

health misinformation; search engine; internet search; information boxes; knowledge graph boxes; misinformation; health information; Microsoft; internet; data; symptoms; results; users; medical; Bing; USA; linear; logistic; regression; web; ads; behavior

XSL•FO
RenderX

Introduction

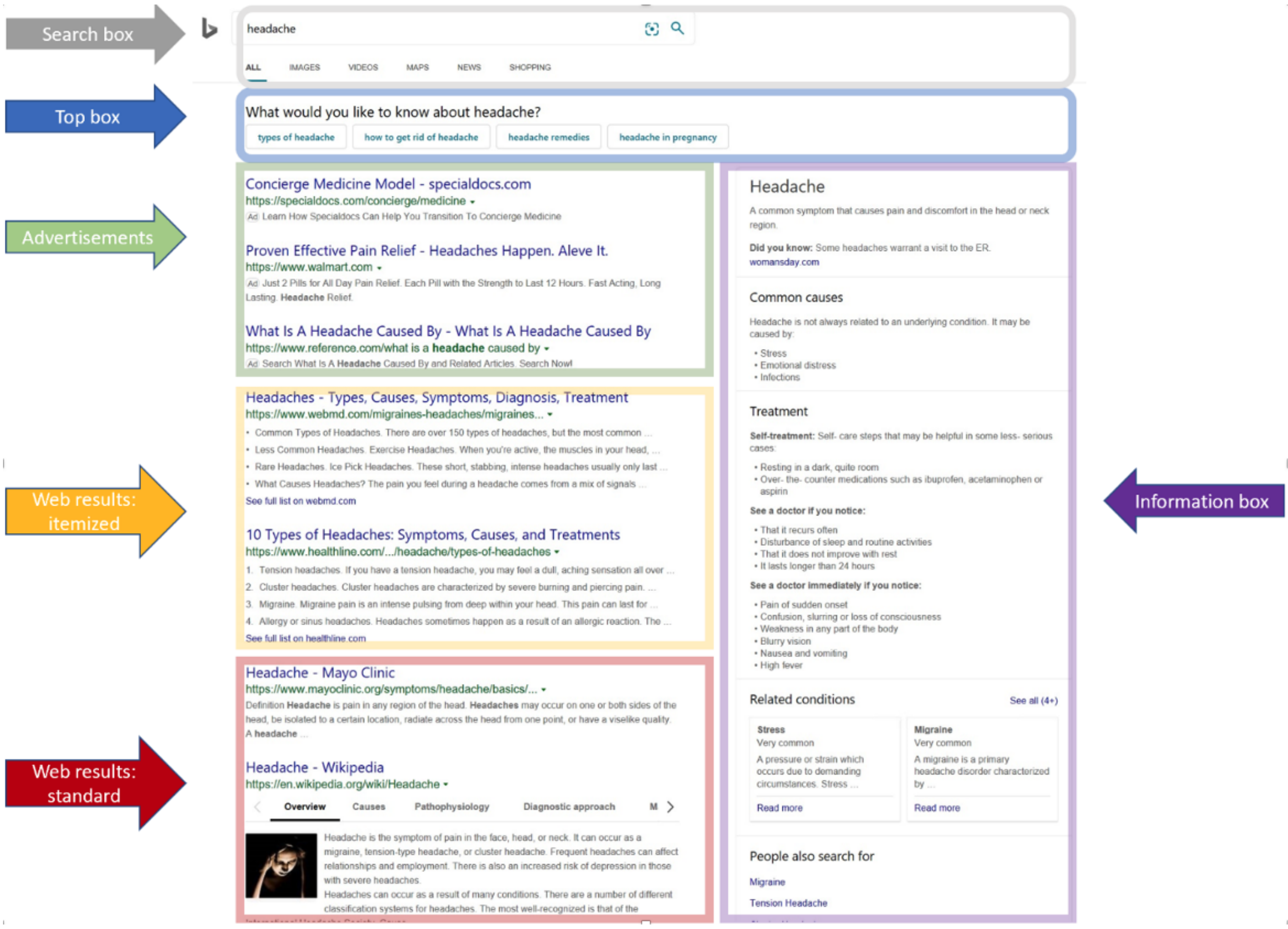
A general-purpose internet search engine is the first stop for most people who experience a health symptom and are seeking information about it [1-3]. The search engine results page (SERP), provided by search engines, generally includes a variety of page elements. These include the standard search results with a URL and a summary or snippet. Additionally, the search results page may include other page elements such as videos, advertisements, recent news stories, and in the case of health-related searches, a health information box [4,5].

Health information boxes (info boxes), also known as health knowledge graph boxes, information cards, or task panes, were created at major search engines about 10 years ago—at Bing in 2010 and at Google in 2012 [6]. They were developed to address health information gaps and misinformation for commonly searched symptoms that might arise from search results alone [4]. Info boxes are typically presented in the right-hand side of a SERP and are available in addition to what is available from the standard search results (as seen in Figure 1). Info boxes could balance the information presented in search results that might otherwise lead a user to, for example, overworry about a symptom (eg, headache) based on standard search results alone [7]. The information in info boxes is provided by the search engine from sources they deem trustworthy (eg, Mayo Clinic and Wikipedia) and may have additional reviews from an internal health team [7].

Few prior studies have sought to understand how individuals who are seeking health information navigate SERPs and their various page elements such as standard search results, ads, or videos (exceptions include [4,5,8]). However, understanding how users interact with page elements is a fundamental question in information retrieval, with implications for understanding search quality and interface design. In the case of symptom search, these have implications for health knowledge acquisition, methods of addressing information gaps and misinformation, as well as future health-seeking behaviors, potentially. Past studies have found that a search engine's sorting and ranking criteria can directly influence engagement, user effort, as well as health beliefs and attitudes [2,8]. The salience on search results page may also affect the decision to present to health services [2].

Despite their long-standing existence and ubiquity, only one study could be identified that had examined the role of info boxes. A study by Ludolph et al [4] found that experimentally developed and manipulated info boxes (termed knowledge graph boxes in the study), which were shown as part of a web-based survey, could positively affect a participants' vaccination-related knowledge and attitudes. No study that we are aware of has previously sought to understand the effects of info boxes using real-world data or in the context of health symptom searches.

Figure 1. Typical search engine results page for "headache" with multiple page elements displayed, including info box, ads, itemized web results, and standard web results.



XSL•FO
RenderX

This study sought to investigate how users searching for common health-related symptoms with Bing interacted with info boxes and other page elements using real-world data collected from anonymized Bing users. The research question under investigation was whether curated content on health symptoms as presented in the info boxes affected health-seeking behavior by Bing users, and to what extent info boxes and other page elements attended to and used in the SERPs were compared with other page elements.

Methods

We compiled a list of the 19 most common medical symptoms queried on Bing by users in the United States between September and November 2019 from a longer list of 195 symptoms originally compiled from Wikipedia in a prior study [9]. The list was refined to remove 2 items identified by our team as not being symptoms (ie, childbirth and weight loss). The remaining list was comprised of 17 symptoms as follows: anxiety, back pain, bleeding, constipation, cough, cramp, depression, diarrhea, fever, headache, itch, pain, paralysis, rash, wound, swelling, and tremor.

To obtain the sample of searches for these symptoms, deidentified data on symptom searches made on Bing in the United States during September 2019 were extracted. We also extracted information about the interaction of the users with the search results page (that is explained in more detail in the following paragraph). Our sample was comprised of a total of 33,872 searches for the 17 symptoms, encompassing 28,552 unique users. We limited our sample to the first search of users in order to have a sample where each search was independent. Thus, 28,552 searches were included for analysis and comprise the final sample. The distribution of symptoms among searches is shown in Table 1.

Search-related information on each user included the following: page elements shown to the user on the SERP, clicks on any of the displayed elements on the SERP, and the time that the mouse pointer spent on each of the elements, previously shown to be a marker for attention [10].

Table 1. The number of searches per symptom in the analyzed data set.

Symptoms	Number of searches
Anxiety	6686
Back pain	1576
Bleeding	136
Constipation	2558
Cough	674
Cramp	100
Depression	5485
Diarrhea	5784
Fever	899
Headache	682
Itch	395
Pain	1371
Paralysis	295
Rash	927
Swelling	236
Tremor	345
Wound	403
Total	28,552

A typical SERP had the following page elements: advertisements, which are created by external parties who pay whenever they are clicked; a health information box or “info box,” which contains Bing-curated health information; and two areas where algorithmic search answers are shown—one containing standard web results described by several sentences of text or snippets and the other containing an itemized web result with a summary of information. The standard web results are ranked based on the result predicted to be most relevant (rank=1 for highest position on the page) to least relevant (rank=8), though not all results may be displayed on the first page.

In addition, a typical SERP may have a top box, which helps disambiguate the user’s intent by offering more focused search options or providing information such as dictionary definitions. Other elements that are sometimes displayed on the page include video results and news. However, as data for these elements (ie, top box, video elements, and news) were not readily available for extraction, these page elements were excluded from our analysis. Thus, we restricted our analysis to the

XSL•FO
RenderX

following page elements: ads, info boxes, standard web results, and itemized web results. It is noteworthy that not all page elements are shown to each user for a given search, and results displayed may depend on factors such as the size of their browser window. Figure 1 shows a sample symptom SERP, displaying its different page elements.

We manually coded the characteristics of the info boxes associated with these commonly searched symptoms. In order to display the info boxes for coding, the symptom name was typed into the Bing search engine with a fresh private (ie, incognito) window using the Microsoft Explorer web browser. Info boxes were coded for reading ease (using the Flesch Reading Ease score, with higher scores on a scale of 0-100 indicating greater ease of reading). They were also coded for whether the info box shows related searches (eg, common causes and treatment) or provides information on related conditions.

In addition, we manually coded the characteristics of ads and the standard search results. For this, the 20 most commonly displayed ads and search results associated with each of the symptoms in September 2019 were identified and manually coded by a single coder. Ads and search results were scored for reading level (using the Flesch Reading Ease score) and coded for the type of information offered (eg, informational or product advertisement). A random subsample of 50 ads and 50 web results were independently coded by a second coder for type of information—the most subjective of the codes.

User engagement with the elements on the page was measured as the time spent on each of the page elements (eg, ads, info box, itemized web results, and standard web results) and whether itemized web results and standard web results were clicked. Times were measured by monitoring whether the mouse pointer of the user was hovering over an element [10]. The total time on a page included the entire time that the user spent with a search result, including any returns to it following a visit to one of the search results.

Descriptive statistics were tabulated for engagement metrics (eg, seconds on page elements and clicks on page elements). Linear regression was used to analyze the correlation between the time spent on different elements of the page, as a function of the elements shown on the page. Logistic regression was used to analyze the association between page characteristics, info box characteristics, the characteristics of standard web results, and (separately) those of ads on clicks on standard web results or ads. This analysis was conducted at the level of a standard web result or ad. We did not analyze clicks on itemized web results, as clicks on them were the least common. We did not analyze clicks on info boxes, as many of them did not have links, and therefore, clicks were rare.

Results

For the subsample analyzed, kappa statistics for the agreement between the coders was generally good for the type of information in ads (κ=0.60) and standard web results (κ=0.44). Among the 28,552 symptom searches of unique individuals analyzed, the number of searches ranged by symptom type from 55 searches for cramps to 7459 searches for anxiety. Users searching for symptoms encountered SERPs with multiple page elements, including standard web results (n=24,034, 84%), itemized web results (n=23,354, 82%), ads (n=13,171, 46%), and info boxes (n=18,215, 64%; Table 2).

When all the 4 elements of the page (ie, info box, ads, itemized web results, and standard web results) were shown to users, 41% (2039) of them went on to click on some elements in the SERP, with the remainder not clicking on anything. Users clicked on standard web results most often (ie, n=4612, 19%), and they clicked on ads 12% (n=1633) of the time. They clicked on itemized web results least often (n=1798, 8%).

On average, users spent 22 (SD=26) seconds on the SERP once the results were shown to them, with 24% (n=1182) spending 30 seconds or more on it. As Table 2 demonstrates, users who saw all page elements spent 25% (7.1 s) of their time on the info box, 23% (6.1 s) on standard web results, 20% (5.7 s) on ads, and 10% (10 s) on itemized web results, with significantly more time on info boxes compared to other elements and the least amount of time on itemized web results (sign test; all pairwise comparisons are statistically significant; P<.001).

Based on manual coding, the info boxes were found to have the following characteristics: the average Flesch Reading Ease score of info boxes was 46 (SD 17; range 6-69); common causes and treatment of the symptom were shown in 76% (n=13) of the info boxes; the info boxes contained a list of related conditions in 71% (n=12) of the cases, and related searches were shown for all but one symptom (diarrhea); the most common data source (as stated in the info boxes) for the information in the info boxes was Focus Medica (n=14), with the remainder citing Wikipedia as their data source (n=3).

The time spent on the info box was modeled using linear regression, as a function of the coded characteristics of the info box. The model fit was R^2=0.016 (P<.001; n=17,255), meaning that the characteristics of the info box (eg, reading ease, showing related conditions, and showing related searches) is associated with time spent on the info box, but the characteristics explain only a small amount of the variance in time. That said, the appearance of related conditions and ease of reading were significantly associated with longer time, whereas related searches were correlated with shorter time.

Table 3 shows a model of the time spent on an element (in seconds), as a function of whether the other elements of the page were visible. As the model shows, there is a weak correlation between the time spent and the visibility of other elements. Longer time spent on the info box is most strongly associated with the display of itemized web results and ads. Longer time spent on itemized web results is most strongly associated with the display of ads and info box and negatively correlated with the display of itemized web results. Longer time spent on standard web results is associated most strongly with the display of itemized web results and ads.

Table 4 shows logistic regression models for predicting clicks on individual standard web results and ads, taking into account the characteristics of the info box, the characteristics of the page, and those of the standard web result, or the ad. Characteristics of the info boxes include the following: whether

XSL•FO
RenderX

they were shown; and if so, whether related conditions were shown; whether related searches were shown; and the reading ease of these boxes. Characteristics of the page include the time the mouse pointer hovers over ads, info boxes, and the two types of web links, as well as the number of elements in each type. The attributes of the web results and ads include whether they were informational or advertisements, their reading ease, the rank at which they were shown on the page (1 being the highest rank), and the time that the mouse pointer hovered over the link.

As can be seen in Table 4, when an info box was shown, an info box showing related conditions was associated with higher likelihood of clicks on ads. Related searches and reading ease were negatively correlated with clicks on ads. None of the parameters of info boxes were associated with clicks on standard web results.

Being shown more ads was associated with more clicks on ads, but it was unrelated to clicking on standard web results, while more standard web results shown were associated with fewer clicks on those results or ads. Standard web results with informational content were less likely to be clicked.

Table 2. Statistics of page elements during symptom searches (N=28,552).

Page elements	Visible to the user, n (%)	Clicks on visible elements, n (%)	Time spent (when all elements are shown), seconds (%)
Ads	13,171 (46)	1633 (12)	5.7 (20)
Info box	18,215 (64)	N/A[a]	6.1 (25)
Itemized web results	23,354 (82)	1798 (8)	2.7 (10)
Standard web results	24,034 (84)	4612 (19)	7.1 (23)

[a]N/A: not applicable. Info boxes are not usually clicked, and therefore, this number is not given.

Table 3. Model for predicting time spent on different elements of the page, as a function of the elements shown on the page. Numbers shown are model slopes.

Page elements	Model R^2	Elements shown			
		Ads	Info boxes	Itemized web results	Standard web results
Ads	0.001	—[a]	–0.020[b]	–0.700	–0.012[b]
Info boxes	0.037	1.413	—	1.654	2.189
Itemized web results	0.009	0.495	0.734	—	–0.343
Standard web results	0.014	1.143	0.961	2.121	—

[a]Not applicable.

[b]Slopes that are not statistically significant (at $P<.05$, with Bonferroni correction).

XSL•FO
RenderX

Table 4. Logistic regression models of clicks on individual standard web results and on individual ads in cases where the information box (info box) was shown.

Characteristics	Standard web results (n=23,776), OR[a] (95% CI)	Ads (n=16,667), OR (95% CI)
Info box		
Info box shows related conditions	1.281 (1.053-1.557)	1.331[b] (1.107-1.599)
Info box shows related searches	0.985 (0.839-1.156)	0.634[b] (0.559-0.718)
Info box's reading ease	0.997 (0.994-1.000)	0.996[b] (0.994-0.998)
Page		
Ad's rank	1.007 (0.983-1.031)	0.988 (0.973-1.005)
Number of ads shown	1.015 (0.998-1.032)	1.058[b] (1.031-1.086)
Number of itemized web results shown	0.989 (0.956-1.023)	0.928 (0.874-0.984)
Number of standard web results shown	0.901[b] (0.884-0.918)	0.881[b] (0.862-0.901)
Time spent on ads	1.009[b] (1.005-1.013)	0.999 (0.994-1.003)
Time spent on info boxes	0.995 (0.989-1.000)	0.996 (0.988-1.003)
Time spent on itemized web results	0.997 (0.990-1.004)	1.007 (0.999-1.016)
Time spent on standard web results	0.992 (0.986-0.997)	1.004 (0.999-1.009)
Standard web result or ad		
Type of information (informational)	0.789[b] (0.693-0.899)	0.905 (0.785-1.044)
Reading ease of elements	1.000 (0.999-1.002)	1.002 (1.000-1.004)
Time spent on standard web results or ads	1.023 (1.005-1.041)	1.005 (0.999-1.009)

[a]OR: odds ratio.

[b]Ratios are statistically significant (at *P*<.05, with Bonferroni correction).

Discussion

Principal Findings

For people experiencing health symptoms, search engines have become a dominant way of initially making sense of that experience [1,3]. As such, understanding how individuals who are seeking information about health symptoms navigate different types of page elements, including info boxes, on SERPs is paramount.

This study of 28,552 unique Bing users searching for 17 common symptoms found that users searched most often for information on anxiety and least often for information on cramps. In doing those searches, users spent on average 22 seconds observing the SERP and encountered SERPs with a complex mix of ads, standard web results, itemized web results, and info boxes. Standard web results and itemized web results were most common in SERPs, and ads and info boxes were present fairly less often, about half of the time. The variation observed in what users saw was likely because of their specific search, the device they used to browse them (eg, screen size differences, with smaller screens displaying fewer content elements), and user behavior, in cases where the user did not scroll down to the location of that element.

When all page elements of the SERP were visible (ie, info box, ads, standard web results, and itemized web results), users spent the most time observing info boxes. This represents the first real-world evidence that info boxes are serving the purpose that they were designed to do, namely, presenting health information in a more user-friendly format compared to standard web results. Users may prefer info boxes over other types of SERP elements because they simplify the information and manage information overload.

Furthermore, info box characteristics were found to be associated with a decreased likelihood of clicking on ads, but they had no effect on standard web results. This implies that a well-designed info box—one that is higher on reading ease and shows related searches—may reduce the likelihood that those searching for health symptom information will be steered to commercial websites. As such, designers of info boxes may wish to carefully consider their design elements and ensure that the reading level is as low as possible. Furthermore, given their importance, search engine companies may wish to pretest their content with users or test out variations in order to optimize them.

The strength of this study is that it provides the first real-world data on symptom searches on search engines, and how users interact with info boxes. It includes real-world stimuli and data from real users searching on Bing. As this study is the first of its type, future studies are needed to confirm these findings as well as take them further by examining the real-world implications of SERPs for symptom searches. For example, studies could examine how info boxes affect future

XSL•FO
RenderX

decision-making about whether to seek out medical care or pursue various treatment options.

Weaknesses of this study include the following: although we were able to examine multiple page elements from SERPs, we were not able to access the types of page elements presented to users; For example, we did not have access to data on top boxes that simplify search or in videos shown to users; future studies should strive to include these other data types. Additionally, the list of 17 symptoms investigated was generated from a longer list of 195 symptoms compiled by Wikipedia, which may be less reliable than other types of data on symptoms, such as population-level survey data.

Conclusions

SERPs for symptom searches often include info boxes that are attended to by users. Info box characteristics may influence future web searching. Future studies are needed to further explore the utility of info boxes, how to optimize them, and their influence on real-world treatment-seeking behaviors.

Acknowledgments

The authors would like to thank Megan Hatheway for her help with manually coding search results.

Conflicts of Interest

EYT is an employee of Microsoft, owner of Bing. LCA declares no conflicts of interest.

References

1. Hochberg I, Allon R, Yom-Tov E. Assessment of the frequency of online searches for symptoms before diagnosis: analysis of archival data. J Med Internet Res 2020 Mar 06;22(3):e15065 [FREE Full text] [doi: 10.2196/15065] [Medline: 32141835]
2. Mueller J, Jay C, Harper S, Davies A, Vega J, Todd C. Web use for symptom appraisal of physical health conditions: a systematic review. J Med Internet Res 2017 Jun 13;19(6):e202 [FREE Full text] [doi: 10.2196/jmir.6755] [Medline: 28611017]
3. Fox S, Duggan M. Health online 2013. California Health Center Foundation. URL: https://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/PIP_HealthOnline.pdf [accessed 2022-09-01]
4. Ludolph R, Allam A, Schulz PJ. Manipulating Google's knowledge graph box to counter biased information processing during an online search on vaccination: application of a technological debiasing strategy. J Med Internet Res 2016 Jun 02;18(6):e137 [FREE Full text] [doi: 10.2196/jmir.5430] [Medline: 27255736]
5. Wu Z, Sanderson M, Cambazoglu BB, Croft W, Scholer F. Providing direct answers in search results: a study of user behavior. 2020 Presented at: CIKM '20; Oct 19-23; Virtual Event Ireland p. 1635-1644. [doi: 10.1145/3340531.3412017]
6. Dormel L. Google search pages now highlight health information. Fast Company. URL: https://www.fastcompany.com/3042225/google-search-pages-now-highlight-health-information [accessed 2022-02-01]
7. White RW, Horvitz E. Cyberchondria. ACM Trans Inf Syst 2009 Nov 01;27(4):1-37. [doi: 10.1145/1629096.1629101]
8. Allam A, Schulz PJ, Nakamoto K. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating Google output. J Med Internet Res 2014 Apr;16(4):e100 [FREE Full text] [doi: 10.2196/jmir.2642] [Medline: 24694866]
9. Yom-Tov E, Gabrilovich E. Postmarket drug surveillance without trial costs: discovery of adverse drug reactions through large-scale analysis of web search queries. J Med Internet Res 2013 Jun 18;15(6):e124 [FREE Full text] [doi: 10.2196/jmir.2614] [Medline: 23778053]
10. Youngman B, Yom-Tov E. Anxiety and information seeking: evidence from large-scale mouse tracking. In: Proceedings of the 2018 World Wide Web Conference. 2018 Presented at: WWW '18; April 23-27, 2018; Lyon, France p. 753-762. [doi: 10.1145/3178876.3186156]

Abbreviations

SERP: search engine results page

XSL•FO
RenderX

Edited by C Baur; submitted 14.02.22; peer-reviewed by K Ahmed, R Khajouei; comments to author 08.06.22; revised version received 12.07.22; accepted 20.08.22; published 14.09.22.

Please cite as:
Abroms LC, Yom-Tov E
The Role of Information Boxes in Search Engine Results for Symptom Searches: Analysis of Archival Data
JMIR Infodemiology 2022;2(2):e37286
URL: https://infodemiology.jmir.org/2022/2/e37286
doi:10.2196/37286
PMID:37113445

©Lorien C Abroms, Elad Yom-Tov. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 14.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

COVID-19–Related Health Inequalities Induced by the Use of Social Media: Systematic Review

Yi Shan[1], Prof Dr; Meng Ji[2], PhD; Wenxiu Xie[3], MPhil; Xiaomin Zhang[4], PhD; Harrison Ng Chok[5], PhD; Rongying Li[6], BM; Xiaobo Qian[6], BEng; Kam-Yiu Lam[3], PhD; Chi-Yin Chow[3], PhD; Tianyong Hao[6], Prof Dr

[1]Nantong University, Nantong, China
[2]University of Sydney, Sydney, Australia
[3]City University of Hong Kong, Hong Kong, China
[4]Macquarie University, Sydney, Australia
[5]Western Sydney Local Health District, Parramatta, Australia
[6]South China Normal University, Guangzhou, China

Corresponding Author:
Yi Shan, Prof Dr
Nantong University
No. 9, Seyuan Rd., Nantong University
Nantong, 226019
China
Phone: 86 15558121896
Email: victorsyhz@hotmail.com

Abstract

Background: COVID-19–related health inequalities were reported in some studies, showing the failure in public health and communication. Studies investigating the contexts and causes of these inequalities pointed to the contribution of communication inequality or poor health literacy and information access to engagement with health care services. However, no study exclusively dealt with health inequalities induced by the use of social media during COVID-19.

Objective: This review aimed to identify and summarize COVID-19–related health inequalities induced by the use of social media and the associated contributing factors and to characterize the relationship between the use of social media and health disparities during the COVID-19 pandemic.

Methods: A systematic review was conducted on this topic in light of the protocol of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement. Keyword searches were performed to collect papers relevant to this topic in multiple databases: PubMed (which includes MEDLINE [Ovid] and other subdatabases), ProQuest (which includes APA PsycINFO, Biological Science Collection, and others), ACM Digital Library, and Web of Science, without any year restriction. Of the 670 retrieved publications, 10 were initially selected based on the predefined selection criteria. These 10 articles were then subjected to quality analysis before being analyzed in the final synthesis and discussion.

Results: Of the 10 articles, 1 was further removed for not meeting the quality assessment criteria. Finally, 9 articles were found to be eligible and selected for this review. We derived the characteristics of these studies in terms of publication years, journals, study locations, locations of study participants, study design, sample size, participant characteristics, and potential risk of bias, and the main results of these studies in terms of the types of social media, social media use–induced health inequalities, associated factors, and proposed resolutions. On the basis of the thematic synthesis of these extracted data, we derived 4 analytic themes, namely health information inaccessibility–induced health inequalities and proposed resolutions, misinformation-induced health inequalities and proposed resolutions, disproportionate attention to COVID-19 information and proposed resolutions, and higher odds of social media–induced psychological distress and proposed resolutions.

Conclusions: This paper was the first systematic review on this topic. Our findings highlighted the great value of studying the COVID-19–related health knowledge gap, the digital technology–induced unequal distribution of health information, and the resulting health inequalities, thereby providing empirical evidence for understanding the relationship between social media use and health inequalities in the context of COVID-19 and suggesting practical solutions to such disparities. Researchers, social media, health practitioners, and policy makers can draw on these findings to promote health equality while minimizing social media use–induced health inequalities.

XSL•FO
RenderX

(JMIR Infodemiology 2022;2(2):e38453) doi:10.2196/38453

KEYWORDS

systematic review; social media use; health inequalities; COVID-19; mobile phone

Introduction

Background

Currently, the focus of web use has shifted from primarily unidirectional information-seeking to web-based interaction, information sharing, and collaboration [1]. “The increased use of Web 2.0...provides potential opportunities to engage people in health-related issues, stimulate an active role in their health care, connect them with others and evidence-based interventions, and create social action focused on the social determinants of health disparities,” thereby offering underserved and underrepresented populations potential access to essential health information resources and social support for addressing health care issues [2]. *Social media* and *social networking* started being increasingly used to depict the intrinsic characteristics of tools, apps, and functions on Web 2.0 [2]. Compared with traditional media (eg, newspapers, magazines, television, and radio), social media offer easy access to information that can be distributed to larger audiences more rapidly and cost-effectively [3,4]. “The rapid adoption of the Internet and computing technologies by all sectors of modern society has made them an indispensable part of our daily work and life” [5]. Popular social media platforms (eg, Facebook, Twitter, and web-based health community forums) have been applied by health service providers to promote health and facilitate community engagement [6-9]. Social media has been widely and frequently adopted to disseminate information, especially during a crisis or emergency [10]. Ever since the outbreak of COVID-19, diversified social media platforms have been serving as prioritized resorts to publicize COVID-19–related information to the public worldwide owing to the vast number of users [4,10].

Social media can enhance target populations’ access to health services and facilitate information flow and service uptake, but little agreement has been reached on the best practices of social media [8,11] because social media are by no means problem free [10]. The first concern relates to equal access to social media. Given that social media require the use of smart devices, such as smartphones, computers, and laptops, to access the internet, a barrier is imposed on those unable to access these devices. Even among those with such access, the differences in language and computer literacy cause disparities in the quantity and quality of information they receive [5]. Besides, the lack of gatekeeping in social media, the immediate communication of scientific information from discovery to dissemination without calibration, and the public’s nonscientific background have all caused the generation and spread of misinformation, especially during the pandemic [12], posing a great threat to people’s health because preventive and protective practices were compromised by such misinformation.

COVID-19–related health inequalities were reported in a recent study [13], showing the failure in public health and communication. “Health inequality is the generic term used to designate differences, variations, and disparities in the health achievements of individuals and groups” [14], closely associated with social, economic, and environmental disadvantages [15]. Inequalities in health care service access have been well documented [16,17]. Studies investigating the contexts and causes of these disparities pointed to the contribution made by communication inequality or poor health literacy and information access to engagement with health care services [18-20]. However, no study exclusively dealt with health inequalities induced by the use of social media.

Objective

The objective of this review was two-fold: (1) to identify and summarize COVID-19–related health inequalities induced by the use of social media and the associated contributing factors and (2) to characterize the relationship between the use of social media and health inequalities during the COVID-19 pandemic. This review can thus inform researchers, social media and health practitioners, and policy makers, who can therefore make joint efforts to take full advantage of social media to promote health equality while minimizing social media use–induced health inequalities [21].

Methods

Overview

This review was conducted and reported in light of the protocol of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [22]. The methods of the review process and the selection criteria were predefined.

Literature Search

The Medical Subject Headings terms we used for this study were “social media,” “COVID-19,” “SARS-CoV-2,” and “coronavirus.” The keyword search strategy was “(social media [Title/Abstract]) AND (COVID-19* [Title/Abstract] OR SARS-CoV-2 [Title/Abstract] OR coronavirus [Title/Abstract]) AND (equal* [Title/Abstract] OR inequal* [Title/Abstract]).” On March 27, 2022, we conducted keyword searches to retrieve articles concerned with health inequalities induced by social media–related factors in multiple databases: PubMed (which includes MEDLINE [Ovid] and other subdatabases), ProQuest (which includes APA PsycINFO, Biological Science Collection, and others), ACM Digital Library, and Web of Science, without any year restriction. In total, 670 publications were retrieved. Among them, 442, including duplicates, other document types, and non-English papers, were first removed. The articles analyzed and synthesized in this review were selected from the remaining 228 articles based on the predefined selection criteria.

Selection Criteria

Publication Information

No limit was put on the publication year in the keyword searches for relevant literature. No restriction was imposed on the age

XSL•FO
RenderX

of the target populations. The selected articles had to be written in English. The articles needed to be research papers published in journals or presented at conferences. Other document types (eg, reviews, abstracts, editorials, workshop summaries, perspectives, opinions, diagnosis methods, and study protocols) were excluded [23]. Studies undertaken in any country were considered.

Population

The target population was any group in the public worldwide who experienced social media use–induced health inequalities during the repeated resurgences of COVID-19.

Health Inequalities

The health inequalities discussed and summarized in this review could be any aspect related to any health issues, mental or physical. The health inequalities could be experienced anywhere worldwide, so long as they were induced by social media use and related to COVID-19. All studies satisfying these inclusion criteria were selected for the review.

Social Media

Social media under discussion in this review referred to ways of sharing information, opinions, images, videos, etc, using the internet, especially social networking sites, including WeChat, WhatsApp, Facebook, Twitter, web-based health community forums, etc.

Comparator

The comparator could be any form of health inequalities induced by social media. Publications with no comparison were also included because the aim of this review was not to determine the relative degrees of social media–induced health inequalities but to scrutinize the current status of social media–induced health inequalities experienced by people worldwide during the COVID-19 pandemic.

Outcomes

The outcomes of the selected studies we considered were as follows: participants' physical and mental health inequalities induced by the use of social media and the associated contributing factors.

Study Design

The designs of eligible studies were quantitative, qualitative, or mixed methods approaches adopted for investigating the outcomes mentioned above. Pilot studies and case studies were included because both types of studies could shed light on the study outcomes above.

Study Selection

Microsoft Excel was used to manage the retrieved articles and collect data from them. The selection of eligible studies was performed in 3 rounds. In the first round, duplicates, non-English articles, and other document types were all excluded. In the second round, 6 reviewers (YS, XQ, RL, YC, XW, and TS) reviewed titles and abstracts independently against the selection criteria. Any disagreements were settled through discussion among these reviewers and consultation with another 2 reviewers (MJ and WX). In the third round, 2 reviewers (MJ and YS) reviewed the full texts of the remaining articles to further identify eligible studies drawing on the selection criteria. The PRISMA flowchart of the screening and full-text review was produced by WX.

Quality Assessment

To verify the relevance and methodological solidity of the selected studies, we evaluated the study purpose, literature review, methodology, results obtained, risk of biases in terms of sampling, outcome measures, and conclusions of the selected studies using a modified version of the quality assessment scale adapted from a recent study [23], which was based on the critical review forms of Critical Review Form—Qualitative Studies and Critical Review Form—Quantitative Studies [24,25]. Specifically, 10 questions, presented in Textbox 1, were used to assess the quality of the selected studies. 1 and 0 meant a *yes* answer and a *no* answer to any of the 10 questions, respectively. The maximum quality score for each study was 10. Any study whose quality score was below 6 was excluded from the review.

Textbox 1. Quality assessment scale of the selected studies.

1. Was the purpose stated clearly?
2. Was relevant literature reviewed?
3. Was the sample described in detail?
4. Was the sample size justified?
5. Were the outcome measures reliable?
6. Was the intervention described in detail?
7. Were results reported in terms of statistical significance?
8. Were the analysis methods appropriate?
9. Was clinical importance reported?
10. Were conclusions appropriate given the study methods and results?

XSL•FO
RenderX

Data Extraction and Synthesis

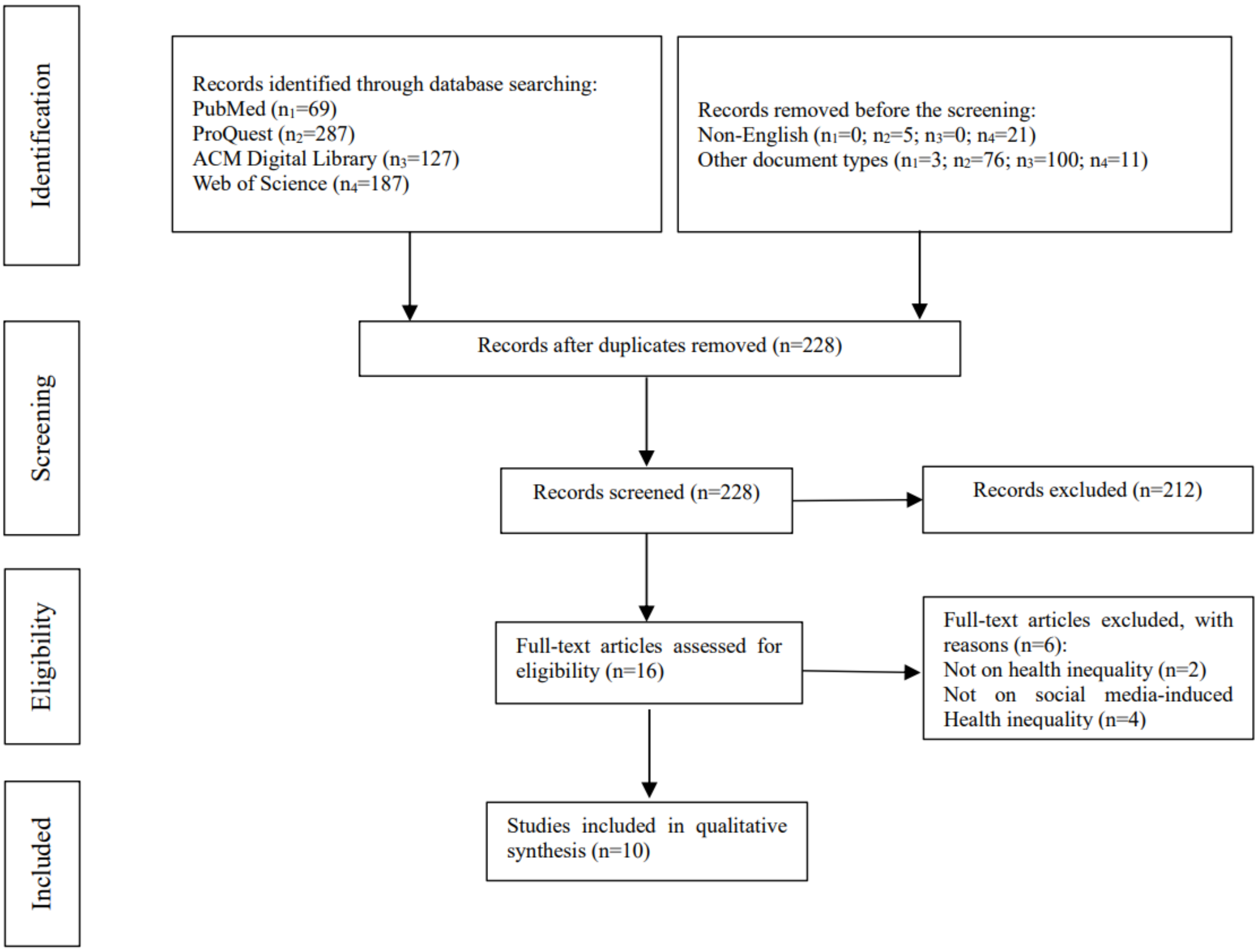
Two reviewers (MJ and YS) extracted data from the eligible articles meeting the quality standard by following a standardized form, in which data items included first author's name and reference, publication year, country, target population, sample size, study design, data collection methods, social media forms, types of health inequalities, social media–related factors for health inequalities, comparator (if applicable), and recommended resolutions.

Results

Study Selection

In the first round of selection, 442 articles (including 226 duplicates, 26 non-English articles, and 190 articles of other document types) were removed. In the second round, 212 articles were excluded because of the violation of at least 1 item in the selection criteria. In the third round, 6 articles were removed from the remaining 16 articles because they were not concerned with health inequalities (2/6, 33%) or social media use–induced health inequalities (4/6, 67%). Therefore, 10 studies were found eligible and subjected to quality assessment. The selection flowchart is shown in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the selection of eligible studies.



Qualitative Analysis

Table 1 shows that except for the study by Shaw [26], of the 10 included articles, 9 (90%) met the inclusion criteria in this systematic review. So, we finally chose these 9 qualified studies [3,27-34] for qualitative synthesis. The most prominent problem of these 9 selected studies was that they did not report the clinical importance (item 9). Besides, 44% (4/9) of studies failed to describe the interventions in detail (item 6) and report results in terms of statistical significance (item 7). According to Zhou and Parmanto [23], the cutoff score for any studies that were qualified for inclusion in a systematic review was 6 (out of a total score of 10). Although failing to meet some of the quality assessment criteria, the 9 studies [3,27-34] were finally included in the qualitative analysis for this systematic review.

XSL•FO
RenderX

Table 1. Quality assessment of eligible studies based on Textbox 1 (N=10).

Study	Items of quality assessment										Score, n (%)
	1	2	3	4	5	6	7	8	9	10	
Wang et al [3], 2021	Y[a]	Y	Y	Y	Y	Y	Y	Y	N[b]	Y	9 (90)
Dai et al [27], 2021	Y	Y	N	Y	Y	Y	N	Y	N	Y	7 (70)
Almusawi et al [28], 2021	Y	Y	Y	N	Y	Y	Y	Y	N	Y	8 (80)
Zeng et al [29], 2020	Y	Y	Y	N	Y	Y	Y	Y	N	Y	9 (90)
Gallagher et al [30], 2021	Y	Y	N	Y	Y	Y	N	Y	Y	Y	8 (90)
Blevins et al [31], 2021	Y	Y	N	Y	Y	N	N	Y	N	Y	6 (60)
Shaw [26], 2020	Y	Y	N	N	N	N	N	N	N	Y	3 (30)
Wade et al [32], 2021	Y	Y	N	Y	Y	N	N	Y	N	Y	6 (60)
Ambelu et al [33], 2021	Y	Y	Y	Y	Y	N	Y	Y	N	Y	8 (80)
Wagner et al [34], 2021	Y	Y	Y	N	Y	N	N	Y	N	Y	6 (60)

[a]Y: study meets the standard of an item of quality assessment.
[b]N: study does not meet the standard of an item of quality assessment.

Study Characteristics

Publication Years

Of the 9 finally selected papers, 8 (89%) studies [3,27,28,30-34] were published in 2021 and 1 (11%) [29] in 2020.

Journals

All the 9 selected articles were published in peer-reviewed journals. Each of the following journals contained 11% (1/9) of selected studies: *BMC Public Health* [3], *Disability & Society* [27], *Research in Developmental Disabilities* [28], *Journal of Medical Internet Research* [29], *Journal of Racial and Ethnic Health Disparities* [32], *Annals of General Psychiatry* [33], and *Health Communication* [34]. Of these, 2 papers [30,31] were retrieved from *Social Media + Society*.

Study Locations

Of the 9 studies, 3 (33%) studies were undertaken in China [3,27,29], 3 (33%) in the United States [30-32], 1 (11%) in Kuwait and Saudi Arabia [28], 1 (11%) in Ethiopia [33], and 1 (11%) in Germany [34].

Locations of Study Participants

We originally intended to identify specific locations of study participants to reveal the health inequalities potentially existing between different areas reported in the 9 selected studies. However, only 33% (3/9) of these studies mentioned the places where the participants were located: city, countryside, or town [3]; Hubei Province, China [27]; and historically Black colleges and universities [32].

Study Design

A total of 67% (6/9) of studies were case studies [3,29-31,33,34] and the other 33% (3/9) were cohort studies [27,28,32]. The studies collected data from participants using web-based cross-sectional surveys and questionnaires [3,28,33], semistructured interviews [27,32,34], a web-based survey (extraction of data from web-based data sets) [29,31], and a mixed methods approach [27,30] (a WeChat ethnography research, participant observation, and semistructured interviews [27]; identifying COVID-19–related content through a keywords-based approach, introducing a measure of sustained amplification, and undertaking a qualitative hand-coding [30]).

Sample Size

In the 67% (6/9) of case studies [3,29-31,33,34], the sample sizes were 981, 1215, 1401, 212, 445, 929, and 22. In the 33% (3/9) of cohort studies [27,28,32], the sample sizes were 110, 190, and 21. It should be noted that in 22% (2/9) of these studies, the samples were 1215 tweets from 134 Weibo accounts [29] and 212,445 Twitter tweets from 137,746 unique users [31]. Most of the studies (6/9, 67%) [3,27,29-31,33] satisfied the standard of an appropriate number of participants (between 150 and 200) proposed by Dunbar [35]. The remaining studies (3/9, 33%) [28,32,34] used small sample sizes below this standard.

Participant Characteristics

Only 56% (5/9) of studies reported the age of the participants, primarily focusing on those aged >16 years [3], ≤60 years [28], ≥18 years [32], 30 to 34 years [33], and 19 to 80 years [34]. With the information provided in the studies, it was impossible to calculate the average age of the participants, and these 56% (5/9) of studies clearly stated the sex of the participants: male and female. Of the 9 studies, 4 (44%) studies [3,29,31,34] had the public as participants; 2 (%) studies [27,28] chose people with disabilities as informants; the other 3 (33%) studies investigated elites [30], college students [26], and educated people with internet access [33]. Merely 22% (2/9) of studies [30,32] reported the races of the participants.

Potential Risk of Bias

Various types of potential risks of bias were identified in the 9 studies. A total of 44% (4/9) of studies had a small sample size [27,28,32,34]. In all, 33% (3/9) of studies [3,29,33] reported an uneven size (Table 2). In addition, 11% (1/9) of studies [30] mentioned the bias of the lack of comparison between different types of social media and 11% (1/9) of studies [31] referred to

XSL•FO
RenderX

weighing all edges equally as a potential bias. Moreover, 44% (4/9) of studies [3,28,32,33] reported that participants were either predominantly male or female, so there was sex bias in these studies. Besides, the study by Zeng and Li [29] pointed out another bias: only using descriptive statistics and content analysis and failing to investigate the psychology and behavior of the audience. These characteristics of the 9 selected studies are summarized in Table 2.

XSL•FO
RenderX

Table 2. Characteristics of the selected studies (N=9).

Study	Journal	Study location	Location of participants	Study design	Study method	Sample size, n	Participant characteristics	Potential bias
Wang et al [3], 2021	*BMC Public Health*	China	City, countryside, or town	Case study	Cross-sectional web-based survey	981	Male and female; aged >16 years; all levels of education; students, workers, farmers, self-employed, employed in enterprises or institutions, retired, unemployed, and other	Uneven sample composition, which is mainly urban residents, young people, and people with a college education or above
Dai and Hu [27], 2021	*Disability & Society*	China	Hubei Province, China	Cohort study	WeChat ethnography research; participant observation; semistructured interviews	190	People with disabilities	Small sample size
Almusawi et al [28], 2021	*Research in Developmental Disabilities*	Kuwait and Saudi Arabia	—[a]	Cohort study	A cross-sectional electronic survey; questionnaire	110	People with hearing loss and no hearing loss; male and female; aged ≤60 years; unemployed, student, employed (non–health care and health care); primary, middle, high school, diploma, bachelor, and postgraduate	Small sample size
Zeng and Li [29], 2020	*Journal of Medical Internet Research*	China	—	Case study	A survey based on data extraction from Weibo accounts	1215 Weibo tweets from 134 sample accounts	The public	Samples not including county-level administrative regions; only evaluating the government Weibo accounts; only using descriptive statistics and content analysis and failing to investigate the psychology and behavior of the audience
Gallagher et al [30], 2021	*Social Media + Society*	United States	—	Case study	A mixed methods approach	1401	Elites of various demographic populations	The lack of comparison with information crowdsourcing on other platforms like Facebook, Reddit, YouTube, WhatsApp, TikTok, etc
Blevins et al [31], 2021	*Social Media + Society*	United States	—	Case study	A survey based on the COVID-19 Twitter data set	212,445 tweets	137,746 unique users	Weighing all edges equally
Wade et al [32], 2021	*Journal of Racial and Ethnic Health Disparities*	United States	Historically Black Colleges and Universities	Cohort study	In-depth interviews; quantitative surveys	21	Students enrolled during the spring 2020 semester; aged ≥18 years; male and female; Black American, Black foreign born, and White American	Small sample size
Ambelu et al [33], 2021	*Annals of General Psychiatry*	Ethiopia	—	Case study	A web-based cross-sectional survey; questionnaire	929	Educated Ethiopian population having access to the internet; male and female; aged 30-34 years	Only sampling communities who could read and write in English and had internet access; only studying the acute psychological impact and possibly being not generalized to subacute and long-term psychological complications

XSL•FO
RenderX

Study	Journal	Study location	Location of participants	Study design	Study method	Sample size, n	Participant characteristics	Potential bias
Wagner and Reifegerste [34], 2021	*Health Communication*	Germany	—	Case study	Semistructured qualitative interviews	22	Aged 19-80 years; male and female; the frequency of interpersonal communication about health topics (low to very high) and the extent of digital media use for interpersonal communication purposes (low to very high)	Small sample size

[a]Not available.

Main Results of the Selected Studies

Overview

The types of social media, social media-use–induced health inequalities, associated factors, and proposed resolutions are presented in Table 3. Through this table, we intended to compare the main research results of the 9 selected studies.

XSL•FO
RenderX

Table 3. Types of social media, social media-use–induced health inequalities, social media–related factors for health inequalities, and proposed resolutions to social media-use–induced health inequalities reported in the 9 selected studies.

Study	Types of social media	Social media use–induced health inequalities	Social media–related factors for health inequalities	Proposed resolutions to social media-use–induced health inequalities
Wang et al [3], 2021	Internet	Different levels of health knowledge related to COVID-19 among groups with different education levels; the digital health knowledge gap	• The use of traditional media, including newspapers, radio, and television failed to improve knowledge levels; different internet media use: web-based media use expanded the COVID-19 knowledge gap between groups with varying education levels	• Improving the pertinence in communication ways and contents; building authoritative scientific knowledge communication platforms; developing internet media literacy and scientific literacy
Dai and Hu [27], 2021	Media coverage on COVID-19 on multiple platforms and avenues, for example, live streaming of government press conferences and reports in digital media	People with disabilities inadequate, accessible information on COVID-19 compared with people with no disabilities	• Gaps between policies and practices regarding the digital accessibility infrastructure: • The accessible web information for people with disabilities focused little on applicable information to meet the individual needs of people with disabilities during the pandemic; • Sign language interpreters are commonly nonexistent in official press conferences and television news on COVID-19; • The newly developed website for disseminating information on COVID-19 also lacks accessibility design and remains inaccessible to the communities with hearing or visual disability	• A self-initiated and volunteer-driven Disability Support Network; the formulation of comprehensive and inclusive communication strategies for people with disabilities, which fully consider multiple dimensions of information for people with disabilities, including formats, content, and situations; government and public service sectors taking more proactive measures to provide inclusive communications and information in emergencies for people with disabilities
Almusawi et al [28], 2021	Social media	Disparities in the use of health information sources	• Participants with hearing loss mainly relied on social media, while the group with no hearing loss relied mainly on official government sources; low health literacy preventing the group with no hearing loss from accessing web-based health information.	• Bridging the gap in health literacy for individuals with hearing loss was essential in policy and practice to ensure equal access to health care and universal compliance with health directives at the population level: • The use of social media and unstandardized dialectic writing on the web • Different modes of disseminating information such as written information and QR codes linking to web-based videos in sign language
Zeng et al [29], 2021	Sina Weibo	Disparities in health information released on Weibo between the eastern region and the central and western regions in China; misinformation on COVID-19 information and prevention and treatment	• Governments' low willingness and ability to use government Weibo accounts; the passives state of the social media operations of public health authorities in China; Centers for Disease Control government Weibo accounts inform the public of the latest developments of the epidemic but fail to respond to public inquiries and a large amount of misinformation during the epidemic promptly.	• Governments in the central and western regions learned from similar experiences of neighboring governments; governments maintain their social media activity and update daily information frequently; governments use social media as a channel to release public health information and transmit health information to the public promptly
Gallagher et al [30], 2021	Twitter accounts (known as crowdsourced elites)	Twitter accounts receiving disproportionate attention for COVID-19 content during the pandemic; disparity between sustained and episodic amplification of COVID-19 information	• Crowdsourced elites varying across demographics in terms of race, geography, and political alignment; different subpopulations preferentially amplifying elites that are demographically similar to them; different subpopulations crowdsourcing different types of elite accounts, such as journalists, elected officials, and medical professionals, in different proportions	• Using the disproportionate voice of crowdsourced COVID-19 elites on the web to equitably promote public health information and mitigate misinformation across the networked public.

XSL•FO
RenderX

Study	Types of social media	Social media use–induced health inequalities	Social media–related factors for health inequalities	Proposed resolutions to social media–use–induced health inequalities
Blevins et al [31], 2021	Twitter	Misinformation on COVID-19 information and prevention and treatment	• Specific actors and networked communities on Twitter spread false information; key voices amplified COVID-19 misinformation on Twitter during the 2020 worldwide pandemic	• As networked societies become more accustomed to relying on information from varying sources on social media outlets and other cyberspaces (even for critical medical knowledge), the implications of how they interpret and apply that information in physical spaces was a significant consideration.
Wade et al [32], 2021	Twitter, YouTube, and Google engine	Misinformation on COVID-19 and associated precautions	• Students trust social media sources over government organizations such as the Centers for Disease Control and Prevention and World Health Organization	• Universities should consider implementing programs to aid in navigating social media for information-gathering, considering the high probability of misinformation
Ambelu et al [33], 2021	Facebook, Twitter, Zoom, etc	Those receiving information from social media have significantly higher odds of experiencing psychological distress	• Misinformation and myths about the COVID-19 pandemic bombarding social media, which strengthened groundless stress about COVID-19 among the population	• Developing an intervention plan to intervene in the psychological distress in the population, mainly targeting those groups who received information from social media
Wagner, and Reifegerste [34], 2021	Facebook, Twitter, Zoom, Facetime, etc	Misinformation on COVID-19 influencing others and providing misleading orientation	• Information-seeking and orientation-seeking practices in and through communication via social media	• Examining people's information-seeking and orientation-seeking practices in and through communication about pandemic-related media coverage could help us judge the importance of (constructive) media coverage and, ultimately, contribute to understanding the processes hindering and fostering public health compliance

Types of Social Media

A total of 33% (3/9) of studies [3,27,28] did not mention specific social media forms contributing to COVID-19–related health inequalities: internet [3], social media [28], and digital media [27]. The remaining 67% (6/9) of studies referred to concrete social media used: Sina Weibo [29]; Twitter [30,31]; Twitter, YouTube, and Google engine [32]; and Facebook, Twitter, Zoom, etc [33,34].

Types of Social Media Use–Related Health Inequalities and Associated Factors

Overall, 4 broad types of social media use–induced health inequalities were revealed in the 9 studies: disparities in the access to COVID-19–related health information [3,27-29]; misinformation regarding COVID-19 and associated precautions [29-32,34]; Twitter accounts receiving disproportionate attention for COVID-19 content during the pandemic [30]; and those obtaining information from social media having significantly higher odds of experiencing psychological distress [33].

In the study by Wang et al [3], in the context of traditional media (eg, newspapers, radio, and television) failing to improve knowledge levels, people with different educational backgrounds acquired different amounts of COVID-19–related health knowledge through the use of the internet, leading to health inequalities during the COVID-19 pandemic. Such COVID-19–related health inequalities created the digital health knowledge gap between individuals with diverse education levels, which influenced them differently in terms of COVID-19–related health behaviors and medical decisions. In contrast, 22% (2/9) of studies [27,28] did not consider the study participants' education but rather dealt with the COVID-19–related health inequalities between people with disabilities and people without disabilities. A study by Dai and Hu [27] described disabled people's inadequate, accessible information about COVID-19 in comparison with nondisabled people owing to the gaps between policies and practices regarding the digital accessibility infrastructure. Specifically, the accessible information on the web designed for people with disabilities focused little on applicable information to meet their needs during the pandemic; official press conferences and television news on COVID-19 failed to use sign language interpreters to inform people with disabilities of the latest situations of the pandemic. There was no accessibility design on the newly developed website for disseminating information on COVID-19, making it inaccessible to people with hearing or visual disabilities. The study by Almusawi et al [28] was similar to the study by Dai and Hu [27]. Similar to the study by Dai and Hu [27], the study by Almusawi et al [28] was also concerned with people with disabilities, specifically participants with hearing loss. However, unlike the study by Dai and Hu [27], the study by Almusawi et al [28] focused on the reliance on different information sources: people with hearing loss mainly relied on social media while people with no hearing loss mainly relied on official government sources. Besides, unlike studies by Wang et al [3] and Dai and Hu [27], the study by Almusawi et al [28] touched upon the hearing participants' low

XSL•FO
RenderX

health literacy, which prevented them from accessing health information on the web.

Misinformation was the most prevalent topic in the 9 selected studies. Of these, 56% (5/9) of studies [29-32,34] dealt with COVID-19 misinformation on various social media platforms and related precautions. The study by Zeng and Li [29] discussed disparities between East and Central China and West China in health information and misinformation released on a popular social medium in China named Sina Weibo. The contributors were government public health authorities' inadequate willingness and ability to use government Weibo accounts, their inactive operation of social media, and the Weibo accounts' failure to respond to public inquiries and huge amounts of COVID-19–related misinformation promptly. In contrast, 4 studies [30-32,34] investigated COVID-19–related misinformation on Twitter, Facebook, Zoom, YouTube, Google, Facetime, etc, which led to negative outcomes of COVID-19 prevention and treatment, psychological problems, and misleading orientation. The underlying factors included purposeful or purposeless amplification of COVID-19 information, people's preference for social media over government organs, and people's orientation-seeking.

Proposed Resolutions to Health Inequalities

Table 3 shows that although the proposed resolutions to COVID-19–related health inequalities are mostly specific to each of the 9 selected studies, what they have in common is intervention on the part of different players including government public health authorities [3,27-30,33], university authorities [32], and scientific communities [31,34]. The intervention measures are concerned with the establishment of relevant platforms [3,27], the development of related programs [3,32,33], the improvement of communication strategies [3,27-30], and the investigation of information-seeking, information application, and information orientation practices [31,34].

Discussion

Principal Findings and Implications

The findings on COVID-19–related health inequalities induced by the use of social media and recommended resolutions reported in the 9 studies were classified into 4 categories and discussed in the following subsections. Meanwhile, the relevant implications of each category were discussed.

Health Information Inaccessibility–Induced Health Inequalities and Proposed Resolutions

With advances in new media and IT, it is of great value to study the COVID-19–related knowledge gap and digital technology–induced unequal health information distribution [3], which were caused by the "Digital Divide" [36], that is, the gaps in the access to and use of the internet among different social groups leading to knowledge gaps [36]. Access gaps may not necessarily breed COVID-19 knowledge gaps because people use the internet media as the most frequent and dependent way to acquire COVID-19–related information [3]. In this case, what induced COVID-19–related health inequalities was use gaps: disparities in intensity, behavior, content, literacy, and other aspects when using the internet media [37,38]. To address the health inequalities caused by the access and use gaps of social media, the following resolutions were proposed: (1) improving the pertinence in the ways and contents of social media–based communication; (2) building social media platforms for authoritative scientific COVID-19 knowledge communication; and (3) developing social media literacy and science literacy of the public [3].

Compared with the people with no disabilities, people with disabilities faced more barriers when accessing health information during the COVID-19 pandemic for two main reasons: (1) the government's commitment to information accessibility was not always fulfilled, leading to the neglect of the needs for information in people with disabilities; and (2) the newly established web sites for COVID-19 information dissemination lacked accessibility design, thereby being inaccessible to people with disabilities, especially the people with hearing or visual disabilities [27,28]. To eliminate these health inequalities, Dai and Hu [27] proposed a self-initiated and volunteer-driven Disability Support Network, which fully considered various dimensions of information in terms of formats, content, and situations for people with disabilities, and government authorities and public service sectors taking more proactive steps to provide inclusive communications and information in emergencies for people with disabilities on social media. Besides, it is necessary to bridge the gap in health literacy for people with hearing loss using social media and web-based unstandardized dialectic writing and adopting different ways of disseminating information linking to web-based videos in sign language, to ensure equal access to health care and universal compliance with health directives at the population level [28].

The use of social media by public health authorities (eg, Center for Disease Control and Prevention) helped popularize daily health information through Weibo accounts, especially during COVID-19 [29]. However, the high dropout rates of Weibo accounts in some areas and the unequal distribution of Weibo accounts between the eastern region and the middle and western regions caused health inequalities among people in terms of access to helpful information on epidemic prevention and control [29]. The passives state and low willingness and ability of social media operations of Chinese public health authorities and their failure to respond to public inquiries and large amounts of misinformation on social media during the epidemic promptly made misinformation on social media even more rampant, causing even greater health inequalities. Governments in the central and western regions need to learn from the similar experiences of governments in the eastern region, maintain their social media activity by updating daily information frequently, and use social media to release public health information to the public promptly [29].

Misinformation-Induced Health Inequalities and Proposed Resolutions

The role of social media in breeding misinformation attracted the attention and aroused the concern of Wang et al [3]. Social media make it very easy for misinformation and fake news about COVID-19 to spread to the public [31]. A good case is a misinformation on Hydroxychloroquine on Twitter.

XSL•FO
RenderX

Interestingly, Donald Trump and his supporters turned out to be the most influential actors in advocating hydroxychloroquine as an effective treatment for coronavirus on Twitter [31]. People's trust in social media sources over government organs (eg, Center for Disease Control and Prevention and World Health Organization) [32] and their preferred information-seeking and orientation-seeking practices via social media [34] made misinformation on social media platforms even more unconstrained. Misinformation regarding the pandemic frequently appears on social media platforms, serving as a source of health risk [39,40].

Therefore, the media outlet should be more responsible for monitoring health message dissemination [41]. An effective way of countering misinformation lies in the gatekeeping of incorrect information, which helped to fight against the spreading of misinformation during COVID-19 owing to its ability to mediate and fact-check the accuracy of the contents [42]. Another way is to use the disproportionate voice of the crowdsourced elites on the web because people crowdsourced a small set of accounts on social media when clear information about COVID-19 prevention and protection was missing [30]. To mitigate misinformation in college students, university authorities need to consider carrying out programs to aid them in navigating social media for information-seeking and considering the high probability of misinformation [32].

"As networked societies become more accustomed to relying on information from varying sources on social media outlets and other cyberspaces (even for critical medical knowledge), the implications of how they interpret and apply that information in physical spaces is a significant consideration" [31]. Therefore, investigating people's information- and orientation-seeking practices in and through social media–based communication about COVID-19–related social media coverage [34] can help find practical approaches to minimize misinformation on social media, which most possibly caused health inequalities during the pandemic. What is needed to combat COVID-19–related misinformation on social media is (1) a keen sense of responsibility and the capability to think critically before sharing any information regarding the SARS-CoV-2 virus on social media platforms [31] and (2) a rational judgment of which social media sites are trustworthy and the ability to read and interpret health information on social media critically [32].

Disproportionate Attention to COVID-19 Information and Proposed Resolutions

Gallagher et al [30] studied the Twitter accounts receiving disproportionate attention during the COVID-19 crisis and the variation across demographics, finding that the public crowdsourced journalists, media outlets, and politicians more than epidemiologists, public health officials, and medical professionals. COVID-19–related health inequalities may arise owing to (1) crowdsourced elites varying across demographics in terms of race, geography, and political alignment; (2) different subpopulations preferentially amplifying elites who are demographically similar to them; and (3) different subpopulations crowdsourcing different elite accounts (eg, journalists, elected officials, and medical professionals) in different proportions [30]. Paradoxically, by working with COVID-19 elites, epidemiologists, public health officials, and medical professionals to popularize scientifically informed health guidelines and debunk misinformation, it is most likely to leverage the crowdsourcing potential of social media to achieve more health equality [30].

Higher Odds of Social Media–Induced Psychological Distress and Proposed Resolutions

In the context of a severe public health emergency, the public depends heavily on media coverage to stay informed [34]. COVID-19 has bred a massive "infodemic" [43] where various social media bombarded people with misinformation and myths about the COVID-19 pandemic, which intensified their groundless anxiety and stress about COVID-19 [33]. As social media exposure was associated with anxiety [44], it is necessary to develop an intervention plan to intervene in people's psychological distress, especially targeting those who predominantly received COVID-19–related health information on social media platforms [33]. People, especially those experiencing greater psychological distress, need to exercise extreme caution when deriving information on COVID-19 from social media and better use information delivered by the World Health Organization's "infodemics" team [45]. Moreover, they are encouraged to communicate with others about social media coverage of COVID-19 health information to understand better and evaluate pandemic-related information [34].

Limitations

This systematic review has some limitations. First, 2 databases, Embase and CINAHL, were not used for retrieving relevant studies owing to our inaccessibility to these databases, possibly making some related studies unidentified from the literature. This is to the detriment of the comprehensive synthesis of the principal findings reported in extant studies. Besides, some principal findings were likely to have low generalizability, considering that some social media use–induced health inequalities and the associated factors and recommended resolutions were reported in only one selected article. Moreover, we failed to compare the principal findings of this review with other systematic reviews, for this review was the first one concerning this topic. Finally, there was no protocol for how to report social media use–induced health inequalities when this review was performed, so certain reporting biases may be involved in this review. Future research will benefit from developing a reporting protocol for evaluating studies on social media use–induced health inequalities based on current frameworks.

Conclusions

This was the first systematic review seeking (1) to identify and summarize COVID-19–related health inequalities induced by social media and the associated contributing factors and (2) to characterize the relationship between the use of social media and health disparities during the COVID-19 pandemic. The findings synthesized from the selected studies highlighted the great value of studying the COVID-19–related knowledge gap and the digital technology–induced unequal health information distribution and the resulting health inequalities, providing knowledge about the relationship between social media use and

XSL•FO
RenderX

health inequalities regarding health knowledge and precautions against COVID-19. The 4 categories of COVID-19–related health inequalities induced by the use of social media and the associated contributors and recommended resolutions summarized in this review can provide some empirical evidence for developing practical solutions to help solve the health inequalities caused by social media use in the context of the repeated resurgences of the pandemic and future public health emergencies and crises. Informed by this review, researchers, social media and health practitioners, and policy makers can join hands to take full advantage of social media to promote health equality while minimizing social media use–induced health inequalities [21].

Conflicts of Interest

None declared.

References

1. Adults on Social Network Sites, 2005-2009. Pew Research Center and American Life Project. 2009 Oct 8. URL: https://www.pewresearch.org/internet/2009/10/08/adults-on-social-network-sites-2005-2009/ [accessed 2022-03-28]
2. Gibbons MC, Fleisher L, Slamon RE, Bass S, Kandadai V, Beck JR. Exploring the potential of Web 2.0 to address health disparities. J Health Commun 2011;16 Suppl 1:77-89. [doi: 10.1080/10810730.2011.596916] [Medline: 21843097]
3. Wang H, Li L, Wu J, Gao H. Factors influencing COVID-19 knowledge-gap: a cross-sectional study in China. BMC Public Health 2021 Oct 09;21(1):1826 [FREE Full text] [doi: 10.1186/s12889-021-11856-9] [Medline: 34627200]
4. Xiong CL. Opportunities, challenges, and countermeasures of healthy transmission in COVID-19 outbreak. Public Commun Sci Technol 2021;13(03):73-76. [doi: 10.16607/j.cnki.1674-6708.2021.03.026]
5. Becker S, Crandall MD, Fisher KE, Kinney B, Landry C, Rocha A. Opportunity for all how the American public benefits from Internet access at U.S. libraries. Institute of Museum and Library Services. 2010 Mar. URL: https://files.eric.ed.gov/fulltext/ED510740.pdf [accessed 2022-03-28]
6. Heldman AB, Schindelar J, Weaver JB. Social media engagement and public health communication: implications for public health organizations being truly “social”. Public Health Rev 2013 Jun 3;35(1):13. [doi: 10.1007/BF03391698]
7. McCosker A. Networks of advocasscy and influence- peer mentors in beyondblue's mental health forums. Social Innovation Research Institute. 2017. URL: https://www.researchgate.net/publication/318441131_Networks_of_Advocacy_Influence_Peer_Mentors_in_beyondblue's_Mental_Health_Forums_2017 [accessed 2022-04-01]
8. Roundtree A. Dialogic of social media in healthcare settings: text mining the rules, attitudes, and behaviors of health organizations and the public. Am Commun J 2018;20(1):1-18.
9. Shi J, Poorisat T, Salmon CT. The use of social networking sites (SNSs) in health communication campaigns: review and recommendations. Health Commun 2018 Jan;33(1):49-56. [doi: 10.1080/10410236.2016.1242035] [Medline: 27858464]
10. González-Padilla DA, Tortolero-Blanco L. Social media influence in the COVID-19 pandemic. Int Braz J Urol 2020 Jul;46(suppl.1):120-124 [FREE Full text] [doi: 10.1590/S1677-5538.IBJU.2020.S121] [Medline: 32550706]
11. Naslund JA, Aschbrenner KA, McHugo GJ, Unützer J, Marsch LA, Bartels SJ. Exploring opportunities to support mental health care using social media: a survey of social media users with mental illness. Early Interv Psychiatry 2019 Jun;13(3):405-413 [FREE Full text] [doi: 10.1111/eip.12496] [Medline: 29052947]
12. Viswanath K, Lee EW, Pinnamaneni R. We need the lens of equity in COVID-19 communication. Health Commun 2020 Dec;35(14):1743-1746. [doi: 10.1080/10410236.2020.1837445] [Medline: 33106029]
13. Chen JT, Waterman PD, Krieger N. COVID-19 and the unequal surge in mortality rates in Massachusetts, by city/town and ZIP Code measures of poverty, household crowding, race/ ethnicity, and racialized economic segregation. Harvard Center for Population and Development Studies. 2020. URL: https://search.bvsalud.org/global-literature-on-novel-coronavirus-2019-ncov/resource/fr/grc-740025 [accessed 2022-04-01]
14. Kawachi I, Subramanian SV, Almeida-Filho N. A glossary for health inequalities. J Epidemiol Community Health 2002 Sep;56(9):647-652 [FREE Full text] [doi: 10.1136/jech.56.9.647] [Medline: 12177079]
15. The Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. Phase I report: Recommendations for the framework and format of Healthy People 2020. U.S. Department of Health and Human Services. 2008 Oct 28. URL: http://www.healthypeople.gov/sites/default/files/PhaseI_0.pdf [accessed 2022-03-28]
16. Lurie N, Dubowitz T. Health disparities and access to health. JAMA 2007 Mar 14;297(10):1118-1121. [doi: 10.1001/jama.297.10.1118] [Medline: 17356034]
17. Marmot M. Social justice, epidemiology and health inequalities. Eur J Epidemiol 2017 Jul;32(7):537-546 [FREE Full text] [doi: 10.1007/s10654-017-0286-3] [Medline: 28776115]
18. Chen X, Hay JL, Waters EA, Kiviniemi MT, Biddle C, Schofield E, et al. Health literacy and use and trust in health information. J Health Commun 2018;23(8):724-734 [FREE Full text] [doi: 10.1080/10810730.2018.1511658] [Medline: 30160641]

XSL•FO
RenderX

19. Estrada E, Ramirez AS, Gamboa S, Amezola de Herrera P. Development of a participatory health communication intervention: an ecological approach to reducing rural information inequality and health disparities. J Health Commun 2018;23(8):773-782. [doi: 10.1080/10810730.2018.1527874] [Medline: 30307838]
20. Heredia NI, Lee M, Reininger BM. Exposure to a community-wide campaign is associated with physical activity and sedentary behavior among Hispanic adults on the Texas-Mexico border. BMC Public Health 2017 Nov 16;17(1):883 [FREE Full text] [doi: 10.1186/s12889-017-4893-4] [Medline: 29145821]
21. Wallace C, McCosker A, Farmer J, White C. Spanning communication boundaries to address health inequalities: the role of community connectors and social media. J Appl Commun Res 2021 Jun 21;49(6):632-650. [doi: 10.1080/00909882.2021.1934513]
22. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
23. Zhou L, Parmanto B. Reaching people with disabilities in underserved areas through digital interventions: systematic review. J Med Internet Res 2019 Oct 25;21(10):e12981 [FREE Full text] [doi: 10.2196/12981] [Medline: 31654569]
24. Letts L, Wilkins S, Law M, Stewart D, Bosch J, Westmoreland M. Critical Review Form - Qualitative Studies (Version 2.0). McMaster Unviersity. 2007. URL: https://www.unisa.edu.au/siteassets/episerver-6-files/global/health/sansom/documents/icahe/cats/mcmasters_qualreview_version2-0.pdf [accessed 2019-06-20]
25. Law M, Stewart D, Pollock N, Letts L, Bosch J, Westmoreland M. Critical Review Form - Quantitative Studies. McMaster University. 1998. URL: https://www.unisa.edu.au/siteassets/episerver-6-files/global/health/sansom/documents/icahe/cats/mcmasters_quantitative-review.pdf [accessed 2019-06-20]
26. Shaw DM. Invisible enemies: coronavirus and other hidden threats. J Bioeth Inq 2020 Dec;17(4):531-534 [FREE Full text] [doi: 10.1007/s11673-020-10015-w] [Medline: 32840841]
27. Dai R, Hu L. Inclusive communications in COVID-19: a virtual ethnographic study of disability support network in China. Disabil Soc 2022;37(1):3-21. [doi: 10.1080/09687599.2021.1933388]
28. Almusawi H, Alasim K, BinAli S, Alherz M. Disparities in health literacy during the COVID-19 pandemic between the hearing and deaf communities. Res Dev Disabil 2021 Dec;119:104089 [FREE Full text] [doi: 10.1016/j.ridd.2021.104089] [Medline: 34624721]
29. Zeng R, Li M. Social media use for health communication by the CDC in mainland China: National Survey Study 2009-2020. J Med Internet Res 2020 Dec 02;22(12):e19470 [FREE Full text] [doi: 10.2196/19470] [Medline: 33151892]
30. Gallagher RJ, Doroshenko L, Shugars S, Lazer D, Foucault Welles B. Sustained online amplification of COVID-19 elites in the United States. Soc Media Soc 2021 Jun 25;7(2):205630512110249. [doi: 10.1177/20563051211024957]
31. 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 Jun 23;7(2):205630512110249. [doi: 10.1177/20563051211024977]
32. Wade J, Poit ST, Lee A, Ryman S, McCain D, Doss C, et al. Navigating a pandemic: a qualitative study of knowledge, sources of information, and COVID-19-related precautions taken by HBCU students. J Racial Ethn Health Disparities 2022 Jan 15:1-12 [FREE Full text] [doi: 10.1007/s40615-021-01210-5] [Medline: 35032009]
33. Ambelu A, Birhanu Z, Yitayih Y, Kebede Y, Mecha M, Abafita J, et al. Psychological distress during the COVID-19 pandemic in Ethiopia: an online cross-sectional study to identify the need for equal attention of intervention. Ann Gen Psychiatry 2021 Mar 25;20(1):22 [FREE Full text] [doi: 10.1186/s12991-021-00344-4] [Medline: 33766076]
34. Wagner A, Reifegerste D. "The part played by people" in times of COVID-19: interpersonal communication about media coverage in a pandemic crisis. Health Commun 2021 Oct 13:1-8. [doi: 10.1080/10410236.2021.1989786] [Medline: 34645317]
35. Dunbar RI. Neocortex size as a constraint on group size in primates. J Human Evol 1992 Jun;22(6):469-493. [doi: 10.1016/0047-2484(92)90081-j]
36. Riccardini F, Fazio M. Measuring the digital divide. In: Proceedings of the IAOS Conference on Official Statistics and the New Economy. 2002 Presented at: IAOS '02; June 10-14, 2002; London, UK p. 27-29.
37. Jung JY, Qiu JL, Kim YC. Internet connectedness and inequality: beyond the “divide”. Commun Res 2001 Aug;28(4):507-535. [doi: 10.1177/009365001028004006]
38. Rogers EM. The digital divide. Convergence 2001 Dec;7(4):96-111. [doi: 10.1177/135485650100700406]
39. 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]
40. Cuello-Garcia C, Pérez-Gaxiola G, van Amelsvoort L. Social media can have an impact on how we manage and investigate the COVID-19 pandemic. J Clin Epidemiol 2020 Nov;127:198-201 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.06.028] [Medline: 32603686]
41. Zhao E, Wu Q, Crimmins EM, Ailshire JA. Media trust and infection mitigating behaviours during the COVID-19 pandemic in the USA. BMJ Glob Health 2020 Oct;5(10):e003323 [FREE Full text] [doi: 10.1136/bmjgh-2020-003323] [Medline: 33037063]

XSL•FO
RenderX

42. Lwin MO, Lee SY, Panchapakesan C, Tandoc E. Mainstream news media's role in public health communication during crises: assessment of coverage and correction of COVID-19 misinformation. Health Commun 2021 Jun 23:1-9. [doi: 10.1080/10410236.2021.1937842] [Medline: 34157919]
43. Naeem SB, Bhatti R. The Covid-19 'infodemic': a new front for information professionals. Health Info Libr J 2020 Sep;37(3):233-239 [FREE Full text] [doi: 10.1111/hir.12311] [Medline: 32533803]
44. Neria Y, Sullivan GM. Understanding the mental health effects of indirect exposure to mass trauma through the media. JAMA 2011 Sep 28;306(12):1374-1375 [FREE Full text] [doi: 10.1001/jama.2011.1358] [Medline: 21903818]
45. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. World Health Organization. 2020 Mar 11. URL: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 [accessed 2022-04-01]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by T Mackey; submitted 03.04.22; peer-reviewed by L HongMin, SF Tsao, C Zielinski; comments to author 18.07.22; revised version received 27.08.22; accepted 22.10.22; published 15.11.22.

Please cite as:
Shan Y, Ji M, Xie W, Zhang X, Ng Chok H, Li R, Qian X, Lam KY, Chow CY, Hao T
COVID-19–Related Health Inequalities Induced by the Use of Social Media: Systematic Review
JMIR Infodemiology 2022;2(2):e38453
URL: https://infodemiology.jmir.org/2022/2/e38453
doi:10.2196/38453
PMID:36420437

©Yi Shan, Meng Ji, Wenxiu Xie, Xiaomin Zhang, Harrison Ng Chok, Rongying Li, Xiaobo Qian, Kam-Yiu Lam, Chi-Yin Chow, Tianyong Hao. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 15.11.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

COVID-19 Messaging on Social Media for American Indian and Alaska Native Communities: Thematic Analysis of Audience Reach and Web Behavior

Rose Weeks[1], MPH; Sydney White[1], MPH; Anna-Maria Hartner[1], MSc; Shea Littlepage[1], MSPH; Jennifer Wolf[2], MPH; Kristin Masten[1], MPH; Lauren Tingey[1], MSW, MPH, PhD

[1]Center for Indigenous Health, Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States
[2]Project Mosaic, Denver, CO, United States

Corresponding Author:
Rose Weeks, MPH
Center for Indigenous Health
Department of International Health
Johns Hopkins Bloomberg School of Public Health
415 N Washington Street, 5th Floor
Baltimore, MD, 21231
United States
Phone: 1 443 287 4832
Email: rweeks@jhu.edu

Abstract

Background: During the COVID-19 pandemic, tribal and health organizations used social media to rapidly disseminate public health guidance highlighting protective behaviors such as masking and vaccination to mitigate the pandemic's disproportionate burden on American Indian and Alaska Native (AI/AN) communities.

Objective: Seeking to provide guidance for future communication campaigns prioritizing AI/AN audiences, this study aimed to identify Twitter post characteristics associated with higher performance, measured by audience reach (impressions) and web behavior (engagement rate).

Methods: We analyzed Twitter posts published by a campaign by the Johns Hopkins Center for Indigenous Health from July 2020 to June 2021. Qualitative analysis was informed by in-depth interviews with members of a Tribal Advisory Board and thematically organized according to the Health Belief Model. A general linearized model was used to analyze associations between Twitter post themes, impressions, and engagement rates.

Results: The campaign published 162 Twitter messages, which organically generated 425,834 impressions and 6016 engagements. Iterative analysis of these Twitter posts identified 10 unique themes under theory- and culture-related categories of framing knowledge, cultural messaging, normalizing mitigation strategies, and interactive opportunities, which were corroborated by interviews with Tribal Advisory Board members. Statistical analysis of Twitter impressions and engagement rate by theme demonstrated that posts featuring culturally resonant community role models (P=.02), promoting web-based events (P=.002), and with messaging as part of Twitter Chats (P<.001) were likely to generate higher impressions. In the adjusted analysis controlling for the date of posting, only the promotion of web-based events (P=.003) and Twitter Chat messaging (P=.01) remained significant. Visual, explanatory posts promoting self-efficacy (P=.01; P=.01) and humorous posts (P=.02; P=.01) were the most likely to generate high–engagement rates in both the adjusted and unadjusted analysis.

Conclusions: Results from the 1-year Twitter campaign provide lessons to inform organizations designing social media messages to reach and engage AI/AN social media audiences. The use of interactive events, instructional graphics, and Indigenous humor are promising practices to engage community members, potentially opening audiences to receiving important and time-sensitive guidance.

(JMIR Infodemiology 2022;2(2):e38441) doi:10.2196/38441

XSL•FO
RenderX

KEYWORDS

COVID-19; American Indian or Alaska Native; social media; communication; tribal organization; community health; infodemiology; Twitter; online behavior; content analysis; thematic analysis

Introduction

Background

The COVID-19 pandemic has disproportionately affected American Indian and Alaska Native (AI/AN) peoples from a health, economic, and spiritual perspective. In August 2020, rates of confirmed COVID-19 cases among AI/AN peoples were 3.5 times higher than non-Hispanic White populations [1]. AI/AN peoples are more likely to live in multigenerational households, making social distancing challenging [2]. Further, AI/AN individuals are more likely to have preexisting medical conditions that amplify the risk of severe COVID-19 disease, such as obesity and diabetes [3]. Such health inequities are rooted in hundreds of years of Western aggression, ranging from genocide to forced institutionalization (ie, boarding schools) that removed Indigenous peoples from protective cultural practices and perpetuated continuing oppression and socioeconomic inequities [4-7]. During the COVID-19 pandemic, health systems starved by years of federal underspending were called upon to treat a flood of cases of the novel virus [2,8]. Communities that have come together during hardship in sacred ceremony since time immemorial were urged and often mandated by tribal law to stay home, with traditional wisdom keepers at risk for severe disease [8].

Despite these layered challenges, tribal and urban Indian organizations showed remarkable agility and resilience in initiating and promoting mitigation measures such as curfews and social distancing orders that many adjacent non-AI/AN communities implemented briefly or not at all [8]. By the spring of 2021, when access to COVID-19 vaccines became widespread, uptake among AI/AN peoples was the highest of any US racial group, although there were variations across regions and tribal lands [9]. This high acceptance has been attributed to Indigenous values, including solidarity and respect for elders and other culture-bearers threatened by COVID-19 [10]. Innovative and highly varied approaches in delivering and encouraging vaccination were also successful within AI/AN communities [8-11].

To increase confidence in vaccines and other pandemic mitigation strategies, tribes and AI/AN organizations used culturally tailored messaging strategies. Communication campaigns highlighted cultural strengths such as reverence for elders and community members using slogans such as “Be a Good Relative” and “For the Love of Our People” [10,12]. Such campaigns often used social media to disseminate guidance and foster connectedness. Social media also aided in countering the marginalization and erasure of AI/AN peoples, sometimes omitted as a distinct population in national communications about the pandemic’s effects [12]. Prior to the pandemic, social media had provided a sense of power and control over Indigenous identities [13,14]. Now, forums such as the Facebook groups Social Distance Powwow and American Indian COVID-19 Resources and Response have helped participants celebrate traditional skills such as beading and dancing to cope with pandemic losses during a time of social isolation [15,16].

Pandemic-era communication campaigns targeting AI/AN communities used social media to disseminate guidance, meet community needs, and help people stay connected to protective culture and community [17]. Campaigns used proven public health communication strategies, such as engaging trusted leaders to deliver culturally adapted messaging [18]. However, evidence-based guidance on using social media to raise awareness about public health measures was limited at the start of the COVID-19 pandemic, particularly with regard to AI/AN communities. Social media outreach has substantial benefits, especially in the context of a rapidly evolving pandemic, by allowing for immediacy and the ability to forge rapid connections, build rapport with audiences, and dispel rumors by providing accurate information [19,20]. Limitations include the need to monitor channels for harmful misinformation—for instance, in negative comments, which can influence viewers’ opinions [21,22]. There is some evidence that well-designed social media campaigns may deliver a range of behavior change components, such as social support, observational learning, instructions on how to perform a behavior, and prompts or cues to practice a behavior [23]. Public health campaigns using social media have been linked with public health impact such as increases in human papillomavirus vaccination coverage and uptake of pre-exposure prophylaxis prescriptions [24,25]. To compete in dense and rapidly changing social media environments rife with misinformation, public health organizations must design social media campaigns using best practices and compare evidence about what works to reach and engage web-based audiences with protective health messaging [26].

Johns Hopkins Center for Indigenous Health Campaign

The Johns Hopkins Center for Indigenous Health (CIH; The Bloomberg School’s Center for American Indian Health was renamed Center for Indigenous Health in September 2022) launched a COVID-19 communications campaign at the start of the pandemic in March 2020 and, over the next several months, established a social media presence to inform and connect tribes, urban Indian organizations, and community members with reliable, culturally adapted communication on evidence-based measures to slow the spread of the virus. CIH convened a Tribal Advisory Board (TAB) and engaged AI/AN colleagues based in Arizona, Maryland, Minnesota, New Mexico, and California to guide the social media campaign [27]. The campaign included hundreds of social media posts consisting of graphics, fact sheets, videos, and slideshows, covering topics including physical distancing, mental health, isolation and quarantine, masking, and vaccination and using a variety of tones and message styles. To acknowledge a collective perception of adversity, some posts used insider humor packaged as colorful memes that could be self-deprecating, satiric, or

XSL•FO
RenderX

refreshingly silly; such messages were reviewed by the TAB prior to distribution, ensuring the voice, tone, and terms were appropriate for and understood by AI/AN peoples across the country. All social media materials were made available in Microsoft Word–based toolkits including graphics and supporting captions, available for download at a public resource library [28]. Applying guidance from the TAB, CIH's social media campaign aimed to frame COVID-19 health information with accessible and engaging content featuring Indigenous illustrations and languages across Facebook, Instagram, and Twitter, which is the focus of this study. Although we could not limit our post reach to AI/AN-identifying audiences, to maximize visibility to AI/AN Twitter users, we mentioned leading national organizational accounts on most posts and used hashtags popular with AI/AN Twitter users.

On Twitter, posts reached increasingly larger audiences throughout the campaign, with some messages organically reaching tens of thousands of people. In November 2020 and May 2021, two Twitter Chats (live, open-discussion, and time-bound Twitter campaign events) were organized around, first, Native American Heritage Month, and second, the rollout of COVID-19 vaccines for people of all ages in the United States. Such events reached a relatively large audience, but other posts shared during the campaign achieved 100 or fewer impressions. The divergence in audience reach and impressions throughout the 1-year campaign demonstrated a need to better characterize the relationship between message characteristics and audience reach and engagement. This analysis aimed to describe the correlation of content themes with Twitter post performance in a health campaign aiming to inform AI/AN communities.

Methods

Ethical Considerations

This study was reviewed by the Johns Hopkins University Institutional Review Board, which concluded it was not human subjects research since the study encompassed (1) key informant interviews, involving information from individuals about something other than themselves and disclosing no personal opinions; and (2) secondary data analysis.

Data Source

To examine trends in social media messaging, our analysis focused on Twitter posts shared by CIH from July 1, 2020, to June 30, 2021—a 1-year campaign. During this time frame, CIH published 162 original campaign-related posts. Twitter analytics data were extracted from CIH's Twitter account, and the analysis reviewed impressions (the number of times a given tweet is viewed); engagements (the number of times a user interacts with a tweet through retweets, favorites, replies, link clicks, hashtag clicks, mention clicks, and media views); and a summary indicator of engagement rate (ER), which measures the number of engagements a tweet has per impression.

To better understand and contextualize themes across Twitter posts, 10 key informant interviews with members of CIH's TAB were conducted. The TAB was made up of AI/AN and allied health communication professionals from various regions. Meeting twice monthly, the TAB provided guidance on AI/AN public health priorities, reviewed health communications content produced by CIH, and sought to ensure that campaign content was culturally appropriate and relevant across Indian Country. Thus, their feedback shaped all content produced for the Twitter campaign analyzed in this paper.

Key informant interviews with TAB members were conducted by a member of the study team as part of a separate evaluation of the TAB procedures using a semistructured interview guide. Participants were female professionals serving in communications and outreach roles for tribal nations and other organizations serving AI/AN peoples, representing 12 tribes and 10 tribal-serving organizations across various regions; all but 2 identified as AI/AN. Interviews were audio recorded, transcribed, and analyzed by a member of the study team familiar with the data. Although interviews largely focused on participants' experience in the TAB, passages related to social media strategy were compiled and applied to inform our analysis of Twitter posts.

To guide the analysis and interpretation of results, we used the Health Belief Model (HBM), a theory adapted to influence health behaviors that has been used in diverse cultural contexts since the 1950s [29]. The theory's 6 constructs include risk susceptibility, risk severity, benefits to action, barriers to action, self-efficacy, and cues to action; the campaign messages aimed to leverage nearly all components [30]. In addition to themes from TAB interviews, risk communication guidance emphasizing the benefits of 2-way communication through social media was also applied to content analysis [19].

Thematic Coding

Twitter posts were iteratively coded by theme using both deductive content derived from the HBM and TAB interviews and inductive codes based on emergent themes. Thematic coding classified all Twitter posts into categories of similar messaging strategies. In this study, 2 members of the study team each independently coded a subset of 50 posts, initially using deductive themes and then creating additional inductive codes as appropriate. Codes were iteratively refined and combined to create overarching categories through discussion. After 3 revisions, the codebook was finalized, and 2 members of the study team coded all posts, resolving each discrepancy through reflexive discussion. These themes were then used as variables in the analysis of audience reach and engagement.

Data Cleaning, Exploration, and Analysis in R Statistical Software

Data obtained from CIH's Twitter account were first entered into R statistical software (version 4.0.3; R Foundation for Statistical Computing) [31]. An initial data exploration stage included data cleaning, in which variables were recategorized and examined for missingness. Data exploration was completed for several variables, including partner tagging, time of day, year and month, and type of post. All date-time variables were parsed to include only the month and year. "High" and "low" ER or impressions were classified through percentiles, in which all posts in the 75th percentile or above in either outcome were classified as "high" and those below the 75th percentile were

XSL•FO
RenderX

classified as "low." Initial descriptive statistics and figures were then used to examine the counts and distribution of posts across the variables of interest using the *dplyr* and *ggplot* packages [32,33]. Possible confounders and a priori variables were evaluated and selected for further analysis. Odds ratios were then calculated to examine the association between theme and impressions or ER; a generalized linear model was used to calculate the adjusted odds ratios. The odds ratios were adjusted for time of year (month and year), which was an a priori variable, to account for several factors over the year-long campaign: an increase in followers over time, a gradual increase in impressions per post, and a decline in average ER per post. Tweets of a particular type where n=1 were excluded from the analysis.

Results

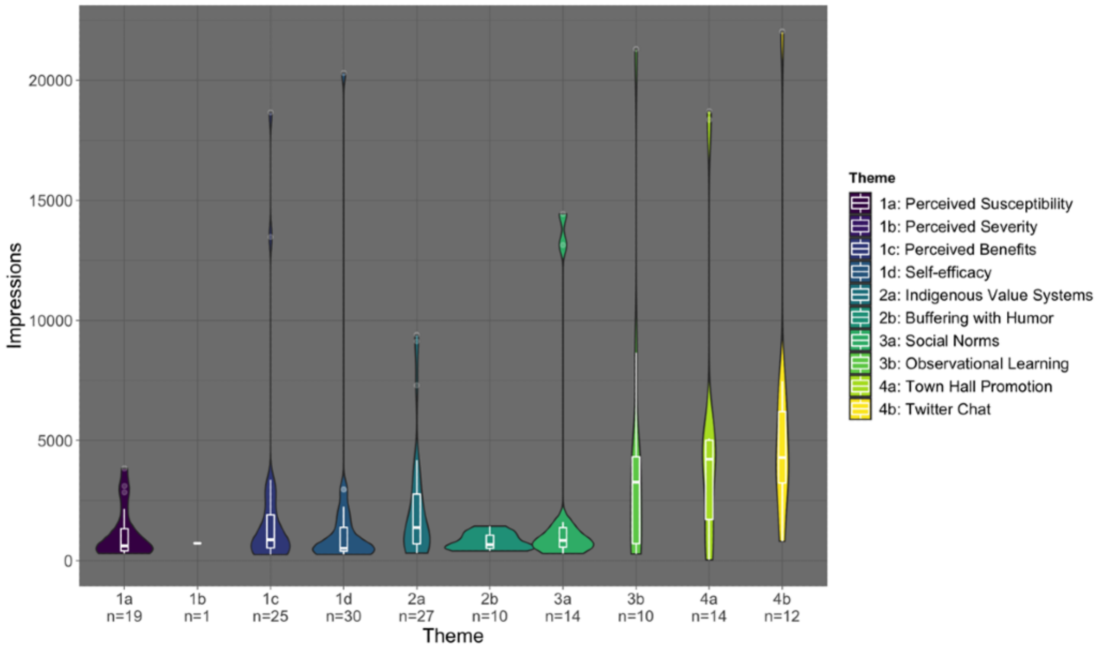
Thematic Analysis

The process of coding Twitter posts led to 4 overarching categories, as seen in Multimedia Appendix 1. Of the 162 tweets, 75 (46.3%) were categorically coded as *Framing Knowledge*, 37 (22.8%) as *Cultural Messaging*, 24 (14.8%) as *Normalizing COVID Mitigation Strategies*, and 26 (16%) as *Interactive Opportunities*. Under these 4 categories, the data revealed 10 themes: *Perceived Susceptibility, Perceived Severity, Perceived Benefits, Self-Efficacy, Indigenous Value Systems, Humor, Social Norms, Observational Learning, Event Promotion*, and *Twitter Chat* (see Multimedia Appendix 1). Definitions and examples of tweets coded within each category and theme are shown in Multimedia Appendix 1.

Engagement and Reach Analysis by Theme

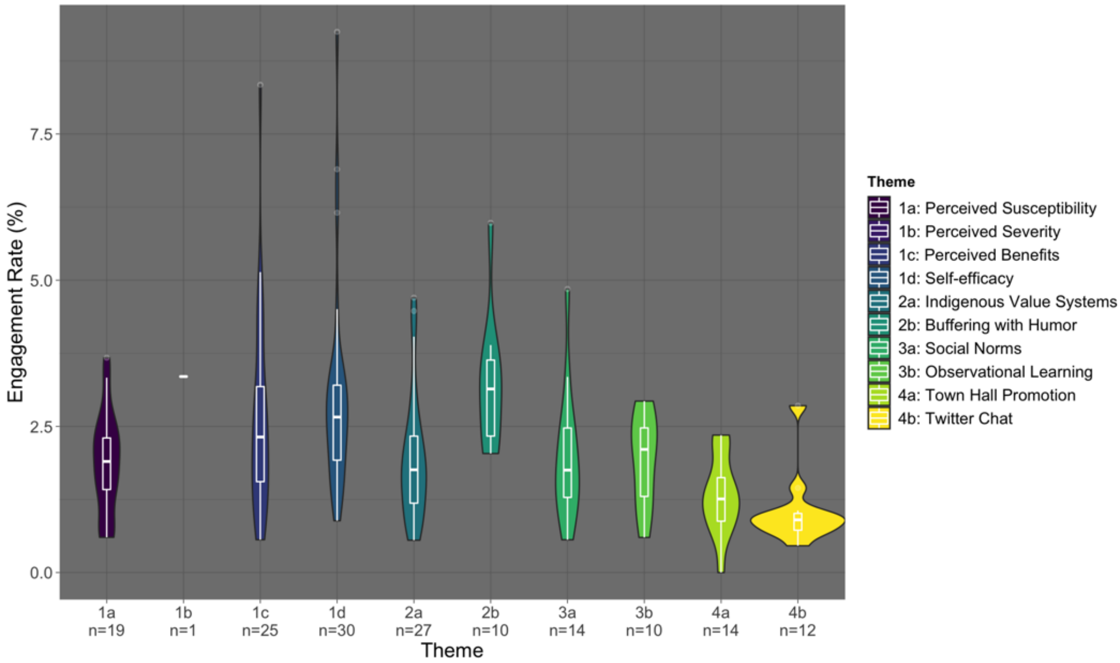
CIH's Twitter account had 900 followers as of February 23, 2021—near the middle of the campaign—which increased to 1200 followers by its end. Throughout the campaign, posts organically generated 425,834 impressions and 6016 engagements. On average, each post received 2628 impressions and 37 engagements, with an average ER of 2.2%. Figures 1 and 2 display the distribution of impressions and ER by the post theme, highlighting initial summary statistics and the density of the distribution. In our data exploration phase, several variables—partner tagging, time of day, and type of post—were not found to be significant and were thus excluded from consideration in the adjusted analysis.

Figure 1. Tweet impressions by social media post theme classification (n=162).



XSL•FO
RenderX

Figure 2. Tweet engagement rate by social media post theme classification (n=162).



Post Themes and Impressions

Table 1 describes post themes and associations with impressions, before and after an adjustment for the date the posts were published. Prior to adjusting for month of posting, posts coded with the theme *Observational Learning* had 5.01 (95% CI 1.36-20.57) times the odds of achieving high impressions than other posts. After adjusting, *Event Promotion* posts had 6.79 (95% CI 1.75-32.27) times the odds of being among the top 75% of tweets by impressions than other posts, and *Twitter Chat* messages had 15.94 (95% CI 3.12-138.42) times the odds of achieving high impressions. The post with the greatest number of impressions (n=22,039) was a *Twitter Chat* welcome message about COVID-19 vaccinations in AI/AN communities (Figure 3). The post with the second highest number of impressions (n=21,309) featured a video with a Navajo traditional healer speaking about her decision to get vaccinated against COVID-19, coded as *Observational Learning* (Figure 3). The reverse trend was observed for posts that were coded as *Self-Efficacy* and *Social Norms*, with these posts having 0.04 (95% CI 0.002-0.31) times and 0.17 (95% CI 0.002-0.73) times the odds of achieving high impressions than other themes.

Table 1. Unadjusted and adjusted odds of high impressions by theme.

Theme	Unadjusted OR[a] (95% CI)	Unadjusted *P* value[b]	Adjusted OR (95% CI)[c]	Adjusted *P* value[b]
1a. Perceived Susceptibility	0.31 (0.05-1.16)	.13	0.70 (0.09-3.70)	.70
1b. Perceived Severity	N/P[d]	N/A[e]	N/P	N/A
1c. Perceived Benefits	0.51 (0.14-1.46)	.35	0.57 (0.14-1.95)	.40
1d. Self-efficacy	0.08 (0.004-0.39)	.001	0.04 (0.002-0.31)	.01
2a. Indigenous Value Systems	1.04 (0.38-2.58)	.94	2.85 (0.82-10.38)	.10
2b. Humor	N/P	N/A	N/P	N/A
3a. Social Norms	0.47 (0.07-1.81)	.33	0.17 (0.002-0.73)	.03
3b. Observational Learning	5.01 (1.36-20.57)	.02	3.40 (0.86-15.18)	.09
4a. Event Promotion	6.53 (2.10-22.53)	.002	6.79 (1.75-32.27)	.003
4b. Twitter Chat	19.19 (4.76-129.15)	<.001	15.94 (3.12-138.42)	.01

[a]OR: odds ratio.

[b]Significant level at *P*<.05.

[c]Adjusted for time of year, see methodology for further details.

[d]N/P: not possible, as the small sample size for these categories leads to 0 values that make the values infinite.

[e]N/A: not applicable.

XSL•FO
RenderX

Figure 3. Clockwise from top left: the top-ranked post by impressions was coded as Twitter Chat; the second highest post by impressions was coded as Observational Learning; a top-ER post showing a cartoon Native American individual chasing a COVID-19 particle with a vaccine syringe was coded as Humor; and the highest ranked post by ER was coded as Self-Efficacy. ER: engagement rate.

Post Theme and Engagement Rate

Posts thematically coded under *Self-Efficacy* and *Humor* were statistically more likely to generate a high ER, with *Self-Efficacy* posts having 2.95 (95% CI 1.27-6.84) times and *Humor* posts having 5.43 (95% CI 1.43-20.70) times the odds of being in the high percentile for ER (Table 2). The post with the highest ER (9%) explained how to wear masks to protect against COVID-19, with simple graphics illustrating masks offering poor protection, such as bandanas, and masks offering good protection, such as disposable surgical masks (Figure 3). An example of a high-ER *Humor* post (ER 6%) was an AI/AN-drawn cartoon of a man with a feather chasing a SARS-CoV-2 particle with a giant syringe, under a headline that read, “Don’t Stop Now—It’s on the Run!” (Figure 3). The post themes associated with having a higher number of impressions in our adjusted analysis, *Event Promotions* and *Twitter Chats*, were not more likely to generate higher ER.

XSL•FO
RenderX

Table 2. Unadjusted and adjusted odds of high engagement by theme.

Theme	Unadjusted OR[a] (95% CI)	Unadjusted *P* value[b]	Adjusted OR (95% CI)[c]	Adjusted *P* value[b]
1a. Perceived Susceptibility	0.31 (0.05-1.16)	.13	0.29 (0.06-1.36)	.12
1b. Perceived Severity	N/P[d]	N/A[e]	N/P	N/A
1c. Perceived Benefits	1.85 (0.72-4.51)	.19	1.99 (0.80-4.99)	.14
1d. Self-efficacy	2.84 (1.22-6.55)	.01	2.95 (1.27-6.84)	.01
2a. Indigenous Value Systems	0.63 (0.20-1.66)	.37	0.61 (0.21-1.74)	.35
2b. Humor	5.01 (1.36-20.57)	.02	5.43 (1.43-20.70)	.01
3a. Social Norms	0.47 (0.07-1.81)	.33	0.48 (0.10-2.28)	.36
3b. Observational Learning	0.72 (.11-3.04)	.69	0.74 (0.15-3.67)	.72
4a. Event Promotion	N/P	N/A	N/P	N/A
4b. Twitter Chat	0.25 (0.01-1.35)	.19	0.23 (0.03-1.84)	.17

[a]OR: odds ratio.

[b]Significant level at *P*<.05.

[c]Adjusted for time of year, see methodology for further details.

[d]N/P: not possible, as the small sample size for these categories leads to 0 values that make the values infinite.

[e]N/A: not applicable.

Discussion

Principal Findings

This study thematically analyzed Twitter posts from a COVID-19 communications campaign prioritizing AI/AN audiences to understand how best to reach and engage audiences with pandemic mitigation guidance. The study organized posts into 4 categories and 10 themes that integrated the HBM and risk communication guidance with Indigenous cultural values such as solidarity and humor. On average, posts that highlighted interactive opportunities to learn about and discuss pandemic and cultural issues were likely to reach more people but were not associated with higher engagement. Posts highlighting cultural role models such as traditional healers and web-based influencers often reached high numbers of people, although this finding did not remain significant after adjustment. In contrast, posts that highlighted instructional content with simple graphics or used insider humor to convey pandemic-related guidance were more likely to create high ER but, on average, reached fewer accounts.

In all, 99 (61.1%) out of 162 posts were coded with themes from the HBM, demonstrating that this theory was a strong fit for coding posts. An additional 37 (22.8%) posts were coded with cultural themes, appealing to traditional identities that are vital to the well-being of AI/AN peoples, as demonstrated by their protective effect on binge substance use, suicide attempts, and other major health risks [34,35]. Within these cultural themes, 10 posts were thematically coded as *Humor*. Finally, the *Interactive Opportunities* theme highlighted 2-way communication aiming to augment community connectedness and increase public health transparency during an uncertain time [36].

In our statistical analysis of theme by impressions, we found that posts highlighting web-based events were more likely to achieve a higher reach. Twitter Chats have been a successful strategy to build community for other public health organizations, practitioners, and health advocacy groups [37-39]. The high number of impressions for *Twitter Chat* messaging in this campaign could reflect interaction with other AI/AN–serving organizations as these messages were promoted and shared with a larger audience base. Thus, encouraging 2-way communication in our campaign was successful in reaching more users even if individual users were not as likely to engage directly with campaign posts.

In our study, posts highlighting role models (ie, *Observational Learning*) were not associated with high impressions after adjusting for the date of posting. This finding runs in contrast to evidence from a variety of community contexts, including a campaign reaching tribal audiences, that posts sharing personal stories from trusted messengers can successfully engage audiences [24,40-44]. However, given the importance of storytelling in AI/AN communities, we feel that highlighting trusted role models is critical to successful communication campaigns and warrants future implementation and evaluation within AI/AN contexts.

Audience engagement varied substantially during the campaign, with a reported average ER of 2.2% and a median ER of 1.9% during the 1-year campaign. Although engagement metrics vary by social media platform, industry, and topic, an industry source estimates that nonprofit posts on Twitter average an ER of 0.05% [45]. There is very limited evidence related to Twitter campaigns targeting AI/AN community members, but an evaluation of an AI/AN-oriented obesity prevention campaign observed that social media posts "generated little involvement and response," and a campaign addressing kidney donation found community members did not engage with Twitter messages during the campaign [46,47]. More research is needed to inform the development of social media content to ensure adequate reach and engagement in AI/AN communities across a range of issues.

XSL•FO
RenderX

Thematically, humorous posts and those with instructional graphics were more likely to spur web-based activity. Other evaluations have found that visual concepts positively affect engagement behaviors, including in AI/AN-focused social media campaigns [48,49]. Our finding that humorous posts were more likely to achieve high ER is bolstered by evidence that scientific visualizations in humorous form can improve knowledge acquisition and problem-solving skills [50,51]. Further, humor has special salience in AI/AN communities as a source of resilience to hardship. Individuals have a great deal of institutional mistrust and may use humor in code switching [52,53]. Across a variety of Indigenous cultures worldwide, humor demonstrates cultural understanding, whereas speaking familiarly and using terms recognized to be of Indigenous origin, such as "Stoodis" and "Skoden"—as the CIH campaign did—can build confidence that messages are coming from within the community [54]. The strategic use of humor in social media messaging also builds upon AI/AN oral traditions such as "Trickster" stories [54]. Humor can also be a powerful tool for building trust in health care relationships [55,56]. Humorous social media content, therefore, has an important role in culturally competent communication strategies, especially around sensitive health topics. Using humor may destigmatize disparities, stimulate discussion, and prompt care-seeking where appropriate. Social media campaigns aiming to reach AI/AN audiences should consider using audience-tailored humor to convey empathy and humility while ensuring cultural appropriateness.

Over the course of the 1-year campaign, the average number of impressions generated per post increased and the average ER declined, potentially showing that the larger audiences the campaign reached later were on average less likely to take action to share or amplify a message they saw. This finding may reflect natural tension between impressions and ER, due to their reciprocal relationship; outside of COVID-19, other campaigns have also found a trade-off between impressions and engagement [41]. This finding also may reflect changes in public sentiment over time; as the pandemic became less novel, community members may have felt less urgency to reshare guidance. Increased burnout and COVID-19 fatigue may have reduced engagement with pandemic-related guidance.

Our findings may be instructive for others seeking to promote culturally tailored content for AI/AN audiences on social media. Social media activity should be integrated into multimodel communication campaigns designed to reach all community members, including those on tribal lands where broadband internet limitations persist [57,58]. Increasing access to high-speed internet in rural AI/AN communities will contribute to the increased relevance of social media communication in the future [58,59].

Although there are differences across social media networks, findings should be relatable to campaigns conducted across numerous platforms. For example, campaigns on Facebook or Instagram incorporating humorous content and highly visual, instructional guidance may be successful in achieving higher performance metrics than those using other types of content, such as posts explaining the benefits of public health measures, which may seem too conventional to engage savvy web-based audiences. This analysis focused on Twitter as the most consistent CIH social media platform throughout the pandemic. Campaign performance metrics on Facebook were highly variable during the time period and thus differences in reach and engagement may not be attributable to the salience of particular thematic content with our audience [60,61]. Meanwhile, CIH's Instagram account, being newly established, saw lower performance measures than on Facebook or Twitter. Future research should focus on whether Twitter-based findings remain consistent across other platforms, especially with variation in user characteristics across different platforms [62].

Limitations

We sought to explore social media metrics during the peak of the COVID-19 pandemic and thus, focused on a limited data set to capture reach and engagement to reflect this unique period of time. Our relatively small data set produced parameters with broad confidence intervals, which limits the strength of the quantitative findings. We integrated only 1 confounding factor in the adjusted analysis, date of posting, and other confounders may be unaccounted for, although other factors we reviewed did not seem to affect performance metrics.

Although the CIH campaign achieved nearly a half-million impressions, the total number of AI/AN peoples in the United States is 9.7 million [63]. The prevalence of social media use among AI/AN peoples is likely similar to that among the general population at around 70% [62,64]. Therefore, the campaign's Twitter posts did not reach a significant proportion of AI/AN social media users in the United States. Additionally, due to data privacy around social media users, it is impossible to verify that those reached by CIH posts were AI/AN. By engaging and cross-promoting content with prominent AI/AN users and organizations, such as through Twitter Chats, we assume that a large proportion of users reached were AI/AN. However, this is a limitation inherent to all studies using social media analytics data for publicly targeted campaigns.

Finally, given the vast diversity across 574 federally recognized tribes and urban AI/AN communities, our findings may not be widely applicable across all AI/AN audiences. The TAB that supervised the development of the campaign and informed this thematic analysis was representative of a variety of tribes and regions but was almost entirely made up of early-to-midcareer professional women. The perspective of other stakeholders such as male leaders may be underrepresented.

Conclusions

AI/AN communities have been disproportionately affected by the COVID-19 pandemic. Social media offered a medium to rapidly provide public health guidance and foster cultural connectedness to counteract the isolation and marginalization of Indigenous experiences within the pandemic. Awareness campaigns using social media can benefit from integrating effective strategies to reach and engage increasingly active AI/AN audiences on platforms such as Twitter. In a 1-year social media campaign to disseminate guidance on COVID-19, posts highlighting opportunities for web-based discussion were, on average, likely to reach larger audiences. Humorous tweets and posts with simple, instructional graphics were 2 leading ways

XSL•FO
RenderX

to engage audiences by demonstrating humility and promoting confidence in public health guidance as well as encouraging the adoption of preventive behaviors. Further analysis across other social media platforms is needed to inform organizations and tribes seeking to disseminate public health guidance to AI/AN communities.

Acknowledgments

We thank all members of our Tribal Advisory Board who have joined us in creating health resources for tribal communities throughout the COVID-19 pandemic. Our advisors have represented a range of tribes and organizations and bring their expertise to create better health communications for American Indian and Alaska Native peoples. We thank the Walmart Foundation for their financial support for the COVID-19 communication campaign.

Land Acknowledgment: We humbly acknowledge that Johns Hopkins University's Baltimore campus is located on the traditional and contemporary homelands of Indigenous peoples. The campus resides on unceded lands of the Piscataway and Susquehannock peoples. We give thanks to the past, present, and future stewards of this land and respect all tribal nations' sovereignty and right to self-determination.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table illustrating the thematic coding of tweets with examples.
[DOCX File , 16 KB - infodemiology_v2i2e38441_app1.docx]

References

1. Disparities in COVID-19 illness. Centers for Disease Control and Prevention. 2020 Dec 10. URL: https://www.cdc.gov/coronavirus/2019-ncov/community/health-equity/racial-ethnic-disparities/increased-risk-illness.html [accessed 2022-03-23]
2. The National Congress of American Indians calls for more attention to COVID-19 impacts to Indian Country. National Congress of American Indians. 2020 Mar 18. URL: https://www.ncai.org/news/articles/2020/03/18/the-national-congress-of-american-indians-calls-for-more-attention-to-covid-19-impacts-to-indian-country [accessed 2022-01-02]
3. Disparities. Indian Health Service. 2019 Oct. URL: https://www.ihs.gov/newsroom/factsheets/disparities/ [accessed 2022-01-02]
4. Zephier Olson MD, Dombrowski K. A systematic review of Indian boarding schools and attachment in the context of substance use studies of Native Americans. J Racial Ethn Health Disparities 2020 Feb;7(1):62-71. [doi: 10.1007/s40615-019-00634-4] [Medline: 31452147]
5. Breathett K, Sims M, Gross M, Jackson EA, Jones EJ, Navas-Acien A, American Heart Association Council on Epidemiology and Prevention, Council on Quality of Care and Outcomes Research, Council on Cardiovascular and Stroke Nursing, Council on Clinical Cardiology, Council on Lifestyle and Cardiometabolic Health. Cardiovascular health in American Indians and Alaska Natives: a scientific statement from the American Heart Association. Circulation 2020 Jun 23;141(25):e948-e959 [FREE Full text] [doi: 10.1161/CIR.0000000000000773] [Medline: 32460555]
6. Godfrey TM, Cordova-Marks FM, Jones D, Melton F, Breathett K. Metabolic syndrome among American Indian and Alaska Native populations: implications for cardiovascular health. Curr Hypertens Rep 2022 May 18;24(5):107-114. [doi: 10.1007/s11906-022-01178-5] [Medline: 35181832]
7. Warne D, Lajimodiere D. American Indian health disparities: psychosocial influences. Soc Personal Psychol Compass 2015 Oct 05;9(10):567-579. [doi: 10.1111/spc3.12198]
8. Doshi S, Jordan A, Kelly K, Solomon D. The COVID-19 response in Indian Country. Center for American Progress. 2020 Jun 18. URL: https://www.americanprogress.org/article/covid-19-response-indian-country/ [accessed 2022-02-25]
9. Hill L, Artiga S. COVID-19 vaccination among American Indian and Alaska Native people. Kaiser Family Foundation. 2021 Apr 09. URL: https://www.kff.org/racial-equity-and-health-policy/issue-brief/covid-19-vaccination-american-indian-alaska-native-people/ [accessed 2022-01-02]
10. Haroz EE, Kemp CG, O'Keefe VM, Pocock K, Wilson DR, Christensen L, et al. Nurturing innovation at the roots: the success of COVID-19 vaccination in American Indian and Alaska Native communities. Am J Public Health 2022 Mar;112(3):383-387. [doi: 10.2105/AJPH.2021.306635] [Medline: 35196058]
11. Foxworth R, Redvers N, Moreno MA, Lopez-Carmen VA, Sanchez GR, Shultz JM. COVID-19 vaccination in American Indians and Alaska Natives - lessons from effective community responses. N Engl J Med 2021 Dec 23;385(26):2403-2406. [doi: 10.1056/NEJMp2113296] [Medline: 34919356]
12. Nagle R. Native Americans being left out of US coronavirus data and labelled as 'other'. The Guardian. 2020 Apr 24. URL: https://www.theguardian.com/us-news/2020/apr/24/us-native-americans-left-out-coronavirus-data [accessed 2022-03-23]

XSL•FO
RenderX

13. Kral I. Youth media as cultural practice: remote Indigenous youth speaking out loud. Aust Aborig Stud 2011 Jan 01;1:4-16 [FREE Full text] [doi: 10.3316/INFORMIT.245205197181798]
14. Rice ES, Haynes E, Royce P, Thompson SC. Social media and digital technology use among Indigenous young people in Australia: a literature review. Int J Equity Health 2016 May 25;15:81 [FREE Full text] [doi: 10.1186/s12939-016-0366-0] [Medline: 27225519]
15. Social Distance Powwow. Facebook. URL: https://www.facebook.com/groups/socialdistancepowwow/ [accessed 2022-03-23]
16. American Indian COVID-19 Resources & Responses. Facebook. URL: https://www.facebook.com/groups/259818782083254/ [accessed 2022-03-23]
17. Kuhn N, Sarkar S, White LA, Hoy J, McCray C, Lefthand-Begay C. Decolonizing risk communication: Indigenous responses to COVID-19 using social media. J Indig Soc Dev 2020 Nov 03;9(3):193-213 [FREE Full text]
18. Lumpkins CY, Goeckner R, Hale J, Lewis C, Gunville J, Gunville R, et al. In our sacred voice - an exploration of tribal and community leader perceptions as health communicators of disease prevention among American Indians in the plains. Health Commun 2022 Aug;37(9):1180-1191. [doi: 10.1080/10410236.2021.2008108] [Medline: 34949125]
19. Crisis emergency risk communication: social media and mobile media devices. Centers for Disease Control and Prevention. 2014. URL: https://emergency.cdc.gov/cerc/ppt/CERC_Social%20Media%20and%20Mobile%20Media%20Devices.pdf [accessed 2022-11-07]
20. Hornmoen H, Backholm K. Social media use in crises and risks: an introduction to the collection. In: Hornmoen H, Backholm K, editors. Social Media Use in Crisis and Risk Communication. Bingley, United Kingdom: Emerald Publishing Limited; Oct 1, 2018:1-12.
21. Choi W, Stvilia B. Web credibility assessment: conceptualization, operationalization, variability, and models. J Assoc Inf Sci Technol 2015 May 13;66(12):2399-2414. [doi: 10.1002/asi.23543]
22. Kim H, Han JY, Seo Y. Effects of Facebook comments on attitude toward vaccines: the roles of perceived distributions of public opinion and perceived vaccine efficacy. J Health Commun 2020 Feb 01;25(2):159-169. [doi: 10.1080/10810730.2020.1723039] [Medline: 32026756]
23. Simeon R, Dewidar O, Trawin J, Duench S, Manson H, Pardo Pardo J, et al. Behavior change techniques included in reports of social media interventions for promoting health behaviors in adults: content analysis within a systematic review. J Med Internet Res 2020 Jun 11;22(6):e16002 [FREE Full text] [doi: 10.2196/16002] [Medline: 32525482]
24. 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;20(8):959-973. [doi: 10.1080/14760584.2021.1949292] [Medline: 34192985]
25. Tesoriero JM, Yuan Y, Newport R, O'Grady T, Cotroneo R, Stevens L, et al. Assessing the impact of a PrEP Aware Week campaign on PrEP prescription fills in NYS. J Public Health Manag Pract 2022 Sep;28(5):463-468. [doi: 10.1097/PHH.0000000000001554] [Medline: 35867501]
26. Vaccine misinformation management field guide. United Nations Children's Fund. 2020 Dec. URL: https://vaccinemisinformation.guide [accessed 2022-11-07]
27. Knowledge center. Tribal Advisory Board Map. Johns Hopkins Center for Indigenous Health. URL: https://cih.jhu.edu/what-we-do/resources/tribal-advisory-board-map/ [accessed 2022-03-23]
28. Resource Library. Johns Hopkins Center for Indigenous Health. URL: http://cih.jhu.edu/knowledge-center/resources/ [accessed 2022-11-07]
29. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
30. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010 Mar 01;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
31. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: https://www.r-project.org/ [accessed 2022-11-07]
32. Wickham H, François R, Henry L, Müller K. dplyr: a grammar of data manipulation. Tidyverse. 2022. URL: https://dplyr.tidyverse.org [accessed 2022-11-16]
33. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag; 2016.
34. Tingey L, Cwik MF, Rosenstock S, Goklish N, Larzelere-Hinton F, Lee A, et al. Risk and protective factors for heavy binge alcohol use among American Indian adolescents utilizing emergency health services. Am J Drug Alcohol Abuse 2016 Nov;42(6):715-725. [doi: 10.1080/00952990.2016.1181762] [Medline: 27315008]
35. Garroutte EM, Goldberg J, Beals J, Herrell R, Manson SM, AI-SUPERPFP Team. Spirituality and attempted suicide among American Indians. Soc Sci Med 2003 Apr;56(7):1571-1579. [doi: 10.1016/s0277-9536(02)00157-0] [Medline: 12614706]
36. Murthy BB, LeBlanc TT, Vagi SJ, Avchen RN. Going viral: the 3 Rs of social media messaging during public health emergencies. Health Secur 2021 Feb;19(1):75-81 [FREE Full text] [doi: 10.1089/hs.2020.0157] [Medline: 33372847]
37. Rabarison KM, Croston MA, Englar NK, Bish CL, Flynn SM, Johnson CC. Measuring audience engagement for public health Twitter chats: insights from #LiveFitNOLA. JMIR Public Health Surveill 2017 Jun 08;3(2):e34 [FREE Full text] [doi: 10.2196/publichealth.7181] [Medline: 28596149]

XSL•FO
RenderX

38. Thomas TH, Nauth-Shelley K, Thompson MA, Attai DJ, Katz MS, Graham D, et al. The needs of women treated for ovarian cancer: results from a #gyncsm Twitter chat. J Patient Cent Res Rev 2018 Apr 26;5(2):149-157 [FREE Full text] [doi: 10.17294/2330-0698.1592] [Medline: 31413999]
39. Carroll CL, Bruno K, Ramachandran P. Building community through a #pulmcc Twitter chat to advocate for pulmonary, critical care, and sleep. Chest 2017 Aug;152(2):402-409. [doi: 10.1016/j.chest.2017.03.003] [Medline: 28302497]
40. Buffard J, Papasava A. A quantitative study on the impact of emotion on social media engagement and conversion. J Digit Soc Media Mark 2020 Mar 01;7(4):1-21 [FREE Full text]
41. Loft LH, Pedersen EA, Jacobsen SU, Søborg B, Bigaard J. Using Facebook to increase coverage of HPV vaccination among Danish girls: an assessment of a Danish social media campaign. Vaccine 2020 Jun 26;38(31):4901-4908. [doi: 10.1016/j.vaccine.2020.04.032] [Medline: 32362529]
42. Ramirez A, Despres C, Chalela P, Weis J, Sukumaran P, Munoz E, et al. Pilot study of peer modeling with psychological inoculation to promote coronavirus vaccination. Health Educ Res 2022 Mar 23;37(1):1-6 [FREE Full text] [doi: 10.1093/her/cyab042] [Medline: 35048119]
43. Dutta-Bergman MJ. Theory and practice in health communication campaigns: a critical interrogation. Health Commun 2005;18(2):103-122. [doi: 10.1207/s15327027hc1802_1] [Medline: 16083406]
44. Britt RK, Britt BC, Anderson J, Fahrenwald N, Harming S. "Sharing Hope and Healing": a culturally tailored social media campaign to promote living kidney donation and transplantation among Native Americans. Health Promot Pract 2021 Nov;22(6):786-795. [doi: 10.1177/1524839920974580] [Medline: 33267677]
45. Cucu E. 2022 Social media industry benchmarks - know exactly where you stand in your market. Social Insider. 2022 Feb 24. URL: https://www.socialinsider.io/blog/social-media-industry-benchmarks/ [accessed 2022-09-07]
46. Gittelsohn J, Jock B, Poirier L, Wensel C, Pardilla M, Fleischhacker S, et al. Implementation of a multilevel, multicomponent intervention for obesity control in Native American communities (OPREVENT2): challenges and lessons learned. Health Educ Res 2020 Jun 01;35(3):228-242 [FREE Full text] [doi: 10.1093/her/cyaa012] [Medline: 32413105]
47. Patten CA, Lando H, Resnicow K, Decker PA, Smith CM, Hanza MM, et al. Developing health communication messaging for a social marketing campaign to reduce tobacco use in pregnancy among Alaska Native women. J Commun Healthc 2018 Jul 16;11(4):252-262 [FREE Full text] [doi: 10.1080/17538068.2018.1495929] [Medline: 31548863]
48. Patten CA, Lando H, Resnicow K, Decker PA, Smith CM, Hanza MM, et al. Developing health communication messaging for a social marketing campaign to reduce tobacco use in pregnancy among Alaska Native women. J Commun Healthc 2018 Jul 16;11(4):252-262 [FREE Full text] [doi: 10.1080/17538068.2018.1495929] [Medline: 31548863]
49. Schreiner M, Fischer T, Riedl R. Impact of content characteristics and emotion on behavioral engagement in social media: literature review and research agenda. Electron Commer Res 2019 May 10;21(2):329-345. [doi: 10.1007/s10660-019-09353-8]
50. Su LYF, McKasy M, Cacciatore MA, Yeo SK, DeGrauw AR, Zhang JS. Generating science buzz: an examination of multidimensional engagement with humorous scientific messages on Twitter and Instagram. Sci Commun 2021 Dec 31;44(1):30-59. [doi: 10.1177/10755470211063902]
51. Farinella M. The potential of comics in science communication. J Sci Commun 2018 Jan 23;17(01):Y01. [doi: 10.22323/2.17010401]
52. Saunkeah B, Beans JA, Peercy MT, Hiratsuka VY, Spicer P. Extending research protections to tribal communities. Am J Bioeth 2021 Oct;21(10):5-12 [FREE Full text] [doi: 10.1080/15265161.2020.1865477] [Medline: 33449863]
53. Guadagnolo BA, Cina K, Helbig P, Molloy K, Reiner M, Cook EF, et al. Medical mistrust and less satisfaction with health care among Native Americans presenting for cancer treatment. J Health Care Poor Underserved 2009 Feb;20(1):210-226 [FREE Full text] [doi: 10.1353/hpu.0.0108] [Medline: 19202258]
54. Hinzo AM, Clark LS. Digital survivance and Trickster humor: exploring visual and digital Indigenous epistemologies in the #NoDAPL movement. Information, Communication & Society 2019 Feb 17;22(6):791-807. [doi: 10.1080/1369118x.2019.1573911]
55. Dean RA. Native American humor: implications for transcultural care. J Transcult Nurs 2003 Jan 29;14(1):62-65. [doi: 10.1177/1043659602238352] [Medline: 12593272]
56. Kalbfleisch PJ. Effective health communication in Native populations in North America. J Lang Soc Psychol 2009 May 04;28(2):158-173. [doi: 10.1177/0261927x08330607]
57. Curtis ME, Clingan SE, Guo H, Zhu Y, Mooney LJ, Hser Y. Disparities in digital access among American rural and urban households and implications for telemedicine-based services. J Rural Health 2022 Jun 06;38(3):512-518. [doi: 10.1111/jrh.12614] [Medline: 34355427]
58. Duarte ME, Vigil-Hayes M, Zegura E, Belding E, Masara I, Nevarez JC. As a squash plant grows: social textures of sparse internet connectivity in rural and tribal communities. ACM Trans Comput Hum Interact 2021 Jun 30;28(3):1-16. [doi: 10.1145/3453862]
59. Sanders CK, Scanlon E. The digital divide is a human rights issue: advancing social inclusion through social work advocacy. J Hum Rights Soc Work 2021 Mar 19;6(2):130-143 [FREE Full text] [doi: 10.1007/s41134-020-00147-9] [Medline: 33758780]
60. Oremus W, Alcantara C, Merrill JB, Galocha A. How Facebook shapes your feed. Washington Post. 2021 Oct 26. URL: https://www.washingtonpost.com/technology/interactive/2021/how-facebook-algorithm-works/ [accessed 2022-03-02]

XSL•FO
RenderX

61. Stepanov A. Changes to News Feed in 2021. Meta. 2021 Dec 20. URL: https://about.fb.com/news/2021/12/changes-to-news-feed-in-2021/ [accessed 2022-03-02]
62. Demographics of social media users and adoption in the United States. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2022-11-07]
63. Jones N, Marks R, Ramirez R, Ríos-Vargas M. 2020 Census illuminates racial and ethnic composition of the country. United States Census Bureau. 2021 Aug 12. URL: https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html [accessed 2022-03-02]
64. Rushing SC, Stephens D. Use of media technologies by Native American teens and young adults in the Pacific Northwest: exploring their utility for designing culturally appropriate technology-based health interventions. J Prim Prev 2011 Aug;32(3-4):135-145. [doi: 10.1007/s10935-011-0242-z] [Medline: 21805055]

Abbreviations

AI/AN: American Indian or Alaska Native
CIH: Johns Hopkins Center for Indigenous Health
ER: engagement rate
HBM: Health Belief Model
TAB: Tribal Advisory Board

Edited by T Mackey; submitted 05.04.22; peer-reviewed by A Calac, W Buente, N Hu; comments to author 18.07.22; revised version received 30.09.22; accepted 22.10.22; published 25.11.22.

Please cite as:
Weeks R, White S, Hartner AM, Littlepage S, Wolf J, Masten K, Tingey L
COVID-19 Messaging on Social Media for American Indian and Alaska Native Communities: Thematic Analysis of Audience Reach and Web Behavior
JMIR Infodemiology 2022;2(2):e38441
URL: https://infodemiology.jmir.org/2022/2/e38441
doi:10.2196/38441
PMID:36471705

©Rose Weeks, Sydney White, Anna-Maria Hartner, Shea Littlepage, Jennifer Wolf, Kristin Masten, Lauren Tingey. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 25.11.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Unmasking the Twitter Discourses on Masks During the COVID-19 Pandemic: User Cluster–Based BERT Topic Modeling Approach

Weiai Wayne Xu[1], PhD; Jean Marie Tshimula[2], PhD; Ève Dubé[3,4], PhD; Janice E Graham[5], PhD; Devon Greyson[6], PhD; Noni E MacDonald[5], MD; Samantha B Meyer[7], PhD

[1]Department of Communication, University of Massachusetts Amherst, Amherst, MA, United States
[2]Department of Computer Science, Université de Sherbrooke, Sherbrooke, QC, Canada
[3]Axe maladies infectieuses et immunitaires, Centre de Recherche du CHU de Québec, Laval University, Quebec City, QC, Canada
[4]Direction des risques biologiques et de la santé au travail, Institut National de Santé Publique du Québec, Quebec, QC, Canada
[5]Department of Pediatrics, Dalhousie University, Halifax, NS, Canada
[6]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada
[7]School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Weiai Wayne Xu, PhD
Department of Communication
University of Massachusetts Amherst
650 N Pleasant St
Amherst, MA, 01003
United States
Phone: 1 (413) 545 1311
Email: weiaixu@umass.edu

Abstract

Background: The COVID-19 pandemic has spotlighted the politicization of public health issues. A public health monitoring tool must be equipped to reveal a public health measure's political context and guide better interventions. In its current form, infoveillance tends to neglect identity and interest-based users, hence being limited in exposing how public health discourse varies by different political groups. Adopting an algorithmic tool to classify users and their short social media texts might remedy that limitation.

Objective: We aimed to implement a new computational framework to investigate discourses and temporal changes in topics unique to different user clusters. The framework was developed to contextualize how web-based public health discourse varies by identity and interest-based user clusters. We used masks and mask wearing during the early stage of the COVID-19 pandemic in the English-speaking world as a case study to illustrate the application of the framework.

Methods: We first clustered Twitter users based on their identities and interests as expressed through Twitter bio pages. Exploratory text network analysis reveals salient political, social, and professional identities of various user clusters. It then uses BERT Topic modeling to identify topics by the user clusters. It reveals how web-based discourse has shifted over time and varied by 4 user clusters: conservative, progressive, general public, and public health professionals.

Results: This study demonstrated the importance of a priori user classification and longitudinal topical trends in understanding the political context of web-based public health discourse. The framework reveals that the political groups and the general public focused on the science of mask wearing and the partisan politics of mask policies. A populist discourse that pits citizens against elites and institutions was identified in some tweets. Politicians (such as Donald Trump) and geopolitical tensions with China were found to drive the discourse. It also shows limited participation of public health professionals compared with other users.

Conclusions: We conclude by discussing the importance of a priori user classification in analyzing web-based discourse and illustrating the fit of BERT Topic modeling in identifying contextualized topics in short social media texts.

(JMIR Infodemiology 2022;2(2):e41198) doi:10.2196/41198

KEYWORDS

infoveillance; data analytics; Twitter; social media; user classification; COVID-19

XSL•FO
RenderX

Introduction

Background

The COVID-19 pandemic is a crisis that has taken millions of lives, devastated the world economy, and disrupted almost every aspect of human society. Mask wearing is one of the few early and effective nonpharmaceutical interventions to curb the spread of the virus [1,2]. However, public health efforts to mandate or recommend mask wearing have been met with public skepticism [3], and in some cases, outright resistance. This could be a result of mixed messaging—early in the pandemic, some public health institutions (eg, World Health Organization and US Centers for Disease Control and Prevention) and media advised against mask wearing, citing concerns regarding mask shortage for health care workers and the efficacy of masks [4]. It could also have resulted from widespread unproven medical claims from many conservative media outlets and influencers [5]. The effectiveness of mask wearing to prevent transmission of SARS-CoV-2 has been much debated as the scientific literature has evolved rapidly, and messages from official government and medical advisory bodies have been mixed since the early days of this pandemic; it is also likely that fierce antimask sentiment more closely reflects deeply rooted anti-Asian racism and xenophobia [6], as well as populist and nativist resentments [7]. Populist leaders and parties sought to politicize mask wearing calling the public health response to the pandemic government overreach and a conspiracy [8]. Armed protests against mask wearing were held across US cities. National surveys demonstrate a clear link between political-right partisanship and Christian-nationalist ideologies and resistance to government-mandated COVID-19 restrictions [9]. Understanding the political context within which public health measures and messaging are being implemented is critical to maximizing the success of attempts to protect population health. Infoveillance based on web-based discourse provides ways to understand the political nature and implications of public health issues. Although there is a growing body of infoveillance studies that leverage the latest digital analytic tools to document and compare public health discourse, we notice several gaps. This project seeks to present an improved infoveillance framework to understand public health discourse varied by political and apolitical groups. This paper was organized as follows. We first situate the case study of mask wearing in the context of medical populism, followed by the introduction of infoveillance. We then proceed to 2 existing gaps in the existent literature, leading to our proposed computational framework.

Medical Populism

There are growing calls to study the politicization of public health issues to understand competing interests and ideologies in public health measures. The COVID-19 pandemic presents an interest case of *medical populism* [7,10,11], defined as "as a political style based on performances of public health crises that pit 'the people' against the dangerous others, which consists of 'the establishment'" [11]. A common thread in populism is the dichotomy between virtuous people and the elite or establishment, which is perceived as corrupt [12]. In the medical populism regarding Ebola, HIV, drug addiction [11], and the antivaccination movement [13], the medical and scientific communities are framed as elites to be blamed and distrusted. Recent surveys show that populist ideology is associated with a higher degree of distrust in political and scientific institutions, leading to a heightened acceptance of COVID-19–related conspiracy theories [14], with such distrust associated with a lower level of education, health literacy, and the use of logic thinking [15,16]. Such distrust of elites and institutions is fertile ground for those peddling alternative and unproven medicines such as hydroxychloroquine, which have been endorsed by populist leaders including Donald Trump and Jair Bolsonaro [17]. Medical populism breeds disinformation and misinformation, which is made worse by viral transmission on social media [18]. Although not unique to the COVID-19 pandemic—misinformation was rampant during the flu pandemic [19], as well as Zika [20] and Ebola [18] outbreaks—the level of politicization and social media involvement led the World Health Organization to establish a task force on the infodemic [21], and some experts call the COVID-19 pandemic the first true social media infodemic [22].

Infoveillance and the COVID-19 Pandemic

Social media provides the public with fodder for civic deliberations and actions. A wealth of research theorizes social media's role as a mediated public sphere or the nexus of networked societies [23]. This trove of social media data, indicative of public attention, attitudes, and actions, can be readily tapped into for infoveillance. Infoveillance is a methodological framework that uses large-scale digital behavioral data to monitor outbreaks and public perceptions of public health issues [24-26]. There are successful implementations of infoveillance in past epidemic outbreaks, including Ebola [27], Zika [28,29], and H1N1 influenza [30].

Our review of the growing body of infoveillance research since the COVID-19 outbreak revealed 3 common themes. First, studies using data from the early stage of the COVID-19 outbreak aim to detect linguistic and content features in social media texts that are predictive of COVID-19 symptoms [31,32]. This approach is in line with traditional infoveillance projects, such as the pioneer, albeit flawed, Google Flu Trends, which became a famous example of "big data hubris" after initially appearing to predict influenza prevalence faster than traditional public health surveillance methods [33].

Second, as public conversations broadened, later studies used latent Dirichlet allocation (LDA) topic modeling to reach beyond mere mention-counting to identifying themes in web-based discourse. Chandrasekaran et al [24] identified COVID-19–related economic impacts, virus spreads, treatment and recovery, impact on the health care sector, and government's response. Abd-Alrazaq et al [34] identified themes surrounding the origin of the virus; the impacts of COVID-19 on people, countries, and the economy; and mitigation and prevention. Similarly, Wahbeh et al [35] identified topics in digital texts that revolve around actions and recommendations, misinformation, knowledge, the health care system, symptoms and illness, immunity, testing, and infection and transmission. Although most studies relied on Twitter data, a few used Weibo, the Chinese microblogging site [32,36-38]. Weibo data revealed uncertainty and changing attitudes about the COVID-19

XSL•FO
RenderX

pandemic by the Chinese public in the early days of the outbreak [36]. Prior works also examined user sentiments [24,35,39]. Zhou et al [40] exemplify this, using Weibo data to monitor Chinese public response to lockdowns and how negative sentiments such as panic evolved.

Third, most prior works examine general discourses. Al-Ramahi et al [3] identify major themes in the antimask discourse, including constitutional rights and freedom of choice; COVID-19–related conspiracy theory, population control, and big pharma; and fake news, fake numbers, and fake pandemic. Relatedly, Doogan et al [41] tracked public responses to mask wearing and social distancing across 6 countries, finding that attention paid to public health measures correlated with case numbers. These studies, along with computational text analysis of news content from traditional media [42], contribute to the growing understanding of the interplay between public health and public opinion.

Gaps in Infoveillance Studies: Ideology and Identity Politics

Audience segmentation is a popular method of understanding the complexity and diversity of the user ecosystem in web-based discourse. In general information science studies, data-driven personas play a vital role in predicting and aggregating user behaviors [43]. The data-driven personas approach includes using various social media data streams and interaction patterns to cluster users based on demographic factors and interests [44]. The approach also applies to the public health domain, such as using survey data to generate psychological and demographic profiles of the public in adopting COVID-19 recommendations [45]. In the infoveillance literature, there has been some research on discourse by different users across geographic locations and with different health profiles [46,47] as well as in various health care sectors [35]. However, these are limited attention to politically and socially defined user clusters including those associated with medical populism. A few exceptions include the study by Walter et al [48] of the Twitter discourse on vaccines. The authors used unsupervised machine learning and network analysis to identify politically different “thematic personas” and subsequently analyzed content by each thematic persona. This study took a similar approach, albeit with new analytic tools, to explore the political nature of public health discourse and users who participate in the discourse. This entails moving beyond the general discourse and focusing on specific user groups that differ by politics and interests.

Internet users, such as offline publics, commonly seek support and influence by forming close-knit and like-minded communities. We borrowed the term *issue publics* from the general social science literature to refer to web-based user clusters connected through common backgrounds, hobbies, interests, and ideologies [23]. Users connect not only through social media following and follower linkages but also, more broadly, through symbolic connective actions such as hashtags [49]. For instance, users who identified with a social cause or political party use shared hashtags (eg, #ChinaVirus or #KungFlu) as a form of expression, resistance, and solidarity building. Hashtags connect ideologically similar causes and weave disconnected local concerns and identities into a global narrative [49-51].

Previous studies view these politically connected user groups as *ad hoc issue publics* [23], *networked counterpublics* [50,51], or *countercoalitions* [52]. These user groups form quickly in response to developing news, emergent social movements, or long-held belief and social identities. They are decentralized, geographically distributed, and marked by coordinated sharing and discussions [23]. They consist of different institutional and individual stakeholders across public spheres, characterized by various levels of internal coordination and committed participation [52]. Terminology aside, the assumption is straightforward: the digital space is a web-based public square consisting of different user groups who have competing interests and ideologies. To understand how publics perceive public health measures, one must extract and triangulate discourse from each specific user group (ie, issue public).

Gaps in the Current Infoveillance Studies: From LDA to BERT Topic Modeling

Current infoveillance studies overwhelmingly use LDA and sentiment scoring [53-54]. The reliance on LDA is not surprising, given it is the most popular and widely used topic model [55-58]. LDA is a probabilistic model that discovers latent topics in a text corpus and can be trained using collapsed Gibbs sampling [55,59,60]. Specifically, LDA assumes K underlying topics, each of which is a distribution over a fixed vocabulary. Although LDA is reputed to yield promising results in modeling text corpora [61], it fundamentally suffers from several shortcomings, including difficulty in setting the parameter *k*, which refers to the number of topics to yield semantically meaningful results, a deficiency in handling short texts [58], in capturing the contextual meaning of sentences [58], as well as its inability to model topic correlations and the evolution of topics over time [62].

To overcome these limitations, the new generation of topic models [56,57,61] use pretrained representations such as BERT to enable topic modeling (1) to consider contextual meaning of sentences for supporting the results to match the adequate topics and (2) to include more features for efficiently modeling topic correlations and topic evolution over time. Recent pretrained contextualized representations such as BERT have pushed the state of the art in several areas of natural language processing due to their ability to expressively represent complex semantic relationships from being trained on massive data sets. BERT is a bidirectional transformer-based pretrained contextual representation using masked language modeling objective and next sentence prediction tasks [62]. The significant advantage of BERT is that it simultaneously gains the context of words from both left and right context in all layers. To this end, BERT uses a multilayer bidirectional transformer encoder, where each layer contains multiple attention heads.

It is important to note that BERT is one of the latest unsupervised topic modeling techniques that seek to improve upon the traditional LDA approach. An alternative technique, the Analysis of Topic Model Networks (ANTMN), applies community-detection algorithms in network analysis to cluster LDA-generated topics [63]. ANTMN is a fitting tool for

XSL•FO
RenderX

revealing framing in web-based and news discourse and has been used in studying public health discourse [64]. Another alternative is the semantic network–based classification algorithm textnets [65], which first uses LDA to cluster corpus into topics and then applies community-detection algorithms to categorize topics into network clusters. Although ANTMN and textnets are much-improved tools compared with the traditional LDA, we opted for BERT because BERT can reveal longitudinal topic trends, which is a feature not available in textnets, making BERT ideal for studying the ebbs and flows of specific topics in web-based discourse over time.

Research Questions

This paper used topic modeling with BERT to overcome the incompatibility between traditional LDA methods and short texts (eg, tweets) and track topical evolutions longitudinally. In addition, we investigated discourses and topics unique to different user groups (ie, issue publics). This approach aimed to understand the role of political ideologies and political groups in defining the public health discourse.

Research question 1: How did English language Twitter discourse on masks and mask wearing change over the course of 2020?

Research question 2: How did English language Twitter discourse on masks and mask wearing vary across issue publics?

Methods

With the focus on distinct user groups (ie, issue publics) and the state-of-art BERT Topic modeling application, this paper sought to present an infoveillance workflow consisting of data collection, data cleaning, and user classification and topic modeling.

Data Collection

This study uses a large-scale COVID-19 Twitter corpus provided by Georgia State University's Panacea Lab [66]. The corpus contains publicly available tweets from the Twitter Stream application programming interface (API) with the following keywords: "COVD19," "CoronavirusPandemic," "COVID-19," "2019nCoV," "CoronaOutbreak," "coronavirus," and "WuhanVirus." We used a modified Python script to hydrate all COVID-19–related tweets sent between January 1 and December 31, 2020, based on the tweet IDs provided in the public data set.

Data Cleaning

To track longitudinal trends associated with changes in developments related to the COVID-19 pandemic, our research team divided the COVID-19 data chronologically. Stage 1 spans the period from January 1 to April 30, 2020, including the early outbreak in China, subsequent travel restrictions and lockdowns first by China and later western democracies, and the emergent shortage of personal protective equipment (PPE). The cutoff date of stage 1 corresponds to when some major US states, such as Texas and Florida, started to relax public health measures in an effort to reopen the economy. Stage 2 spans May 1 to August 31, 2020, including events such as recommended or mandated mask wearing, armed protests against public health measures in the United States, controversial remarks about the COVID-19 pandemic by politicians, and the worsening pandemic in the English-speaking world. Stage 3 spans September 1 to December 31, 2020, during which significant political events in the United States include President Trump's October contraction of COVID-19, the November presidential election, and the national vaccination campaigns. To identify the mask-related discourse, the following keyword filters were used: *mask*, *face cover*, *facecover*. To achieve computational efficiency (running BERT Topic models on a large corpus is time-consuming), we only kept English language tweets that received at least 1 retweet by other users to focus on tweets that are actually promoted to a wider audience. We also excluded tweets sent by users with blank Twitter user bio pages (to be explained in the *User Classification* section).

User Classification

To identify the user classifications, we applied the k-means clustering algorithm [67] to Twitter users' bio descriptions to classify users based on expressed identities and interests. With the focus on clusters of users who have expressed common interests and identities, users who had blank Twitter bios (0.54% of the total users) were excluded from the analysis. Although this exclusion may affect the representativeness of the discourse under study, we argue that users who use a common set of hashtags and terms in Twitter bios are more engaged (topically, socially, or politically) in this digital public square.

The k-means clustering algorithm was applied to yield 10 clusters. The algorithm put users who used similar words or phrases in Twitter bios in the same group, with the number of clusters (ie, 10) and the size of each cluster determined by the k-means algorithm. Researchers manually inspected the clustering output and removed 2 clusters mostly associated with news media and official sources (eg, Centers for Disease Control and Prevention, city and county governments) due to this study's focus on citizens' discourse. The remaining 8 clusters were reduced to 4 clusters based on topical similarity and political affiliation. The *general* cluster includes users whose Twitter bio descriptions indicate various social, professional interests, affiliations, and identities without sign of political affiliation. The *conservative* cluster includes users who use keywords and hashtags that indicate their conservative ideologies and support of the Trump administration (eg, #maga, #kag, #2a, or #prolife). The *progressive* cluster includes users who use hashtags and keywords reflecting a progressive ideology (eg, LGBTQI, Democrat, #BidenHarris2020, #Biden2020, or #BlackLivesMatter). Finally, the public health cluster includes users affiliated with the health care sector and public health research, as indicated by keywords such as *healthcare*, *science*, *epidemiologist*, *professor*, and *radiologist*. To provide descriptive findings on characteristics of each user cluster, we used a short text–classification algorithm called *textnets* to produce network visualizations of various phrases and hashtags used in Twitter bios [65]. This algorithm applies network analysis to natural language processing, providing an alternative to topic modeling for analyzing short texts such as Twitter bios. This approach can show latent identities, interests, and movements with which users in a particular cluster identify.

XSL•FO
RenderX

BERT Topic Modeling

We removed all English stop words in the data set, using the natural language toolkit. We noticed a ubiquitous presence of the words “mask,” “covering,” “cover,” “face cover,” and “face mask” in the learned topics because our data set contained mask-related discourse. Practically, these words were noisy and degraded the performance of proper topics and hindered the interpretability of results. To overcome this problem, we extended the natural language toolkit vocabulary by adding these words and removed them in our data set. To identify potential topics within the mask-related discourse, we applied the BERTopic, a BERT-based topic modeling Python library. BERTopic extracts document embeddings using a pretrained BERT model. We used the BERT topic model, which comprises 12 layers, 12 attention heads, and 110 million parameters, to enable BERTopic to produce document embeddings to detect semantic similarity between sentences. BERTopic leverages BERT embeddings and a class-based term frequency–inverse document frequency to create dense clusters to detect unique topics. In addition, BERTopic generates the topic representations at each timestamp for each topic. The traditional LDA topic modeling requires a predefined k (the number of topics) for algorithms to cluster corpus around k topics [68]. BERTopic does not require a predefined k, reducing the need for various iterations of model fine-tuning.

Ethics Approval

As we are using a publicly archived data set and no personally identifiable information is included nor published, we deem this research outside the purview of the institutional review board. Nevertheless, we have taken extra caution when analyzing each cluster’s user profiles to ensure that the reported data are aggregated and anonymous.

Results

Overview

With the mask-related keywords applied as filters, the raw data set includes 1,061,686 unique tweets by 648,528 unique users in stage 1, includes 1,060,987 tweets by 576,274 unique users in stage 2, and includes 678,474 unique tweets by 359,561 unique users. Among them, stage 1 had 171,271 English language unique tweets that were retweeted at least once by 115,349 users; stage 2 had 234,997 unique English tweets by 137,426 users, and stage 3 produced 129,089 tweets by 76,443 users. As noted earlier, we also excluded tweets sent by users with blank Twitter bio pages. The final tweet data set before the user-classification scheme and BERT Topic modeling were applied included 163,378 tweets by 109, 097 users in stage 1, included 224,830 tweets by 129,830 users in stage 2, and included 123,843 tweets by 72,495 users in stage 3. This result focuses on tweets sent by the 4 identified user clusters: general, progressive, conservative, and public health. Figure 1 shows the tweet volumes by each distinct user community over time. There is a marked peak in tweet volume on April 30, 2020, across user clusters. The time corresponds to prominent US politicians’ mask-wearing practices, such as then Vice President Mike Pence’s mask wearing on April 30 when visiting a factory and his widely criticized maskless visit to Mayo Clinic on April 28.

To carry out topical analysis by user clusters, we ran BERT Topic models for each previously identified user cluster. Note that in the topic models some tweets were found to have no coherent theme and thus assigned to the unclassified topic −1 (the nonthematic). Following the common practice suggested by the authors of BERTopic, we did not include such nonthematic tweets in the final analysis. Table 1 shows the number of nonthematic tweets in each user cluster and the calculated ratios of nonthematic tweets. The nonthematic ratios vary across user clusters and stages. This shows the potential limitations of this topic modeling approach in that it leaves out some percentages of the corpus due to incongruent themes. Nevertheless, the approach reveals the most salient part of the corpus with distinct themes. After identifying topics, the authors manually inspected the topics based on example tweets and created topical labels that describe the major themes in the tweets.

XSL•FO
RenderX

Figure 1. Volumes of topically classified tweets over time.

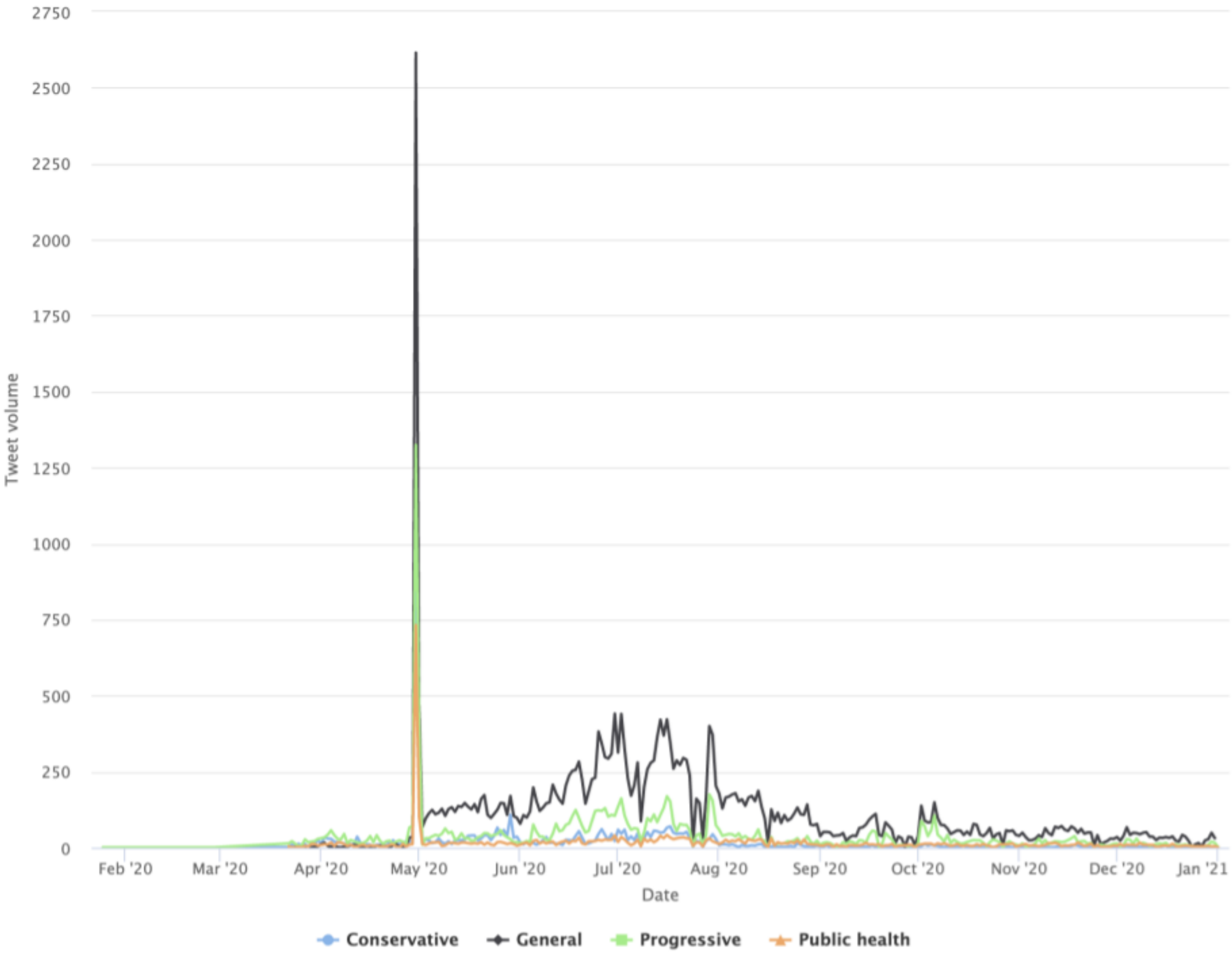


Table 1. Tweet count by user clusters.

User cluster	Stage	Number of nonthematic tweets	Total number of tweets included for modeling	Nonthematic tweet ratio
The conservative cluster	1	1041	3094	0.34
The general cluster	1	11,304	31,364	0.36
The progressive cluster	1	1210	4377	0.28
The public health cluster	1	764	1414	0.54
The conservative cluster	2	1565	4711	0.33
The general cluster	2	20,057	43,281	0.46
The progressive cluster	2	3475	10,462	0.33
The public health cluster	2	0	2300	0
The conservative cluster	3	129	430	0.30
The general cluster	3	5309	12,077	0.44
The progressive cluster	3	1120	3539	0.32
The public health cluster	3	0	983	0

Topics in the Conservative Cluster

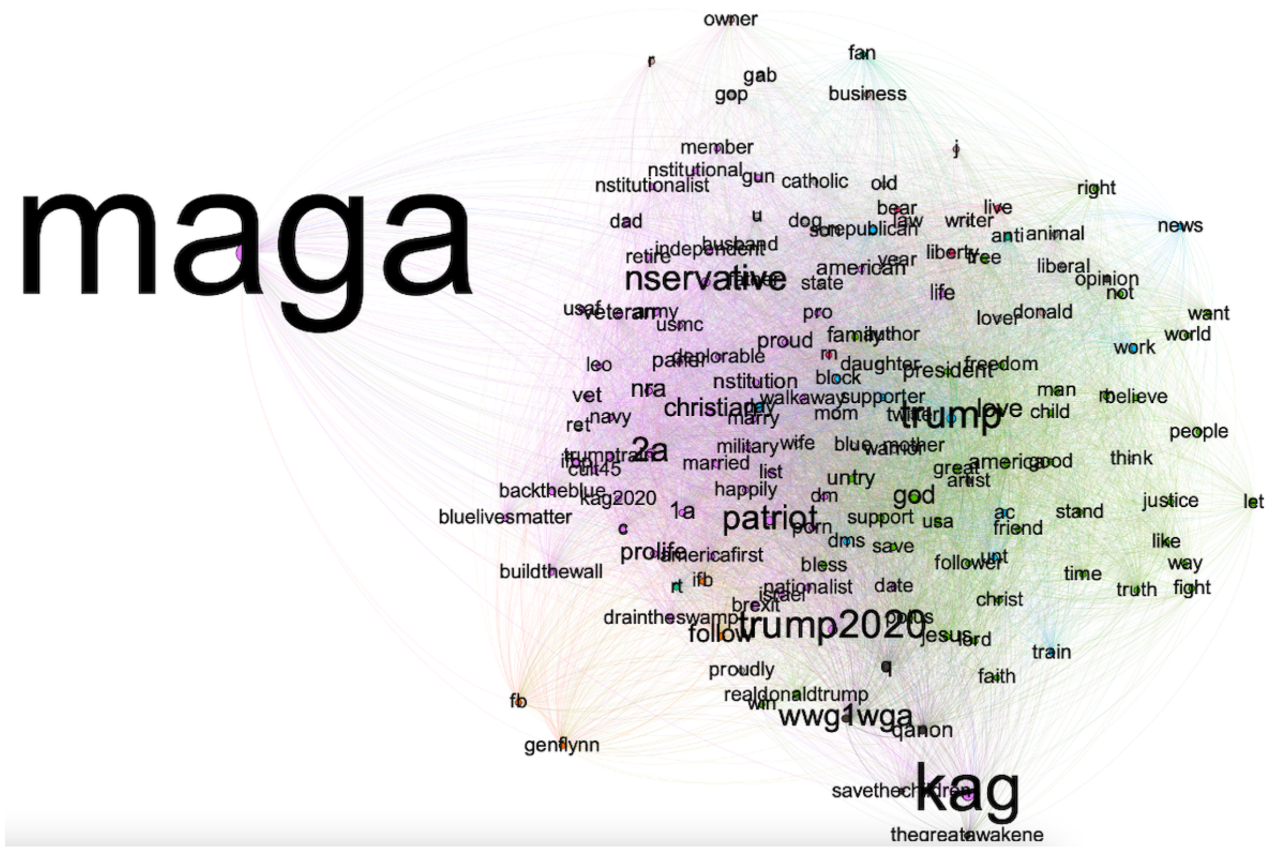
The conservative cluster consists of users whose Twitter bios include keywords such as *maga,kag*, *trump2020*, *trump*, *conserve*, *patriot,wwg1wga*, *2a*, *god*, *Christian*, *nra*, *prolif*, *qanon*, *1a*, *american*, *constitute*, *veteran*, *jesus*, *proud*, *country*, *presid*, *buildthewal*, *America*, *parler*, *militari*, *famili*, *kag2020*, *vet*, *draintheswamp*, *marri*, *deplor*, *q*, *americafirst*, *usa*, *backtheblu*, *wife*, *freedom*, *back*, *truth*, *retir*, *ifb*, *trumptrain*, *walkaway*, *dms*, etc. These words indicate their alliance with Donald Trump's campaigns, conservative causes, and religious identities. The cluster also seems US-centric, given that the most central keywords from Twitter bios are associated with US politics. Also notable is the cluster's tie with the fringe and

XSL•FO
RenderX

cult-like QAnon movement. The cluster produced 8235 tweets in stage 1, with approximately 33% of the tweets classified as nonthematic and not included in the following results. Among the tweets included and assigned topics, there are 3600 unique users. Many users (as identified by unique Twitter user IDs) no longer had an accessible Twitter bio in August 2022 (2408 out of 3600), suggesting either that they deleted their accounts or that Twitter suspended their accounts for suspicious activities. Note that tweets sent by suspended accounts and bots, albeit inauthentic, need to be included in the analysis because of their potential polarizing effects. Among those with valid Twitter bios, their follower counts range from 476,284 to 0, with a median of 3860.

The short text–classification algorithm, *textnets* algorithm, scans all key terms that appear at least twice on the conservative users' bios and creates a cooccurrence-based semantic network, as seen in Figure 2. Figure 2 shows the top 154 terms ranked by betweenness centrality, a network indicator of key terms' salience in the entire corpus. Colors in the network graphs indicate distinct thematic clusters. The network shows the central role of Trump-related terms, the purple-colored populist political movements (eg, #BacktheBlue, #AmericanFirst, #DrainTheSwamp, or #WWG1WGA), and the green-colored conservative evangelical community. Table 2 shows the 30 most mentioned locations in the Twitter profiles of this cluster. Note that the data contain user entries on the location fields of their Twitter profile pages. The location information is raw and unstandardized. Specifically, some users may enter detailed cities and states, whereas others may provide general terms such as *United States* or *Planet*. Some could even provide fake or user-created terms to convey one's politics and ideologies. Such terms include *Real America* or *Hell*. Therefore, the summary statistics about user location entries should be interpreted with caution. Nevertheless, the top entries in the field suggest that users are primarily based in the United States, notably in the most populous US states.

Figure 2. Central terms in Twitter bios of the conservative cluster.



XSL•FO
RenderX

Table 2. Top user-entered location information in the conservative cluster.

Location	Value of location
United States	406
Florida, United States	67
California, United States	67
Texas, United States	66
Georgia, United States	30
Virginia, United States	25
Michigan, United States	21
Texas	19
North Carolina, United States	19
Florida	19
Arizona, United States	19
New York, United States	17
Pennsylvania, United States	17
California,	16
Las Vegas, Nevada	15
San Diego, California	14
Missouri, United States	14
Tennessee, United States	13
South Carolina, United States	13
Kentucky, United States	12
Colorado, United States	12
Ohio, United States	11
Pacific Northwest	11
Louisiana, United States	11
Alabama, United States	11
Colorado,	11
Washington DC	10
Phoenix, Arizona	10

Figure 3 shows the 10 most salient (denoted by colors) topics in the conservative cluster in stage 1. The less-salient topics were still included in the visualization but grayed out. A small spike was found in early April concerning PPE shortage. The topic's popularity was overtaken in late April by a broad array of topics suggesting users' distrust of institutions and resistance to lockdown measures. This includes the third most prominent topic, labeled as "Distrust, plandemic, anti-lockdown," which peaked in late April 2020 (green colored in Figure 3). One redacted tweet in the topic reads "the who are globalists and are just playing their game f them do not trust they," which seems to capture the ethos of many similar tweets that display anger toward politicians and the elite. The antiestablishment sentiment is echoed by the fourth most prominent topic labeled "Anti-media and antielite," which peaked around the same time. The distrust of mainstream media such as CNN is exemplified by this redacted tweet "cnns lemon not holding coronavirus briefings part of the plan for you to think that this is over." In a similar vein, the fifth-most prominent topic (ie, Doubting COVID-19 death and antilockdown) shows doubts about official statistics on COVID-19 death, as reflected in this tweet example: "...more cancer patients will now die in england because the covid lockdown scared the patients from going to..." and "this is why cause of death is listed as covid even if someone dies of a heart attack it skews the numbers so noon." The timing of the aforementioned prominent topics also corresponds to the widely reported armed antilockdown protests in the US state of Michigan throughout mid and late April. This specific event is captured by the topic labeled "COVID and anti-lockdown protest in Michigan." In addition, much of the focus in late April was centered around the Chinese state's cover-up of the virus in the early days of the pandemic. One such tweet reads, "China knew of virus ability to spread but kept silent for days leaked documents..." Overall, topics prevalent at this stage align with the widely reported Conservatives' defiance of mask policies and their strong criticism of China in handling the pandemic.

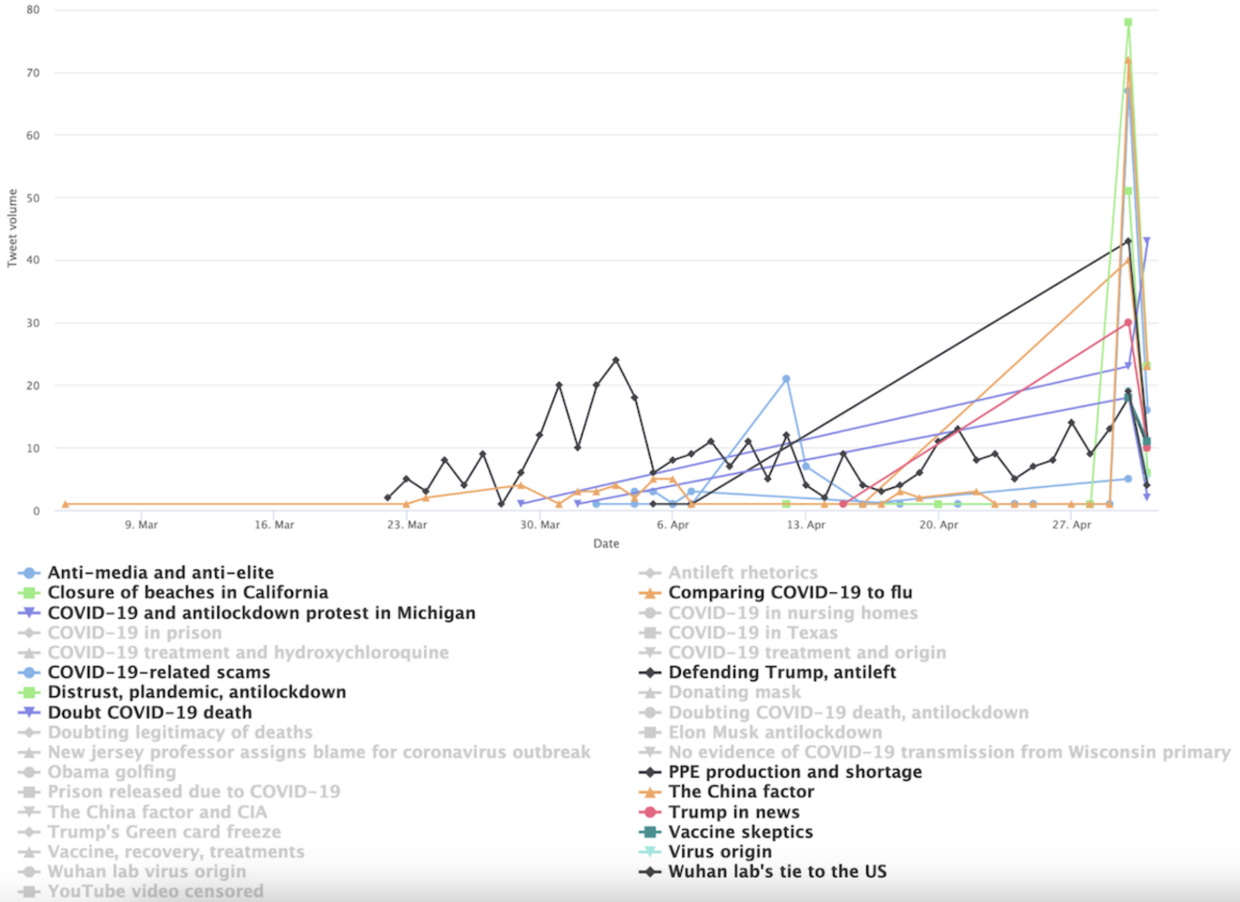
XSL•FO
RenderX

The less-prominent topics (the grayed-out topic labels), albeit comparatively small in tweet size, nevertheless shows a diverse range of concerns and interest among the conservative users, such as the alleged laboratory origin of the virus, alternative treatments such as hydroxychloroquine, and skepticism over vaccines.

The conservative cluster's topics are distinct in stage 2 (Figure 4). Notably, there are common topical clusters about risks associated with mask wearing and the effectiveness of mask wearing in preventing COVID-19. This topical cluster includes topics such as "Mask, risk, mask efficacy" and "Mask offers little protection" as well as the topic labeled "The science behind mask." Example tweets for these topics include "these face masks will not provide any protection against covid or other viruses or contaminants" and "a cloth mask is as effective fighting covid as a tube sock is preventing pregnancy." Disputes over masking are also seen in other prominent topics, such as "Mask-wearing dispute," which contains users' complaints of having to wear masks for grocery shopping and attending medical appointments. The topic labeled "Mask rules in business entities" includes tweets such as "sheeple are wearing masks like obedient sheep and now stores like walmart require a mask i feel like i am in Orwellian..." which is a clear indication of the users' resistance to mask wearing. Aside from the 10 most prominent topics, some grayed-out topics (the less-prominent ones by tweet volumes) show spikes and appear politically related. One such spike occurred on May 29, 2020, in relation to Dr Fauci, the lead member of the White House Coronavirus Task Force in 2020, and his stand on masks, which conservative users viewed as inconsistent. One tweet assigned to the topic reads "coronavirus minneapolis faucifraud watch fauci tell you not to wear a mask flashbackfriday" and another that reads "so old fauci was right wearing a mask is useless coronavirus can still pass between face mask wearers even."

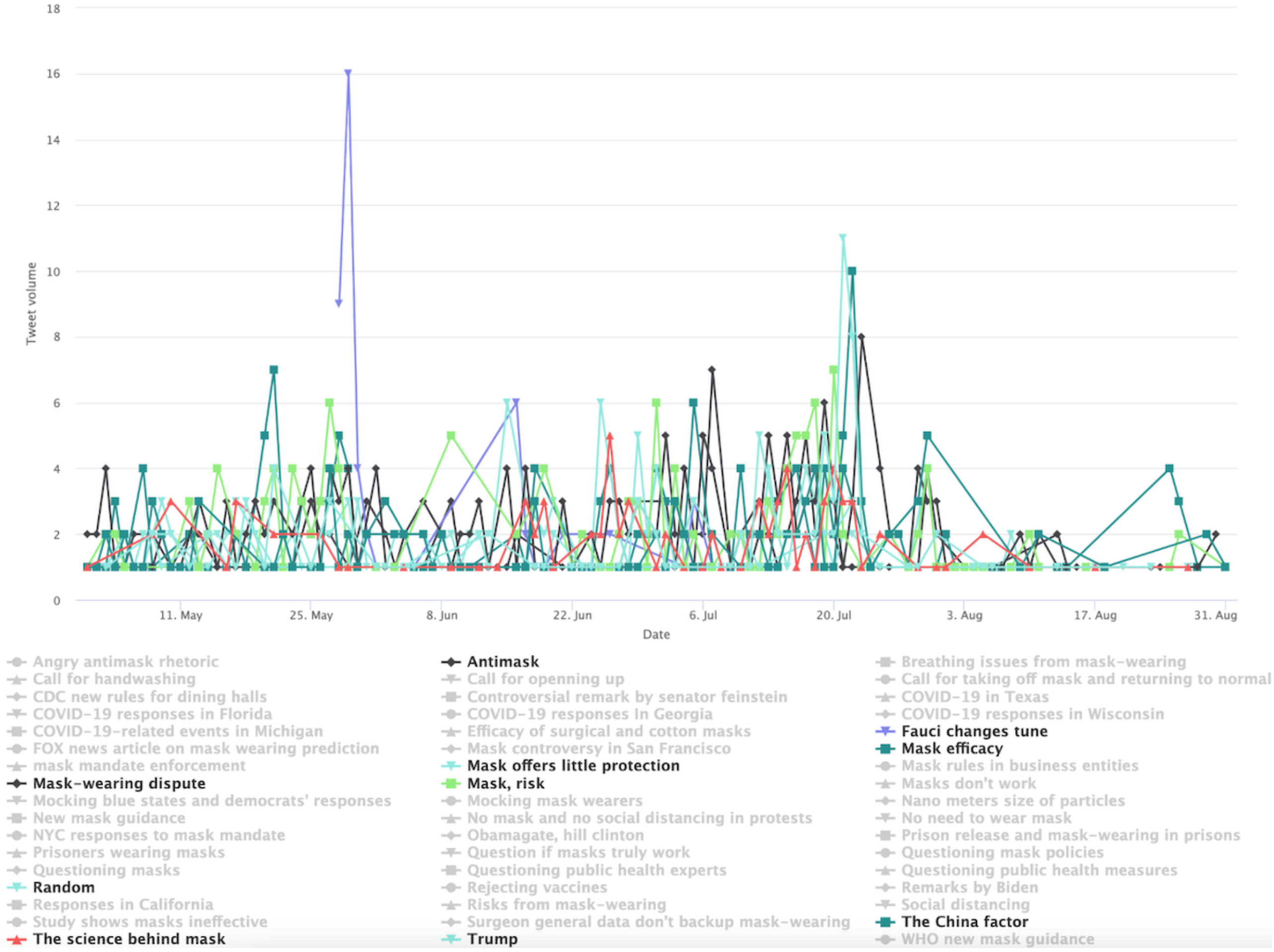
By stage 3 (Figure 5), the interest in masks by the conservative cluster seems to have dwindled (judging by the sheer tweet volume). Early September was marked by the Conservatives' focus on Trump and mask wearing, whereas by the later months, the general COVID-19 discourse became more prevalent.

Figure 3. Top topics in the conservative cluster during stage 1.



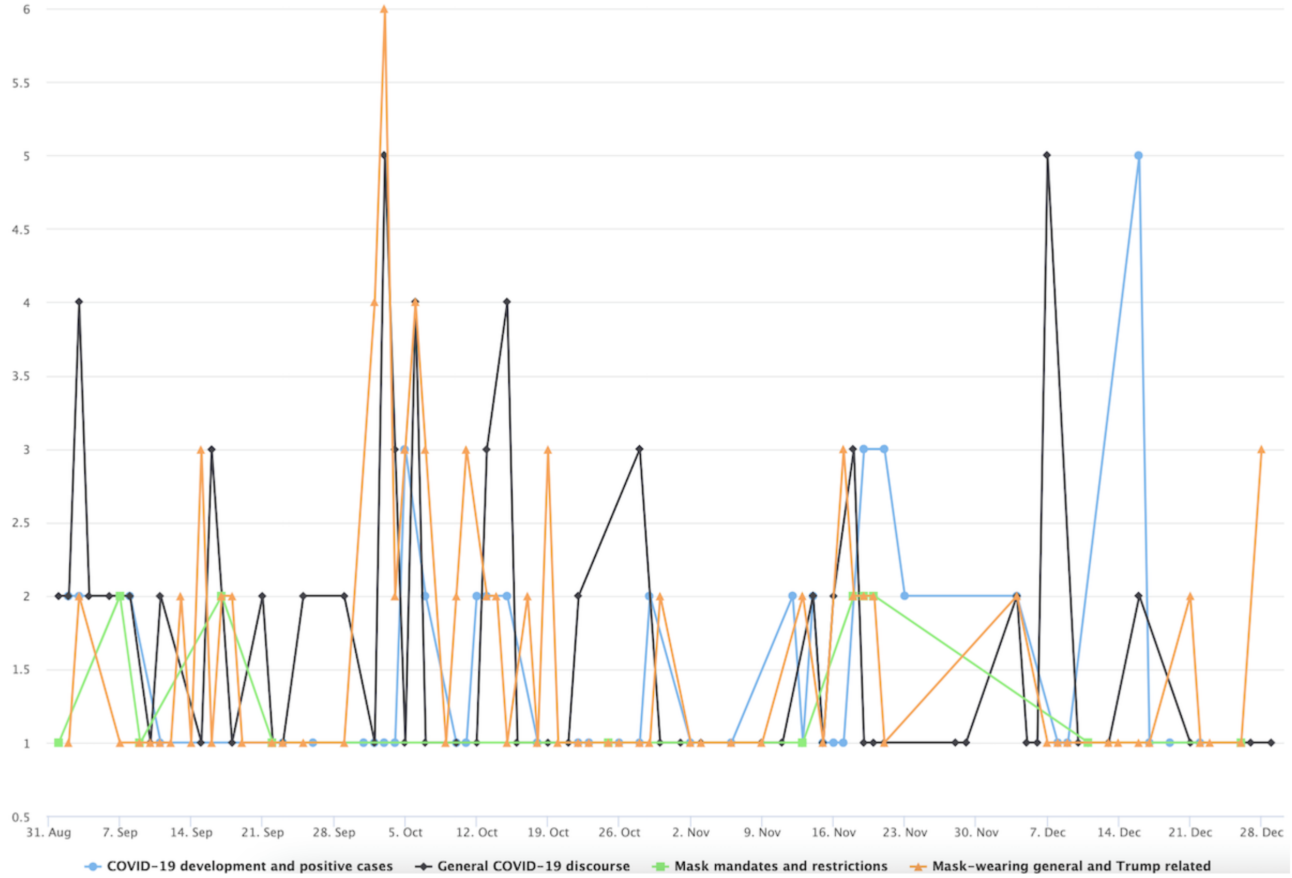
XSL•FO
RenderX

Figure 4. Top topics in the conservative cluster during stage 2.



XSL•FO
RenderX

Figure 5. Top topics in the conservative cluster during stage 3.



Topics in the Progressive Cluster

The progressive cluster is distinguishable by marker words on bios such as *resist*, *fbr*, *blm*, *dms*, *trump*, *voteblu*, *theresist*, *democrat*, *list*, *vote*, *love*, *anim*, *lover*, *biden2020*, *proud*, *block*, *follow*, *liber*, *votebluenomatterwho*, *bidenharris2020*, *mom*, *bluewav*, *retir*, *blue*, *dog*, *pleas*, *equal*, *bluewave2020*, *dm*, *polit*, *junki*, *resist*, *fbr*, *news*, *media*, *human*, *social*, *tweet*, etc. Similar to the conservative cluster, the progressive cluster is centered around US politics and the 2020 election in particular. The cluster produced 18,378 tweets, with approximately 32% of tweets classified as nonthematic and excluded from the final analysis. Among the included tweets, there are 6991 unique users; 1499 of them had invalid Twitter bios in August 2022. The users' follower size ranges from 5,503,681 to 0, with a median of 4410.

The textnets algorithm (Figure 6) shows 2 distinct clusters: one is tied to progressive social movements such as Black Lives Matter and the Biden Campaign and the second cluster indicates opposition to Trump. Data from the location field (Table 3) indicate US users primarily, particularly in major US metropolitans.

Figure 7 shows the progressive cluster's topics in stage 1, such as the conservative cluster, the most prominent topic at this stage concerns the shortage of PPE. The conversation about this topic picked up in early April and it peaked in late April. Example tweets include "the government s emergency stockpile of respirator masks gloves and other medical supplies is running low and..." and "...with having to reuse now either the n masks or the gowns or even the gloves that they are asking us to reuse..." The second-most prominent topic is concerned with COVID-19–related deaths, which registered the biggest peak in the graph on April 30. Some example tweets include "remember when the president fell asleep for a whole month and then poof k dead..." and "all medical workers took an oath and are dying because they believe their oath djt took an oath all of congress..." The third- and fifth-most prominent topics contain criticism of then President Trump and Vice President Pence. One example tweet reads, "trump is losing his mind over reports he is losing his mind this is do or die for trump expect everything..." Another tweet reads, "pence flouts mayo clinic policy by touring coronavirus testing facility without a mask pence defended his actions..." Other prominent topics include outbreaks in different states, and distinct locales (eg, meat processing plants and nursing homes), treatments and vaccines, etc. Similar to the conservative cluster, the China factor and the virus origin were brought up but not to the level of prominence of the Conservatives.

XSL•FO
RenderX

Figure 6. Central terms in Twitter bios of the progressive cluster.

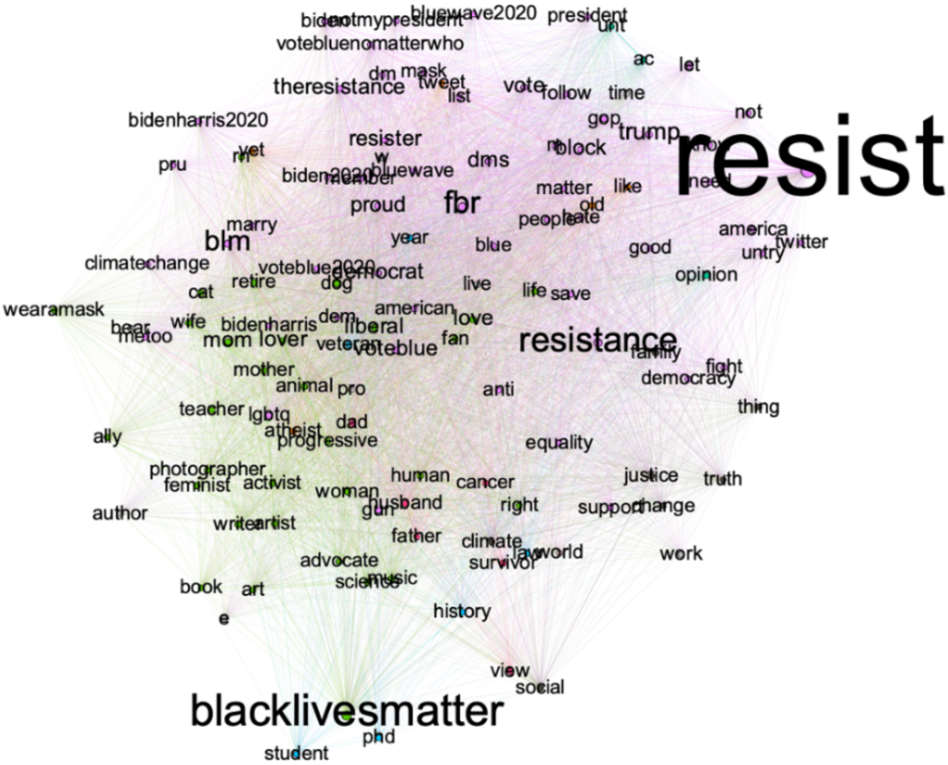
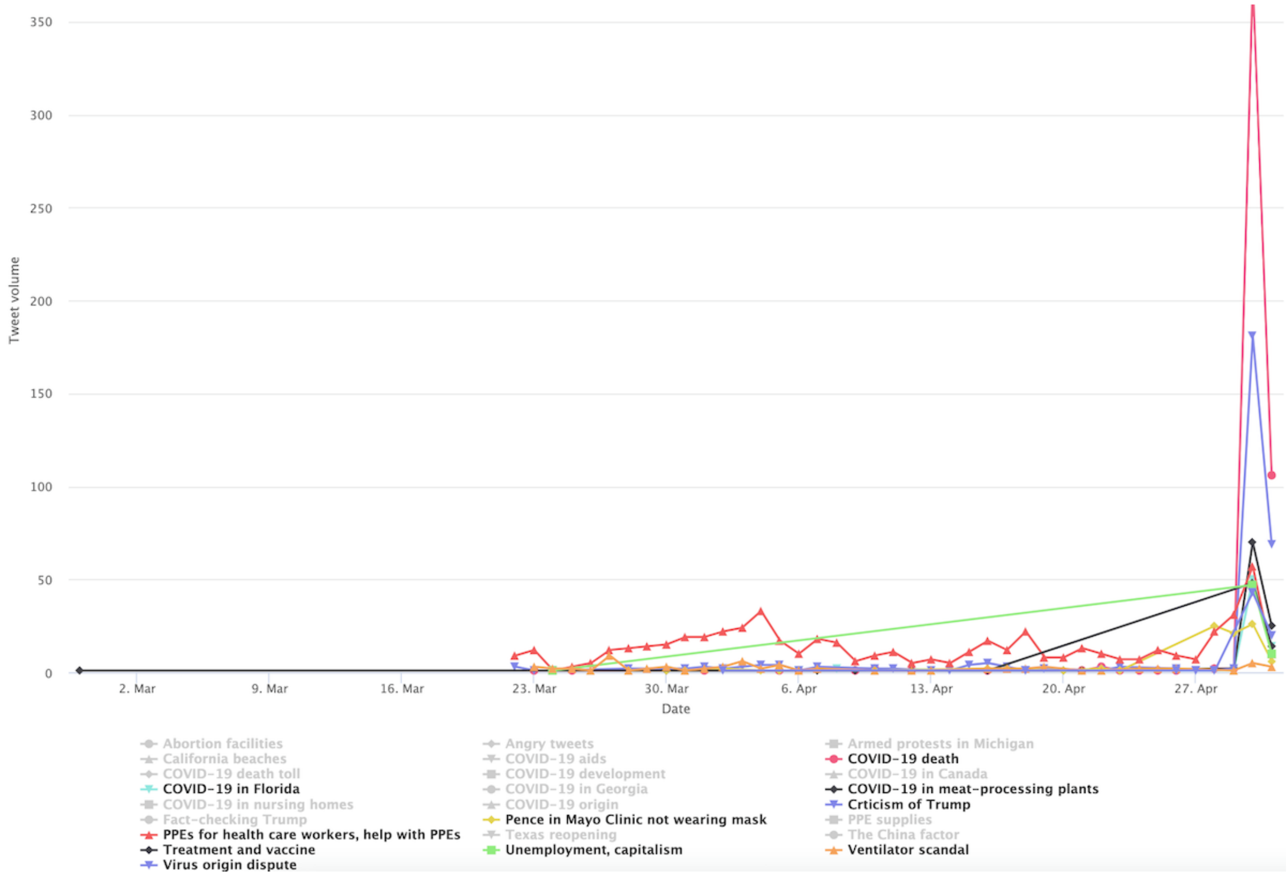


Table 3. Top user-entered location information in the progressive cluster.

Location	Value of location
United States	431
California, United States	123
Los Angeles, California	84
Florida, United States	77
Texas, United States	59
New York	49
Texas	42
New York, United States	39
Florida	39
Chicago, Illinois	39
Pennsylvania, United States	38
Earth	38
Ohio, United States	35
New Jersey, United States	34
Washington DC	31
Canada	31
California	30
Dallas, Texas	29
Colorado, United States	27
Virginia, United States	26
San Francisco, California	26
Portland, Oregon	26
North Carolina, United States	26
Oregon, United States	25
Atlanta, Georgia	25
Arizona, United States	25
Seattle, Washington	24
Minnesota, United States	24
Maryland, United States	24

XSL•FO
RenderX

Figure 7. Top topics in the progressive cluster during stage 1.

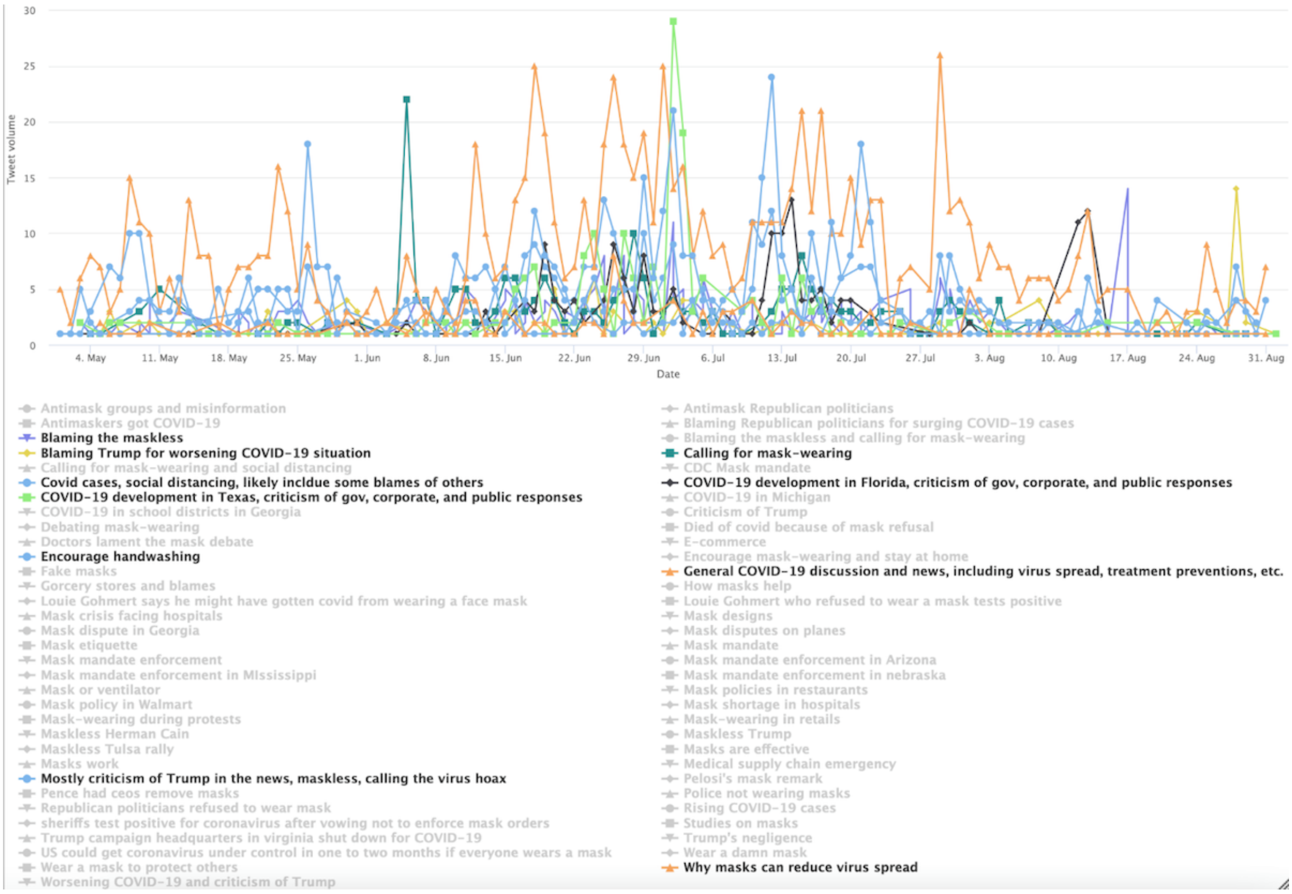


Entering stage 2 (Figure 8), although general conversations and news sharing about COVID-19 spread dominated the tweets, a notable share of tweets criticized Trump for his negligence in COVID-19 responses and the maskless population for spreading the virus. Example tweets include, “to all the people who think the coronavirus is a hoax humor the scientists and the woke people and wear your god damn mask.” Another reads, “coronavirus cases and covid deaths spike following trump s maskless rally in phoenix more doctors need.” The volume of criticism tweets ebbs and flows, likely reflecting events on the ground. The cluster also includes tweets calling for mask wearing and advocating for the effectiveness of masks. One example reads, “coronavirus cases in florida today please stay safe floridians everybody please wear a mask.” The progressive users’ topical interest also seems to reflect evolving COVID-19 spread across the US states. For instance, a spike was registered on July 2, 2020, following rising cases in Texas.

In stage 3 (Figure 9), the progressive cluster’s conversations were consistently dominated by Trump-related topics, critical of his administration. One notable topic is “COVID spreads and discussions of covid denialism,” which peaked around October 2, 2020. This marked the day when then President Trump tested positive for COVID-19. We also identified a few spikes in volume in topics related to calling for mask wearing and discussions of COVID-19 denialism.

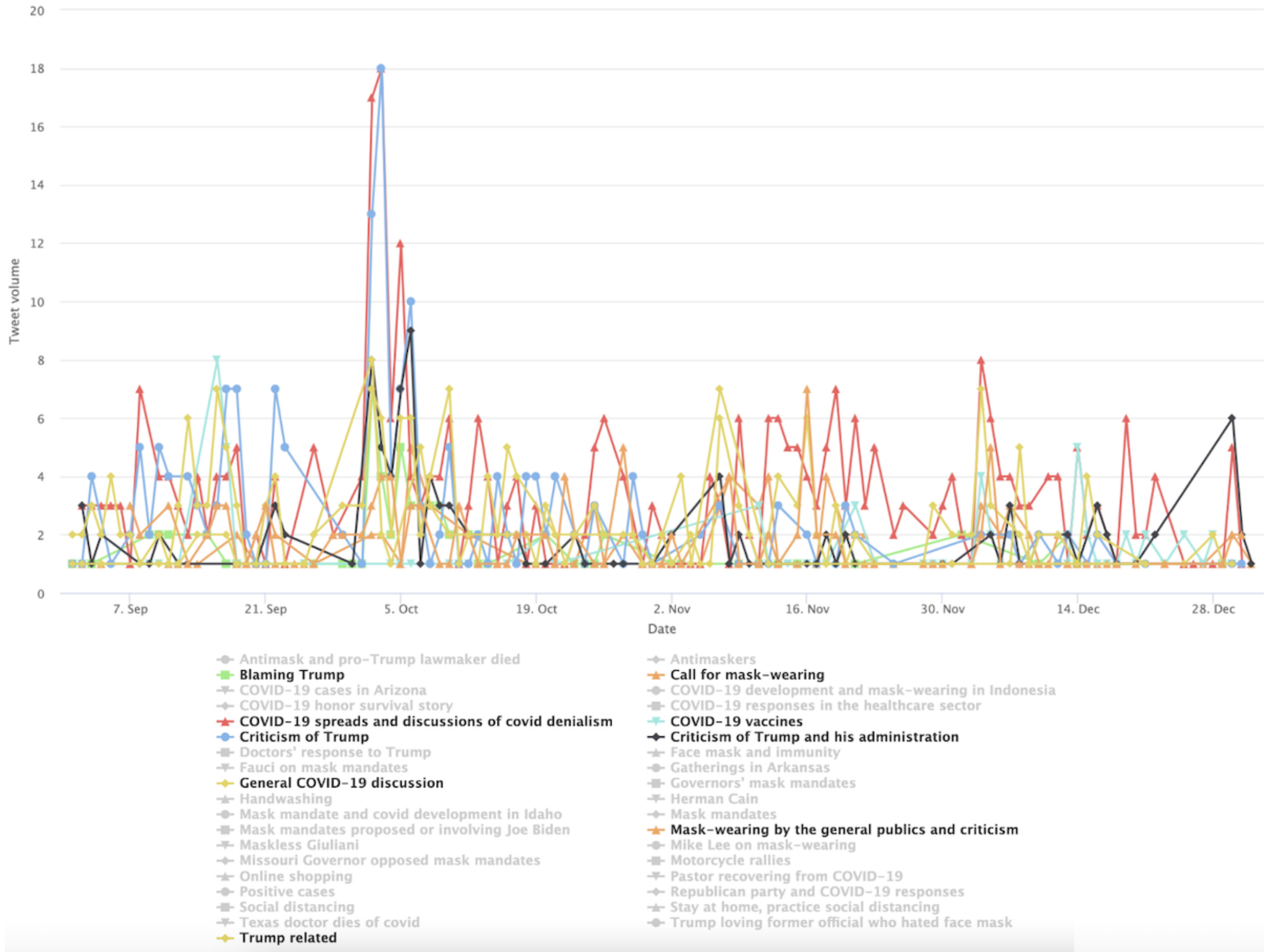
XSL•FO
RenderX

Figure 8. Top topics in the progressive cluster during stage 2.



XSL•FO
RenderX

Figure 9. Top topics in the progressive cluster during stage 3.



Topics in the General Cluster

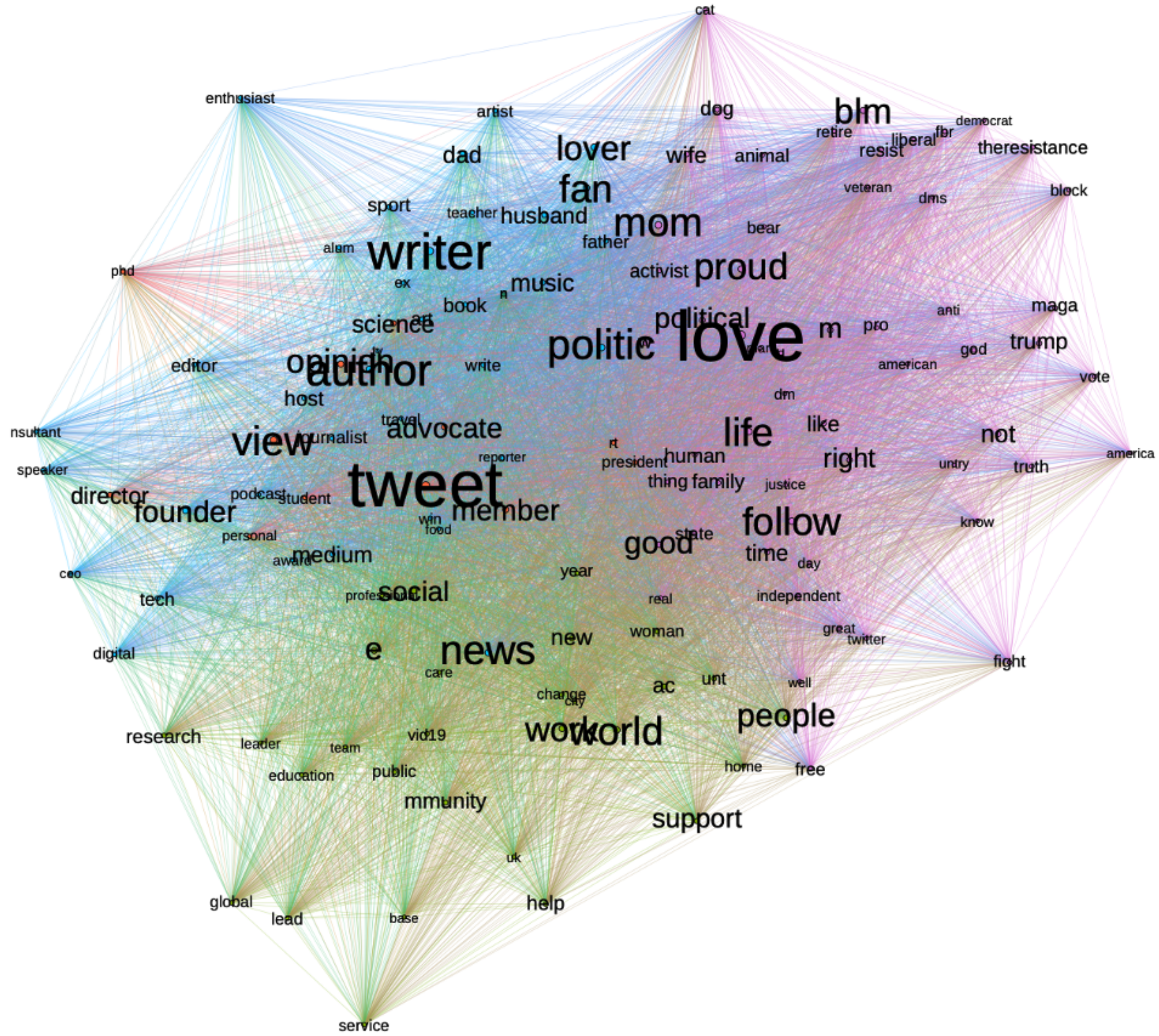
The general cluster consists of users whose Twitter bios contain but are not limited to the following keywords: *tweet, love, world, author, view, work, support, follow, people, former, proud, theresist, writer, us, covid19, research, fan, mom, polit, junki, resist, fbr, news, media, human, social, tweet, right, report, music, opinion*. As the keywords suggest, these users, although they could be interested in politics, do not feature strong partisanship through Twitter profiles. This cluster produced 86,722 tweets, with about 42.3% classified as nonthematic in the topic modeling. There are 33,364 unique users, among which 6973 users did not have a valid and accessible Twitter bio in August 2022. Notably, some accounts affiliated with news media and international organizations were classified into this category (notably World Economic Forum, UNICEF, MSNBC, the partisan influencer Ben Shapiro, and China's state media CGTN), despite our efforts to exclude a distinct media-affiliated cluster. This means that the general cluster includes both average citizens and some affiliated media. The follower count ranges from 13,280,615 to 0, with a median of 4410.

Figure 10 shows the top 134 key terms ranked by betweenness centrality. For the keywords to be included in the textnets clustering, they must appear at least 15 times. Figure 10 shows a cluster (blue) based on social and professional roles, a cluster (green) based on news, and a cluster (purple) that contains keywords related to Trump and his campaign. However, it should be noted that Trump-related keywords are not as central as they appear in the Twitter bios of the conservative cluster. The top entries in the location fields (Table 4) show that users in the cluster are primarily based in the United States, residing in major metropolitan areas.

Unlike the previous 2 clusters, which are visibly political in Twitter bios and in mask-related tweets, the users in the general cluster indicated various social and professional identities and lifestyles but with peripheral mentions of politics. Therefore, this user cluster is considered less politically inclined than the previous 2 clusters. Their apolitical nature is reflected in mask-related tweets in stage 1 (Figure 11). Although the most prominent topic is about Trump's responses to COVID-19, other topics do not seem to have a clear partisan slant. Such topics include showing appreciation and support and calling for donations. Example tweets include "we are truly grateful to our heroes in this covid pandemic ayekoo staysafe flattenthecurve" and "thank you to everyone that donated to the covid donation drive for navajo nation today you give me hope." General users also seem to pay attention to economic impacts and the loss of lives. The roles of China were brought up but less saliently than previously mentioned top topics.

XSL•FO
RenderX

Figure 10. Central terms in Twitter bios of the general user cluster.



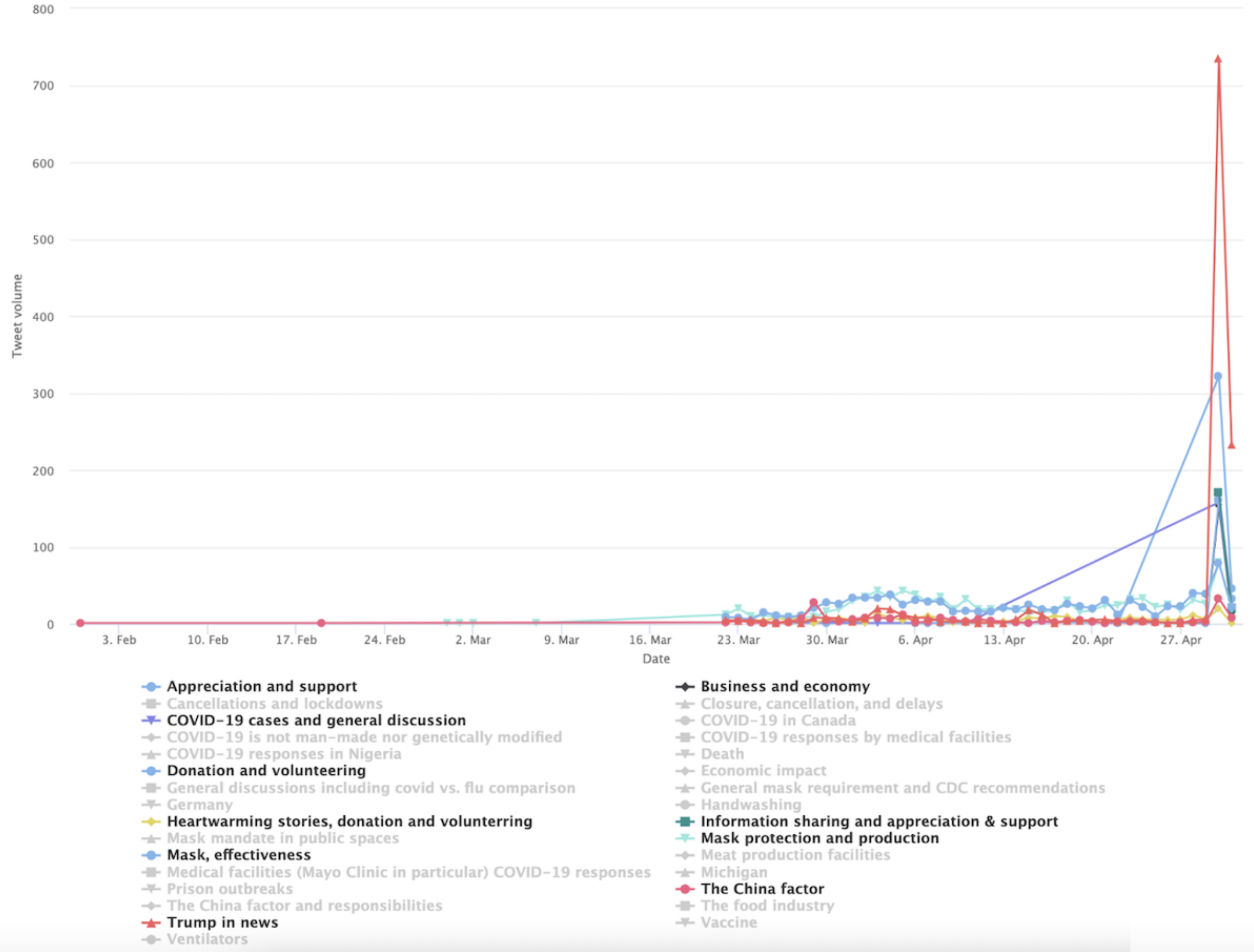
XSL•FO
RenderX

Table 4. Top user-entered location information in the general cluster.

Location	Value of location
United States	1758
Washington DC	486
California, United States	417
Los Angeles, California	403
New York	322
Texas, United States	292
Florida, United States	279
London, England	267
London	265
Canada	241
Chicago, Illinois	214
United Kingdom	380
New York, United States	186
Global	177
Boston, Massachusetts	168
Atlanta, Georgia	161
Houston, Texas	153
Worldwide	151
Texas	151
Earth	146
New York	144
San Francisco, California	142
India	140
Seattle, Washington	139
California	139
Toronto, Ontario	137
Austin, Texas	135
Nigeria	129

XSL•FO
RenderX

Figure 11. Top topics in the general cluster during stage 1.



Entering stage 2 (Figure 12), the general users' topics became much more diverse. Although the most prominent topic, labeled "General discussions about masks," does not seem partisan leaning, the second-most prominent topic is related to news coverage of Trump. The spike on June 20, 2020, corresponds to Trump's campaign rally in Tulsa, Oklahoma. The spike on July 12 corresponds to the timing when Trump was seen wearing mask in public for the first time. The spike around July 20 corresponds to Trump's endorsement of masks on Twitter and in media appearances. Other prominent topics include calling for masking and handwashing and blaming antimaskers. However, such topics were overshadowed by the Trump-related topic.

In stage 3 (Figure 13), the cluster's conversation was more general, following several topics identified in previous stages. Such topics include a call for mask wearing and handwashing, general discussion and news sharing about COVID-19 cases, as reflected by the prominence of the topics labeled *COVID cases and development*, which ebbs and flows throughout stage 3. However, Trump-related topics registered several spikes.

XSL•FO
RenderX

Figure 12. Top topics in the general cluster during stage 2. PPE: personal protective equipment.

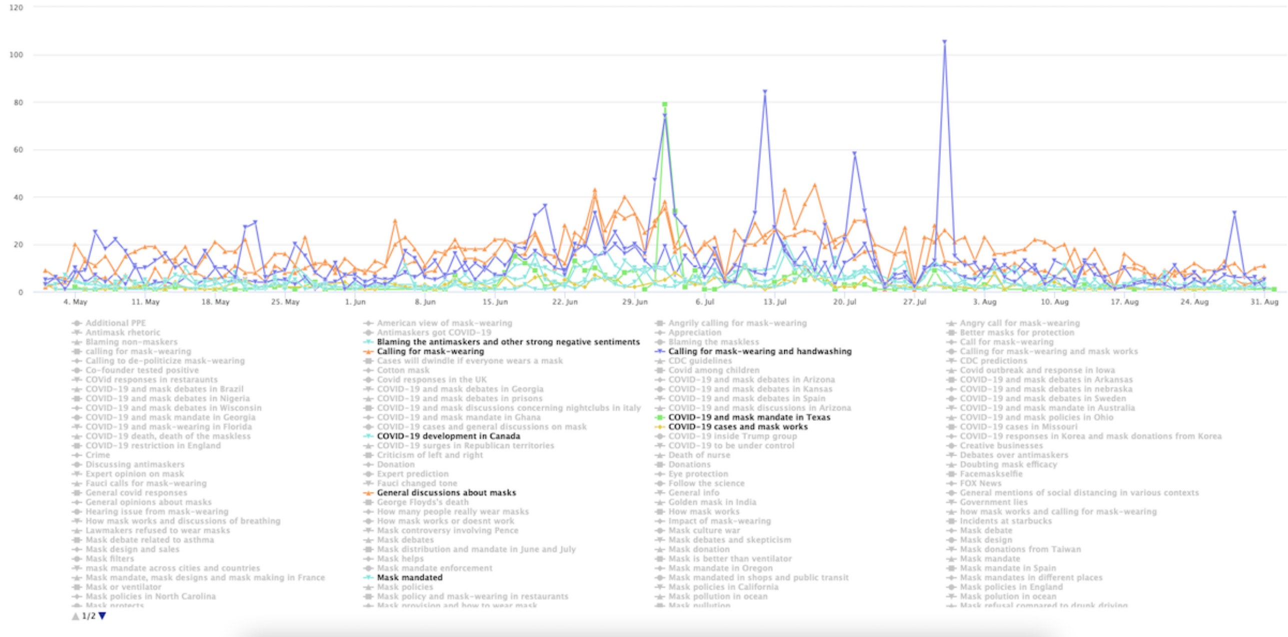
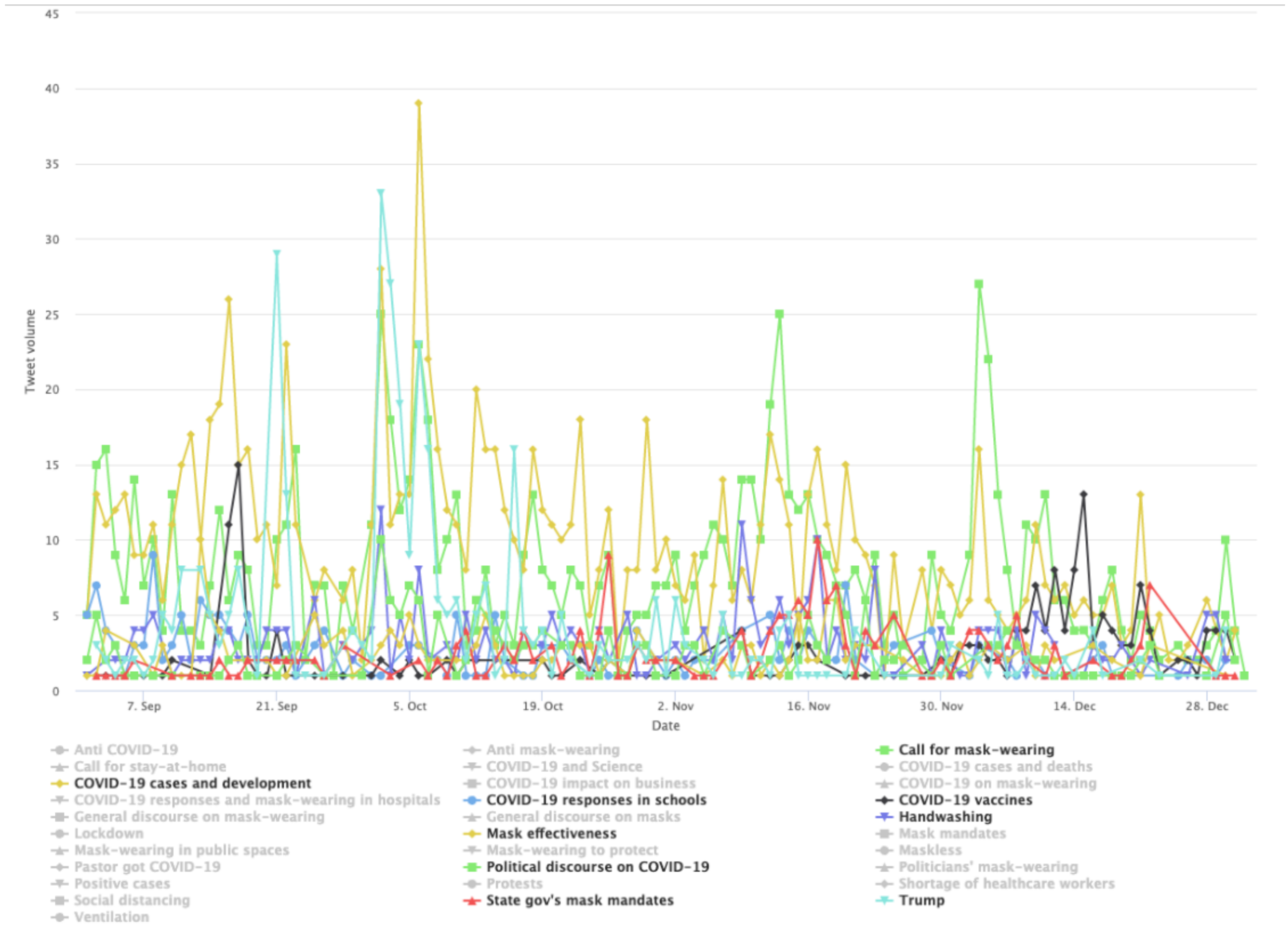


Figure 13. Top topics in the general cluster during stage 3.



Topics in the Public Health Cluster

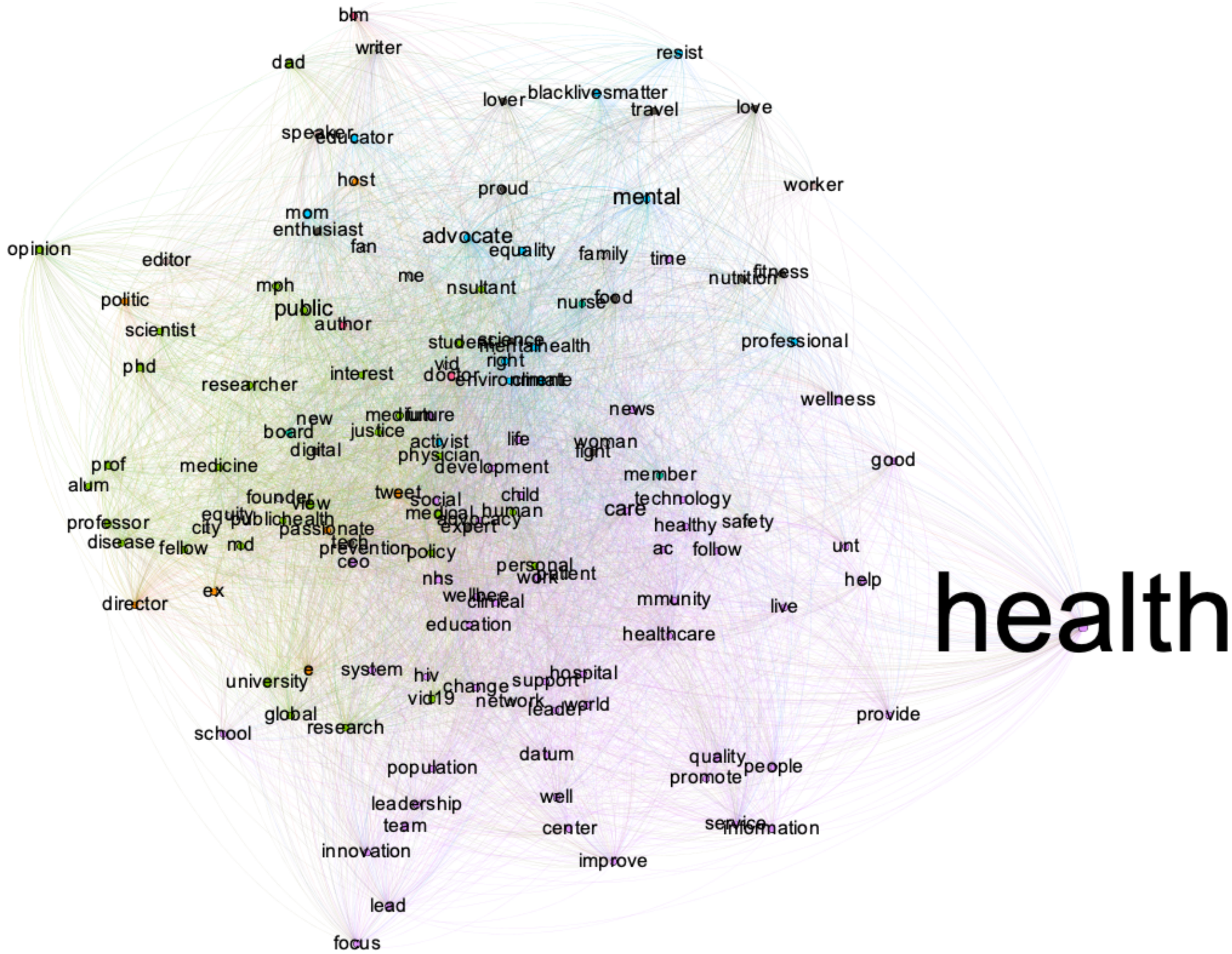
Users in the public health cluster are defined by the following bio keywords: *health*, *care*, *public*, *advoc*, *mental*, *covid19*, *research*, *global*, *scienc*, *center*, *community*, *improv*, *view*, *tweet*, *polici*. This cluster produced a total of 4697 tweets, with 16.2% classified as nonthematic. The unique user count is 2165, with 13% of them having no valid Twitter bios in August 2022. The follower size ranges from 11,703,587 to 3, with a median of 4413. The textnets algorithm shows clusters by health care specialties and fields, and users seem to be predominantly related to the health care sector. Figure 14 shows the top 137

XSL•FO
RenderX

key terms (which appeared at least twice in the users' bios) by betweenness centrality. Top entries on the location fields (Table 5) show a more geographically diverse of users compared with other clusters.

The public health cluster sent fewer tweets than other user clusters, and the cluster produced fewer topics. Early on, their tweets were about showing appreciation (Figure 15) and discussing mask effectiveness such as this tweet "...asymptomatic covid carriers have led the to reconsider its guidelines for who should wear masks..." In stage 2 (Figure 16), the cluster produced a more diverse set of topics, with general news sharing about masks being the most prominent, followed by a cluster of topics that call for handwashing and mask wearing. A similar set of topics were found for stage 3 (Figure 17), centered around a call for mask wearing and handwashing and general discussions about the COVID-19 pandemic development.

Figure 14. Central terms in Twitter bios of the public health cluster.



XSL•FO
RenderX

Table 5. Top user-entered location information in the public health cluster.

Location	Value of location
Washington DC	45
United States	39
Los Angeles, California	23
London, England	21
Canada	19
Global	19
Chicago, Illinois	19
Toronto, Ontario	16
Nigeria	16
London	16
Ann Arbor, Michigan	16
United Kingdom	15
Geneva, Switzerland	13
Boston, Massachusetts	12
California	12
Worldwide	11
United Kingdom	11
Philadelphia, Pennsylvania	11
Seattle, Washington	10
San Diego, California	10
New York	10
Houston, Texas	10
London, United Kingdom	10
Columbus, Ohio	10
Austin, Texas	10
California, United States	10
Toronto, Canada	9
Washington DC	9
New York, United States	9

XSL•FO
RenderX

Figure 15. Top topics in the public health cluster during stage 1.

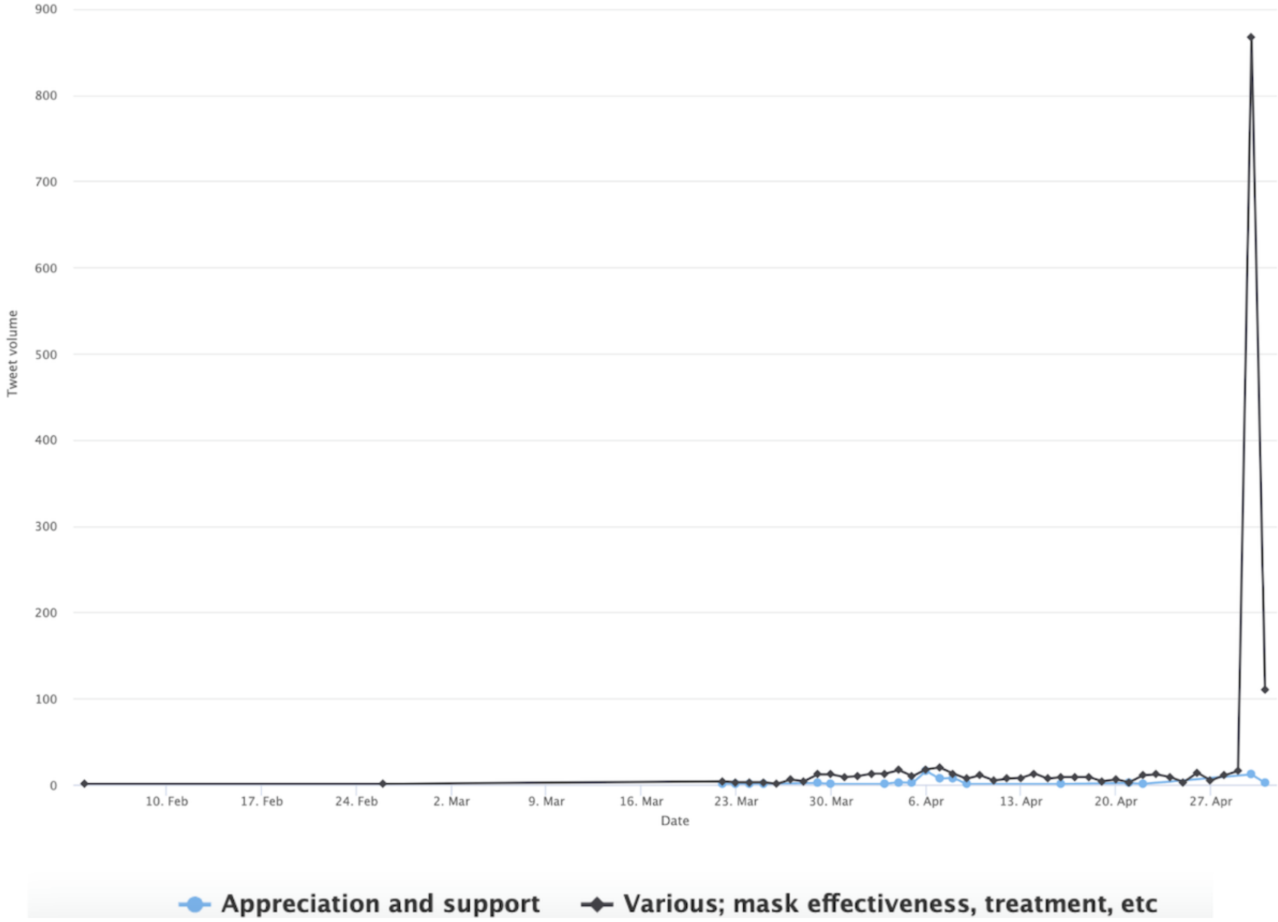
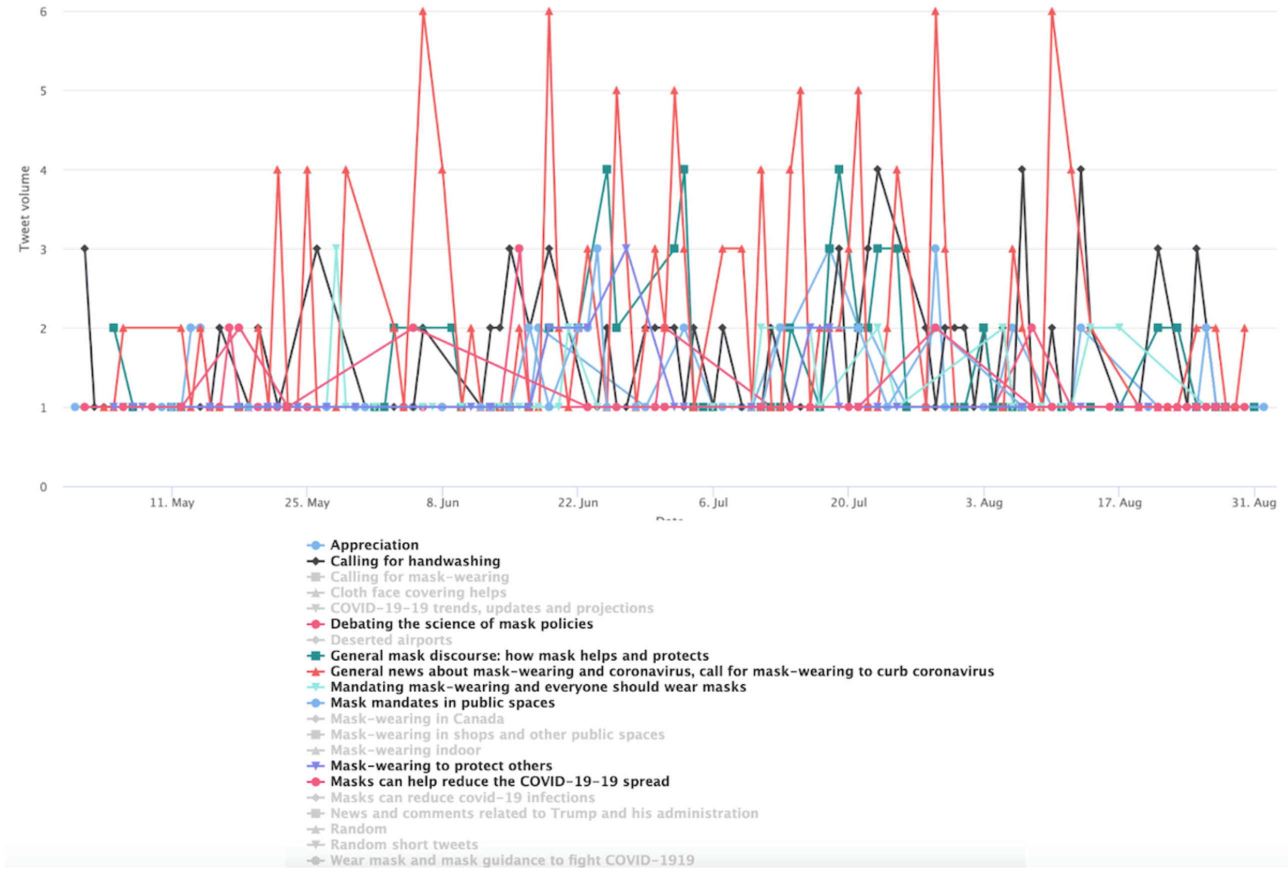
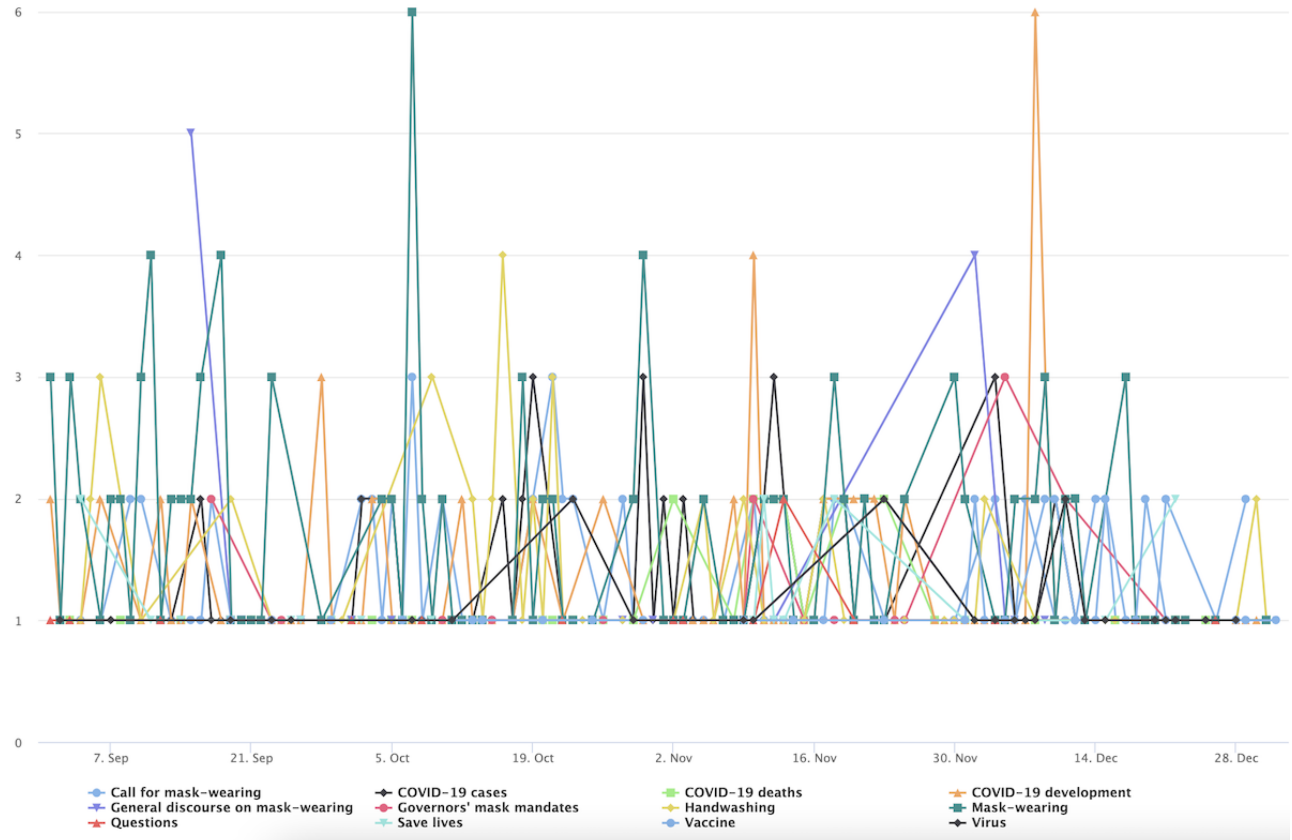


Figure 16. Top topics in the public health cluster during stage 2.



XSL•FO
RenderX

Figure 17. Top topics in the public health cluster during stage 3.



Discussion

Principal Findings

First, our findings echo the importance of a priori user classification in analyzing web-based discourse. As illustrated in the work by Walter et al [48], user clusters can be detected by the type of content they produce on the web. Nevertheless, different from the prior work, which assigns users to clusters by topics, we consider users' active expressions of social, political, and professional identities on social media profiles as the base for clustering. This ensures that we can compare how users' discussed topics vary by their expressed identities. As expected, in our data set, users varied by the level of political interest and the spectrum of conservative-to-progressive ideology. In support of prior research identifying issue publics based on distinct political and social identity–related expressions on social media profiles, we found that users in the mask discourse also come from both ends of the political spectrum. Some users were visibly politically motivated, as indicated by mobilizational and identarian hashtags (eg, #kag and #maga). It should be noted that, although politically motivated tweets were plentiful, they remained a minority. By comparison, the general user cluster (those that do not have explicit political expressions on Twitter bio descriptions) constituted the largest cluster. The participation from users in the public health sector was less prominent, implying that much of the public discourse was contributed by either laymen or politically minded individuals rather than public health experts. This finding might point to an expert gap in public health messaging. This finding echoes what is found in previous studies of Twitter discourse concerning alternative treatments of COVID-19. Previous studies show that mainstream medical experts and institutions were less influential than partisan figures [69,70]. Arguably, public health experts' lesser degree of influence could result from politically motivated public distrust in light of medical populism or the absence of public health voices in this important public sphere. Nevertheless, given the increasingly political nature of mask wearing and mask policies, scientific rather than political voices were much needed in the public sphere. Our findings echo other research that has shown that public policies are politicized in civic discourse [71].

Second, topics did vary by different user clusters. Mask policies have become a sharp point of division between the political left and right in many western democracies. Such divisions in our study mapped onto different topical focuses between the progressive and conservative user clusters. One focused on the criticism of the Trump administration, and the other showed cynicism and skepticism toward public health experts. One attended to the impacts of lockdown, whereas the other tweeted more about COVID-19–induced death. Our topic models broadly reflect the policy preferences and ideological variations in response to mask policies. Equally important to note that political topics also emerged in the general users' discourse. In particular, the public attention paid to elected officials and their masking practices. This shows how politicians' behavior could potentially drive or divert public attention to and away from important public health measures. To relate to the concept of medical populism, which has been studied in the context of vaccination and pandemic, our topic model revealed potentially populist discourse that pits people against the elite. This is

XSL•FO
RenderX

specifically revealed in the conservative users' dismissive attitude toward public health experts such as Dr Fauci and the US National Institutes of Health and mainstream media that many view as left-leaning. Although populist-sounding topics did emerge, we caution that they were not the most prominent by tweet volume. To recap, our model was able to pick up critical signals (emergent topics or changes of topics) that should be analyzed further to evaluate public health efforts.

Third, although much of the discourse focused on the impact of COVID-19 and politics behind mask policies, some part of the discourse did appear to focus on the science of mask wearing. All user clusters tweeted about the effectiveness of mask wearing. Identifying these topics is critical because further qualitative analyses can be performed on this specific set of tweets to understand the users' sources, cited studies, and evidence. Findings could be particularly revealing in tweets from the politically motivated users and the general users because some topics appear to question the effectiveness of mask wearing. Our study showed that, methodologically, our model can pick up signals that may point to important public health discourse that needs to be fact-checked.

Finally, much of the discourse fluctuated with significant political developments that involved then President Trump, the early outbreaks in China, and the controversy surrounding the Wuhan laboratory. For public health monitoring, this again illustrated that public acceptance of public health measures did not occur in a vacuum but interacted with political events on the ground. Our implemented model was able to map out topical evolution over time, thus factoring in how external events influence web-based discourse.

Limitations

Readers are advised to review the findings with several limitations in mind. First, our sample selection left out tweets that were not retweeted. The nonretweeted may be less influential in message spread but are a significant part of the web-based discourse. In other words, our sample choice may have overlooked a broader discourse on the topic. Second, some tweets may have contained hyperlinked content or embedded images. Public responses to mask-related policies could well be reflected in this embedded content rather than in the plain-tweet text. The clustering based on Twitter bios also left out users who did not explicitly used Twitter bios to express social, professional, and political identities, as well as those whose accounts were deleted or suspended. In addition, a certain percentage of tweets are unclassified (the nonthematic) by BERT. This might be the inherent result of user classification based on Twitter bios. We also caution readers that bots were potentially present in the discourse, although their presence might be minimal. This is because we studied only original tweets (as opposed to retweeted content), and typical bots exclusively retweet others' content without producing original content. Nevertheless, bot traffic should be distinguished from the genuine citizen-generated Twitter conversations. At the time of the writing, the popular opensource bot-detection tools (eg, *tweetbotornot* and *tweetbotornot2*) were experiencing technical issues due to updates to the Twitter API and the proprietary Botometer presents significant cost barriers for analyzing many Twitter users. We alternatively calculated the ratios of users whose profiles were either deleted or suspended by Twitter, which could give us a glimpse into the potential bot traffic in the corpus. All the factors mentioned above may limit the representativeness of the study finding. We call for future studies to investigate the embedded content and to study tweets. We also deem it is important to study the discourse by regions and narratives. Future studies should compare authentic discourse on Twitter to inauthentic discourse (propagated by bots and trolls) and to media discourse (produced by news media accounts on Twitter). More importantly, we call for comparative methodological work to evaluate various text-classification schemes when applied to infoveillance. Although this study focuses on BERT Topic modeling, whether BERT models do outperform other novel text-classification schemes (such as textnets and ANTMN) is an unanswered question. In addition, future works can compare multiple user-classification schemes, which include the bio-based classification and the thematic personas classification [48].

Comparison With Prior Work

This work builds upon the existing infoveillance work that uses web-based behavioral data to track public health measures and messaging. This work has the following novelties. It is one of the few studies that has specifically looked at the web-based discourse on mask and mask wearing. It also improved the existing infoveillance framework by conducting a priori user classification and using the BERT Topic modeling, which is optimized for short texts.

Conclusions

This study improves upon the current infoveillance frame that relies mostly on LDA topic modeling and sentiment analysis. We argue that researchers must first conduct a proper identity and interest-based user classicization to reveal topics emergent in the web-based discourse. This step is lacking in many prior works. We then point out the weakness of the traditional LDA modeling and resort to the much-improved BERTopic. The BERT Topic modeling is optimized for short texts and can reveal longitudinal changes of topics. This implementation has resulted in a more gradient picture of the social media discourse on the issue of mask wearing.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, van der Westhuizen HM, et al. An evidence review of face masks against COVID-19. Proc Natl Acad Sci U S A 2021 Jan 26;118(4):e2014564118 [FREE Full text] [doi: 10.1073/pnas.2014564118] [Medline: 33431650]
2. MacIntyre CR, Chughtai AA. A rapid systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients. Int J Nurs Stud 2020 Aug;108:103629 [FREE Full text] [doi: 10.1016/j.ijnurstu.2020.103629] [Medline: 32512240]
3. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public discourse against masks in the COVID-19 era: infodemiology study of Twitter data. JMIR Public Health Surveill 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: 10.2196/26780] [Medline: 33720841]
4. Molteni M, Rogers A. How Masks Went From Don't-Wear to Must-Have. Wired. 2020 Jul 2. URL: https://www.wired.com/story/how-masks-went-from-dont-wear-to-must-have/ [accessed 2022-11-28]
5. 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]
6. Cai W, Burch AD, Patel JK. Swelling Anti-Asian Violence: Who Is Being Attacked Where. The New York Times. 2021 Apr 3. URL: https://www.nytimes.com/interactive/2021/04/03/us/anti-asian-attacks.html [accessed 2022-11-28]
7. Lasco G. Medical populism and the COVID-19 pandemic. Glob Public Health 2020 Oct;15(10):1417-1429. [doi: 10.1080/17441692.2020.1807581] [Medline: 32780635]
8. Schradie J. "Give me Liberty or Give me Covid-19": anti-lockdown protesters were never Trump puppets. Commun Public 2020 Dec 08;5(3-4):126-128. [doi: 10.1177/2057047320969433]
9. Perry SL, Whitehead AL, Grubbs JB. Culture wars and COVID‐19 conduct: christian nationalism, religiosity, and Americans' behavior during the coronavirus pandemic. J Sci Study Relig 2020 Jul 26;59(3):405-416. [doi: 10.1111/jssr.12677]
10. Buštíková L, Baboš P. Best in Covid: populists in the time of pandemic. Politics Governance 2020 Dec 17;8(4):496-508. [doi: 10.17645/pag.v8i4.3424]
11. Lasco G, Curato N. Medical populism. Soc Sci Med 2019 Jan;221:1-8. [doi: 10.1016/j.socscimed.2018.12.006] [Medline: 30553118]
12. Mede NG, Schäfer MS. Science-related populism: conceptualizing populist demands toward science. Public Underst Sci 2020 Jul;29(5):473-491 [FREE Full text] [doi: 10.1177/0963662520924259] [Medline: 32515652]
13. Żuk P, Żuk P, Lisiewicz-Jakubaszko J. The anti-vaccine movement in Poland: the socio-cultural conditions of the opposition to vaccination and threats to public health. Vaccine 2019 Mar 07;37(11):1491-1494. [doi: 10.1016/j.vaccine.2019.01.073] [Medline: 30755366]
14. 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]
15. Eberl JM, Huber RA, Greussing E. From populism to the "plandemic": why populists believe in COVID-19 conspiracies. J Elect Public Opin Parties 2021 Jun 15;31(sup1):272-284. [doi: 10.1080/17457289.2021.1924730]
16. 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]
17. Casarões G, Magalhães D. The hydroxychloroquine alliance: how far-right leaders and alt-science preachers came together to promote a miracle drug. Rev Adm Pública 2021 Feb;55(1):197-214. [doi: 10.1590/0034-761220200556]
18. Oyeyemi SO, Gabarron E, Wynn R. Ebola, Twitter, and misinformation: a dangerous combination? BMJ 2014 Oct 14;349:g6178. [doi: 10.1136/bmj.g6178] [Medline: 25315514]
19. Larson HJ. The biggest pandemic risk? Viral misinformation. Nature 2018 Oct;562(7727):309. [doi: 10.1038/d41586-018-07034-4] [Medline: 30327527]
20. Carey JM, Chi V, Flynn DJ, Nyhan B, Zeitzoff T. The effects of corrective information about disease epidemics and outbreaks: evidence from Zika and yellow fever in Brazil. Sci Adv 2020 Jan;6(5):eaaw7449 [FREE Full text] [doi: 10.1126/sciadv.aaw7449] [Medline: 32064329]
21. Zarocostas J. How to fight an infodemic. Lancet 2020 Feb 29;395(10225):676 [FREE Full text] [doi: 10.1016/S0140-6736(20)30461-X] [Medline: 32113495]
22. Hao K, Basu T. The coronavirus is the first true social-media "infodemic". MIT Technology Review. 2020 Feb 12. URL: https://www.technologyreview.com/2020/02/12/844851/the-coronavirus-is-the-first-true-social-media-infodemic/ [accessed 2022-11-28]
23. Bruns A, Burgess J, Highfield T, Kirchhoff L, Nicolai T. Mapping the Australian networked public sphere. Soc Sci Comput Rev 2010 Sep 21;29(3):277-287. [doi: 10.1177/0894439310382507]
24. 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]
25. 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]

XSL•FO
RenderX

26. 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]
27. van Lent LG, Sungur H, Kunneman FA, van de Velde B, Das E. Too far to care? Measuring public attention and fear for Ebola using Twitter. J Med Internet Res 2017 Jun 13;19(6):e193 [FREE Full text] [doi: 10.2196/jmir.7219] [Medline: 28611015]
28. 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]
29. Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, et al. Zika in Twitter: temporal variations of locations, actors, and concepts. JMIR Public Health Surveill 2017 Apr 20;3(2):e22 [FREE Full text] [doi: 10.2196/publichealth.6925] [Medline: 28428164]
30. 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]
31. Sarker A, Lakamana S, Hogg-Bremer W, Xie AA, Al-Garadi MA, Yang YC. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. J Am Med Inform Assoc 2020 Aug 01;27(8):1310-1315 [FREE Full text] [doi: 10.1093/jamia/ocaa116] [Medline: 32620975]
32. 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 May 28;22(5):e19421 [FREE Full text] [doi: 10.2196/19421] [Medline: 32452804]
33. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
34. 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]
35. Wahbeh A, Nasralah T, Al-Ramahi M, El-Gayar O. Mining physicians' opinions on social media to obtain insights into COVID-19: mixed methods analysis. JMIR Public Health Surveill 2020 Jun 18;6(2):e19276 [FREE Full text] [doi: 10.2196/19276] [Medline: 32421686]
36. Li J, Xu Q, Cuomo R, Purushothaman V, Mackey T. Data mining and content analysis of the Chinese social media platform Weibo during the early COVID-19 outbreak: retrospective observational infoveillance study. JMIR Public Health Surveill 2020 Apr 21;6(2):e18700 [FREE Full text] [doi: 10.2196/18700] [Medline: 32293582]
37. Wang J, Zhou Y, Zhang W, Evans R, Zhu C. Concerns expressed by Chinese social media users during the COVID-19 pandemic: content analysis of Sina Weibo microblogging data. J Med Internet Res 2020 Nov 26;22(11):e22152 [FREE Full text] [doi: 10.2196/22152] [Medline: 33151894]
38. Zhang S, Pian W, Ma F, Ni Z, Liu Y. Characterizing the COVID-19 infodemic on Chinese social media: exploratory study. JMIR Public Health Surveill 2021 Feb 05;7(2):e26090 [FREE Full text] [doi: 10.2196/26090] [Medline: 33460391]
39. 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]
40. 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]
41. Doogan C, Buntine W, Linger H, Brunt S. Public perceptions and attitudes toward COVID-19 nonpharmaceutical interventions across six countries: a topic modeling analysis of Twitter data. J Med Internet Res 2020 Sep 03;22(9):e21419 [FREE Full text] [doi: 10.2196/21419] [Medline: 32784190]
42. 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 Mar 04;26(3):161-173. [doi: 10.1080/10810730.2021.1899344] [Medline: 33787462]
43. McGinn J, Kotamraju N. Data-driven persona development. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Apr Presented at: CHI '08; April 5-10, 2008; Florence, Italy p. 1521-1524. [doi: 10.1145/1357054.1357292]
44. An J, Kwak H, Jung S, Salminen J, Admad M, Jansen B. Imaginary people representing real numbers: generating personas from online social media data. ACM Trans Web 2018 Nov 06;12(4):1-26. [doi: 10.1145/3265986]
45. Haupt MR, Weiss SM, Chiu M, Cuomo R, Chein JM, Mackey T. Psychological and situational profiles of social distance compliance during COVID-19. J Commun Healthc 2022 Feb 01;15(1):44-53. [doi: 10.1080/17538068.2022.2026055]
46. Gao Y, Xie Z, Li D. Electronic cigarette users' perspective on the COVID-19 pandemic: observational study using Twitter data. JMIR Public Health Surveill 2021 Jan 05;7(1):e24859 [FREE Full text] [doi: 10.2196/24859] [Medline: 33347422]
47. van Draanen J, Tao H, Gupta S, Liu S. Geographic differences in cannabis conversations on Twitter: infodemiology study. JMIR Public Health Surveill 2020 Oct 05;6(4):e18540 [FREE Full text] [doi: 10.2196/18540] [Medline: 33016888]
48. Walter D, Ophir Y, Jamieson KH. Russian Twitter accounts and the partisan polarization of vaccine discourse, 2015-2017. Am J Public Health 2020 May;110(5):718-724. [doi: 10.2105/AJPH.2019.305564] [Medline: 32191516]

XSL•FO
RenderX

49. Xu WW. Mapping connective actions in the global alt-right and Antifa counterpublics. Int J Commun 2020;14:1070-1091. [doi: 10.1163/9789004444836_003]
50. Jackson SJ, Foucault Welles B. Hijacking #myNYPD: social media dissent and networked counterpublics. J Commun 2015 Nov 05;65(6):932-952. [doi: 10.1111/jcom.12185]
51. Kuo R. Racial justice activist hashtags: counterpublics and discourse circulation. New Media Soc 2016 Aug 12;20(2):495-514. [doi: 10.1177/1461444816663485]
52. Adam S, Häussler T, Schmid-Petri H, Reber U. Coalitions and counter-coalitions in online contestation: an analysis of the German and British climate change debate. New Media Soc 2019 Jun 19;21(11-12):2671-2690. [doi: 10.1177/1461444819855966]
53. Boon-Itt S, Skunkan Y. Public perception of the COVID-19 pandemic on Twitter: sentiment analysis and topic modeling study. JMIR Public Health Surveill 2020 Nov 11;6(4):e21978 [FREE Full text] [doi: 10.2196/21978] [Medline: 33108310]
54. 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]
55. Blei DM, Lafferty JD. Dynamic topic models. In: Proceedings of the 23rd international conference on Machine learning. 2006 Jun Presented at: ICML '06; June 25-29, 2006; Pittsburgh, PA, USA p. 113-120. [doi: 10.1145/1143844.1143859]
56. Peinelt N, Nguyen D, Liakata M. tBERT: topic models and BERT joining forces for semantic similarity detection. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020 Presented at: ACL '20; July 5-10, 2020; Virtual p. 7047-7055. [doi: 10.18653/v1/2020.acl-main.630]
57. Bianchi F, Terragni S, Hovy D, Nozza D, Fersini E. Cross-lingual contextualized topic models with zero-shot learning. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. 2021 Apr 19 Presented at: EACL '21; April 19-23, 2021; Virtual p. 1676-1683. [doi: 10.18653/v1/2021.eacl-main.143]
58. Dieng AB, Ruiz FJ, Blei DM. Topic modeling in embedding spaces. Trans Assoc Comput Linguist 2020 Dec;8:439-453. [doi: 10.1162/tacl_a_00325]
59. Tang J, Meng Z, Nguyen X, Mei Q, Zhang M. Understanding the limiting factors of topic modeling via posterior contraction analysis. In: Proceedings of the 31st International Conference on International Conference on Machine Learning. 2014 Jun Presented at: ICML '14; June 21-26, 2014; Beijing, China p. 190-198.
60. Blei DM, Lafferty JD. Correlated topic models. In: Proceedings of the 18th International Conference on Neural Information Processing Systems. 2005 Presented at: NeurIPS '05; December 5-8, 2005; Vancouver, Canada.
61. Angelov D. Top2Vec: distributed representations of topics. arXiv 2020 Aug 19.
62. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019 Presented at: NAACL '19; June 2-7, 2019; Minneapolis, MN, USA p. 4171-4186. [doi: 10.18653/v1/N19-1423]
63. Walter D, Ophir Y. News frame analysis: an inductive mixed-method computational approach. Commun Methods Meas 2019 Jul 23;13(4):248-266. [doi: 10.1080/19312458.2019.1639145]
64. Lattimer TA, Tenzek KE, Ophir Y, Sullivan SS. Exploring Web-based Twitter conversations surrounding national healthcare decisions day and advance care planning from a sociocultural perspective: computational mixed methods analysis. JMIR Form Res 2022 Apr 13;6(4):e35795 [FREE Full text] [doi: 10.2196/35795] [Medline: 35416783]
65. Bail CA. Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media. Proc Natl Acad Sci U S A 2016 Oct 18;113(42):11823-11828 [FREE Full text] [doi: 10.1073/pnas.1607151113] [Medline: 27694580]
66. Banda JM, Tekumalla R, Wang G, Yu J, Liu T, Ding Y, et al. A large-scale COVID-19 Twitter chatter dataset for open scientific research—an international collaboration. Epidemiologia 2021 Aug 05;2(3):315-324. [doi: 10.3390/epidemiologia2030024]
67. MacQueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. 1965 Presented at: MSP '65; June21-July 18, 1965; Berkeley, CA, USA p. 281-297.
68. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. J Mach Learn Res 2003 Jan 3;2:993-1022.
69. Haupt MR, Li J, Mackey TK. Identifying and characterizing scientific authority-related misinformation discourse about hydroxychloroquine on twitter using unsupervised machine learning. Big Data Soc 2021 May 06;8(1):205395172110138. [doi: 10.1177/20539517211013843]
70. 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 Society 2021 Jun 23;7(2):205630512110249. [doi: 10.1177/20563051211024977]
71. Żuk P, Żuk P. Right-wing populism in Poland and anti-vaccine myths on YouTube: political and cultural threats to public health. Glob Public Health 2020 Jun;15(6):790-804. [doi: 10.1080/17441692.2020.1718733] [Medline: 31964228]

XSL•FO
RenderX

Abbreviations

ANTMN: Analysis of Topic Model Networks
API: application programming interface
LDA: latent Dirichlet allocation
PPE: personal protective equipment

Edited by T Mackey; submitted 18.07.22; peer-reviewed by M Haupt, Y Ophir; comments to author 15.08.22; revised version received 26.09.22; accepted 02.11.22; published 09.12.22.

Please cite as:
Xu WW, Tshimula JM, Dubé È, Graham JE, Greyson D, MacDonald NE, Meyer SB
Unmasking the Twitter Discourses on Masks During the COVID-19 Pandemic: User Cluster–Based BERT Topic Modeling Approach
JMIR Infodemiology 2022;2(2):e41198
URL: https://infodemiology.jmir.org/2022/2/e41198
doi:10.2196/41198
PMID:36536763

©Weiai Wayne Xu, Jean Marie Tshimula, Ève Dubé, Janice E Graham, Devon Greyson, Noni E MacDonald, Samantha B Meyer. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 09.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Platform Effects on Public Health Communication: A Comparative and National Study of Message Design and Audience Engagement Across Twitter and Facebook

Nic DePaula[1], PhD; Loni Hagen[2], PhD; Stiven Roytman[3], BSc; Dana Alnahass[4], BSc

[1]School of Information Sciences, Wayne State University, Detroit, MI, United States
[2]School of Information, University of South Florida, Tampa, FL, United States
[3]Department of Radiology, University of Michigan, Ann Arbor, MI, United States
[4]School of Medicine, Wayne State University, Detroit, MI, United States

Corresponding Author:
Nic DePaula, PhD
School of Information Sciences
Wayne State University
42 W Warren Ave
Detroit, MI, 48202
United States
Phone: 1 313 577 1825
Fax: 1 313 577 7563
Email: ndepaula@wayne.edu

Abstract

Background: Public health agencies widely adopt social media for health and risk communication. Moreover, different platforms have different affordances, which may impact the quality and nature of the messaging and how the public engages with the content. However, these platform effects are not often compared in studies of health and risk communication and not previously for the COVID-19 pandemic.

Objective: This study measures the potential media effects of Twitter and Facebook on public health message design and engagement by comparing message elements and audience engagement in COVID-19–related posts by local, state, and federal public health agencies in the United States during the pandemic, to advance theories of public health messaging on social media and provide recommendations for tailored social media communication strategies.

Methods: We retrieved all COVID-19–related posts from major US federal agencies related to health and infectious disease, all major state public health agencies, and selected local public health departments on Twitter and Facebook. A total of 100,785 posts related to COVID-19, from 179 different accounts of 96 agencies, were retrieved for the entire year of 2020. We adopted a framework of social media message elements to analyze the posts across Facebook and Twitter. For manual content analysis, we subsampled 1677 posts. We calculated the prevalence of various message elements across the platforms and assessed the statistical significance of differences. We also calculated and assessed the association between message elements with normalized measures of shares and likes for both Facebook and Twitter.

Results: Distributions of message elements were largely similar across both sites. However, political figures ($P<.001$), experts ($P=.01$), and nonpolitical personalities ($P=.01$) were significantly more present on Facebook posts compared to Twitter. Infographics ($P<.001$), surveillance information ($P<.001$), and certain multimedia elements (eg, hyperlinks, $P<.001$) were more prevalent on Twitter. In general, Facebook posts received more (normalized) likes (0.19%) and (normalized) shares (0.22%) compared to Twitter likes (0.08%) and shares (0.05%). Elements with greater engagement on Facebook included expressives and collectives, whereas posts related to policy were more engaged with on Twitter. Science information (eg, scientific explanations) comprised 8.5% (73/851) of Facebook and 9.4% (78/826) of Twitter posts. Correctives of misinformation only appeared in 1.2% (11/851) of Facebook and 1.4% (12/826) of Twitter posts.

Conclusions: In general, we find a data and policy orientation for Twitter messages and users and a local and personal orientation for Facebook, although also many similarities across platforms. Message elements that impact engagement are similar across platforms but with some notable distinctions. This study provides novel evidence for differences in COVID-19 public health

XSL•FO
RenderX

messaging across social media sites, advancing knowledge of public health communication on social media and recommendations for health and risk communication strategies on these online platforms.

(JMIR Infodemiology 2022;2(2):e40198) doi:10.2196/40198

KEYWORDS

platform effects; COVID-19; social media; health communication; message design; risk communication; Twitter; Facebook; user engagement; e-government

Introduction

Background

Social media have become integral tools for public health messaging and online communication of health and risk information worldwide [1-3]. As of 2021, in the United States, 72% of adults and 84% of those aged 18-29 years say they use at least 1 social media site [4,5] and the sites are widely adopted by public health agencies [3,6,7]. On social media, public health messages can be shared by users, widening message reach. The public may also like and comment on agency messages, and agencies may directly reply to public comments. Although there are opportunities for public health messaging on these sites, there are also challenges. These sites have been sources of misinformation, especially concerning the COVID-19 pandemic [8,9, 10] and antivaccination propaganda [11,12]. The targeted marketing of health-harming products, such as e-cigarettes [13], has also been problematic. Nevertheless, given their prevalence, public health agencies need to understand the dynamics of these sites to better promote health behavior.

There is ample research on social media use by public health agencies [2,7,14,15,16]. However, studies are generally conducted on one site or another, either Facebook or Twitter. Although studies in other domains abound exploring the distinct affordances or characteristics of different social media sites [17-20], there are few studies examining user engagement with public health messages [13,21] and no analyses of the actual messages posted by public health agencies across social media platforms. Despite the lack of such comparative studies, it is important to understand the media effects, or at least the differences across sites. Studies often use the term "social media" broadly when they only investigate a single platform. However, the stark differences across some platforms are now well researched [22-24], and there has been an explicit call for addressing social media affordances in health communication research [25]. This study thus makes a novel contribution to the literature by comparing public health messaging and audience engagement across two of the most popular platforms in public health communication.

Public Health Message Design and Audience Engagement

Research on public health messaging on social media has focused on 2 broad areas: (1) the *content* and *purposes* of messages and (2) *audience* (or *user*) *engagement* with the messages. Analyses of message content have focused on "themes," such as "closures," "risk factors," "case updates," "reassurance," and others, in various pandemic and crisis contexts [26,27], including the COVID-19 pandemic [7,28]. Analyses of message purposes have discussed the goals of "to inform," "call to action" [28], increase "self-efficacy" [29], "fight misinformation" [30], and others. However, there is a lack of formalization of message design elements and little consideration for the more objective textual elements of messages, including relevant content, such as the speaker, audience, and types of images in the messages. To address this shortcoming, in this study, we adopted a framework of *textual and media message design elements* that identify the various objective characteristics of the text—focusing on the content, not on the purpose—which may be useful for multiple health and risk communication scenarios and related research [31].

Audience or user engagement on social media is often formalized in the platform via a *Like* button, a *Share* button, and a *Comment* function, the content or count of which is appended to the message. Facebook also offers other sentiments or reactions to be expressed that are formalized as buttons and counts (ie, *love*, *care*, *ha-ha*, *wow*, *sad*, and *angry*). Although social media reactions to messages may not directly relate to behavioral intent or actual behavior change, analyses of this engagement provide some insight into public interest in and acceptance of the messaging [25] and may therefore help improve message strategies and message design, what others have termed evidence-based science communication [13]. There is a downside to an overreliance on user engagement as the ultimate goal of social media communication, since user engagement is biased toward positive emotional or high arousal content [23,32, 33]. However, these metrics at least provide some evidence of the quality or success of health promotion and information campaigns on these platforms and can be used to increase message reach [13].

Platform Effects on Health and Risk Messages

Although studies in the social media literature recognize the distinct *affordances*—the functions or *action possibility* [25]—of these technologies, previous studies on COVID-19 lack a study of message elements across the most popular platforms: Facebook and Twitter [13,25]. Although they are similar, Facebook and Twitter share some key differences. On Facebook, connections of people are bidirectional and termed as "friends." On Twitter, they are unidirectional; individuals may follow others without being followed by them. This makes Twitter a more public and open platform. However, Facebook is a more popular site, with a marketplace, event calendars, and pages that can be unidirectionally followed [34]. On both Facebook and Twitter, individuals may make posts that include text, hyperlinks, and photos or videos, but the text length of a post is restricted on Twitter to 280 characters. They both have a *newsfeed* that presents users with posts of their friends, or those

XSL•FO
RenderX

followed, the organization of which is determined by the platform algorithms [35].

In practice, Facebook is more widely adopted than Twitter across all demographic groups [34]. Twitter has been used as a "news media" [36] and is associated with political news [37]. Twitter has been found to be more used for public information [38], whereas Facebook is used for "shared identities" [24] and "social interaction" [39] and is associated with higher levels of privacy concern and bonding social capital [22]. A recent study of user engagement with antismoking messages found that the message theme (ie, health/appearance/addiction, money, or family) has no impact on the click-through rate (CTR) of messages, but Facebook had the highest and lowest CTR levels and on average higher CTRs than the same messages on Twitter [13], showing that users on Facebook generally engage more than users on Twitter. However, messages on Twitter had a higher website CTR than those in any other platform, indicating that Twitter users are more likely to go to and scroll through the website linked to in the messages [13]. The literature thus supports the notion of Facebook as more of a social interaction platform, whereas Twitter is more of a news-oriented platform.

Research Objectives and Summary

For this study, we aim to assess differences in public health message design elements and audience engagement with the various message elements across Twitter and Facebook regarding COVID-19 during 1 year of the pandemic. We therefore ask the following research questions (RQs):

- RQ1. How do public health message design elements differ across Twitter and Facebook?
- RQ2. How does audience engagement with public health message elements differ across Twitter and Facebook?

In the following sections, we describe the methods of the study, the results, and the discussion in relation to the literature and provide evidence-based policy recommendations for better-targeted health communication strategies.

Methods

Data Collection and Sampling

We identified 11 major federal health agencies in the United States associated with infection prevention and control [40], the major public health agency of each of the 50 US states (plus Washington, DC), and the major local public health agency of each of the largest city/county in the 50 states. We then searched for the official account of these agencies on Twitter and Facebook, as well as their own website. Not all of the largest city/county public health agencies of the states had a Facebook or Twitter presence. From the list of agencies identified, we retrieved all COVID-19–related posts generated in 2020. This period enables an analysis of messages from the beginning of the pandemic through several waves. We then searched for any of the following strings anywhere in any of the posts of all identified agencies: *ncov*, *covid*, *corona*, *pandemic*, or *sars-cov*. To retrieve these posts, we used the standard Twitter application programming interface (API) and the Facebook API via Crowdtangle [41]. Note that the terms "post" and "message" are used here interchangeably. Unless otherwise specified, the term "post" refers to *original posts* and not *retweets* (shared posts) or *replies* (comments on other posts).

On Twitter, we identified 11 federal accounts (with a total of COVID-19–related original posts and retweets), 48 state accounts (with a total of 40,716 posts and retweets), and 33 local accounts (with a total of 20,164 posts and retweets) that matched the criteria. On Facebook, we identified 10 federal accounts (with a total of 3592 posts), 49 state accounts (with a total of 34,930 posts), and 38 local accounts (with a total of 14,356 posts) that matched the criteria. On Facebook, it is more difficult to differentiate original posts from shared posts; the figures just reported for Facebook include both. This data set of all COVID-19–related posts from all identified agencies in 2020 was called the *population data set*.

For manual content analysis, we used a stratified random sampling technique where we sampled 900 posts from Twitter and 900 posts from Facebook proportional to the amount of posts made by agency level (ie, local vs state vs federal), the *sample data set*. The rationale for the sampling was based on similar studies and generating a manageable number of posts to manually code. For example, Reuter et al [13] analyzed a total of 1275 antismoking health messages posted across 3 social media platforms, and Slavik et al [15] used 501 tweets for content analysis of Canadian public health agencies' messages on Twitter. We should note that for Facebook, our sampling strategy only focused on posts that were shorter than 340 characters (which may include relatively long hyperlinks). This was intended to provide a data set more comparable to Twitter posts, which are restricted to 280 characters (where hyperlinks may be shortened). After removing nonrelated posts, reply posts, and shared posts, or posts without any discernible content, our *final sampledata set* consisted of a total of 1677 (93.2%) posts (826, 49.3%, original Twitter posts and 851, 50.7%, original Facebook posts) that were coded. For Twitter, this included 82 (9.9%) federal posts, 482 (58.4%) state posts, and 262 (31.7%) local posts. For Facebook, this included 60 (7.1%) federal posts, 560 (65.8%) state posts, and 231 (27.1%) local posts. Multimedia Appendix 1 presents the sampled accounts.

Coding Framework

We adapted an existing framework [31] for the analysis of health and risk communication social media message elements. The framework is based on theories of text analysis [31,42, 43] and social media studies in health and crisis communication [7,15,28,29], including image use in risk communication [44]. These are interdisciplinary studies in the health communication, health informatics, and crisis communication literature. The framework focuses on message elements that are more objective compared to the abstract (eg, "open and transparent message" [45]) and metaphorical (eg, "fighting misinformation" [30]) categories used in the literature—or assuming everything is a "frame" or "theme" [26,27]. Message elements in this framework are composed of *textual* and *media elements*. The framework integrates message elements into 8 major dimensions: *speech function*, *topic*, *threat focus*, *type of resource*, *audience*, *speaker*, *rhetorical tactic*, and *media*. Each of these dimensions includes more granular *message features* (or *elements*). Tables 1 and 2 introduce definitions and examples

XSL•FO
RenderX

of the textual and media elements, respectively. The framework is not exhaustive and could be reduced or expanded, as needed. It is conceived for relatively short social media posts, since the analysis focuses on the clause or sentence level, and therefore lengthier documents would be largely more complex to analyze. Further details of the framework and the elements are provided in Multimedia Appendix 2.

XSL•FO
RenderX

Table 1. Definitions and examples of message elements: textual.

Textual element	Definition	Example
Speech function		
Representative	Clause in declarative form, describing a behavior, state, or event	“#COVID19 can be spread by people who do not have symptoms”
Directive	A sentence that directs, commands, or mandates an action, especially via an imperative sentence	“Continue to wear masks” OR “Donate blood.”
Question	A rhetorical question or question prompt	“Are you looking for work? We are hiring!”
Expressive	Expression of sentiment by the message speaker (eg, sadness, appreciation)	“Thank you, #EMS heroes, for staying strong”
Request	Request to participate in research, volunteer, or means to reach an agency	“Call us for questions at this number”
Topic		
Protection	Information about what to do to prevent or treat the issue	“Disinfect things you and your family touch frequently”
Policy	Actions, policies, or programs of officials, government agencies, or related entities	“Multnomah County is almost ready for reopening schools.”
Surveillance	Statistics or data about prevalence (eg, cases/deaths)	“Yesterday, there were 85 new deaths”
Science	Describes or explains a cause, mechanism, or symptom of the issue	“there is no evidence that produce can transmit #COVID19”
Emergent	Event of emergency concern or immediate priority	“Travelers: DON'T book air travel to NY for just a few days”
Resource type		
Interactive	Interactive service, such as question-and-answer (Q&A) with policy makers or watching live	“FDA will host a virtual Town Hall on 3D printed swabs”
Material	Testing sites, financial assistance, vaccine provision	“Use our map to find locations for vaccination sites.”
Corrective	Correction of a rumor, misinformation, or pointing to related resources	“A death previously reported in Warren was incorrect, and has been removed.”
Focus and audience		
Group	Refers to a demographic group (eg, adults, Hispanics) or a vulnerable population	“Cancer patients are among those at high risk of serious illness from a COVID19 infection.”
Secondary	Consequences of or issues directly related to the main issue	“Many are feeling stressed because of #COVID19.”
Other language	Message or part of message in another language, including sign language	“Números del #COVID19 en California:”
Speaker		
External	Expert or staff from another agency	“The head of the CDC will speak…”
Political	Mayor, governor, or other political figure	“Watch the Mayor’s updates on…”
Expert	Expert or staff of the agency	“Our own Dr. Elinore will discuss the crisis”
Personality	Nonpolitical or nongovernmental personality, including celebrities or community members	“Juan from Blue Eagles football club speaks about COVID19”
Rhetorical		
Collective	Focus on collective terms to characterize an issue or to address it	“We all need to do our part to combat Covid-19”
Emphasis	Sentence with an explanation point or with all capitalized directive	“WEAR a mask!”
Positive	Positive framing of agency action	“We’re making progress is getting vaccines”
Metaphor	Using metaphors to explain the science or prevention of the issue	“The swiss cheese respiratory virus defense”

XSL•FO
RenderX

Table 2. Definitions and examples of message elements: media.

Media element	Definition	Example
Hyperlink	A long or short web URL	https://twitter.com/...
Hashtag	Any term preceded by a # symbol	#COVID-19 #WearAMask
Text-in-image	Image with additional text not included in the text part of the message	See examples below.
Illustration	Illustration in the image—at least beyond use of a table and colors	
Photograph	Photograph of a person, object, or scene	
Infographic	Image that conveys data or illustrated directives (overrides illustration)	
Video	A video embedded in the message	

Content Analysis

The content analysis consisted of manual binary coding for the presence or lack of each element in a post. As the definition of the categories became apparent, the nature of some definitions made some categories mutually exclusive, especially within each textual or media dimension. For example, a question is, by definition, not a *representative* and not an *expressive*. These coding rules are summarized in Tables 1 and 2 and are further detailed in Multimedia Appendix 2.

A random training sample of 150 posts (75, 50%, from Twitter and 75, 50%, from Facebook) was first retrieved for training and category development. Using these 150 posts, during training, 3 authors updated and defined the message categories. Once this training was accomplished, the 3 authors independently began coding a 20% subsample of the *sample data set*, where at least 2 coders double-coded the same post to calculate the Cohen κ statistic of interrater reliability (IRR).

After obtaining IRR measures, the coders discussed the results. At this point, the results were not perfect and discrepancies in coding existed and needed to be reconciled. In particular, there were issues with the *representative* and *request* speech functions, the *external speaker*, and some of the *rhetorical* dimensions. For example, it was not clear whether a slogan on an image, such as "COVID-19 news update," was to be considered a *representative* sentence. We ultimately agreed on the definitions as shown in Multimedia Appendix 2, but IRR results were ultimately not perfect for all categories. The κ values are provided in Multimedia Appendix 1. After the IRR analyses, we discussed issues identifying the categories and then better defined and narrowed the rules for final coding of the data. In the cases that discrepancies existed across coders, and categories were revised, we re-examined the data based on the revised definitions and obtained agreement among coders. We then set out to code the remaining data. Each coder independently coded approximately 450 posts, producing a final sample data set of 1677 posts for statistical analysis.

Statistical Analyses

To address our first RQ, we calculated the distribution of each message element on Twitter and Facebook and then compared this total across platforms via an independent 2-sample Z-test of proportions, where the null hypotheses assumed that the proportion of each message element is equal on both platforms. Although Z-tests expect normal distributions, and social media phenomena are notoriously not normally distributed, given the relatively large sample of most message elements, we found it reasonable to apply the Z-tests [44].

To address our second RQ, we operationalized audience engagement as normalized frequencies of likes and shares. Other studies have used the CTR to measure audience engagement [13], seemingly nonnormalized tweet counts [15], and regression models where follower count, and other dimensions, are controlled for [45]. The CTR measure used by Reuter et al [13] was not possible for our study since we could not have access to message clicks or actual message views (the total_views field provided by the Facebook API was not reliable and contained missing data; no such measure was provided by the Twitter API). Our approach is simpler than the regression models, but given the focus on a single issue, the random sampling of data across agencies and time, and normalized measure of likes and shares based on an agency's follower count, our approach provides a robust and easy-to-interpret method to test the association between message features and audience engagement.

We calculated a measure of normalized likes (NL_m) as the number of likes of each message "m," divided by the follower count of the account that posted the message. NL_m is the percentage of the agency's follower count that liked the message. Although Facebook includes additional positive and negative measures of audience engagement—namely *love*, *care*, *ha-ha*, *wow*, *sad*, and *angry*—these were not included as part of the NL_m measure to make it more comparable with the single *like* feature of Twitter. Although we considered and analyzed the more negative measures of Facebook sentiment, namely sad and angry, these overly complicated the research and ultimately seemed out of scope, since our aim was to compare Facebook

XSL•FO
RenderX

and Twitter elements. This study thus focused on only likes and shares on Facebook and Twitter, both of which are types of *positive engagement*. Generally, in this study, engagement refers to “liking” or “sharing” a message.

Similar to normalized likes, we created a measure of normalized shares (NS_m) of each message “m.” The NS_m measure, compared to likes, can be more directly considered a *diffusion rate* [46] or *retransmission rate* [7] of a message (or message elements), since it is a direct share by the user to its network. Although messages are not only liked or shared by the followers of an account, the size of an account’s followers largely influences the total engagement with posts of that account [47]. Equations of these normalized like and normalized share measures are provided in Multimedia Appendix 3.

For every message element, we then computed the *mean NS* and *mean NL* of all messages that contained the element, and of all messages that did not contain it, and compared these 2 groups via a 2-tailed independent-samples Wilcoxon-Mann-Whitney (WMW) test, given the skewness of the data and as similar studies have approached it [15]. We considered and discussed $P \leq .05$ as statistically significant.

Results

Data Set Details

Table 3 shows descriptive statistics for the final sample data set in relation to the population data set of COVID-19–related posts.

The list of agency accounts in the sample are in Multimedia Appendix 1. As shown in Table 3, local, state, and federal agencies made a comparable number of Facebook and Twitter posts (these measures do not include shares or retweets). In general, per account, state agencies were more active in posting than local and federal agencies. For example, on Facebook—based on population statistics—state accounts made 712 posts per account (34,930 total posts by 49 accounts), whereas local accounts made 377 posts per account (14,356 total posts by 38 accounts), and federal accounts made 359 posts per account (3592 total posts by 10 accounts). Results were relatively similar for Twitter.

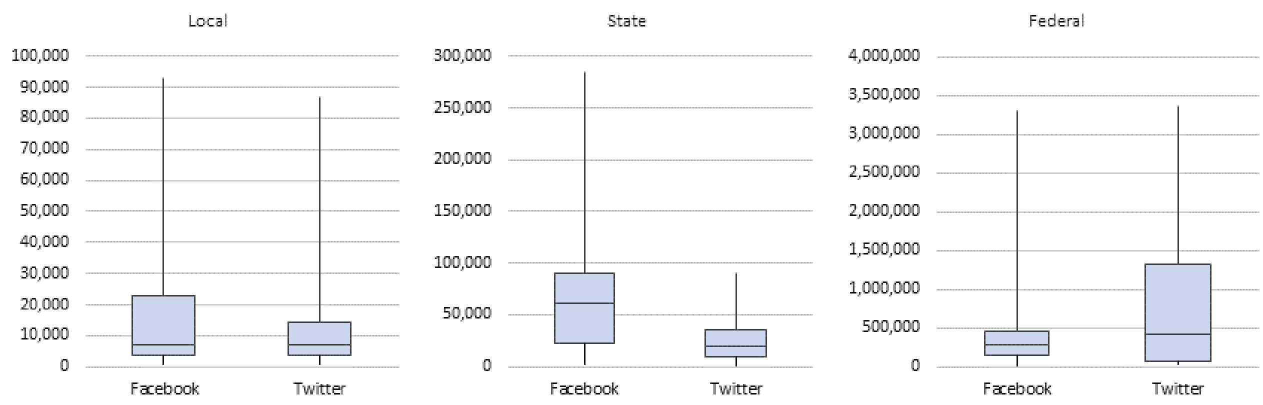
Figure 1 shows the mean and IQR of account followers, separated for local, federal, and agency accounts (based on the sample data set). There were strong variations across local, state, and federal agencies in the distribution of followers and platforms. Not surprisingly, federal agency accounts had the most followers, and state agencies had more followers than local agencies, on average. Federal agencies were more popular (ie, had more followers) on Twitter, whereas state agencies were more popular on Facebook. Local agencies were similarly popular on Facebook and Twitter. Generally, there is great variation in the top quartile of the distribution. Detailed numbers for this box plot can be found in Multimedia Appendix 3.

Table 3. Statistics[a] for the sample data set as a percentage of the population of COVID-19 posts in 2020.

	Local, n/N (%)	State, n/N (%)	Federal, n/N (%)	All, n/N (%)
Facebook accounts	32/38 (84.0)	48/49 (98.0)	9/10 (90.0)	89/97 (92.0)
Twitter accounts	29/33 (88.0)	45/48 (94.0)	9/11 (82.0)	83/92 (90.0)
Facebook total posts	231/14,356 (1.6)	560/34,930 (1.6)	60/3592 (1.7)	851/52,878 (1.6)
Twitter total posts	262/15,421 (1.7)	482/27,866 (1.7)	82/4620 (1.8)	826/47,907 (1.7)

[a]Statistics are for the final sample data set used in content and statistical analyses in relation to the population data set of all COVID-19–related posts from all accounts identified in 2020.

Figure 1. Box plot of IQRs of followers per account across agency levels and platforms.



XSL•FO
RenderX

Platform Effects on Message Design

Table 4 shows the total count of each message element in the coded sample data set as the number of posts in which the element appeared, separately for Facebook and Twitter. Table 4 also provides results from a 2-tailed Z-test that compares whether the proportions are equal across platforms. Results showed that most features are used to a similar extent across platforms. These results provide some validity for the notion that these message features are indeed part of public health and risk communication on social media more broadly. However, we also found some statistically significant differences across the 2 sites. A positive Z-score indicates higher use on Twitter; a negative score indicates higher use on Facebook.

Figure 2 shows the message elements used significantly more or less on Facebook or Twitter, relative to each other, the bars identifying the percentage of posts in which each message element appeared. *External*, *political*, and *expert* actors, along with *video*, *photograph*, and *other language*, were the features more frequently used in Facebook posts compared to Twitter posts. *Policy*, *directive*, *infographic*, *surveillance*, *hyperlink*, and *hashtag* features were used more frequently on Twitter compared to Facebook. *Personality* and *positive* framing features were not included in Figure 2 due to the low sample size. However, *policy* was included in the graph, although at the significance boundary.

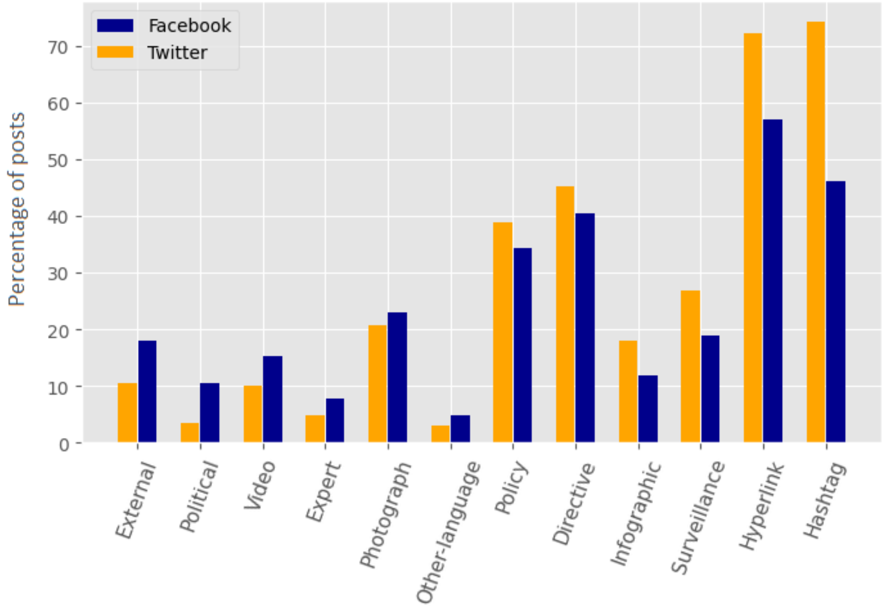
XSL•FO
RenderX

Table 4. Message design elements across Facebook (n=851) and Twitter (n=826) posts.

Message element	Facebook, n (%)	Twitter, n (%)	Z-score	*P* value
Speech function				
Representative	755 (88.7)	722 (87.4)	–0.83	.41
Directive	344 (40.4)	374 (45.2)	2.01	.04
Question	107 (12.5)	96 (11.6)	–0.60	.55
Expressive	79 (9.2)	77 (9.3)	0.03	.98
Request	28 (3.2)	38 (4.6)	1.40	.17
Topic				
Protection	391 (45.9)	395 (47.8)	0.77	.44
Policy	292 (34.3)	321 (38.8)	1.93	.05
Surveillance	160 (18.8)	222 (26.8)	3.94	<.001
Science	73 (8.5)	78 (9.4)	0.62	.53
Emergent	39 (4.5)	26 (3.1)	–1.52	.13
Resource type				
Interactive	192 (22.5)	175 (21.1)	–0.68	.49
Material	112 (13.1)	112 (13.5)	0.24	.81
Corrective	11 (1.2)	12 (1.4)	0.28	.78
Focus and audience				
Group	85 (9.9)	113 (13.6)	2.34	.02
Secondary	73 (8.5)	59 (7.1)	–1.09	.27
Other language	42 (4.9)	25 (3.0)	–1.99	.04
Speaker				
External	153 (17.9)	86 (10.4)	–4.43	<.001
Political	89 (10.4)	28 (3.3)	–5.68	<.001
Expert	66 (7.7)	39 (4.7)	–2.56	.01
Personality	17 (1.9)	5 (0.6)	–2.51	.01
Rhetorical				
Collective	123 (14.4)	105 (12.7)	–1.04	.30
Emphasis	103 (12.1)	81 (9.8)	–1.50	.13
Positive	12 (1.4)	23 (2.7)	1.97	.05
Metaphor	5 (0.5)	2 (0.2)	–1.10	.27
Media				
Hyperlink	485 (56.9)	597 (72.2)	6.54	<.001
Hashtag	392 (46.0)	613 (74.2)	11.76	<.001
Text-in-image	387 (45.4)	343 (41.5)	–1.63	.10
Illustration	235 (27.6)	258 (31.2)	1.63	.10
Photograph	196 (23.0)	170 (20.5)	–1.22	.22
Infographic	101 (11.8)	149 (18.0)	3.55	<.001
Video	130 (15.2)	83 (10.0)	–3.21	<.001

XSL•FO
RenderX

Figure 2. Elements used significantly more on Facebook and significantly more on Twitter.



Platform Effects on Audience Engagement

Tables 5 and 6 show audience engagement with messages containing each specific feature compared to those without the feature, calculated separately for Facebook and Twitter as normalized likes and normalized shares. In general, Facebook had higher engagement of users compared to Twitter. In addition, Facebook users used shares more frequently than likes, while Twitter users liked more frequently than they shared. Facebook posts, on average, were liked by 0.19% of account followers, whereas on Twitter, on average, posts were liked by 0.08% of account followers, a difference of 2.25 times higher for Facebook likes. Regarding sharing, Facebook posts, on average, were shared by 0.22% of account followers, whereas on Twitter, on average, posts were shared by 0.05% of account followers, which is more than a 4.4 times difference. However, these engagement measures do not include other forms of engagement on Facebook (eg, *love*, *care*), as previously discussed under Methods.

Table 5 provides the mean normalized likes of all messages with the feature compared to those without it, along with *P* values for the WMW test comparing these 2 sets. For example, in the Facebook sample, on average, 0.16% of the (count of the) account's followers liked the message that contained a *representative*, whereas 0.26% liked the messages that did not contain a *representative*. Therefore, on Facebook, messages that did not contain a *representative* were liked more than messages that did. However, this was not a statistically significant difference (P=.22). On Twitter, however, on average, 0.08% of the account's followers liked messages that contained a *representative* and 0.05% liked messages that did not contain it, which was a significant difference (P<.001).

Table 6 provides the mean normalized shares of all messages with the feature and those without it, along with *P* values from the WMW test comparing differences between them. Results here can be similarly interpreted as the results in Table 5.

Figure 3 shows the message elements from Tables 5 and 6 that had a significant association with an increase or decrease in normalized likes and shares. Figure 3 shows the percentage points in the increase/decrease associated with the inclusion of the message element. *Expressives* and the use of a *collective* frame in messages were associated with more likes across both platforms. *Surveillance* information as well as *infographics* were also associated with more likes and shares across Facebook and Twitter. References to *material* resources, surprisingly, were generally associated with fewer likes and shares on both platforms. We speculate this may be due to the repeated posts about testing and vaccine sites coded under *material*. Although *political* figures were more present on Facebook compared to Twitter, they were associated with less engagement on both platforms, especially Facebook. *Requests* were particularly popular on Facebook but not significant on Twitter. *Correctives* and *policy information* were associated with higher engagement on Twitter but less so or not significantly on Facebook.

XSL•FO
RenderX

Table 5. Mean percentage of account followers that liked messages with and without specific elements.

Message element	Facebook			Twitter		
	With feature	Without feature	*P* value[a]	With feature	Without feature	*P* value[a]
Speech function						
Representative	0.16	0.26	.22	0.08	0.05	<.001
Directive	0.20	0.15	.01	0.07	0.09	<.001
Question	0.26	0.16	.04	0.05	0.08	<.001
Expressive	0.28	0.16	<.001	0.10	0.08	<.001
Request	0.52	0.16	.05	0.06	0.08	.32
Topic						
Protection	0.18	0.17	.43	0.08	0.08	.02
Policy	0.19	0.17	.03	0.09	0.07	.20
Surveillance	0.13	0.18	.02	0.12	0.07	<.001
Science	0.14	0.18	.41	0.05	0.08	.08
Emergent	0.14	0.17	.26	0.25	0.07	.06
Resource type						
Interactive	0.17	0.17	.20	0.07	0.08	.04
Material	0.05	0.19	<.001	0.05	0.08	<.001
Corrective	0.18	0.17	.49	0.41	0.07	.03
Focus and audience						
Group	0.16	0.17	<.001	0.04	0.09	<.001
Secondary	0.13	0.18	.13	0.06	0.08	.01
Other language	0.10	0.18	.07	0.02	0.08	<.001
Speaker						
External	0.13	0.18	.07	0.06	0.08	.13
Political	0.12	0.18	.01	0.06	0.08	.08
Expert	0.17	0.17	.06	0.06	0.08	.42
Personality	0.22	0.17	.01	0.06	0.08	.30
Rhetorical						
Collective	0.27	0.16	<.001	0.10	0.08	.004
Emphasis	0.29	0.16	.004	0.08	0.08	.10
Positive	0.41	0.17	.12	0.10	0.08	.43
Metaphor	0.41	0.17	.09	0.02	0.08	.26
Media						
Hyperlink	0.15	0.20	<.001	0.07	0.10	<.001
Hashtag	0.19	0.16	.32	0.07	0.10	.01
Text-in-image	0.17	0.17	.01	0.09	0.07	.002
Illustration	0.10	0.20	.03	0.06	0.09	.12
Photograph	0.21	0.16	.08	0.07	0.08	<.001
Infographic	0.20	0.17	<.001	0.12	0.07	<.001
Video	0.21	0.17	.07	0.07	0.08	.09

[a]*P* values refer to the Wilcoxon-Mann-Whitney test of comparing the mean normalized likes for posts containing the feature with those not containing it, separately for Facebook and Twitter.

XSL•FO
RenderX

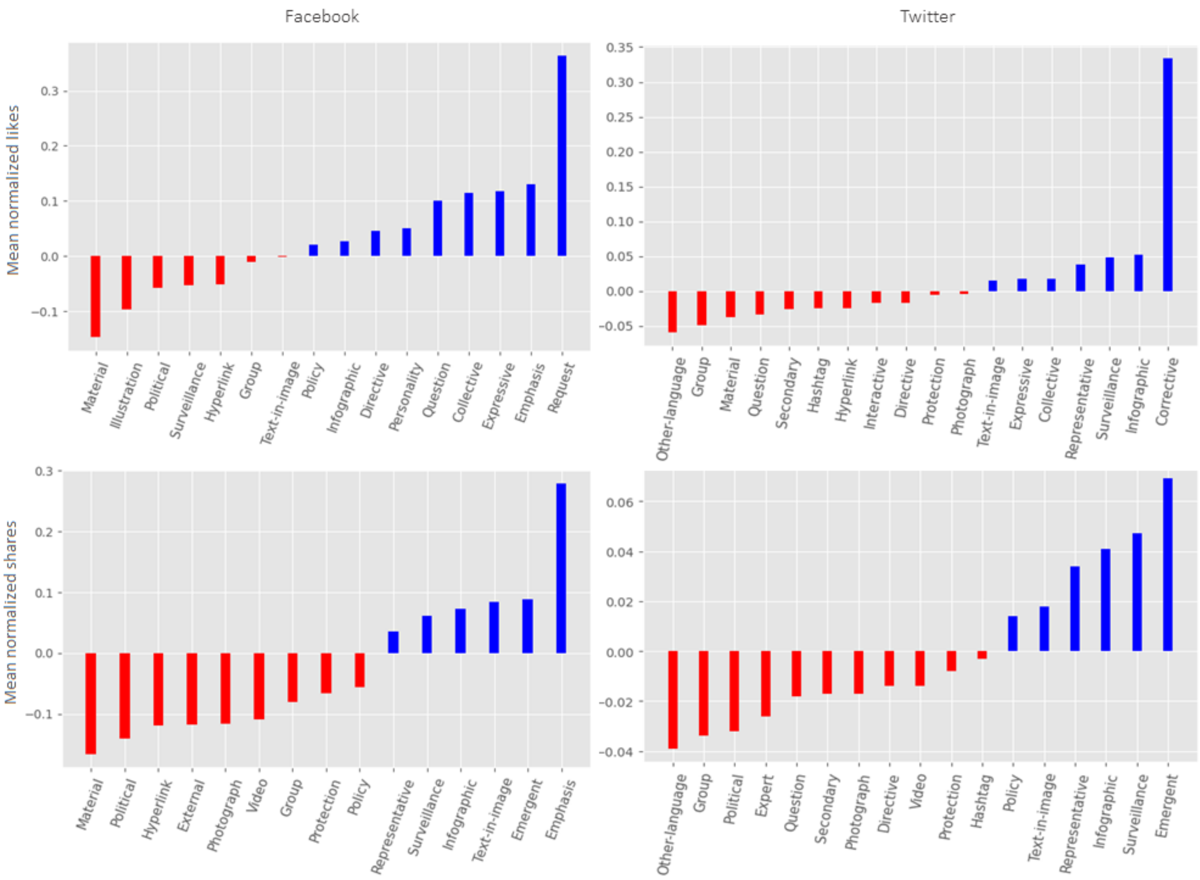
Table 6. Mean percentage of account followers that shared messages with and without specific features.

Message element	Facebook			Twitter		
	With feature	Without feature	*P* value[a]	With feature	Without feature	*P* value[a]
Speech function						
Representative	0.20	0.16	.01	0.06	0.03	<.001
Directive	0.17	0.21	.27	0.05	0.06	<.001
Question	0.22	0.19	.10	0.04	0.06	.003
Expressive	0.29	0.19	.06	0.07	0.05	.06
Request	0.53	0.18	.18	0.05	0.06	.39
Topic						
Protection	0.16	0.23	.03	0.05	0.06	.002
Policy	0.16	0.21	.004	0.06	0.05	.04
Surveillance	0.25	0.18	<.001	0.09	0.04	<.001
Science	0.15	0.20	.36	0.04	0.06	.05
Emergent	0.28	0.19	.04	0.12	0.05	.03
Resource type						
Interactive	0.30	0.17	.34	0.05	0.06	.35
Material	0.05	0.22	<.001	0.04	0.06	.45
Corrective	0.18	0.20	.29	0.18	0.05	.19
Focus and audience						
Group	0.12	0.20	.001	0.03	0.06	<.001
Secondary	0.33	0.18	.19	0.04	0.06	.01
Other language	0.09	0.20	.16	0.02	0.06	.001
Speaker						
External	0.10	0.22	.04	0.05	0.06	.18
Political	0.07	0.21	<.001	0.02	0.06	.01
Expert	0.07	0.21	.26	0.03	0.06	.02
Personality	0.08	0.20	.31	0.03	0.06	.21
Rhetorical						
Collective	0.20	0.19	.07	0.07	0.05	.28
Emphasis	0.44	0.16	.04	0.06	0.05	.31
Positive	0.22	0.20	.34	0.05	0.06	.40
Metaphor	1.63	0.19	.27	0.01	0.06	.14
Media						
Hyperlink	0.14	0.26	.003	0.05	0.06	.08
Hashtag	0.26	0.14	.37	0.05	0.06	.04
Text-in-image	0.24	0.16	<.001	0.07	0.05	<.001
Illustration	0.15	0.21	.40	0.05	0.06	.13
Photograph	0.11	0.22	<.001	0.04	0.06	<.001
Infographic	0.26	0.19	<.001	0.09	0.05	<.001
Video	0.10	0.21	.01	0.04	0.06	.01

[a] *P* values refer to the Wilcoxon-Mann-Whitney test of comparing the mean normalized shares for posts containing the feature with those not containing it, separately for Facebook and Twitter.

XSL•FO
RenderX

Figure 3. Significant changes in likes and shares associated with the inclusion of message element. The blue bars refer to increases and the red bars to decreases in mean normalized likes and mean normalized shares associated with the inclusion of the message element.



Discussion

Principal Findings

This study analyzed 1677 COVID-19–related posts on Facebook and Twitter, by public health agencies across the United States in 2020, and found differences and similarities in the overall use and popularity of these sites in terms of message design elements and audience engagement. Our results show that Facebook posts received 2.25 times more likes and 4.4 times more shares, in general, than posts on Twitter. However, within each platform, messages received more shares than likes within Facebook—as a percentage of account followers that liked or shared the message—whereas on Twitter, measures were more liked than shared.

Our results show that messages on Twitter, compared to Facebook, are significantly more focused on *surveillance information* (eg, data and statistics about the threat), *policy information*, *infographics*, and *hyperlinks*. Moreover, federal agencies are more active and more popular on Twitter compared to Facebook, whereas local and state agencies are more active or more popular on Facebook. We also observe that messages on Facebook, compared to Twitter, have significantly more references to *political figures*, public health *experts*, and (nonpolitical) *personalities* (eg, personal stories or local celebrities) as speakers in the messages. From this, we may conclude a type of *data and policy* orientation for Twitter and a *local and personal orientation* for Facebook.

We observed that data (eg, *infographics*, *surveillance* data) and *policy* information had significant positive associations with audience engagement on Twitter but not at all or not as much on Facebook, further suggesting this data and policy characterization for Twitter. Although Facebook was the platform where political figures and health experts were more highlighted as speakers in the messages, this personalization was generally not associated with higher engagement on both sites. However, we observe that *photographs*, which are often of people, and rhetorical elements, such as a *collective* framing (eg, "we are in this together"), *positive* framing (eg, "we are trying our best"), and *emphasis* (eg, exclamation points), which may trigger sentiment and personal connection, received more or significantly more audience engagement on Facebook but not as much or not at all on Twitter. This further suggests the local and personal orientation for Facebook.

The distribution of message design elements is largely similar across both platforms, suggesting consistency in public health messaging, but with some significant differences between the 2 social media sites studied. Results also show significant associations between message elements and audience engagement, with some expected and surprising differences across platforms. In general—for this health and risk communication scenario—we may thus suggest that Twitter

XSL•FO
RenderX

has more of a data and policy orientation, whereas Facebook has more of a local and personal orientation on the content, which largely follows the literature on social media affordances.

Integration With Existing Literature

Previous studies have examined the characteristics of Facebook in relation to Twitter as 2 of the major social media sites in the United States and in the world today. Generally, studies support the notion that Twitter is more of a "news media" [22,36] for "information dissemination" [38] and for being "quickly informed" [39], while Facebook is more for "shared identities," "photographs" [24], and "social interaction" [39], being associated more with bonding social capital [22]. This distinction between Twitter and Facebook is usually explained as the specific affordances of each site [13,25], which may be related to some of its technical features, such as the more open unidirectional networks of Twitter compared to the bidirectional networks of Facebook [38]. Studies also suggest that certain technical features of a site (eg, focus on visual imagery) may lead to an overall higher audience engagement [13,22].

In this study, we did not analyze whether certain platform features caused the use of specific message elements or whether certain message features caused more or less engagement. However, our results generally support the existing literature that suggests that Facebook, while bigger and more popular across the US adult population, has more of a local and personal orientation, associated with close social interactions. Twitter, in contrast, is both a more active and a popular site for federal agencies, compared to local and state agencies, and both the content and engagement on Twitter point to more of a data and policy orientation. Ultimately, we observe great similarities in message elements and audience engagement across Facebook and Twitter, suggesting a standardization of social media policies and practices across agencies and platforms, and also similarities in user engagement on both Facebook and Twitter.

Contributions to Health Communication Policy

This study provides some evidence for policy recommendations on social media health communication strategies. These recommendations are based on the results of this study, which is focused on COVID-19 communication during the beginning and multiple waves of the pandemic in 2020. Public health agencies and further research need to assess whether these are valid for broader contexts as well.

Recommendation 1

For public health agencies using Facebook, we recommend caution when using political figures and external experts on their messages and instead highlight nonpolitical or nongovernment personalities, such as local celebrities or ordinary individuals who have a special story to tell. We also see an opportunity for greater or at least continued use of emotional expressions on messages and the use of collective frames to generate greater positive engagement.

Our results show that messages on Facebook, compared to Twitter, are significantly more focused on highlighting political figures, as well as internal and external experts. However, political figures and external experts were generally associated with less engagement on Facebook. Personalities, including celebrities or ordinary people (eg, an authentic post of a child from the community), were significantly associated with greater engagement on Facebook but were present in few posts (2%) on Facebook. Ultimately, the use of expressives (ie, expressing emotions) and collective frames (eg, using collective pronouns and focusing on collective issues) were particularly well engaged with on Facebook.

Recommendation 2

For public health agencies using Twitter, we recommend caution on the use of hyperlinks and hashtags on Twitter messages if the goal is to increase message likes and overall message diffusion, but continued use of surveillance information and infographics is recommended. Moreover, we recommend a greater focus on messages containing emergent issues (eg, emergency or timely information), and the use of correctives to address misinformation, because these were both not prevalent but were associated with greater positive engagement.

Our results show that messages on Twitter, compared to Facebook, are significantly more focused on policy and surveillance information and include significantly more hyperlinks and hashtags compared to messages on Facebook. Since the hashtag is a textual construction first popularized on Twitter, this is not surprising. However, both hashtags and hyperlinks were generally associated with less engagement on Twitter. Surveillance information and infographics, however, were generally associated with greater engagement on Twitter. Emergent issues, and correctives, were particularly well engaged with on Twitter. However, correctives were included in a minority of tweets (1.4%). Given that social media is part of a misinformation crisis [48], or *infodemic* [49,50], it is important to consider how public health agencies are addressing misinformation on these environments.

Recommendation 3

For public health agencies using both platforms, we recommend careful use of images in their messages, including photographs, illustrations, and videos, as these were all media types associated with less engagement across both platforms. However, including text-in-image is a reasonable recommendation, since these were associated with greater engagement across platforms.

In general, our results show that not all types of images are similarly engaged with. On both platforms, photographs were significantly associated with fewer shares, whereas infographics were generally associated with greater shares and likes. Although illustrations were associated with fewer likes and shares on both platforms, this negative impact was only significant for Facebook likes. Infographics about the pandemic were associated with higher engagement on both platforms, but they were also largely prevalent. Therefore, the amount of use of these features in this context is likely sufficient. Lastly, text-in-image was generally associated with greater likes and shares on Twitter and greater sharing on Facebook, highlighting the importance of textual and semantic content along with visual content.

XSL•FO
RenderX

Limitations and Future Work

This study intended to show how public health agencies construct their messages across Facebook and Twitter and how users respond to these messages similarly or differently across platforms. To control for aspects of the message topic, we only focused on COVID-19–related messages. COVID-19 is also a major health and risk issue and one that we could expect public health agencies in the country to be communicating about in 2020. However, the focus on COVID-19 puts a limitation on the extent to which we can generalize the findings to health and risk communication more broadly. Moreover, the statistical tests used could be improved with a regression model that assesses and controls for other variables on audience engagement. Nevertheless, our random sampling technique, over multiple kinds of agencies and an entire year, helps us generalize and have confidence in the results.

Health communicators should consider that social media algorithms themselves are problematic as they lead to echo chamber effects [35] and are biased toward hyperactive users [51]. Audience engagement on social media itself should thus be considered with care. The literature generally points to social media engagement as being driven by high emotional content [52], out-group animosity [53], and fear-arousing sensationalism [54]. Simply acquiring more engagement is thus not always appropriate for health and risk communicators. Moreover, there is a chance that social media in government may be used for political purposes [55,56]. Future studies may thus advance this work by examining the *quality* of engagement across platforms, political issues in public health communication, and examining the nature of the comments to public health messages.

There were few posts with personalities featured on Facebook (17/851, 1.9%) and Twitter (5/826, 0.6%) posts. We could thus not properly assess the impact of this message element on engagement. However, celebrities and personal stories can positively influence health behavior and may be further studied in this context [54,57]. In addition, analyses of fear appeals, distinctions between more or less informative (or scientific) messages, or the use of storytelling, could have improved this study. Some message features need better definition to increase reliability, including representatives and requests. The category of representatives and its results here should be considered with caution, since it is the broadest category of the framework and had a low κ. In all, future research may gain from refining the framework categories, further examining the use of celebrities or personal stories, and the relationship between fear-appeals or other rhetorical strategies on different levels and qualities of user engagement.

Conclusion

In general, we find a data and policy orientation for Twitter messages and users and a local and personal orientation for Facebook, although also many similarities across both platforms. Message elements that impact engagement are similar across both platforms but with some notable distinctions. This study provides novel evidence for differences in COVID-19 public health messaging on social media, advancing health communication research and recommendations for health and risk communication strategies.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sampled accounts and κ values.
[DOCX File , 38 KB - infodemiology_v2i2e40198_app1.docx]

Multimedia Appendix 2
Detailed framework description and coding rules.
[DOCX File , 562 KB - infodemiology_v2i2e40198_app2.docx]

Multimedia Appendix 3
Sample statistics and analyses.
[DOCX File , 21 KB - infodemiology_v2i2e40198_app3.docx]

References

1. Centers for Disease Control and Prevention (CDC). Crisis and Emergency Risk Communication (CERC). 2018. URL: https://emergency.cdc.gov/cerc/ [accessed 2022-11-09]
2. Thackeray R, Neiger BL, Burton SH, Thackeray CR. Analysis of the purpose of state health departments' tweets: information sharing, engagement, and action. J Med Internet Res 2013 Nov 11;15(11):e255 [FREE Full text] [doi: 10.2196/jmir.3002] [Medline: 24217361]
3. Hou Q, Zhao Y, Su X, Rong R, Situ S, Cui Y. Using Sina-Weibo microblogs to inform the development and dissemination of health awareness material about Zika virus transmission, China, 2016-17. PLoS One 2022 Jan 27;17(1):e0261602 [FREE Full text] [doi: 10.1371/journal.pone.0261602] [Medline: 35085257]

XSL•FO
RenderX

4. Shearer E, Matsa K. News Use Across Social Media Platforms 2018. 2018. URL: https://www.pewresearch.org/journalism/2018/09/10/news-use-across-social-media-platforms-2018/ [accessed 2022-11-10]
5. Hughes A, Wojcik S. 10 Facts about Americans and Twitter. 2022. URL: https://www.pewresearch.org/fact-tank/2019/08/02/10-facts-about-americans-and-twitter/ [accessed 2022-11-10]
6. Alonso-Cañadas J, Galán-Valdivieso F, Saraite-Sariene L, Caba-Pérez C. Committed to health: key factors to improve users' online engagement through Facebook. Int J Environ Res Public Health 2020 Mar 11;17(6):1814 [FREE Full text] [doi: 10.3390/ijerph17061814] [Medline: 32168792]
7. Sutton J, Renshaw SL, Butts CT. COVID-19: Retransmission of official communications in an emerging pandemic. PLoS One 2020;15(9):e0238491 [FREE Full text] [doi: 10.1371/journal.pone.0238491] [Medline: 32936804]
8. 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]
9. 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]
10. León B, Martínez-Costa MP, Salaverría R, López-Goñi I. Health and science-related disinformation on COVID-19: a content analysis of hoaxes identified by fact-checkers in Spain. PLoS One 2022;17(4):e0265995 [FREE Full text] [doi: 10.1371/journal.pone.0265995] [Medline: 35417493]
11. Jenkins MC, Moreno MA. Vaccination discussion among parents on social media: a content analysis of comments on parenting blogs. J Health Commun 2020 Mar 03;25(3):232-242. [doi: 10.1080/10810730.2020.1737761] [Medline: 32154770]
12. Hernandez RG, Hagen L, Walker K, O'Leary H, Lengacher C. The COVID-19 vaccine social media infodemic: healthcare providers' missed dose in addressing misinformation and vaccine hesitancy. Hum Vaccin Immunother 2021 Sep 02;17(9):2962-2964 [FREE Full text] [doi: 10.1080/21645515.2021.1912551] [Medline: 33890838]
13. Reuter K, Wilson ML, Moran M, Le N, Angyan P, Majmundar A, et al. General audience engagement with antismoking public health messages across multiple social media sites: comparative analysis. JMIR Public Health Surveill 2021 Feb 19;7(2):e24429. [doi: 10.2196/24429]
14. Kothari A, Foisey L, Donelle L, Bauer M. How do Canadian public health agencies respond to the COVID-19 emergency using social media: a protocol for a case study using content and sentiment analysis. BMJ Open 2021 Apr 22;11(4):e041818 [FREE Full text] [doi: 10.1136/bmjopen-2020-041818] [Medline: 33888527]
15. Slavik CE, Buttle C, Sturrock SL, Darlington JC, Yiannakoulias N. Examining tweet content and engagement of Canadian public health agencies and decision makers during COVID-19: mixed methods analysis. J Med Internet Res 2021 Mar 11;23(3):e24883 [FREE Full text] [doi: 10.2196/24883] [Medline: 33651705]
16. Sell TK, Hosangadi D, Trotochaud M. Misinformation and the US Ebola communication crisis: analyzing the veracity and content of social media messages related to a fear-inducing infectious disease outbreak. BMC Public Health 2020 May 07;20(1):550 [FREE Full text] [doi: 10.1186/s12889-020-08697-3] [Medline: 32375715]
17. Yang K, Pierri F, Hui P, Axelrod D, Torres-Lugo C, Bryden J, et al. The COVID-19 infodemic: Twitter versus Facebook. Big Data Soc 2021 May 05;8(1):205395172110138. [doi: 10.1177/20539517211013861]
18. Oz M, Zheng P, Chen G. Twitter versus Facebook: comparing incivility, impoliteness, and deliberative attributes. In: New Media & Society. Thousand Oaks, CA: Sage; Dec 31, 2017:3400-3419.
19. Halpern D, Valenzuela S, Katz J. We face, I tweet: how different social media influence political participation through collective and internal efficacy. J Comput-Mediat Commun 2017;22(6):336. [doi: 10.1111/jcc4.12198]
20. Chu K, Sidhu AK, Valente TW. Electronic cigarette marketing online: a multi-site, multi-product comparison. JMIR Public Health Surveill 2015;1(2):e11 [FREE Full text] [doi: 10.2196/publichealth.4777] [Medline: 27227129]
21. Eriksson M, Olsson E. Facebook and Twitter in crisis communication: a comparative study of crisis communication professionals and citizens. J Contingencies Crisis Manag 2016 Jun 13;24(4):198-208. [doi: 10.1111/1468-5973.12116]
22. Shane-Simpson C, Manago A, Gaggi N, Gillespie-Lynch K. Why do college students prefer Facebook, Twitter, or Instagram? Site affordances, tensions between privacy and self-expression, and implications for social capital. Comput Hum Behav 2018 Sep;86:276-288. [doi: 10.1016/j.chb.2018.04.041]
23. Nikolinakou A, King KW. Viral video ads: emotional triggers and social media virality. Psychol Mark 2018 Jul 24;35(10):715-726. [doi: 10.1002/mar.21129]
24. Spiliotopoulos T, Oakley I. An exploration of motives and behavior across Facebook and Twitter. J Syst Inf Technol 2020 Jul 13;22(2):201-222. [doi: 10.1108/jsit-12-2019-0258]
25. Moreno MA, D'Angelo J. Social media intervention design: applying an affordances framework. J Med Internet Res 2019 Mar 26;21(3):e11014 [FREE Full text] [doi: 10.2196/11014] [Medline: 30912754]
26. Liu B, Kim S. How organizations framed the 2009 H1N1 pandemic via social and traditional media: implications for U.S. health communicators. Public Relat Rev 2011 Sep;37(3):233-244. [doi: 10.1016/j.pubrev.2011.03.005]
27. Lwin M, Lu J, Sheldenkar A, Schulz P. Strategic uses of Facebook in zika outbreak communication: implications for the crisis and emergency risk communication model. Int J Environ Res Public Health 2018 Sep 10;15(9):1974 [FREE Full text] [doi: 10.3390/ijerph15091974] [Medline: 30201929]

XSL•FO
RenderX

28. Wang Y, Hao H, Platt LS. Examining risk and crisis communications of government agencies and stakeholders during early-stages of COVID-19 on Twitter. Comput Human Behav 2021 Jan;114:106568 [FREE Full text] [doi: 10.1016/j.chb.2020.106568] [Medline: 32982038]
29. Vos SC, Sutton J, Yu Y, Renshaw SL, Olson MK, Gibson CB, et al. Retweeting risk communication: the role of threat and efficacy. Risk Anal 2018 Dec 06;38(12):2580-2598. [doi: 10.1111/risa.13140] [Medline: 30080933]
30. Wahbeh A, Nasralah T, Al-Ramahi M, El-Gayar O. Mining physicians' opinions on social media to obtain insights into COVID-19: mixed methods analysis. JMIR Public Health Surveill 2020 Jun 18;6(2):e19276 [FREE Full text] [doi: 10.2196/19276] [Medline: 32421686]
31. DePaula N, Hagen L, Roytman S, Dyson D, Alnahass D, Patel M, et al. A framework of social media messages for crisis and risk communication: a study of the covid-19 pandemic. 2022 Presented at: 55th Hawaii International Conference on System Sciences (HICSS); 2022; Virtual. [doi: 10.24251/hicss.2022.300]
32. Leppert K, Saliterer I, Korać S. The role of emotions for citizen engagement via social media – a study of police departments using Twitter. Gov Inf Q 2022 Jul;39(3):101686. [doi: 10.1016/j.giq.2022.101686]
33. DePaula N, Dincelli E. Information strategies and affective reactions: how citizens interact with government social media content. FM 2018 Apr 01;23(4). [doi: 10.5210/fm.v23i4.8414]
34. Pew Research Center. Social Media Fact Sheet. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2022-11-10]
35. 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]
36. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? 2010 Presented at: 19th International Conference on World Wide Web (WWW 2010); April 26-30, 2010; Raleigh, NC. [doi: 10.1145/1772690.1772751]
37. Shapiro M, Hemphill L. Politicians and the policy agenda: does use of twitter by the U.S. Congress direct New York Times content? Policy Internet 2017;9(1):109-132. [doi: 10.1002/poi3.120]
38. Lin H, Qiu L. Two sites, two voices: linguistic differences between Facebook status updates and tweets. In: Rau PLP, editor. Cross-Cultural Design. Cultural Differences in Everyday Life. CCD 2013. Lecture Notes in Computer Science, vol 8024. Berlin, Heidelberg: Springer; 2013:432-440.
39. Voorveld HAM, van Noort G, Muntinga DG, Bronner F. Engagement with social media and social media advertising: the differentiating role of platform type. J Advert 2018 Feb 13;47(1):38-54. [doi: 10.1080/00913367.2017.1405754]
40. Association for Professionals in Infection Control and Epidemiology. US Government Health Agencies: With Functions Related to Infection Prevention and Control. 2008. URL: https://apic.org/Resource_/TinyMceFileManager/Advocacy-PDFs/outline_of_govt_health_agencies.pdf [accessed 2022-11-10]
41. Miles C. About Us: Learn More about CrowdTangle. URL: http://help.crowdtangle.com/en/articles/4201940-about-us [accessed 2022-11-10]
42. Fairclough N. Analysing Discourse: Textual Analysis for Social Research. London, UK: Routledge; 2003.
43. van DT. Discourse and knowledge: A Sociocognitive Approach. Cambridge, UK: Cambridge University Press; 2014.
44. Vos SC, Cohen E. Using pictures in health and risk messages Internet. Oxf Res Encycl Commun 2017 Aug 22:Online. [doi: 10.1093/acrefore/9780190228613.013.314]
45. Vos SC, Sutton J, Gibson CB, Butts CT. #Ebola: emergency risk messages on social media. Health Secur 2020 Dec 01;18(6):461-472. [doi: 10.1089/hs.2019.0158] [Medline: 33326333]
46. Howell DC. Statistical Methods for Psychology. Boston, MA: Wadsworth Cengage Learning; 1994.
47. Renshaw S, Mai S, Dubois E, Sutton J, Butts C. Cutting through the noise: predictors of successful online message retransmission in the first 8 months of the COVID-19 pandemic. Health Secur 2021;19(1):31-43 [FREE Full text] [doi: 10.1089/hs.2020.0200] [Medline: 33606574]
48. DePaula N, Fietkiewicz KJ, Kaja J, Froehlich TJ, Million AJ, Dorsch I, et al. Challenges for Social Media: Misinformation, Free Speech, Civic Engagement, and Data Regulations. Proceedings of the Association for Information Science and Technology 2018;55(1):665-668. [doi: 10.1002/pra2.2018.14505501076]
49. Fagerland MW. t-tests, non-parametric tests, and large studies--a paradox of statistical practice? BMC Med Res Methodol 2012 Jun 14;12(1):78 [FREE Full text] [doi: 10.1186/1471-2288-12-78] [Medline: 22697476]
50. 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. [doi: 10.2196/23272]
51. Papakyriakopoulos O, Serrano JCM, Hegelich S. Political communication on social media: a tale of hyperactive users and bias in recommender systems. Online Soc Netw Media 2020 Jan;15:100058. [doi: 10.1016/j.osnem.2019.100058]
52. Keib K, Espina C, Lee Y, Wojdynski BW, Choi D, Bang H. Picture this: the influence of emotionally valenced images, on attention, selection, and sharing of social media news. Media Psychol 2017 Oct 05;21(2):202-221. [doi: 10.1080/15213269.2017.1378108]
53. Rathje S, Van Bavel JJ, van der Linden S. Out-group animosity drives engagement on social media. Proc Natl Acad Sci U S A 2021 Jun 29;118(26):e2024292118 [FREE Full text] [doi: 10.1073/pnas.2024292118] [Medline: 34162706]

XSL•FO
RenderX

54. Ali K, Zain-ul-abdin K, Li C, Johns L, Ali AA, Carcioppolo N. Viruses going viral: impact of fear-arousing sensationalist social media messages on user engagement. Sci Commun 2019 May 03;41(3):314-338. [doi: 10.1177/1075547019846124]
55. DePaula N. Supporting the Cause: An Analysis of How Government Agencies Use Twitter Hashtags. Proceedings of the Association for Information Science and Technology 2018;55(5):788-789 [FREE Full text] [doi: 10.1002/pra2.2018.14505501117]
56. DePaula N. Political Ideology and Information Technology in Government Online Communication.. Government Information Quarterly 2022 Aug 20:e101747. [doi: 10.1016/j.giq.2022.101747]
57. Sillence E, Martin R. Talking about decisions: the facilitating effect of a celebrity health announcement on the communication of online personal experiences around decision-making. Health Commun 2020 Nov;35(12):1447-1454. [doi: 10.1080/10410236.2019.1652064] [Medline: 31411044]

Abbreviations

API: application programming interface
CTR: click-through rate
IRR: interrater reliability
RQ: research question
WMW: Wilcoxon-Mann-Whitney

Edited by T Purnat; submitted 09.06.22; peer-reviewed by C Baur, JB Del Rosario; comments to author 28.07.22; revised version received 27.08.22; accepted 08.09.22; published 20.12.22.

Please cite as:
DePaula N, Hagen L, Roytman S, Alnahass D
Platform Effects on Public Health Communication: A Comparative and National Study of Message Design and Audience Engagement Across Twitter and Facebook
JMIR Infodemiology 2022;2(2):e40198
URL: https://infodemiology.jmir.org/2022/2/e40198
doi:10.2196/40198
PMID:36575712

©Nic DePaula, Loni Hagen, Stiven Roytman, Dana Alnahass. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 20.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Investigation of COVID-19 Misinformation in Arabic on Twitter: Content Analysis

Ahmed Al-Rawi[1], PhD; Abdelrahman Fakida[1], MA; Kelly Grounds[1], BA

School of Communication, Simon Fraser University, Burnaby, BC, Canada

Corresponding Author:
Ahmed Al-Rawi, PhD
School of Communication
Simon Fraser University
Schrum Science Centre-K 9653
Burnaby, BC, V5A1S6
Canada
Phone: 1 7787824419
Email: aalrawi@sfu.ca

Abstract

Background: The COVID-19 pandemic has been occurring concurrently with an infodemic of misinformation about the virus. Spreading primarily on social media, there has been a significant academic effort to understand the English side of this infodemic. However, much less attention has been paid to the Arabic side.

Objective: There is an urgent need to examine the scale of Arabic COVID-19 disinformation. This study empirically examines how Arabic speakers use specific hashtags on Twitter to express antivaccine and antipandemic views to uncover trends in their social media usage. By exploring this topic, we aim to fill a gap in the literature that can help understand conspiracies in Arabic around COVID-19.

Methods: This study used content analysis to understand how 13 popular Arabic hashtags were used in antivaccine communities. We used Twitter Academic API v2 to search for the hashtags from the beginning of August 1, 2006, until October 10, 2021. After downloading a large data set from Twitter, we identified major categories or topics in the sample data set using emergent coding. Emergent coding was chosen because of its ability to inductively identify the themes that repeatedly emerged from the data set. Then, after revising the coding scheme, we coded the rest of the tweets and examined the results. In the second attempt and with a modified codebook, an acceptable intercoder agreement was reached (Krippendorff $\alpha \geq .774$).

Results: In total, we found 476,048 tweets, mostly posted in 2021. First, the topic of infringing on civil liberties (n=483, 41.1%) covers ways that governments have allegedly infringed on civil liberties during the pandemic and unfair restrictions that have been imposed on unvaccinated individuals. Users here focus on topics concerning their civil liberties and freedoms, claiming that governments violated such rights following the pandemic. Notably, users denounce government efforts to force them to take any of the COVID-19 vaccines for different reasons. This was followed by vaccine-related conspiracies (n=476, 40.5%), including a Deep State dictating pandemic policies, mistrusting vaccine efficacy, and discussing unproven treatments. Although users tweeted about a range of different conspiracy theories, mistrusting the vaccine's efficacy, false or exaggerated claims about vaccine risks and vaccine-related diseases, and governments and pharmaceutical companies profiting from vaccines and intentionally risking the general public health appeared the most. Finally, calls for action (n=149, 12.6%) encourage individuals to participate in civil demonstrations. These calls range from protesting to encouraging other users to take action about the vaccine mandate. For each of these categories, we also attempted to trace the logic behind the different categories by exploring different types of conspiracy theories for each category.

Conclusions: Based on our findings, we were able to identify 3 prominent topics that were prevalent amongst Arabic speakers on Twitter. These categories focused on violations of civil liberties by governments, conspiracy theories about the vaccines, and calls for action. Our findings also highlight the need for more research to better understand the impact of COVID-19 disinformation on the Arab world.

(JMIR Infodemiology 2022;2(2):e37007) doi:10.2196/37007

XSL•FO
RenderX

KEYWORDS

COVID-19; Arab world; Twitter; misinformation; vaccination; infodemiology; vaccine hesitancy; infoveillance; health information; social media; social media content; content analysis; Twitter analysis

Introduction

Background

The COVID-19 pandemic has occurred in tandem with what is being called an *infodemic*, referring to the large amounts of false and misleading content about the virus being disseminated primarily over social media platforms [1]. Although English COVID-19 disinformation and misinformation research and data sets have received extensive attention [2], content shared in other languages, specifically Arabic, has been neglected [3,4].

Misinformation is not a new phenomenon in the Arab region [5]. Some scholars have focused on the reach and effects of Arabic misinformation during the pandemic [2], while others have included Arabic misinformation as a secondary focus [6]. Nevertheless, there is an urgent need for an examination of the scale of Arabic COVID-19 disinformation. Arabic, for example, is currently Facebook's third-most common language [7] and a language spoken by more than 400 million people [8] in countries with significant social media presence [9]. Hence, this study attempts to fill a gap in the literature by examining social media data retrieved from Twitter to better understand the main online discussions that revolve around COVID-19 misinformation.

There seems to be an abundance of misinformation in non-English discourses that social media companies largely ignore, for they do not often react to viral Arabic COVID-19 misinformation with the same urgency they have for English misinformation. Facebook took some measures to address English COVID-19 misinformation, but the platform applied limited measures to other languages [10]. For example, Facebook fact-checks viral information in Western countries through fact-checking partners, yet such debunking is not applied to the same information when translated to Arabic [10]. Various spoken Arabic dialects pose a challenge to Facebook's algorithms and human moderators, given each dialect's unique vocabulary and historical and cultural contexts [7]. Transcribing Arabic varies between native speakers from different countries, and as Facebook continues to rely on artificial intelligence to moderate content, the platform will continue to misinterpret Arabic posts [7].

There is a pressing need to thoroughly understand and examine the social media platforms where Arabic COVID-19 disinformation spreads. Current studies focus on Facebook pages and groups given their popularity, especially in some Arab countries [11,12]. However, Facebook is not the only platform where Arabic COVID-19 misinformation is spread [8]. Content on Facebook has found its way to Twitter accounts and YouTube channels [10], where other scholars have then analyzed Arabic COVID-19 tweets and hashtags [2,4,6]. Reviewing the use of religious misinformation during the pandemic, it was found that misinformation in YouTube videos can receive millions of views [13]. These same videos were found circulating on messaging platforms, such as WhatsApp. As 1 of the world's most popular messaging platforms and a significant source of COVID-19 misinformation [14], WhatsApp is suggested to be a vital source of rumors, with usage surpassing Twitter in Arab countries, such as Saudi Arabia [9]. Taken together, these studies provide an overview of the online ecosystem that allows Arabic COVID-19 disinformation to spread. Our study attempts to contribute to the existing knowledge by answering the following research question: what are the major topics that are discussed around COVID-19 misinformation in Arabic on Twitter?

Literature Review

Although various disinformation and misinformation narratives in Arabic spread on social media range from "full-throated conspiracy theories to unscientific health advice" [9], our review of the literature found that there is a focus on 3 categories of disinformation. These 3 categories are COVID-19 disinformation, government-targeted disinformation, and religious disinformation. COVID-19 disinformation tends to focus on the symptoms of the virus, its vaccine, alternative remedies, and theories about its origins. Government-targeted disinformation looks at content that addresses specific governments and the health measures that they have taken. Finally, religious disinformation looks at misleading advice given by religious leaders and misinterpretations of religious texts during the pandemic and its potential to cause fear and confusion. In addition to these categories, we also found a smaller body of literature that discussed the origins of the Arabic disinformation, focusing on the regions that it emerges from and the governments that sponsor it. Each of these categories will be discussed later.

COVID-19 Disinformation

Disinformation about the COVID-19 virus and its vaccination has received ample attention from social media users. The disinformation includes false information about the virus and its treatment, the promotion of inaccurate or false claims about home remedies as alternatives to existing treatments, and conspiracy theories about the virus's origins [12]. This disinformation has also been noticeably popular in Iran, with users engaging with it consistently. One popular false remedy that appeared to Iranian users touted alcohol as a cure for COVID-19. This false remedy resulted in at least 27 people dying of alcohol poisoning [15]. This example shows the potential impact of misinformation that caused real-life harm.

In Arab countries, misinformation about topics such as the virus's origins or government-related conspiracies about the vaccine have begun to widely appear on social media as legitimate sources of information. A paper exploring COVID-19 misinformation in Jordan found that different Arabic media outlets amplified conspiracy theories about the pandemic and the most common conspiracy theories amplified tended to focus on the virus's origins [16]. The majority of these theories stated that the virus was either created in a lab as part of a biologic warfare campaign or was caused by the 5G network [17]. In the

XSL•FO
RenderX

paper's sequel, the authors found that these same narratives had persisted, with the theories that the virus came out of a lab remaining the most prominent [17]. The study also showed a growing connection between conspiracy theories and disinformation about the treatment and vaccine. Most of these conspiracy theories claim that the long-term effects of both treatments are unknown and could potentially be dangerous [16]. Like social media content that focuses on the virus and its treatments, conspiracy theory content also poses a significant risk to public health efforts by continuing to propagate antivaccine theories that could lower the overall rates.

Government-Targeted Disinformation

Government-targeted disinformation includes speculation that they are co-opting the pandemic to impose new restrictions on individuals or a manipulation of official government statements to fit a narrative. It is important to note that government-targeted disinformation is different from posts that discuss the Deep State and other conspiracy theories, such as the New World Order. Posts about the Deep State and other groups are based on unfounded claims and about their role in the pandemic and appear in the previous category of COVID-19 disinformation [18]. In contrast, government-targeted disinformation uses actual statements and actions taken by different governments and twists them to shape a false narrative.

There are 2 subcategories of government-targeted disinformation. The first targets the pandemic responses and policies of Arab governments, while the second focuses on non-Arabic governments and falls under the conspiracy theory category. Government-focused disinformation tries to co-opt information and statements to use against the rulers. Most of the content takes the statements of government officials out of context to put their pandemic responses in a negative light [12]. Based on additional analysis, this type of disinformation was the most prevalent type on Egyptian social media, while also having the lowest engagement of all 4 content categories [12]. The observed lower levels of engagement are likely because this category of disinformation rarely attached videos or sources to its posts, meaning that users were less likely to engage with it [12].

The disinformation that utilizes manipulated pandemic-related content has been shown to have some of the highest levels of engagement. Another paper argues that these higher levels of engagement are a result of these posts imitating and co-opting sections of actual news articles to increase their perceived legitimacy [15]. This is echoed by other scholars who found that on Egyptian social media, in particular Facebook, manipulated content was 1 of the most prominent types of disinformation to be shared. Within this category of disinformation, posts that reconfigured existing news articles to agree with a false narrative were the most popular [12]. However, the manipulated content was not limited to existing news articles. One of the more prominent examples shared with Egyptian Facebook users was a clip from a documentary manipulated to show that an asteroid would hit the earth once the COVID-19 pandemic ended [12]. Of all 3 categories, this content has the most significant potential to create panic by using posts that appear legitimate, creating a sense of trust with the user.

Religious Arabic Disinformation

The final category, religious disinformation, tends to take the form of misleading advice, and misinterpretations of religious scriptures are vital to consider, given Islam's role in the region culturally, socially, and politically [13]. Analyzing the growth of religious misinformation during the COVID-19 pandemic in the Arab region, 1 paper highlights the ability of actors in the region to promote their misinformation, endanger public health, and cause confusion and fear [13]. This type of misinformation includes top-down religious misinformation coming from authority figures. In Iran, clerics have been spreading top-down misinformation, while there has also been bottom-up misinformation disseminated by content creators aiming to attract followers and subscribers [13]. These actors promoted fake remedies through religious misinformation and took advantage of the pandemic-induced uncertainty [13]. However, regardless of the original source of disinformation, it still poses a significant risk to public health.

Sources of Arabic Disinformation

The majority of the literature on the topic of COVID-19 conspiracies focuses on the different types of disinformation with which users interact. However, to date, there has been less of an emphasis on the origins of the content. The literature that does explore this tends to focus on content that emerges from state-sponsored campaigns with political motives [11]. The purpose of these campaigns is to shift the focus off a state's poor response at the domestic and international levels. Across the Arab world, such campaigns include religious elements and attempts to weaponize information to blame rivals [3]. Analyzing Arabic COVID-19 Facebook posts, 1 study noted that the content generated in these campaigns targets regional governments using digital marketing firms that leverage "COVID-19 to push geopolitically aligned narratives" [11]. The literature also suggests that there have been coordinated Arabic COVID-19 disinformation operations by Iran and Saudi Arabia and, to a lesser extent, narratives from Egypt and the United Arab Emirates [10].

Looking at Iran, there has been an effort to design and sponsor COVID-19 disinformation campaigns to deflect from their own government's pandemic response, while blaming 1 of their primary adversaries. One paper found that Tehran used 2 narratives to try and undermine the reputation of the United States [19]. The first narrative accused the United States of using sanctions to undermine Tehran's public health response to the virus. The second also targeted the United States and falsely accused them of creating the COVID-19 virus as part of their plans for biological warfare [19]. This was echoed in another study, having found similar narratives blaming the United States for creating the virus and accusing them of developing the virus to attack Shiites and Iran [3]. Both studies also found that Tehran used state-controlled media as its primary method of spreading disinformation, with a lesser degree of social media utilization [3,19].

XSL•FO
RenderX

The study also explored the Saudi COVID-19 disinformation campaigns and found that they also used narratives to attack and undermine their adversaries [3]. It is important to note that prior to the pandemic, different governments in the Arab Gulf had used fake news campaigns to attack opponents following a Saudi Arabia–led blockade of Qatar [20]. Since then, fake news battles in the Arab Gulf have taken a more public angle than other countries in the region [21]. Unlike Iran, the Saudis focused on Qatar but did not create their own disinformation. Instead, the Saudi regime took disinformation created by individuals and amplified it across multiple platforms, with the most appearing on Twitter [3]. Among the amplified narratives, there are 2 that appear to be spreading on a much larger scale. The first narrative stated that Qatar had known about COVID-19 since 2015, and the second was that Qatar was deliberately spreading the virus to damage Saudi's plans to diversify its economy [3]. The Saudis also relied heavily on social media platforms to spread their narratives, unlike Iran, which used its state-controlled media as its primary method. Despite the differences in dissemination, both regimes used their disinformation campaigns to smear their adversaries with narratives accusing them of having inside knowledge of the virus and using that to target and cripple the regimes.

The research conducted in a separate study addresses a less discussed source of Arabic COVID-19 misinformation [10]. The author identified 18 Facebook pages and 10 Facebook groups with a collective following of more than 2.4 million users that share COVID-19 content in Arabic. It found that "Arabic-language conspiracy hubs are masquerading as independent institutions, think tanks and research initiatives and are manipulating COVID-19 data, conducting their own research." [10]. These sources, primarily located in Egypt, spread pandemic conspiracy narratives focusing on the apocalypse and antisemitism. The author of the study points out how the analyzed pages boost COVID-19 misinformation coming from Western countries by translating content and adding Arabic subtitles or voiceovers [10]. Another study also analyzed tweets using the viral conspiracy theory hashtag #FilmYourHospital, which encouraged people to take pictures of empty hospitals to show that the COVID-19 pandemic is a scam [6]. Using social network analysis techniques, the scholars found that "the second and third largest non-English clusters were users tweeting in Arabic" [6]. Taken together, these 2 studies highlight the high rates of Arabic disinformation, in addition to the already prevalent English content.

The consequences of the Arabic disinformation range from diminished trust in governments to potential decreases in vaccination rates and dangers to public health. Although the existing literature is limited, it does give insight into the types of content with which users might interact. Disinformation about the virus and its treatments appears to be the most dangerous of all the existing categories. As 1 author highlights, increasing vaccination hesitancy rates appear in countries engaging with such disinformation [16]. Within the literature discussing the origins of the content itself, Iran and Saudi Arabia have shown a clear preference for being involved in information wars through content that spreads conspiracy theories and targets regional governments [16]. However, a different paper argues that private actors are also motivated by political or financial reasons to spread fake news, such as state actors [21]. Moreover, religious leaders and content creators also took advantage of the panic the pandemic brought and utilized the religion of Islam for their benefit by spreading religious misinformation to gain new subscribers [13]. However, content created by Arabic conspiracy hubs needs more examination as there is a significant gap "in understanding the trends and the intersections with other conspiracy communities online" [10]. These trends and the understanding of content categories reflect the experiences of a limited number of countries and social media platforms.

Methods

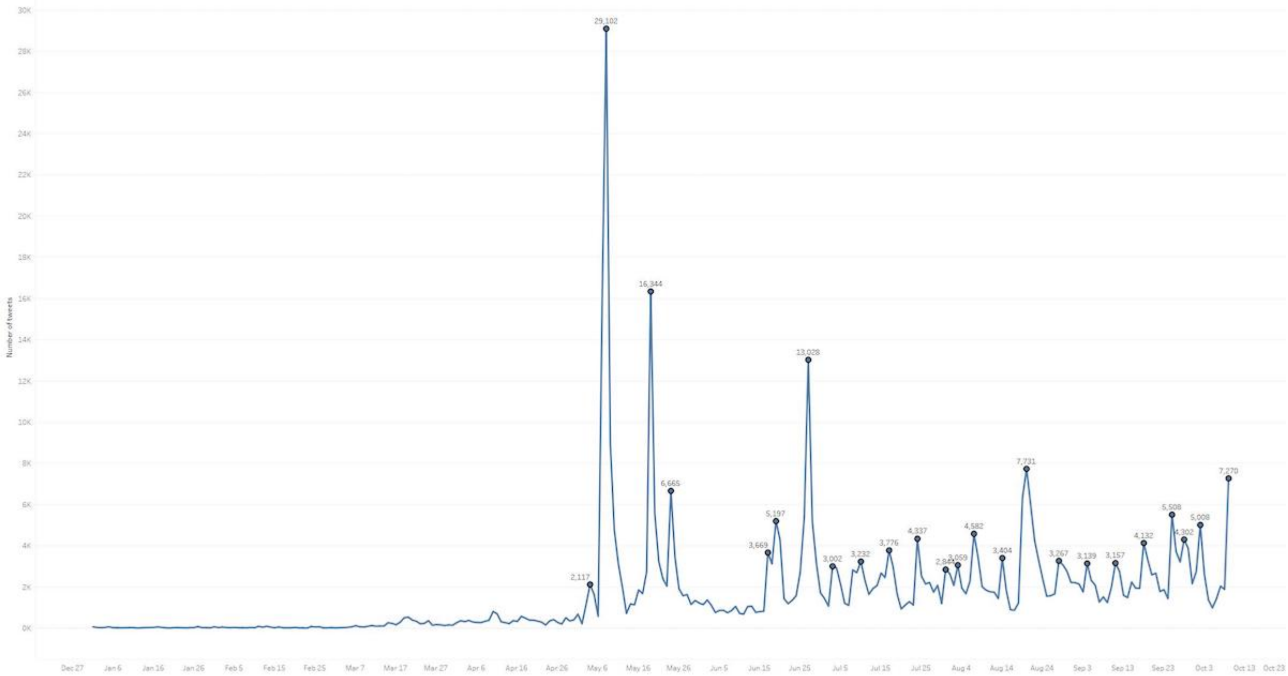
Study Design

In this study, we used content analysis to investigate the most retweeted messages that reference the following 13 hashtags that are mostly used by Arabic-speaking antivaccine communities: #No_to_forced_vaccinations (# _ _), #Medical_freedom (# _), #No_for_vaccinating_children (# _ _), #Say_no_to_the_vaccine (# _), #No_to_vaccines (# _), #Notovaccines (#), #Say_no_to_injections (# _), #No_to_injections (# _), #Notoinjections (#), #Affected_by_the_injections (# _), #Injections_complications (# _), #No_to_the_new_world_order (# _ _ _), and #No_to_human_genetic_mutilation (# _ _ _). To identify these hashtags, we started our search a few days before October 10, 2021, by using 4 generic hashtags: #No_to_forced_vaccinations (# _ _), #Say_no_to_the_vaccine (# _), #No_to_vaccines (# _), and #Notovaccines (#). After collecting all the tweets, we used a Python script to extract the most recurrent hashtags in the data set that helped us identify the popular 13 hashtags often used by this online Arab community.

Using Twitter academic API v2 allowed us to search these hashtags from the beginning on August 1, 2006, until October 10, 2021, which is when the final data set was collected. In total, we found 476,048 tweets, mostly posted in 2021 (see Figure 1, Multimedia Appendix 1). We started from the beginning in 2006 to make sure that similar hashtags were not popular before the pandemic emerged, since this issue could distort the findings of this study. Since the data set is large, we manually analyzed the top 1000 most retweeted messages. We excluded 96 (9.6%) vague and unclear messages from the data set. By vague messages, we refer to unclear and incomprehensible tweets. For example, 1 user said in a tweet a few words like "A short story!" Another tweeted saying only the word "Kuwait." Similar tweets that could not be analyzed were replaced with other tweets from the rest of the data set in order to have a total of 1000 retweeted messages. Our goal was to exclusively collect Arabic hashtags, but there were tweets written partly or fully in other languages, including English.

XSL•FO
RenderX

Figure 1. Frequency of tweets referencing 13 hashtags in 2021.



To identify major categories or topics in the sample data set, 2 coders relied on emergent coding [22] and a codebook was designed inductively that covered the range of issues found in the data set. Emergent coding was selected because of its ability to identify themes that appear repeatedly in a subset of the larger data set. These themes could then be developed into a codebook, which was used to code the entire sample in the data set. The first step of the process followed a qualitative review of 100 tweets (10% of the overall data set) by 2 coders working independently to highlight themes emerging from the data. For example, the coders noticed several tweets focusing on different vaccine conspiracy theories, which led to highlighting such themes and creating subthemes that discuss each conspiracy. Moreover, the coders noticed that a few tweets had specific calls to action that encouraged others to participate in several different activities, such as rallies or protests. We note that some of the themes that emerged were similar to the ones found in previous research that we referenced in our literature review. Based on the themes found, an initial codebook was created. To check the validity of the first designed codebook, the 2 coders individually tested it by examining the same sample of 100 tweets (10% of the overall data set). In the first attempt, intercoder agreement was low, so the 2 coders made several changes to the codebook after some discussion and deliberation. The first main change in the codebook was to reduce topics from 5 to 4, as we realized that 1 category focused on promotional and marketing messages belonged to the third topic. Following this, we renamed the second topic to include all types of calls to action and news related to lockdowns and protests against vaccinations. Finally, the descriptions of vaccine conspiracy theories were slightly revised to better represent what emerged from the Twitter data (Table 1). In the second attempt and with a modified codebook, an acceptable intercoder agreement was reached (Krippendorff α≥.774) [23].

Table 1. The codebook used in analyzing the Twitter data set.

Topic	Description
Infringing on civil liberties or perceived online censorship	• Policies and government-imposed COVID-19 restrictions (eg, quarantine and travel restrictions) and online restrictions
Call for action or news on antilockdown and antivaccine protests	• Joining or organizing a protest or rally or news about protests
Other	• Minor issues or unrelated promotional messages
Vaccine conspiracy theories	• The Deep State or the New World Order planning the pandemic for political gains • Governments or pharmaceutical companies profiting from vaccines and intentionally risking general public health • Attacking heath authorities or official medical news • Mistrusting the efficacy of the vaccine or making false or exaggerated claims about vaccine risks and vaccine-related diseases • Believing that unproven treatments (eg, eating healthy and exercising well) can prevent the virus

XSL•FO
RenderX

Ethical Considerations

Ethics approval was not required for this study because it used publicly available data on social media. SFU's Ethics Board does not require ethics clearance for studies using data from public domains like Twitter.

Results

Topic Details

This study examined Arabic COVID-19 and vaccination disinformation discourse on Twitter. Through an analysis of antivaccination hashtags, we examined the main topics users discuss online to cover a gap in the literature regarding pandemic-related content in Arabic. To answer the study's research question, the data analysis showed that infringement on civil liberties is the topic most engaged with, covering 41.1% (n=483) of the data and gathering 66,835 (41.1%) retweets. The second-most engaging topic is vaccine conspiracy theories, with 40.5% (n=476) of the data analyzed and a total of 62,336 (38.4%) retweets. The scope of vaccine conspiracy theories varied as mistrusting vaccine efficacy (263/476, 55.3%) attracted the most attention, followed by the government's intentional chaos (110/476, 23.1%). Minor conversations focused on attacking health authorities (49/476, 10.3%), the Deep State (40/476, 8.4%), and unproven treatments (14/476, 2.9%). As for the third topic users discuss, tweets pushing a call for action cover 12.7% (n=149) of the tweets that were analyzed, garnering 19,733 (12.1%) retweets, while other unrelated content is the least present topic, with only 5.7% (n=67) of the tweets analyzed and a total of 13,531 (8.3%) retweets (see Table 2). To further understand the results, we delved into each topic to present examples explaining arguments against COVID-19 vaccines found in the data.

Table 2. Frequencies and percentages of the main topics discussed on Twitter.

Topic	Frequency, n/N (%)	Retweets, n/N (%)
Infringing on civil liberties	483/1175 (41.1)	66,835/162,435 (41.1)
Call for action	149/1175 (12.7)	19,733/162,435 (12.1)
Other	67/1175 (5.7)	13,531/162,435 (8.3)
Vaccine conspiracies (n=476, 40.5%; 62,336, 38.4%, retweets)		
Deep State	40/476 (8.4)	4876/62,336 (7.8)
Government's intentional chaos	110/476 (23.1)	13,874/62,336 (22.3)
Attacking health authorities	49/476 (10.3)	6248/62,336 (10.0)
Mistrusting vaccine efficacy	263/476 (55.3)	35,505/62,336 (57.0)
Unproven treatments	14/476 (2.9)	1833/62,336 (2.9)

Infringing on Civil Liberties

This is the most prevalent topic found in the data set: 483 (41.1%) of the tweets analyzed that had 66,835 (41.1%) retweets. Users here focus on topics concerning their civil liberties and freedoms, claiming that governments violated such rights following the pandemic. Particularly, users here denounce government efforts to force them to take any of the COVID-19 vaccines for different reasons. Some users claim their criticism of vaccine mandates stems from a suspicion of the vaccines' safety and effectiveness, while others discuss their fundamental right as individuals to refuse vaccination regardless of whether vaccines are safe. For example, 1 user posted the following message, which was retweeted 680 times:

> *The vaccine prevents contagion and infection. What's on the market currently does not do either. So don't call it a vaccine, but call it "gene therapy," which is closer to reality. Why force people to an experimental and unapproved gene therapy? Its long-term damage is unknown.*

Another user, who self-identifies as a lawyer, said in a tweet retweeted 403 times:

> *For the millionth time and over and over again I personally don't care who took the vaccine and who refuses to take it, my role is only limited to providing legal awareness for those who refuse to take the compulsory vaccine. And that no person or entity has the right to impose the vaccine on a person who doesn't want it, according to the constitution, agreements and laws indicating that.*

We can see here that both tweets refuse a mandatory vaccine but for different reasons. The first user is suspicious of its components, which intersects with our analysis of the fourth topic, which highlights vaccine conspiracy theories. The second user's refusal comes from a legal principle rather than conspiracy theories or disinformation.

In a similar vein, users in this category argue against government restrictions that differentiate between the vaccinated and the unvaccinated. Examples of such rules include limiting travel and access to public spaces to vaccinated individuals only while forcing a workplace vaccine mandate. One of the most popular posts that was retweeted 576 times mentioned the following:

> *When a distinction is made between the vaccinated and the unvaccinated in employment as well as travel and shopping, it is considered indirect coercion and a type of humanitarian crime.*

Here the user goes further in their condemnation by denouncing restrictions that they believe deny the unvaccinated their basic human and civil rights.

XSL•FO
RenderX

Looking at the origins of conspiracies that talk about civil liberties, there are 2 dominant narratives: In relation to the first narrative, the underlying theory is that each subsequent variant of COVID-19 was released by the government to prolong the pandemic and expand its powers [24]. By doing this, governments can allegedly restrict the unvaccinated and isolate them from the general population in a similar method to the concentration camps used in Nazi Germany.

The second narrative, focusing on the Great Reset, is similar to the first one in that governments are allegedly using the pandemic to maintain their powers. However, the Great Reset conspiracy takes this a step further and claims that the COVID-19 pandemic was started by a secret global government to cause a global economic collapse [25]. This will be achieved through the continual lockdowns, which will then allow the secret global government to implement a socialist world government for its own benefit. The Great Reset became a mainstream theory once a proposal from the 2020 Davos summit about a postpandemic reset went viral and became misconstrued as something sinister [25]. When taken together, these narratives have the potential to create doubt in the actions of their governments and to protect their own rights, as was seen in several examples from this study.

Vaccine Conspiracy Theories

The second-most recurring topic deals with tweets spreading vaccine conspiracy theories. This topic covers 40.5% (n=476) of the tweets analyzed, garnering 62,336 (38.4%) retweets. We noted the presence of different theories, from government and pharmaceutical actions and policies to intentionally create chaos and the existence of a network of actors exploiting the pandemic and vaccines to gain power and money, as well as attacks on health authorities, distrust in vaccines, and theories about alternatives to vaccinations. Although users tweeted about a range of different conspiracy theories, mistrusting the vaccine's efficacy, false or exaggerated claims about vaccine risks and vaccine-related diseases, and governments and pharmaceutical companies profiting from vaccines and intentionally risking the general public health, appeared the most.

To further illustrate this finding, 1 of the most retweeted tweets with 607 retweets states:

> *By virtue of my studies as a medical assistant at Kuwait University and because of my specialization in the Department of Medical Laboratories and my work experience in the laboratory of viral diseases in Mubarak Hospital, I read the websites of vaccines manufacturers and discussed with a group of doctors and made my decision not to get vaccinated.*

Although some social media posts express a general hesitancy in vaccines, tweets such as the previous one are more dangerous as they claim to be coming from an expert in the medical or pharmaceutical field. Other users went a step further, claiming that vaccines cause death to anyone getting them. In 1 example, retweeted 471 times, the user argues that a citizen died following a vaccine shot, tagging and attacking Kuwait's former health minister Basel Al Sabah. The tweet says:

> *Dr. Basil @drbaselalsabah. A citizen took the vaccine and suffered a stroke and today she died. Do you still want to force people to be vaccinated! The tragedy is that there was a doctor mocking her condition and considered it a figment of her imagination.*

In fact, COVID-19 vaccines can cause serious side effects in some people, but the reasoning is wrong because of overgeneralization.

Some tweets focus on attacking the Kuwaiti and Arab governments, criticizing how they handled the pandemic and accusing them of intentionally creating chaos. This tweet, which was retweeted 201 times, says:

> *A new heresy invented and approved by the government right away which is to prevent unvaccinated citizens from travelling. The confused government became an expert in creating crises and adding restrictions to citizens. I suggest they change their advisors because they sent the government into a vortex.*

Again, the reasoning suggests that the government is intentionally trying to create chaos and social disruption.

When tracing the origins of vaccine-related conspiracy theories, the Center for Reality and Historical Studies, which is an online content hub that publishes disinformation about COVID-19, has been a reoccurring source of conspiracies. In 2021, the group uploaded a 27-minute video to Facebook, titled *Ask the Experts (COVID-19)*, which featured testimonials from 30 individuals who were credited as being doctors, health experts, and journalists. The key claims made in this video are that there was no pandemic so there could be no vaccine and that the industry skipped animal trials while producing the vaccine, so the harms are unknown; in addition, taking the vaccine will allegedly change your DNA [10]. The video subsequently went viral on Facebook, specifically among Arabic speakers.

A second prominent conspiracy theory about the vaccine is that pharmaceutical companies created COVID-19 to make the global populations ill and, in turn, increase their profits [18]. In the same vein, there are multiple conspiracy theories that believe that Bill Gates or the Gates Foundation created COVID-19 as a way to mass-vaccinate populations with a microchip. This in turn will allow Gates and the world government to track everyone at once [18]. For these conspiracy theories, Gates was singled out largely because of 2015 Ted Talk where he discussed the Ebola outbreak as a precursor of future pandemics, which individuals consider to be foreshadowing for the COVID-19 pandemic. These 2 examples, when taken along with statements from prominent individuals, such as the president of Lebanon and the leader of the Hezbollah [26], who express doubt over taking the vaccine, have created a strong foundation for the spreading of disinformation among users, as is shown in the results of this study.

Calls for Action

The third-most referenced topic includes tweets suggesting calls to action by users. These calls range from protesting to encouraging other users to take action about the vaccine mandate. Additionally, disseminating news about protests

XSL•FO
RenderX

against COVID-19 measures happening worldwide falls under this topic. Overall, this topic covers 12.7% (n=149) of the tweets analyzed, garnering 19,733 (12.1%) retweets. Many tweets discuss a sit-in by individuals opposing the vaccine in Al-Erada Square, a public gathering square in Kuwait City in front of the Kuwaiti National Assembly Building. Historically, this public space has been used for protesting political causes [27], and after the pandemic, it became a popular place for protestors opposing the vaccine [28]. One of the most retweeted posts in this category, with 370 retweets, reads as follows:

> *I will show you the other side of the real sit-in, and unfortunately, last week most of the newspapers portrayed protestors badly. That's why today I was in #Al-Erada_square and filmed with my personal camera a group of interviews that I will upload to my personal account which is not subject to any agenda like that of some newspapers.*

Such tweets suggest the importance of this square as a spot for Kuwaitis opposing vaccines as well as the mainstream news media in Kuwait.

In another example, a user urged vaccinated individuals to speak about any health issues they faced after receiving the vaccine and suggested that several vaccinated individuals are suffering but afraid to talk. In a post retweeted 338 times, 1 Twitter user said:

> *To every person whose healthcare worsened following the vaccine but was told that this has nothing to do with the vaccine…don't be afraid and speak loudly. A committee must be established to collect such cases and study them. People's healthcare is not a game!!*

Similar tweets encourage those against mandatory vaccinations to join protests and actively engage in different forms of peaceful civil disobedience. We also found 1 user promoting fake vaccine passports and offering their WhatsApp number for anyone interested in getting one as a form of action against state vaccination policies.

When tracing the logic behind calls for action, the narratives are not outwardly conspiratorial. These calls for action are rooted in a deep distrust of governments that is compounded by the increased restriction stemming from the pandemic. It is these feelings, along with the invocation of key phrases, that can lead to large-scale calls for action. An example of this process from outside the Middle East can be seen in the communications of 2 Austrian political parties. The People, Freedom, Fundamental Rights (MFG) Party has taken a staunch antivaccination stance and planned multiple protests using key words such as "dictatorship" and "apartheid" to rally large crowds [29]. In a similar vein, the Freedom Party (FPOe) in Austria has also planned large political events, building off a strong opposition to COVID-19 restriction and lockdowns. The party leader, Herbet Kickl, for example, has described Austria's vaccination programs as a genetic experiment, further rallying individuals who may subscribe to similar conspiracy theories [29]. We argue here that the seeds of these ideas and their false claims are similar to what we have studied in Arabic social media posts around COVID-19 disinformation.

Co-opting of Hashtags

Finally, the last topic present in our analysis represents tweets from users aiming to take advantage of antivaccination hashtags to promote irrelevant commercial activities or even reach a wider audience. These tweets, which constituted 67 (5.7%) tweets from the total sample, were not related to the pandemic or vaccines in any sense but were merely marketing and promotional attempts or efforts to expand their online reach by using trending hashtags. For instance, 1 tweet promoted the service of a cleaning company, noting that it uses premium scented cleaning products to remove stains.

Discussion

Principal Findings

This study identified 3 prominent topics that were prevalent amongst Arabic speakers on Twitter. The first topic focused on civil liberties and governments' alleged violations of freedom during the pandemic. Users who posted about this topic tended to express opposition toward vaccine mandates and increased restrictions for those who chose to remain unvaccinated. Those who posted about refusing the vaccine would often cite their legal rights or a lack of trust in the safety of the vaccine. There are also instances of individuals discussing a violation of their human rights due to vaccine mandates. When exploring the logic that underpins these narratives, there is a common theme that governments or pharmaceutical companies are using the pandemic as a way to expand their own powers or profits.

Moving to the category of conspiracy theories about the COVID-19 vaccine, we found that most common theories discussed are about the efficacy of the vaccine and the exaggerated risks that would come with taking it. With both of these narratives, credible individuals who self-identified as either doctors or authorities within the medical community discussed why they are not getting the vaccine and also shared some examples of individuals getting sick or dying after taking the vaccine. We also found instances where individuals accused their governments of intentionally mishandling the pandemic and creating restrictions for their own benefits. The logic used in these examples builds off the mistrust of governments that underpins the majority of the conspiracy theories that have been explored in this study. These theories have been further emboldened by examples of political officials publicly expressing their own hesitations in or rejection of taking the vaccine.

Finally, the third category that we identified is calls for action, in which we found that there are users encouraging others to participate in several different activities, such as rallies or protests. There is an underlying narrative of encouraging others to be brave in the face of intimidation and injustice when participating in these activities. Looking at the logic behind this type of disinformation, we found some similar narratives among far-right parties in the West that co-opted key phrases to build support for their own movements.

Looking at the results of our study, we find that our themes are similar to those identified in several other papers. Specifically, content that spreads rumors and conspiracy theories about

XSL•FO
RenderX

governments and pharmaceutical companies has been identified in separate Twitter and Facebook studies, as well as in Egyptian social media pages [5,12]. Additionally, our category of infringing on civil liberties is similar to social media posts identified in another study that looked at social media posts in the Middle East and the North Africa region [11]. Although our study focused on tweets coming from what seems to be regular users, Grossman et al [11] examined posts disseminated by state-sponsored social media users.

Throughout our study, we noticed that the moderation policies toward COVID-19 misinformation in Arabic on Twitter are not strict, allowing people to post and maintain a variety of conspiracy theories.

Study Limitations

This study was limited to the examination of a few Arabic hashtags on Twitter that imply doubt about the pandemic and its vaccines. Future research needs to focus on mobile apps that are popular in the Middle East, such as Telegram, Signal, and ClubHouse. To better understand the impact of COVID-19 disinformation on the Arab world, the experiences of many countries in the Arab world will need further exploration and reviewing platforms, such as Twitter, YouTube, Instagram, and WhatsApp. It will also be interesting to examine how COVID-19 misinformation in other languages, such as English, is translated, shared, and used by Arabic speakers. Finally, interviews with antivaxxers from the region are lacking, and more surveys based on cross-national comparative research in the Middle East are needed in order to obtain a clearer picture about the perceived views on the efficacy of vaccination and public health measures.

Conclusion

Based on our findings, we were able to identify 3 prominent topics that were prevalent amongst Arabic speakers on Twitter. These categories focused on violations of civil liberties by governments, conspiracy theories about the vaccines, and calls for action. Our findings also highlight the need for more research to better understand the impact of COVID-19 disinformation on the Arab world.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary Figure 1.

[PNG File , 108 KB - infodemiology_v2i2e37007_app1.png]

References

1. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, et al. The COVID-19 social media infodemic. Sci Rep 2020 Oct 06;10(1):16598 [FREE Full text] [doi: 10.1038/s41598-020-73510-5] [Medline: 33024152]
2. Hadj Ameur MS, Aliane H. AraCOVID19-MFH: Arabic COVID-19 Multi-label Fake News & Hate Speech Detection Dataset. Proca Comp Sci 2021;189:232-241. [doi: 10.1016/j.procs.2021.05.086]
3. Jones MO. Disinformation superspreaders: the weaponisation of COVID-19 fake news in the Persian Gulf and beyond. Global Discourse 2020 Nov 01;10(4):431-437. [doi: 10.1332/204378920x16020074088784]
4. Mahlous AR, Al-Laith A. Fake news detection in Arabic tweets during the COVID-19 pandemic. Int J Adv Comp Sci Apps 2021;12(6):776-785. [doi: 10.14569/ijacsa.2021.0120691]
5. Abrahams A. Regional Authoritarians Target the Twittersphere. URL: https:// merip.org/2019/12/regional-authoritarians-target-the-twittersphere/ [accessed 2022-02-21]
6. Gruzd A, Mai P. Going viral: how a single tweet spawned a COVID-19 conspiracy theory on Twitter. Big Data Soc 2020 Jul 20;7(2):205395172093840. [doi: 10.1177/2053951720938405]
7. Debre I, Akram F. Facebook's Language Gaps Weaken Screening of Hate, Terrorism. URL: https://apnews.com/article/the-facebook-papers-language-moderation-problems-392cb2d065f81980713f37384d07e61f [accessed 2022-02-21]
8. Burgees M. New Vaccine Conspiracy Theories Are Going Viral in Arabic. URL: https://www.wired.com/story/new-vaccine-conspiracy-theories-going-viral-arabic/ [accessed 2022-02-21]
9. Alasmari A, Addawood A, Nouh M, Rayes W, Al-Wabil A. A retrospective analysis of the COVID-19 infodemic in Saudi Arabia. Future Internet 2021 Sep 30;13(10):254. [doi: 10.3390/fi13100254]
10. O'Connor C. MENA Monitor: Arabic Covid-19 Vaccine Misinformation Online. URL: https://www.isdglobal.org/isd-publications/mena-monitor-arabic-covid-19-vaccine-misinformation-online/ [accessed 2022-02-21]
11. Grossman S, Khadija H, DiResta R, Kheradpir T, Miller C. Blame It on Iran, Qatar, and Turkey: An Analysis of a Twitter and Facebook Operation Linked to Egypt, the UAE, and Saudi Arabia (TAKEDOWN). URL: https://fsi.stanford.edu/publication/twitter-facebook-egypt-uae-saudi [accessed 2022-02-21]
12. Shehata A, Eldakar M. An exploration of Egyptian Facebook users' perceptions and behavior of COVID-19 misinformation. Sci Technol Libr 2021 May 17;40(4):390-415. [doi: 10.1080/0194262x.2021.1925203]
13. Alimardani M, Elswah M. Online temptations: COVID-19 and religious misinformation in the MENA region. Soc Media Soc 2020 Jul 30;6(3):2056305120948251 [FREE Full text] [doi: 10.1177/2056305120948251] [Medline: 34192037]

XSL•FO
RenderX

14. Gold H, O'Sullivan D. Facebook Has a Coronavirus Problem. It's WhatsApp. URL: https://www.cnn.com/2020/03/18/tech/whatsapp-coronavirus-misinformation/index.html [accessed 2022-02-21]
15. Ben Messaoud M. Social media and the COVID-19 pandemic: the dilemma of fake news clutter vs. social responsibility. JAMMR 2021 Apr 01;14(1):25-45. [doi: 10.1386/jammr_00023_1]
16. 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]
17. Sallam M, Dababseh D, Yaseen A, Al-Haidar A, Taim D, Eid H, et al. COVID-19 misinformation: mere harmless delusions or much more? A knowledge and attitude cross-sectional study among the general public residing in Jordan. PLoS One 2020 Dec 3;15(12):e0243264 [FREE Full text] [doi: 10.1371/journal.pone.0243264] [Medline: 33270783]
18. Lynas M. COVID: Top 10 Current Conspiracy Theories. URL: https://allianceforscience.cornell.edu/blog/2020/04/covid-top-10-current-conspiracy-theories/ [accessed 2022-02-21]
19. Dubowitz M, Ghasseminejad S. Iran‘s Covid-19 disinformation campaign. CTCSent 2020;13(6):48.
20. Ulrichsen K. Qatar and the Gulf Crisis. London, UK: Hurst & Company; 2020.
21. Martin JD, Hassan F. News media credibility ratings and perceptions of online fake news exposure in five countries. Journal Stud 2020 Oct 04;21(16):2215-2233. [doi: 10.1080/1461670x.2020.1827970]
22. Wimmer R, Dominick J. Mass Media Research. 10th edition. New York, NY: Cengage; 2013.
23. Krippendorff K. Krippendorff's alpha. In: Salkind N, editor. Encyclopaedia of Research Design. Vols 1-0. Thousand Oaks, CA: SAGE Publications; 2010.
24. Stewart A. The Great COVID-19 Infodemic: How Disinformation Networks Are Radicalizing Canadians. URL: https://globalnews.ca/news/8450263/infodemic-covid-19-disinformation-canada-pandemic/ [accessed 2022-02-21]
25. BBC. What Is the Great Reset - and How Did It Get Hijacked by Conspiracy Theories?. URL: https://www.bbc.com/news/blogs-trending-57532368 [accessed 2022-02-21]
26. Shaheen K. In the Middle East, Covid Misinformation Is Putting Lives at Risk. URL: https://www.thenationalnews.com/opinion/comment/in-the-middle-east-covid-misinformation-is-putting-lives-at-risk-1.1141385 [accessed 2022-02-21]
27. Westall S. Kuwait Opposition to Stage Joint Demonstration Sunday. URL: https://www.reuters.com/article/us-kuwait-politics-idUSBRE8A61VH20121107 [accessed 2022-02-21]
28. Nasrallah T. Kuwait Deports Jordanian for Protesting against Vaccination. URL: https://gulfnews.com/world/gulf/kuwait/kuwait-deports-jordanian-for-protesting-against-vaccination-1.80440775 [accessed 2022-02-21]
29. Schraer R. Covid: Conspiracy and Untruths Drive Europe's Covid Protests. URL: https://www.bbc.com/news/59390968 [accessed 2022-02-21]

Edited by T Mackey; submitted 02.02.22; peer-reviewed by M Szeto, L Leininger; comments to author 14.05.22; revised version received 02.06.22; accepted 18.07.22; published 26.07.22.

Please cite as:
Al-Rawi A, Fakida A, Grounds K
Investigation of COVID-19 Misinformation in Arabic on Twitter: Content Analysis
JMIR Infodemiology 2022;2(2):e37007
URL: https://infodemiology.jmir.org/2022/2/e37007
doi:10.2196/37007
PMID:35915823

©Ahmed Al-Rawi, Abdelrahman Fakida, Kelly Grounds. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 26.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Negative COVID-19 Vaccine Information on Twitter: Content Analysis

Niko Yiannakoulias[1], PhD; J Connor Darlington[2], MSc; Catherine E Slavik[3], PhD; Grant Benjamin[4], MA

[1]School of Earth, Environment and Society, McMaster University, Hamilton, ON, Canada

[2]School of Geography and Environmental Management, University of Waterloo, Waterloo, ON, Canada

[3]Center for Science Communication Research, School of Journalism and Communication, University of Oregon, Eugene, OR, United States

[4]Department of Economics, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Niko Yiannakoulias, PhD
School of Earth, Environment and Society
McMaster University
1280 Main Street West
L8S4L8
Hamilton, ON
Canada
Phone: 1 905 525 9140 ext 20117
Email: yiannan@mcmaster.ca

Abstract

Background: Social media platforms, such as Facebook, Instagram, Twitter, and YouTube, have a role in spreading anti-vaccine opinion and misinformation. Vaccines have been an important component of managing the COVID-19 pandemic, so content that discourages vaccination is generally seen as a concern to public health. However, not all negative information about vaccines is explicitly anti-vaccine, and some of it may be an important part of open communication between public health experts and the community.

Objective: This research aimed to determine the frequency of negative COVID-19 vaccine information on Twitter in the first 4 months of 2021.

Methods: We manually coded 7306 tweets sampled from a large sampling frame of tweets related to COVID-19 and vaccination collected in early 2021. We also coded the geographic location and mentions of specific vaccine producers. We compared the prevalence of anti-vaccine and negative vaccine information over time by author type, geography (United States, United Kingdom, and Canada), and vaccine developer.

Results: We found that 1.8% (131/7306) of tweets were anti-vaccine, but 21% (1533/7306) contained negative vaccine information. The media and government were common sources of negative vaccine information but not anti-vaccine content. Twitter users from the United States generated the plurality of negative vaccine information; however, Twitter users in the United Kingdom were more likely to generate negative vaccine information. Negative vaccine information related to the Oxford/AstraZeneca vaccine was the most common, particularly in March and April 2021.

Conclusions: Overall, the volume of explicit anti-vaccine content on Twitter was small, but negative vaccine information was relatively common and authored by a breadth of Twitter users (including government, medical, and media sources). Negative vaccine information should be distinguished from anti-vaccine content, and its presence on social media could be promoted as evidence of an effective communication system that is honest about the potential negative effects of vaccines while promoting the overall health benefits. However, this content could still contribute to vaccine hesitancy if it is not properly contextualized.

(JMIR Infodemiol 2022;2(2):e38485) doi:10.2196/38485

KEYWORDS

vaccine acceptance; vaccine hesitancy; Twitter; health communication; COVID-19; social media; infodemiology; misinformation; content analysis; sentiment analysis; vaccine misinformation; web-based health information

XSL•FO
RenderX

Introduction

Major social media platforms, such Facebook, Instagram, Twitter, and YouTube, have been studied for their role in spreading anti-vaccine opinion and misinformation in recent years [1-3]. Evidence suggests that this content may be responsible for lower vaccine coverage in some populations [4-7]. Several processes could explain how content on these platforms influences opinions about vaccines. The simplest explanation is that information directly changes beliefs and behavior; for example, misinformation may lower the acceptance of vaccines by influencing what people believe to be true [8]. Another explanation is that social media is effective at mobilizing the engagement of like-minded individuals that reinforces anti-vaccine perspectives [9]. Another mechanism is psychological reactance—information that appears to threaten freedom of choice (such as vaccine certification) can motivate resistance to recommended behaviors (such as vaccination) [10].

Vaccine hesitancy—a concern toward vaccines that can lead to a delay or rejection of recommended scheduled immunization—is the product of personal experience, knowledge, and structural factors that influence individual trust in the efficacy and safety of vaccines [11,12] and, therefore, could emerge without exposure to misinformation or explicitly anti-vaccine sentiment. For example, the media reporting of genuine adverse reactions to vaccines or infections among vaccinated individuals could increase concerns in the value of vaccination without being factually untrue or expressing an anti-vaccine opinion. In this study, we use the concept of negative vaccine information (NVI) to describe content that characterizes vaccines or vaccination in a negative way without also expressing anti-vaccine sentiment. NVI includes descriptions of possible side effects, vaccine quality concerns, vaccine contamination, statistics on morbidity or mortality associated with vaccination, and negative personal experiences of vaccination.

NVI may be an important part of the communication of vaccine information, since an individual's choice to vaccinate often involves a comparison of health risks to health benefits. Adverse events following vaccination—while uncommon—can occur, and vaccination can be associated with other consequences of personal concern [13]. Provided that vaccination remains a choice, the communication of information about adverse events and vaccine efficacy is necessary to empower people to make personally satisfying decisions. This information may also increase the credibility of the information provider [14], and although it could lead to greater vaccine hesitancy for some people in the short term, it could also increase the trust of public authorities in the long term [15].

Conversely, NVI may be harmful when it lacks contextual background and contributes to greater concern and hesitancy [16]. In the early days of COVID-19 vaccines, the lack of knowledge would have made normal subjective interpretations about vaccine safety and value more challenging [17]. NVI could have been the first information that some people encountered on the web, providing simple accounts of discomfort or adverse reactions that are easier to process than the statistics showing that the balance of evidence is in favor of vaccination. This information, when combined with a number of well-known cognitive biases, may have contributed to some of the vaccine hesitancy during the early period of COVID-19 vaccination [18].

The primary objective of this research was to understand the prevalence and characteristics of NVI on Twitter during the first 4 months of 2021. This time period was chosen because it follows a shift in policy from Twitter and other major social media platforms to address growing concerns about COVID-19 vaccine misinformation [19-21]. There is a consensus that this shift resulted in a decline in anti-vaccine misinformation on these platforms as well as an open question as to whether these changes in policy removed most NVI content. This time period was also marked by the widespread discussion of vaccine brand preference, vaccine nationalism, and early concerns of differences in adverse reaction risk [22], all of which could be a part of various NVI narratives. In this research, we estimated the prevalence of anti-vaccine content and NVI by time, geography, and vaccine developer to understand the degree to which changes in policy by social media outlets may have impacted the availability of both anti-vaccine content and NVI to the public.

Methods

Data Collection

We used the Twitter API to collect data on vaccines and vaccination between December 23, 2020, and April 30, 2021, using the *rtweet* package [23]. An R script was used to automate a search of tweet text, up to 18,000 tweets every 20 minutes, matching the search condition in stage 1 (Table 1). After excluding retweets, this stage yielded a total of 7,827,949 tweets. Next, the text of each tweet was searched to identify and retain only tweets referring to 1 or more of the search terms in stage 2, resulting in 785,107 tweets. In stage 3, each tweet's geocoded country field, location field, and description field were searched, keeping only those tweets with at least one of the stage 3 search terms in at least one of these fields. Tweets were then georeferenced to Canada, the United States, or the United Kingdom based on the geographic information associated with the tweet author's country, location, or description field. These countries were chosen because they initiated vaccination at similar times (in late 2021) and had a sufficient volume of English-language tweets related to vaccine to facilitate analysis. For some tweets, the country, location, and description fields indicated more than 1 country—the United Kingdom, Canada, and the United States—as a location. In these instances, the order of dominance was country, location, and description. For example, if a tweet author had "Canada" in location and "United States" in their description, they were assigned to "Canada." Occasionally, the location and description fields contained more than 1 country. In these cases, the tweets were deleted. In addition, only tweets whose authors had at least 1 follower and only English-language tweets were retained. This process resulted in 217,954 tweets, which was the final sampling frame used in this study.

XSL•FO
RenderX

Table 1. Search criteria.

Stage	Search criteria
Stage 1	(“covid” OR “coronavirus” OR “corona” OR “sars-cov-2” OR “sarscov2”) AND (“vaccine” OR “vaccination” OR “vaccinated” OR “shot” OR “inoculate” OR “inoculation” OR “inoculated” OR “immunize” OR “immunized” OR “immunization”)
State 2	(“astrazeneca” OR “astrazeneca” OR “azd1222” OR “covishield” OR “vaxzervia” OR “oxford-astra-zeneca” OR “oxford-astrazeneca” OR “oxfordastrazeneca” OR “pfizer” OR “tozinameran” OR “BNT162b2” OR “biontech” OR “pfizer-biontech” OR “fosun-biontech” OR “pfizerbiontech” OR “fosunbiontech” OR “moderna” OR “mrna-1273” OR “cx-024414” OR “tak919” OR “cx024414” OR “tak919” OR “mrna1273” OR “sputnik” OR “sputnikv” OR “gam-covid-vac” OR “gamcovidvac” OR “sinopharm” OR “bbibpcorv” OR “bbibpcorv” OR “johnsonandjohnson” OR “johnson&johnson” OR “janssen” OR “ad26.cov2,s” OR “jnj-78436735” OR “ad26covs1” OR “vac31518” OR “sinovac” OR “coronavac” OR “picovacc” OR “covaxin” OR “bbv152” OR “whole-virioninactivated” OR “wholevirioninactivated” OR “bharat” OR “novavax” OR “nvx-cov2373” OR “nvxcov2373” OR “sars-cov-rs” OR “covovax”)
Stage 3	(“Canada” OR “Canadian” OR “British Columbia” OR “Alberta” OR “Saskatchewan” OR “Manitoba” OR “Ontario” OR “Quebec” OR “New Brunswick” OR “Nova Scotia” OR “Prince Edward Island” OR “Newfoundland” OR “Yukon” OR “Northwest Territories” OR “Nunavut” OR “U.S.A.” OR “USA” OR “United States of America” OR “United States” OR “American” OR “Alabama” OR “Alaska” OR “Arizona” OR “Arkansas” OR “California” OR “Colorado” OR “Connecticut” OR “Delaware” OR “Florida” OR “Georgia” OR “Hawaii” OR “Idaho” OR “Illinois” OR “Indiana” OR “Iowa” OR “Kansas” OR “Kentucky” OR “Louisiana” OR “Maine” OR “Maryland” OR “Massachusetts” OR “Michigan” OR “Minnesota” OR “Mississippi” OR “Missouri” OR “Montana” OR “Nebraska” OR “Nevada” OR “New Hampshire” OR “New Jersey” OR “New Mexico” OR “New York” OR “North Carolina” OR “North Dakota” OR “Ohio” OR “Oklahoma” OR “Oregon” OR “Pennsylvania” OR “Rhode Island” OR “South Carolina” OR “South Dakota” OR “Tennessee” OR “Texas” OR “Utah” OR “Vermont” OR “Virginia” OR “Washington” OR “West Virginia” OR “Wisconsin” OR “Wyoming” OR “UK” OR “U.K.” OR “England” OR “Wales” OR “Northern Ireland” OR “Scotland” OR “United Kingdom” OR “British” OR “Scottish” OR “Welsh” OR “English”)

Data Classification

In total, 9000 tweets were randomly sampled (with replacement) from this sampling frame. After training on a separate sample of 200 tweets, all 4 authors read through mutually exclusive subsamples of 8800 tweets and coded every tweet in which the text contained negative information or sentiment about vaccines as “1”; otherwise, the tweets are coded as “0.” This criterion includes statistics or reports of adverse reactions, personal statements about adverse reactions, skepticism toward vaccine developers, policy decisions to stop or delay the administration of vaccines, expressions of concern about the vaccines, and implicitly or explicitly anti-vaccine tweets. After these tweets were coded, the authors went through the tweets coded as “1” and distinguished between those that were implicitly or explicitly anti-vaccine and those that were not. Tweets that were not anti-vaccine were classified as NVI, and the remainder were classified as anti-vaccine. Textbox 1 illustrates examples of anti-vaccine and NVI tweets. Each coder classified a random sample of 300 of the same tweets to compare interrater reliability using Krippendorff α [24]. Finally, after all the above coding was completed, each tweet author was coded into 1 of 5 types: media, medical and health, government, other, or restricted/closed account. Due to an error in data extraction early in the study, tweets were excluded if they occurred outside the period from January 6 to April 30, 2021. The final data set included 7306 tweets exclusive of the 300 tweets used to compare agreement between data coders, which were not used in the analysis.

XSL•FO
RenderX

Textbox 1. Examples of anti-vaccine and negative vaccine information (NVI) tweets.

Anti-vaccine tweets

- “Israelis got facial paralysis after having received the Pfizer Covid vaccine. This vaccine is anything but safe. It’s not Covid which is threatening the public health. It’s the Pfizer vaccine”
- “Please listen and share widely esp. with authorities. Moderna/Pfizer in highly deceptive, harmful medical practice re covid “vaccine” (in fact ‘gene therapy technology)...”
- “These vaccines are not safe for everyone! Do not be peer pressured into destroying your life over this!”
- “An experimental vaccine using experimental technology. And in the case of Moderna, a company with no prior pharmaceutical, much less vaccine track record. Shame on you if you don’t protect the people against mandatory Covid vaccine by employers and businesses.”
- “6 people died after taking the Pfizer vaccine, I didn’t look into Moderna yet. Since There are known treatments for covid, why would anyone want to take a vaccine With unknown long term side effects especially given how extremely high covid’s survival rate is?”

NVI tweets

- “I’m curious of your age bracket. I’m mid sixties and received my 2nd Pfizer on the 3rd. Just the arm pain. My SIL late 30’s contracted covid-19 got the first moderna shot and was down for 2 days. Point being I believe the youth have more side effects bc of better immune sys.”
- “The European Medicines Agency (EMA) said on Wednesday it had found a possible link between AstraZeneca’s coronavirus vaccine and reports of very rare cases of blood clots in people who had received the shot.”
- “Well i guess i join the many that experience side effects of the 2nd vaccine. 12 hours later. Aches, chills and a horrendously sore arm. Started to feel better after a few hours. Now im just sore.”
- “AstraZeneca Concerns Throw Europe’s Covid-19 Vaccine Rollout Into Deeper Disarray”
- “Moderna says possible allergic reactions to COVID-19 vaccine under investigation”

Results

Of the 300 tweets coded by all authors for the purpose of measuring between coder reliability, 79% (n=237) had full agreement across all content coders on NVI (Krippendorff α=0.63) and 80% (n=240) had full agreement across all content coders on anti-vaccine content (Krippendorff α=0.25).

Of the 7306 tweets, 131 (1.8%) were coded as anti-vaccine and 1533 (21%) were coded as NVI. Table 2 shows the anti-vaccine and NVI tweet frequencies by author type. Due to the small number of anti-vaccine tweets and the relatively low level of interrater agreement of anti-vaccine tweets, no further analysis of these data is shown, and all remaining analysis excluded anti-vaccine tweets.

For the 1533 NVI tweets, 37.9% (n=581) originated from the United Kingdom, 49.7% (n=762) originated from the United States, and 12.4% (n=190) originated from Canada. The total number of tweets and percentage of NVI tweets by geography are shown on Figure 1. Pairwise *z* tests of differences in the percentage of NVI tweets in this figure suggest that the apparent difference between Canada and the United Kingdom could be due to chance (*P*=.23), although the differences were statistically significant in the comparison between Canada and the United States (*P*=.01) and between the United Kingdom and the United States (*P*<.001).

Comparisons of NVI tweets across different vaccine developers are shown on Figure 2. The number of tweets varied by developer, but the most noteworthy contrast involved Oxford/AstraZeneca, for which NVI tweets made up almost 35.69% (713/1998) of the content, more than double the percentage of NVI tweets observed for other developers (Moderna: 204/1290, 15.81%; Pfizer-BioNTech: 477/2920, 16.34%; Other: 139/967, 14.37%; *P*<.001 for all pairwise comparisons of Oxford/AstraZeneca with other developers). Figure 3 provides more detail with the percentage of NVI tweets by country and vaccine developer. NVI tweets were more commonly associated with the Oxford/AstraZeneca vaccine than the other vaccine developers for tweet authors in all 3 counties studied. The figure also suggests that a higher proportion of NVI tweets related to Moderna and Pfizer-BioNTech originated in the United Kingdom than in the United States or Canada.

Figure 4 illustrates the proportion of NVI tweets by country, vaccine developer, and month of year. The dotted horizontal lines are the proportions of NVI tweets for the entire study period. These figures illustrate a very similar trend of rising NVI tweets over time associated with the Oxford/AstraZeneca vaccine for Twitter users in all 3 countries. Another noteworthy observation is the uniformly higher proportion of NVI tweets authored by Twitter users in the United Kingdom associated with the Pfizer-BioNTech and Moderna vaccines, although due to the smaller number of Moderna-related tweets authored by UK Twitter users, these proportions have a considerably larger confidence interval. Unlike the Oxford/AstraZeneca vaccine, neither of these observations is accompanied by a clear trend over time. In Canada, it appears that the highest volume of NVI tweets occurred in April for all vaccines.

XSL•FO
RenderX

Table 2. Anti-vaccine and negative vaccine information by account type.

Tweet	Government (n=140), n (%)	Media (n=1755), n (%)	Medical (n=1078), n (%)	Other (n=4032), n (%)	Closed, deleted, or restricted account (n=300), n (%)
Anti-vaccine	0 (0)	3 (0.17)	0 (0)	108 (2.68)	19 (6.33)
Negative vaccine information	19 (13.57)	342 (19.49)	81 (7.51)	1002 (24.85)	89 (29.67)

Figure 1. Frequency of NVI and non-NVI tweets by country. Percentages are the fraction of tweets in a country that have NVI. CA: Canada; NVI: negative vaccine information; UK: United Kingdom; US: United States.

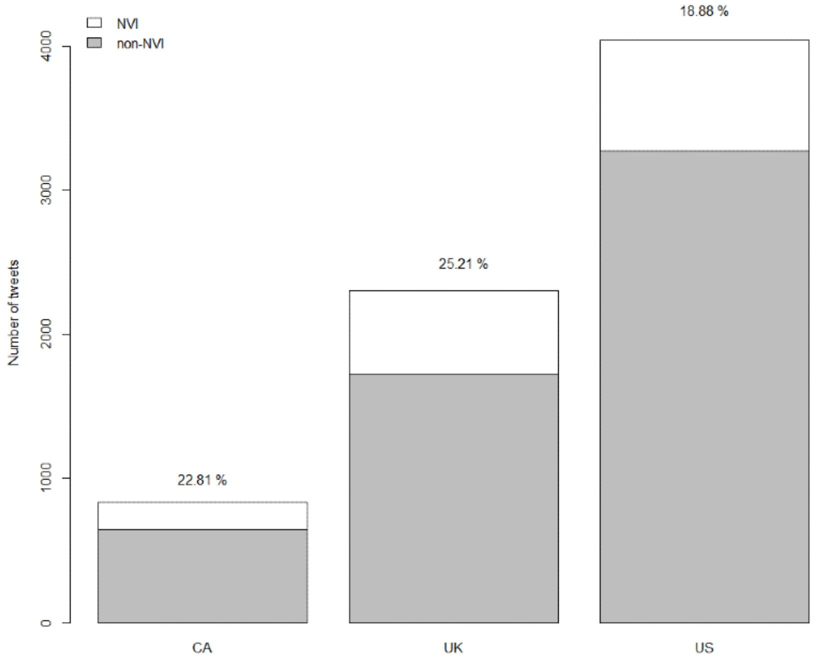
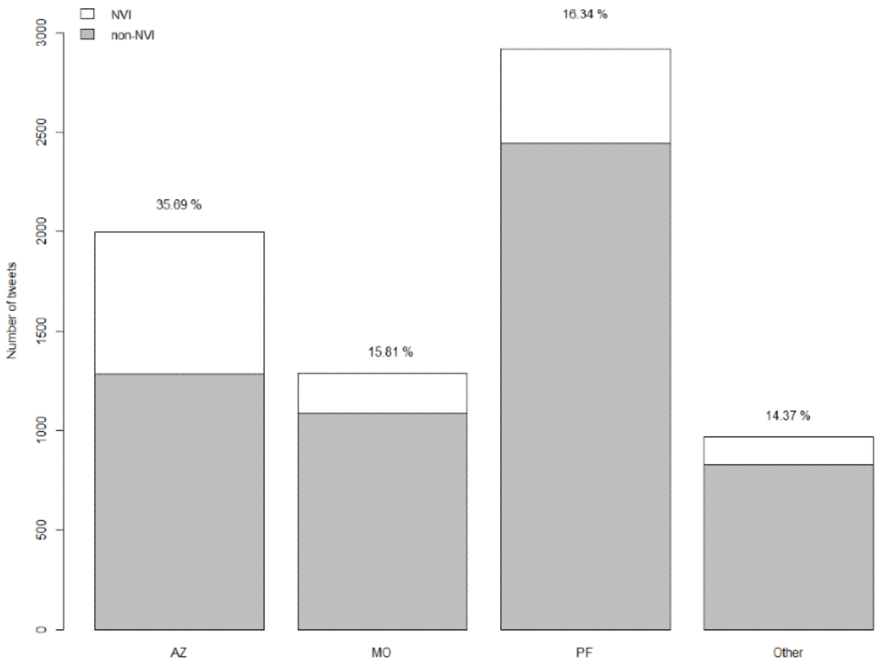


Figure 2. Frequency of NVI and non-NVI tweets by COVID-19 vaccine developers. Numbers inside bars are percentages of tweets that are NVI. AZ: Oxford/AstraZeneca; MO: Moderna; NVI: negative vaccine information; Other: any other COVID-19 vaccine developer; PF: Pfizer-BioNTech.



XSL•FO
RenderX

Figure 3. Proportion of NVI tweets by vaccine developer and country. AZ: Oxford/AstraZeneca; CA: Canada; MO: Moderna; NVI: negative vaccine information; Other: any other COVID-19 vaccine developer; PF: Pfizer-BioNTech; UK: United Kingdom; US: United States.

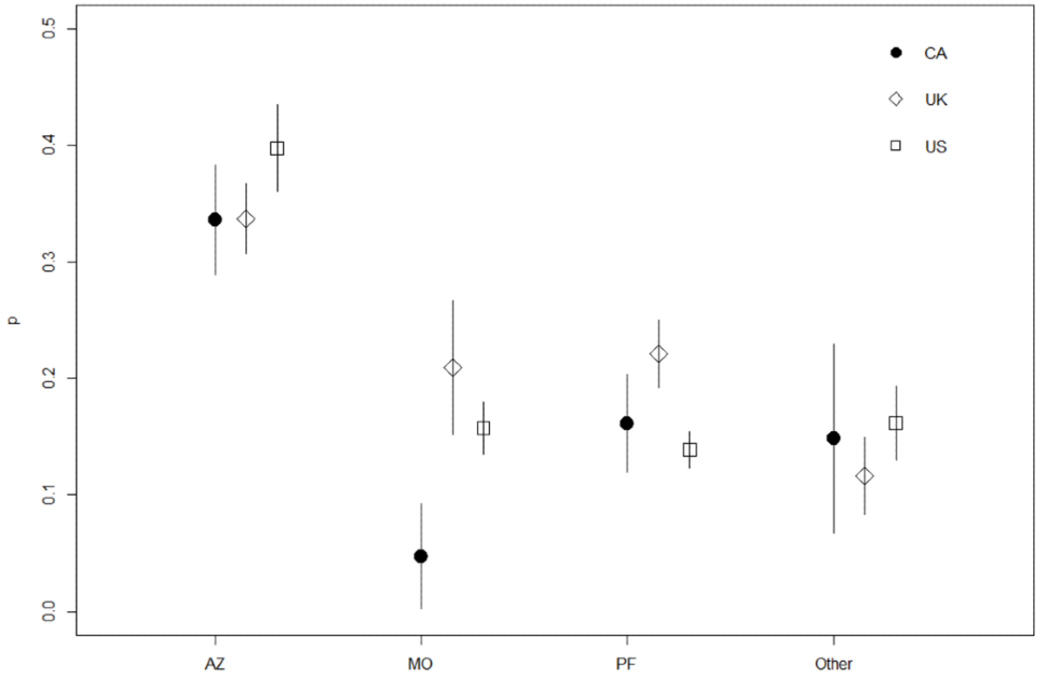
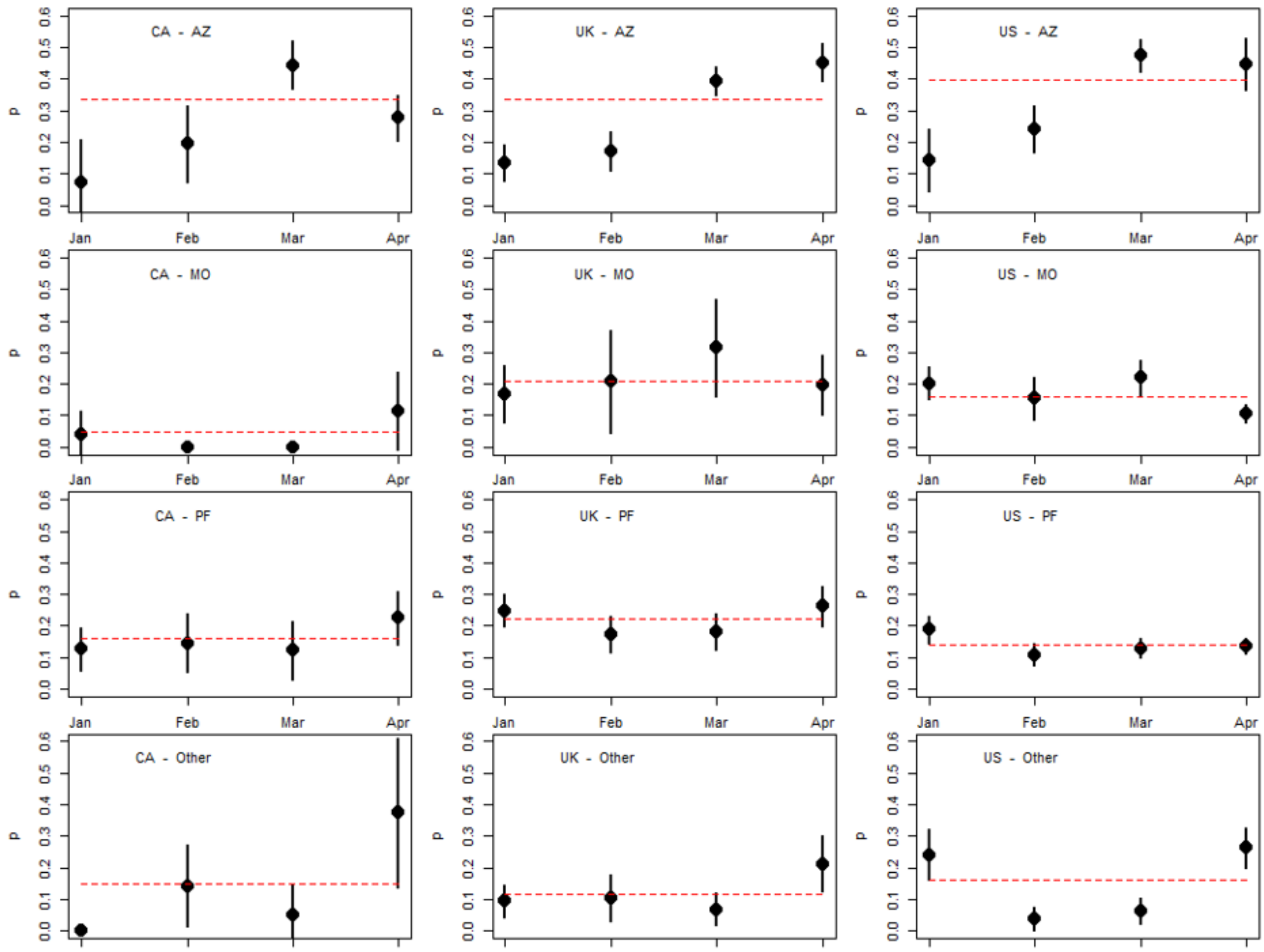


Figure 4. Proportion of NVI tweets by month, country and vaccine developer. Vertical lines are 95% CIs. AZ: Oxford/AstraZeneca; CA: Canada; MO: Moderna; NVI: negative vaccine information; Other: any other COVID-19 vaccine developer; PF: Pfizer-BioNTech; UK: United Kingdom; US: United States.



XSL•FO
RenderX

Discussion

Principal Findings

Our results indicate that less than 2% of vaccine-related tweets contain anti-vaccine content and 21% contain NVI. This finding suggests that very little Twitter content was anti-vaccine in early 2021, which is consistent with the findings of other research [25]. When compared to research on pre–COVID-19 anti-vaccine content on Twitter (which found that anti-vaccine content was closer to 9%) [26], this finding could suggest that the changes in policy in late 2020 did reduce anti-vaccine content. Although we found anti-vaccine content to be rare on Twitter over the study period, NVI tweets were not uncommon and were generated by a broad range of content authors. NVI content was generated by all Twitter content generator groups, making up almost 20% of the content from media sources and almost 14% of the content from government sources.

More than 25% of Twitter content authored in the United Kingdom appeared to be NVI, but in terms of absolute quantity, a plurality of NVI originated from Twitter accounts in the United States. This finding reflects one of the ongoing realities of globalized social media—that content has few barriers—and domestic regulations that attempt content control will only work if they are enforced in the jurisdictions responsible for a large share of the material. Nevertheless, it is difficult to know the reasons for the relatively low percentage of NVI tweets generated in the United States (compared to Canada and the United Kingdom). One explanation is that alternative platforms were more popular for the communication of NVI in the United States, including those with specific political agendas that emerged in the last year. As such, NVI content generators in the United States may have shifted to an alternative platform in the anticipation of changes to Twitter's content policy, resulting in less NVI content on Twitter. It is also possible that Twitter targeted more content authored in the United States than in the United Kingdom or Canada. However, other explanations are possible, and our analysis offers no clear evidence explaining this observation.

In January, the volume of NVI tweets was similar for all vaccines, but as concerns about the safety of the Oxford/AstraZeneca vaccine rose in March of 2021, NVI tweets specific to this vaccine rose for Twitter users in all 3 countries—a finding consistent with other research [27]. Unlike the United States, both Canada and the United Kingdom approved and administered the Oxford/AstraZeneca vaccine for emergency use; however, Twitter users in the United States reported the highest proportion of NVI tweets mentioning the Oxford/AstraZeneca vaccine. Twitter users in the United Kingdom were responsible for more NVI content related to the Moderna and Pfizer-BioNTech vaccines than in Canada or the United States. This finding is noteworthy since Pfizer-BioNTech and Moderna made up a smaller quantity of vaccines acquired for use in the United Kingdom than Oxford/AstraZeneca. This pattern—where less commonly used vaccines are associated with higher NVI—could be explained by the absence of positive public health messaging related to that vaccine. In the United States, for example, public health officials and clinicians would have no reason to make Twitter posts about getting the Oxford/AstraZeneca vaccine, as it was not available for use, which could result in a smaller denominator in the calculation of NVI prevalence and a higher proportion of negative tweets associated with this vaccine.

Overall, our results suggest that a small fraction of COVID-19 vaccine–related tweets included anti-vaccine content, but NVI was relatively common. NVI was authored by all types of Twitter users and varied by geography, time, and vaccine developer. Unlike most anti-vaccine content, NVI could be viewed as a legitimate part of the pro-vaccine information narrative, since its presence may provide information consumers an increased sense of trust about the transparency of vaccine developers and government. Its presence on social media could even be promoted as evidence of an effective communication system that is honest about the potential negative effects of vaccines while promoting the overall health benefits. Indeed, the high level of NVI associated with the Oxford/AstraZeneca vaccine over time could even be seen as an important indicator of openness and transparency as evidence changes over time.

This research provides no insight as to whether NVI on Twitter has any impact on COVID-19 vaccine hesitancy. Some research has suggested that certain types of information presented in the media could increase vaccine hesitancy [7,28], yet other research has suggested that Twitter content has little effect on public opinion or behavior [29,30]. Arguably, if NVI on Twitter or other forms of media is a concern, it is not through its presence but the absence of context required for proper interpretation. Information about adverse reactions is not by itself evidence against the benefits of vaccination, but without context for understanding the balance of risks, it could cause concern that creates or amplifies vaccine hesitancy [16]. Research in risk communication suggests the importance of a foundational knowledge of science and numeracy [31]. Since the availability of NVI is likely to persist in all media, efforts must continue to improve how information is communicated by focusing on individualized risk estimates and visual risk displays [32].

Implications

Content moderation remains a challenge for all media platforms, but unlike most traditional media, social media content is user generated, with the social media exerting little editorial control. Changes in policy in 2020 seem to have impacted the content on social media, but striking the right balance between freedom of expression and content control remains an important challenge. Further discussion of the content moderation process is a critical public service and can help us better understand the social media platforms we use [33].

Research conducted in the early phase of the COVID-19 outbreak [34] had suggested a substantial rise in anti-vaccine content on Twitter even before vaccines were widely available. After the changes in COVID-19–related policy, some Twitter users were banned and some of their content was removed. As reported in recent research, some censored content authors view the censorship as a sign of subterfuge and that social media companies are complicit in a cover-up of the true harms of vaccination [35]. We found that anti-vaccine information was rare on these platforms as the vaccines were rolled out to the

XSL•FO
RenderX

public; however, critiques that all negative information about vaccines has been suppressed is not consistent with the evidence presented in this study. In the early days of vaccination, Twitter was widely used as a platform for sharing information about adverse events associated with vaccination, including content published by official public health sources as well as the media. Moreover, NVI associated with the Oxford/AstraZeneca vaccine is consistent with the general concern that it was associated with more adverse reactions in early 2021, something that would not have been expected if Twitter had universally censored the NVI that could harm the reputation of vaccine manufacturers.

Nonetheless, the presence of NVI may still present a challenge to public health communicators if it results in a net increase in vaccine hesitancy. NVI may underlie several cognitive biases that contribute to vaccine hesitancy [36]. Personal stories about adverse reactions can create a negative impression of the vaccination experience that is easily recalled when making decisions—a form of availability bias [37]. As a social media platform, Twitter is particularly effective at delivering short, easily digested, and impactful messages rather than scientifically informed and data-driven arguments. Early negative impressions about vaccines that were neither anti-vaccine nor misinformation may have had a substantial influence on the prevalence of vaccine hesitancy, particularly in early 2021.

Given that NVI is common and can be viewed as a normal part of the health communication process, eliminating it is neither possible nor desirable. Growing evidence shows that personal narratives (from experts and nonexperts) are effective at engaging social media consumers about health information and may often be more effective than strictly informational guidance [38,39]. On this basis, countering the effects of NVI on vaccine hesitancy may be best addressed on Twitter by offering alternative positive personal narratives about pro-vaccine experiences [40]. Existing research suggests that such pro-vaccination narratives may be more effective when accompanied by video or audio content rather than text alone [41], but further works needs to be done to determine how these messages can be used most effectively.

Limitations

One important limitation to this study is the lack of agreement on anti-vaccine tweets, which had, at best, fair interrater agreement [42]. The text limit for individual tweets can make the meaning and intent of a tweet difficult to interpret, and determining intent is important for classifying tweets as anti-vaccine. It is for this reason that an extensive analysis of anti-vaccine tweets was not presented. Importantly, however, coding did not yielded a large number of anti-vaccine tweets—with 3 of the 4 coders yielding less than 2% of their tweets as anti-vaccine. The share of NVI tweets was similarly uniform, although interrater agreement was not particularly high.

The search criteria used to select tweets for analysis were likely to have excluded relevant tweets from the sampling frame. First, we did not include alternative spellings of vaccine that are sometimes used by the anti-vaccine community. This exclusion very likely led to an underestimation of anti-vaccine tweets in the sampling frame. It is difficult to estimate the effect of excluding these search terms on our analysis, but even if we underestimated by half, it would still leave less than 4% of the tweets as anti-vaccine and would not dramatically change our conclusions. Second, the georeferencing process eliminated a large number of tweets, and it is unclear if this exclusion introduced a bias into the results. It is possible that certain forms of geographic identification that we did not consider—for example, referring to the city a person lives in rather than the country or province/state—may be associated with disposition toward vaccines in some way. Although the authors cannot rule out this possibility, it seems implausible that such an effect would have a large impact in all 3 jurisdictions studied, and it seems reasonable to assume this effect would be small.

Conclusions

Our results suggest that Twitter was not a substantial source of anti-vaccine content in early 2021, but it still contained a large quantity of information that could contribute to vaccine hesitancy. It is important to note, however, that NVI is not unique to social media and can be found in traditional media sources and even public health notifications from government agencies. Therefore, it would be inappropriate to treat all (or even most) NVI as socially deleterious. Moreover, this information (particularly when authored by reputable sources) may have the long-term benefit of increasing trust in public health messaging, as open communication of negative and positive effects could contribute to increase faith in the transparency and honesty of public health messaging.

Conflicts of Interest

None declared.

References

1. Gunaratne K, Coomes EA, Haghbayan H. Temporal trends in anti-vaccine discourse on Twitter. Vaccine 2019 Aug 14;37(35):4867-4871. [doi: 10.1016/j.vaccine.2019.06.086] [Medline: 31300292]
2. Yiannakoulias N, Slavik CE, Chase M. Expressions of pro- and anti-vaccine sentiment on YouTube. Vaccine 2019 Apr 03;37(15):2057-2064. [doi: 10.1016/j.vaccine.2019.03.001] [Medline: 30862365]
3. 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]

XSL•FO
RenderX

4. Dubé E, Gagnon D, Ouakki M, Bettinger JA, Guay M, Halperin S, Canadian Immunization Research Network. Understanding vaccine hesitancy in Canada: results of a consultation study by the Canadian Immunization Research Network. PLoS One 2016 Jun 3;11(6):e0156118 [FREE Full text] [doi: 10.1371/journal.pone.0156118] [Medline: 27257809]
5. Eskola J, Duclos P, Schuster M, MacDonald NE, SAGE Working Group on Vaccine Hesitancy. How to deal with vaccine hesitancy? Vaccine 2015 Aug 14;33(34):4215-4217 [FREE Full text] [doi: 10.1016/j.vaccine.2015.04.043] [Medline: 25896378]
6. Basch CE, Basch CH, Hillyer GC, Meleo-Erwin ZC, Zagnit EA. YouTube videos and informed decision-making about COVID-19 vaccination: successive sampling study. JMIR Public Health Surveill 2021 May 06;7(5):e28352 [FREE Full text] [doi: 10.2196/28352] [Medline: 33886487]
7. Wilson SL, Wiysonge C. Social media and vaccine hesitancy. BMJ Glob Health 2020 Oct 23;5(10):e004206 [FREE Full text] [doi: 10.1136/bmjgh-2020-004206] [Medline: 33097547]
8. 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]
9. 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]
10. Kim H, Seo Y, Yoon HJ, Han JY, Ko Y. The effects of user comment valence of Facebook health messages on intention to receive the flu vaccine: the role of pre-existing attitude towards the flu vaccine and psychological reactance. Int J Advert 2021 Jan 04;40(7):1187-1208. [doi: 10.1080/02650487.2020.1863065]
11. Burke PF, Masters D, Massey G. Enablers and barriers to COVID-19 vaccine uptake: an international study of perceptions and intentions. Vaccine 2021 Aug 23;39(36):5116-5128 [FREE Full text] [doi: 10.1016/j.vaccine.2021.07.056] [Medline: 34340856]
12. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: an overview. Hum Vaccin Immunother 2013 Aug 27;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
13. McLenon J, Rogers MA. The fear of needles: a systematic review and meta-analysis. J Adv Nurs 2019 Jan 11;75(1):30-42. [doi: 10.1111/jan.13818] [Medline: 30109720]
14. Ashwell D, Murray N. When being positive might be negative: an analysis of Australian and New Zealand newspaper framing of vaccination post Australia's No Jab No Pay legislation. Vaccine 2020 Jul 31;38(35):5627-5633 [FREE Full text] [doi: 10.1016/j.vaccine.2020.06.070] [Medline: 32653274]
15. Petersen MB, Bor A, Jørgensen F, Lindholt MF. Transparent communication about negative features of COVID-19 vaccines decreases acceptance but increases trust. Proc Natl Acad Sci U S A 2021 Jul 20;118(29):e2024597118 [FREE Full text] [doi: 10.1073/pnas.2024597118] [Medline: 34292869]
16. Löfstedt R, Way D. Transparency and trust in the European pharmaceutical sector: outcomes from an experimental study. J Risk Res 2014 Jun 12;19(9):1082-1103. [doi: 10.1080/13669877.2014.919517]
17. Reyna VF. Risk perception and communication in vaccination decisions: a fuzzy-trace theory approach. Vaccine 2012 May 28;30(25):3790-3797 [FREE Full text] [doi: 10.1016/j.vaccine.2011.11.070] [Medline: 22133507]
18. Azarpanah H, Farhadloo M, Vahidov R, Pilote L. Vaccine hesitancy: evidence from an adverse events following immunization database, and the role of cognitive biases. BMC Public Health 2021 Sep 16;21(1):1686 [FREE Full text] [doi: 10.1186/s12889-021-11745-1] [Medline: 34530804]
19. COVID-19 medical misinformation policy. YouTube Help. URL: https://support.google.com/youtube/answer/9891785?hl=en [accessed 2021-07-27]
20. COVID-19: Our approach to misleading vaccine information. Twitter Safety. 2020 Dec 16. URL: https://blog.twitter.com/en_us/topics/company/2020/covid19-vaccine [accessed 2021-07-27]
21. Facebook Help Center. URL: https://www.facebook.com/help/230764881494641/ [accessed 2021-08-27]
22. Lowen P. Report 3: do vaccine brand preferences exist? series | COVID-19 vaccine skepticism in Canada. Public Policy Forum. 2021 Mar 16. URL: https://ppforum.ca/publications/report-3-do-vaccine-brand-preferences-exist/ [accessed 2021-08-30]
23. Kearney M. rtweet: collecting and analyzing Twitter data. J Open Source Softw 2019 Oct;4(42):1829. [doi: 10.21105/joss.01829]
24. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Commun Methods and Meas 2007 Apr;1(1):77-89. [doi: 10.1080/19312450709336664]
25. 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]
26. Piedrahita-Valdés H, Piedrahita-Castillo D, Bermejo-Higuera J, Guillem-Saiz P, Bermejo-Higuera JR, Guillem-Saiz J, et al. Vaccine hesitancy on social media: sentiment analysis from June 2011 to April 2019. Vaccines (Basel) 2021 Jan 07;9(1):28 [FREE Full text] [doi: 10.3390/vaccines9010028] [Medline: 33430428]
27. Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J 2022 Jul 09;98(1161):544-550 [FREE Full text] [doi: 10.1136/postgradmedj-2021-140685] [Medline: 34373343]

XSL•FO
RenderX

28. Tran BX, Boggiano V, Nguyen LH, Latkin C, Nguyen H, Tran TT, et al. Media representation of vaccine side effects and its impact on utilization of vaccination services in Vietnam. Patient Prefer Adherence 2018;12:1717-1728 [FREE Full text] [doi: 10.2147/PPA.S171362] [Medline: 30233151]
29. Hong S, Nadler D. Does the early bird move the polls?: the use of the social media tool 'Twitter' by U.S. politicians and its impact on public opinion. 2011 Jun 12 Presented at: dg.o '11: The 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times; June 12-15, 2001; College Park, MD p. 182-186. [doi: 10.1145/2037556.2037583]
30. Mackert M, Kim E, Guadagmo M, Donovan-Kicken E. Using Twitter for prenatal health promotion: encouraging a multivitamin habit among college-aged females. Stud Health Technol Inform 2012;182:93-103. [Medline: 23138084]
31. Reyna VF, Brainerd CJ. The importance of mathematics in health and human judgment: numeracy, risk communication, and medical decision making. Learn Individ Differ 2007 Apr;17(2):147-159. [doi: 10.1016/j.lindif.2007.03.010]
32. Gordon-Lubitz RJ. MSJAMA. Risk communication: problems of presentation and understanding. JAMA 2003 Jan 01;289(1):95. [doi: 10.1001/jama.289.1.95] [Medline: 12503988]
33. Gillespie T. Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media. New Haven, CT: Yale University Press; 2018.
34. 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 Health 2020 Dec 15;14(1):12-19. [doi: 10.1080/17538068.2020.1858222]
35. Leader AE, Burke-Garcia A, Massey PM, Roark JB. Understanding the messages and motivation of vaccine hesitant or refusing social media influencers. Vaccine 2021 Jan 08;39(2):350-356 [FREE Full text] [doi: 10.1016/j.vaccine.2020.11.058] [Medline: 33280856]
36. Azarpanah H, Farhadloo M, Vahidov R, Pilote L. Vaccine hesitancy: evidence from an adverse events following immunization database, and the role of cognitive biases. BMC Public Health 2021 Sep 16;21(1):1686 [FREE Full text] [doi: 10.1186/s12889-021-11745-1] [Medline: 34530804]
37. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974 Sep 27;185(4157):1124-1131. [doi: 10.1126/science.185.4157.1124] [Medline: 17835457]
38. Solnick RE, Chao G, Ross RD, Kraft-Todd GT, Kocher KE. Emergency physicians and personal narratives improve the perceived effectiveness of COVID-19 public health recommendations on social media: a randomized experiment. Acad Emerg Med 2021 Feb 27;28(2):172-183 [FREE Full text] [doi: 10.1111/acem.14188] [Medline: 33263357]
39. Yiannakoulias N, Tooby R, Sturrock S. Celebrity over science? an analysis of Lyme disease video content on YouTube. Soc Sci Med 2017 Oct;191:57-60. [doi: 10.1016/j.socscimed.2017.08.042] [Medline: 28898716]
40. Cawkwell PB, Oshinsky D. Storytelling in the context of vaccine refusal: a strategy to improve communication and immunisation. Med Humanit 2016 Mar 05;42(1):31-35. [doi: 10.1136/medhum-2015-010761] [Medline: 26438615]
41. Shen F, Sheer VC, Li R. Impact of narratives on persuasion in health communication: a meta-analysis. J Advert 2015 May 05;44(2):105-113. [doi: 10.1080/00913367.2015.1018467]
42. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]

Abbreviations

NVI: negative vaccine information

Edited by R Cuomo; submitted 04.04.22; peer-reviewed by D Wawrzuta, N Hu; comments to author 28.06.22; revised version received 29.07.22; accepted 18.08.22; published 29.08.22.

Please cite as:
Yiannakoulias N, Darlington JC, Slavik CE, Benjamin G
Negative COVID-19 Vaccine Information on Twitter: Content Analysis
JMIR Infodemiology 2022;2(2):e38485
URL: https://infodemiology.jmir.org/2022/2/e38485
doi:10.2196/38485
PMID:36348980

©Niko Yiannakoulias, J Connor Darlington, Catherine E Slavik, Grant Benjamin. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 29.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Quantifying Changes in Vaccine Coverage in Mainstream Media as a Result of the COVID-19 Outbreak: Text Mining Study

Bente Christensen[1], MSc; Daniel Laydon[2], PhD; Tadeusz Chelkowski[3], MSc; Dariusz Jemielniak[3], Prof Dr; Michaela Vollmer[2], PhD; Samir Bhatt[2,4], Prof Dr; Konrad Krawczyk[1], PhD

[1]Department of Mathematics and Computer Science, University of Southern Denmark, Odense, Denmark

[2]Department of Infectious Disease Epidemiology, MRC Centre for Global Infectious Disease Analysis, Imperial College London, London, United Kingdom

[3]Department of Management in the Network Society, Kozminski University, Warsaw, Poland

[4]Section of Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Denmark

Corresponding Author:
Konrad Krawczyk, PhD
Department of Mathematics and Computer Science
University of Southern Denmark
Campusvej 55
Odense, 5230
Denmark
Phone: 45 5551122
Email: konradk@imada.sdu.dk

Abstract

Background: Achieving herd immunity through vaccination depends upon the public's acceptance, which in turn relies on their understanding of its risks and benefits. The fundamental objective of public health messaging on vaccines is therefore the clear communication of often complex information and, increasingly, the countering of misinformation. The primary outlet shaping public understanding is mainstream online news media, where coverage of COVID-19 vaccines was widespread.

Objective: We used text-mining analysis on the front pages of mainstream online news to quantify the volume and sentiment polarization of vaccine coverage.

Methods: We analyzed 28 million articles from 172 major news sources across 11 countries between July 2015 and April 2021. We employed keyword-based frequency analysis to estimate the proportion of overall articles devoted to vaccines. We performed topic detection using BERTopic and named entity recognition to identify the leading subjects and actors mentioned in the context of vaccines. We used the Vader Python module to perform sentiment polarization quantification of all collated English-language articles.

Results: The proportion of front-page articles mentioning vaccines increased from 0.1% to 4% with the outbreak of COVID-19. The number of negatively polarized articles increased from 6698 in 2015-2019 to 28,552 in 2020-2021. However, overall vaccine coverage before the COVID-19 pandemic was slightly negatively polarized (57% negative), whereas coverage during the pandemic was positively polarized (38% negative).

Conclusions: Throughout the pandemic, vaccines have risen from a marginal to a widely discussed topic on the front pages of major news outlets. Mainstream online media has been positively polarized toward vaccines, compared with mainly negative prepandemic vaccine news. However, the pandemic was accompanied by an order-of-magnitude increase in vaccine news that, due to low prepandemic frequency, may contribute to a perceived negative sentiment. These results highlight important interactions between the volume of news and overall polarization. To the best of our knowledge, our work is the first systematic text mining study of front-page vaccine news headlines in the context of COVID-19.

(JMIR Infodemiology 2022;2(2):e35121) doi:10.2196/35121

KEYWORDS

data mining; COVID-19; vaccine; text mining; change; coverage; communication; media; social media; news; outbreak; acceptance; hesitancy; understanding; knowledge; sentiment

XSL•FO
RenderX

Introduction

Theoretical models suggest that the herd immunity threshold for SARS-CoV-2 requires at least two-thirds of the population to be immunized through either natural infection or vaccination [1]. Though multiple safe and effective vaccines have been developed [2-4], one significant challenge in achieving pandemic control is “vaccine hesitancy,” which ranges from mistrust to outright refusal of vaccination [5].

Vaccine hesitancy extends beyond COVID-19 and is 1 of the 10 biggest threats to global health according to the World Health Organization (WHO). At its core, vaccine hesitancy is an issue of perception, rooted in the information individuals receive [6].

Social media is an important source of both vaccine information and misinformation. Although vaccine-related tweets are predominantly positively polarized [7], there is also substantial (possibly coordinated) misinformation [8] that contributes to vaccine hesitancy [9]. Further, the volume of tweeted fake news within a given country negatively correlates with its vaccine uptake [10]. Antivaccination supporters on Twitter share more conspiracy theories and make greater use of emotional language than provaccination supporters [11]. Moreover, vaccine discourse is highly politicized [12], and the likelihood of endorsing misinformation is ideologically driven [13,14].

Different sides of vaccine discourse prioritize different objective values: Arguments in favor of vaccines prioritize community, while arguments against vaccines focus on individual freedom [15]. A high proportion of parents' opinions on vaccines expressed online is aggressive, accusatory, or inaccurate [16].

Major news outlets also play an important role in vaccine discourse [17,18]. Although several text mining studies have covered vaccines within specific regions [19-22], to the best of our knowledge, there are no large-scale text mining studies to date of vaccine front-page news headlines that encompass multiple countries focusing specifically on COVID-19.

Here, we analyzed online news media coverage of COVID-19 vaccines. We used text mining analysis to estimate the volume of online vaccine news coverage during 3 time periods: (1) before the COVID-19 pandemic, (2) before the COVID-19 vaccine announcement, and (3) after the COVID-19 vaccine announcement. We used ~28 million front-page headlines collected from 11 different countries with a healthy online news media ecosystem, defined using SimilarWeb traffic and BBC media profiles [23]. Because sentiment toward vaccines is influenced by the context in which they are mentioned, the most frequently mentioned topics were gathered alongside the most frequently mentioned companies and organizations. Our analysis aimed to inform future public health and vaccine communication, with a view to hopefully reducing vaccine hesitancy.

Methods

Curation of a Front-page News Article Database

We analyzed the landing pages from major online news sources (ONSs) in countries with a healthy media ecosystem. The data are fully described in a previous study [23] that focused on front-page news from 172 leading ONSs in 11 countries (Australia, Canada, France, Germany, Ireland, Italy, New Zealand, Russia, Spain, the United Kingdom, the United States) and an international category. The international category contained headlines from ONSs that were internationally distributed (eg, EuroNews or AlJazeera). The data used articles published from July 2015 to April 2021, which covered the following 3 time periods: (1) before the outbreak of COVID-19, (2) during the pandemic before the COVID-19 vaccine announcement, and (3) during the pandemic after the COVID-19 vaccine announcement. We took November 2020 as the cutoff date for the COVID-19 vaccine announcement, as from this point on, the press started covering SARS-CoV-2 vaccines following the announcement by BioNTech and Pfizer. We note this date applies to western countries, which are the subject of our study, and is less applicable globally. The updated data set included a total of 28,709,060 headlines, from which 14,638,278 were in the English language and 14,070,782 were in a language other than English.

Identifying Vaccine Headlines

Keywords were used to identify whether a given headline was vaccine-related. For non-English headlines, keywords were supplied by native speakers. For English headlines, we supplied the keywords ourselves. The keywords used can be found in Table 1.

Non-English headlines were stemmed using SnowballStemmer [24] and case-folded (Table 1) to capture the equivalence class of different forms of words (eg, the German words Impfung, impfen, Impfgegner all map to impf). English headlines were lemmatized using TreeTagger [25], all words were case-folded, and punctuation was removed, whereby words connected by a hyphen were separated into 2 words. English headlines were lemmatized to avoid misclassifications (eg “immunity” understood in a legal rather than a biomedical sense).

The techniques used to identify vaccine headlines varied by language, and we used the same methodology as in our previous work [23]. In French, Italian, Russian, and Spanish, titles and descriptions were tokenized, and if either the title or the description contained at least one keyword, the headline was labeled as a vaccine headline. In English and German, titles and descriptions were kept as strings, and a search was performed for keyword patterns. If a keyword pattern was present, the headline was designated as a vaccine headline (eg, in German, the prefix Impf-). Machine learning translation offers an alternative way to identify vaccine headlines across languages; however, this was beyond the scope of this work.

XSL•FO
RenderX

Table 1. Keywords used to identify the vaccine headlines.

Language	Keywords
English	• vaccin • immunis • immuniz • anti vax • antivax
French	• vaccin • antivaccin • immunis
German	• impf
Italian	• vaccin • antivaccin • immunizz
Russian	• прививк • привива • вакцин • иммунизац • вакцинац
Spanish	• vacun • antivacun • inmuniz

Splitting the Data Into 3 Vaccination-Specific Periods

We divided the data into 3 time periods: (1) the pre-COVID-19 era, (2) during the pandemic before the COVID-19 vaccine announcement, and (3) during the pandemic after the COVID-19 vaccine announcement. This division of the data was based on clear changes within media coverage with respect to vaccines and COVID-19. On January 9, 2020, daily media coverage of the coronavirus began, so we chose this date as the end of the pre-COVID-19 era. We chose November 9, 2020, as the cut-off date separating the prevaccine and after-vaccine announcements. This resulted in the following 3 periods:

1. Before COVID-19: July 2015 to January 8, 2020
2. Before the COVID-19 vaccine announcement: January 9, 2020, to November 9, 2020
3. After the COVID-19 vaccine announcement: November 10, 2020, to April 2, 2021

To identify changes in each period, the relative frequency of vaccines mentioned in the full data set, along with the relative frequency of headlines containing either "COVID-19" or "coronavirus," was calculated at weekly intervals using equation 1.

where $|ONS_{Topic,Week}|$ is the number of headlines on a particular topic in a given week and $|ONS_{Week}|$ is the number of headlines in that same given week. The relative frequency was calculated first with respect to vaccines, where all vaccine-related headlines were included, and second with respect to COVID-19, where all headlines containing either the keyword "coronavirus" or "COVID-19" were included.

Topic Detection of the Vaccine Headlines in the 3 Periods Using BERTopic

Topics were identified for 91 English ONSs using BERTopic. Topics were not identified for the non-English ONSs, as finding the optimal number of topics within non-English ONSs would require languages to be handled separately and would also require in-depth knowledge about each language. BERTopic is a topic modelling technique that uses a combination of transformers and c-TF-IDF to create dense clusters using HDBSCAN, where c-TF-IDF is a class-based TF-IDF that can be used to generate features from text [26]. We chose to use BERTopic as it was previously successful in heterogeneous text mining [27,28] and it offers multiple pretrained models. Additionally, scatterplots of the embeddings of the data from the 3 periods did not show a clear clustering of the headlines, which rules out several other topic detection techniques (please see Figures S1-S3 in Multimedia Appendix 1).

To remove patterns from the text input to BERTopic that could otherwise affect the model, all abbreviations, links, and names referring to the different newspapers were removed. Additionally, the word "news" was removed, along with words containing "immuniz," "immunis," and "vaccin," which were used to extract the vaccine headlines. The phrases "anti vax" and "antivax" were retained, as they refer to resistance toward vaccination.

Text input to BERTopic was normalized to reduce word variation. The headlines were lemmatized using TreeTagger combined with case-folding. TreeTagger is a tool for annotating text with part-of-speech and lemma information using a Markow tagger, which uses a decision tree to obtain reliable estimates. TreeTagger was also used to remove filler words from headlines

XSL•FO
RenderX

by only using words tagged as either a noun (including proper nouns), verb, or adjective and removing words that contained little information about topics.

We employed a 2-step evaluation method to identify the number of clusters reflecting the most common topics (Section 1 in Multimedia Appendix 1). The pseudocode for this is illustrated in Figure 1. Evaluating topic similarity (step 2) was performed manually, as 2 topics might deal with the same subject but contain several seemingly different keywords or word combinations, which would make the model split them into 2 topics instead of 1 topic. Therefore, the decision of how to continue from step 2 was likewise done manually.

Figure 1. Pseudocode for the 2-step evaluation method to identify the number of clusters reflecting the most common topics.

```
Algorithm 1:

    Input: Normalized English headlines.

    Output: The largest clusters within the headlines.

1.  Initiate a BERTopic using selected parameters
    If Number of topics < 20 then
            Repeat step 1 using different parameters
    Else
            Continue to step 2
2.  For m, n ∈ the 20 largest clusters, m ≠ n then
            If m and n are too similar in their keywords and synonyms then
                    Repeat step 1 using different parameters
            Else
                    Continue to step 3
3.  For k, j ∈ the 50 largest clusters then
            If k and j are too similar in their keywords and synonyms then
                    Repeat step 1 using different parameters
            Else
                    Return The found topics
```

Named Entity Recognition of Vaccine Headlines Using SpaCy

Named entity recognition (NER) identifies and categorizes words (or strings of words) for an entity, where an entity can be the name of a person, organization, location, or work of art. We used NER to determine the companies and organizations that were mentioned frequently in the context of vaccination. NER was performed on both English and non-English data using SpaCy with different pipelines depending on the language. SpaCy is an advanced natural language processing tool that is able to perform NER on multiple different languages using statistical models. Therefore, it uses previous training and predictions to decide whether a word or collection of words is a named entity and which kind of entity it most likely is [29]. Pipelines were chosen according to the reported accuracy by SpaCy. In all cases, the most accurate pipeline was used, which were en_core_web_trf, de_core_news_lg, fr_core_news_lg, it_core_news_lg, ru_core_news_lg, and es_core_news_lg. The 2 first letters in each pipeline refer to the language for which it was trained.

Entities such as "AstraZeneca-Oxford" or "Pfizer-BioNTech" were split to count as separate entities. The occurrences of "Johnson and Johnson" and "J&J" were altered to "Johnson & Johnson."

Individual entities were enumerated using case-folded entities. We created 2 bar plots (see Multimedia Appendix 1), one containing the 30 most frequently occurring named entities from English ONSs and another containing the 30 most frequently named entities from non-English ONSs.

Frequent N-grams With Respect to the Different Vaccine Manufacturers

Changes in sentiment toward vaccination before and after the COVID-19 vaccine announcement were determined by assessing 7 frequently occurring vaccine manufacturers found using NER. A data set containing English headlines for each vaccine manufacturer was created, which was then assessed with respect to frequent bigrams and trigrams (referred to as n-grams henceforth). The lemmatized headlines created for the topic detection were used for this purpose.

For all vaccines and periods, the 50 most frequent n-grams were assessed. In some cases, a combination of 2 bigrams, with almost the same count as a trigram, would combine to give that trigram. For instance, the bigrams (*food, drug*) and (*drug, administr*) combined give the trigram (*food drug administr*). This was caused by "Food and Drug Administration" in some cases being referred to as "Food and Drug Authority" or "Food and Drug Association." Such bigrams were removed, keeping only the trigrams. Similar bigrams were excluded for "Food and Drug Administration," "Centers for Disease Control," and "European Medicines Agency." Additionally, "FDA," "CDC," "NIH," "WHO," and "EMA" were commonly occurring abbreviations among the frequent words with respect to some vaccines, which were added to the number of occurrences of "Food and Drug Administration," "Center for Disease Control," "National Institute of Health," "World Health Organization," and "European Medicines Authority," respectively. Other abbreviations such as "NHS," "HHS," and "PHE" were assessed with respect to frequent bigrams and trigrams. Likewise, if bigrams occurred the same number of times as a trigram containing the bigram, the bigram was removed.

XSL•FO
RenderX

Sentiment Analysis of the Vaccine Headlines of 3 Periods Using VADER

We performed sentiment analysis on English-language headlines using VADER [30]. Before assessing sentiment values, each headline's raw score was calculated using the positive and negative sentiment values in equation 2:

$$Raw_{score}=Positive_{score} - Negative_{score} \quad \textbf{(2)}$$

The extent of negative or positive sentiment polarization varied between ONSs and over time. Therefore, a comparison of sentiment toward vaccines between the periods and ONSs on the raw sentiment values would not show whether a change in sentiment toward vaccines was due to an overall change in sentiment or, instead, due to a change in sentiment specifically toward vaccines. Therefore, to enable comparison of the periods and between the ONSs, each sentiment value for a vaccine headline was adjusted according to the overall average sentiment in the given ONS. The adjustment was done using the VADER sentiment values (either raw or compound, denoted by $S_{ONS,Topic,Period}$), subtracting the mean sentiment value for the same ONS, with respect to nonvaccine headlines in the same period (either raw or compound, denoted by).

This is referred to as the relative sentiment skew (RSS) and is given in equation 3:

where $ONS_{Topic,Period}$ is the collection of headlines of a given topic for a given ONS in a specific period, is the collection of headlines not pertaining to that topic for that same ONS in all periods, h is a single headline, and $sent(h)$ is the sentiment value of h, while is the number of headlines not in the given topic for that same ONS in all periods. In this case, the topic in equation 3 is vaccines. The raw scores were used to RSS each headline, with respect to the 3 periods. These were illustrated in line plots, in which the cumulative frequency showed the proportion of negative and positive RSS values of a certain smaller value. Because of the nuanced nature of the news, we applied the same manual checks here as in our previous work to make sure sentiment annotations were correct [23].

Results

Of the 14,638,278 English-language headlines identified over all 3 data periods, 83,395 (0.6%) were found to be vaccine-related using the keywords defined in Table 1. Dividing these with respect to the 3 periods gave the following number of vaccine headlines within each period: (1) before COVID-19: 11,361; (2) before the COVID-19 vaccine announcement: 17,112; (3) after the COVID-19 vaccine announcement: 54,922.

Large Increase in Ratio of Vaccine Headlines With the Rollout of COVID-19 Vaccines

We calculated the percentage of vaccine coverage within newspaper headlines for each week within each time period of data collection, plotted in Figure 2. Before the pandemic, the percentage of vaccine headlines was low (0.1% across 172 ONSs). With the COVID-19 outbreak in early 2020, the proportion of vaccine headlines increased to an average of 4%.

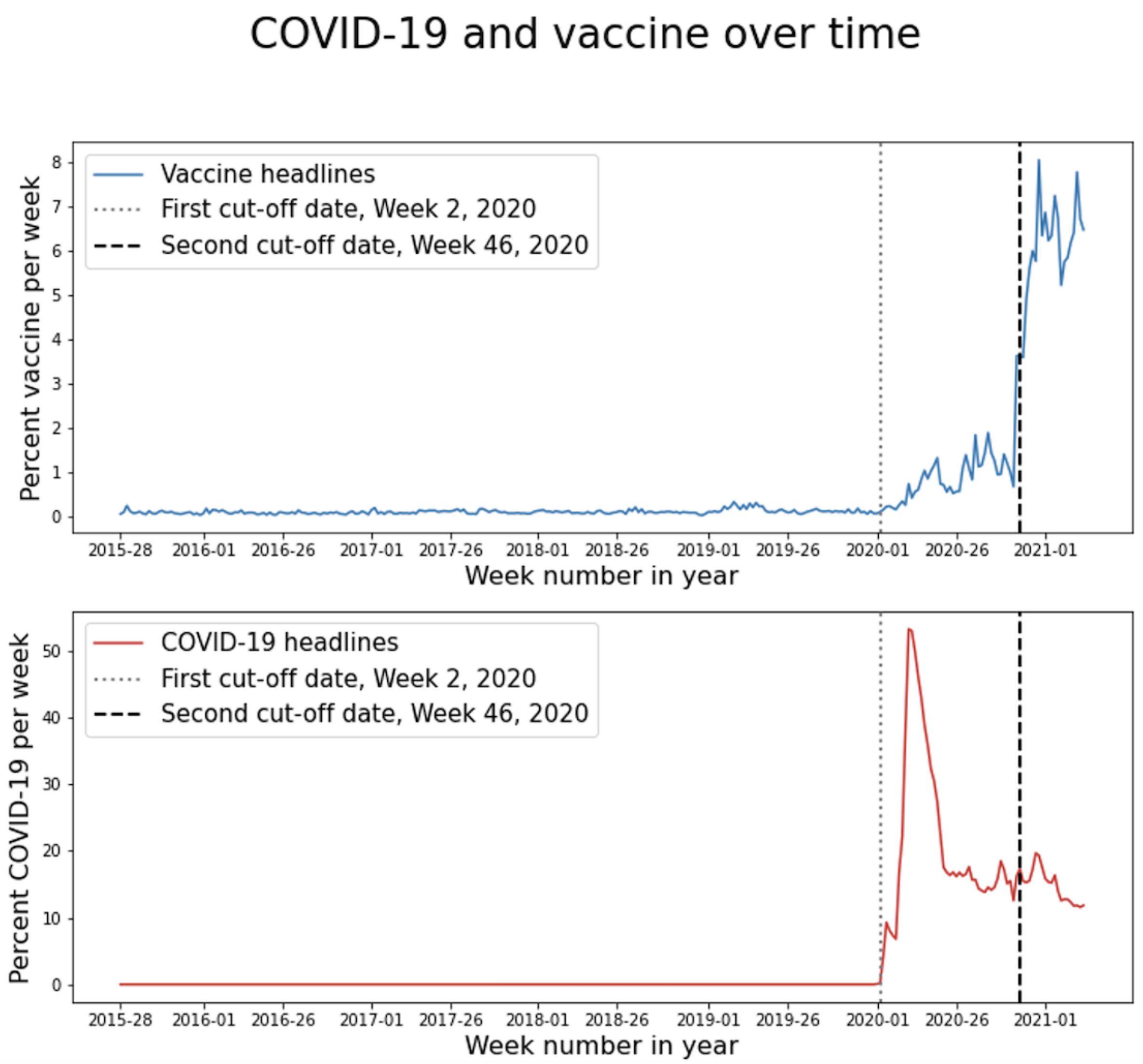
Increased reporting on vaccines during the second period coincided with the advent of COVID-19 reporting. The 10 most common topics in vaccine coverage in the 3 periods are shown in Figure 3. Causal connections cannot be established, as the COVID-19 coverage reached one-quarter of all front-page coverage with nuanced associations with reported topics [23]. Unsurprisingly, the most common vaccine-related topics during the second and third time periods were related to the pandemic. Although COVID-19 increased vaccine news coverage, coverage of COVID-19 was not directly correlated with that of vaccine coverage (Figure 2).

Rather than dropping to a stable level, as COVID-19 headlines did (Figure 2), the proportion of vaccine headlines increased from week 45 to week 47 of 2020 to between 6% and 8% and remained at this level until April 2, 2021. This increase is linked to the Pfizer and BioNTech press release on November 9, 2020, which reported 90% effectiveness in preventing COVID-19, paving the way for the rollout in the United Kingdom beginning on December 2, 2020.

Relative frequencies of vaccine headlines were calculated for each period and each country (Figure 4). Relative frequencies for each country were similar, with very limited attention toward vaccines before the pandemic and a steep rise after the introduction of the first SARS-CoV-2 vaccine.

XSL•FO
RenderX

Figure 2. Percentage of headlines mentioning (A) vaccines and (B) "COVID-19" or "coronavirus" in the mainstream media over time, with the first and second cut-off dates (dotted and dashed vertical lines, respectively).

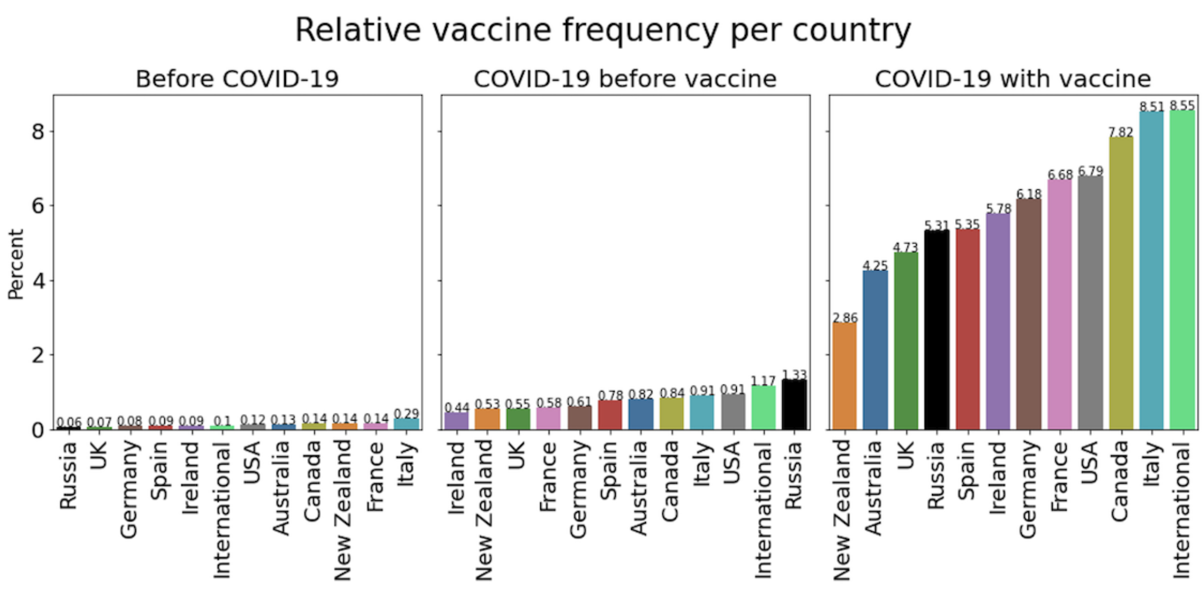


XSL•FO
RenderX

Figure 3. The 10 most common topics within vaccine-related articles during the 3 time periods. Purple cells highlight topics directly related to COVID-19, while red cells highlight topics that occur during more than one period. Notice that "Russia and COVID-19" is colored purple even though it occurs in multiple periods. EU: European Union; HPV: human papillomavirus.

	Before COVID-19	COVID-19 before vaccine	COVID-19 with vaccine
Topic 1	Influenza, 11.0%	Donald Trump, 9.3%	COVID-19, 3.7%
Topic 2	Measles, 7.1%	Vaccine development, 8.8%	Economic, 2.6%
Topic 3	Diseases mainly found in tropical climate, 5.3%	Oxford and AstraZeneca, 5.0%	Donald Trump, 2.2%
Topic 4	Unvaccinated children, 4.2%	Chinese hackers sought to steal vaccine research, 4.9%	Brexit and export ban from EU, 1.9%
Topic 5	Meningitis, 3.6%	Russia and COVID-19, 3.8%	Reject vaccine, 1.7%
Topic 6	Social media and misinformation, 3.5%	Economic, 2.9%	Health care, 1.5%
Topic 7	Ebola/Congo, 3.1%	Treating COVID-19, 2.5%	Vaccine rollout COVID-19, 1.3%
Topic 8	Antivaccination and new vaccine law California, 2.7%	Influenza, 2.4%	Russia and COVID-19, 1.2%
Topic 9	HPV, 2.2%	Sick children, 1.6%	Reopening of schools, 1.2%
Topic 10	Measles outbreak Brooklyn, 2.1%	Social media and misinformation, 1.2%	Vaccine production India, 1.2%

Figure 4. Relative vaccine frequency for each country including the international online news sources for each of the 3 periods: (A) before COVID-19, (B) before the COVID-19 vaccine announcement, and (C) after the COVID-19 vaccine announcement.



Majority of Vaccine Reporting Had Positive Sentiment Polarization With the Outbreak of COVID-19 as Opposed to the Prepandemic Era

Figure 5 shows the VADER sentiment scores for vaccine-associated headlines within each time period. The increased frequency of vaccine reporting during the pandemic led to an increase in the absolute number of negatively polarized articles, from 6698 in 2015-2019 to 28,552 in 2020-2021. Overall, however, polarization during the pandemic was majority positive (38% negatively polarized) as opposed to the prepandemic period, when 57% of articles were negatively polarized. Figure 3 suggests that the difference in sentiment between pre-COVID-19 and post-COVID-19 vaccine coverage could be associated with COVID-19 coverage. This could be because COVID-19 became the dominant topic globally, accounting for one-quarter of all news during the pandemic.

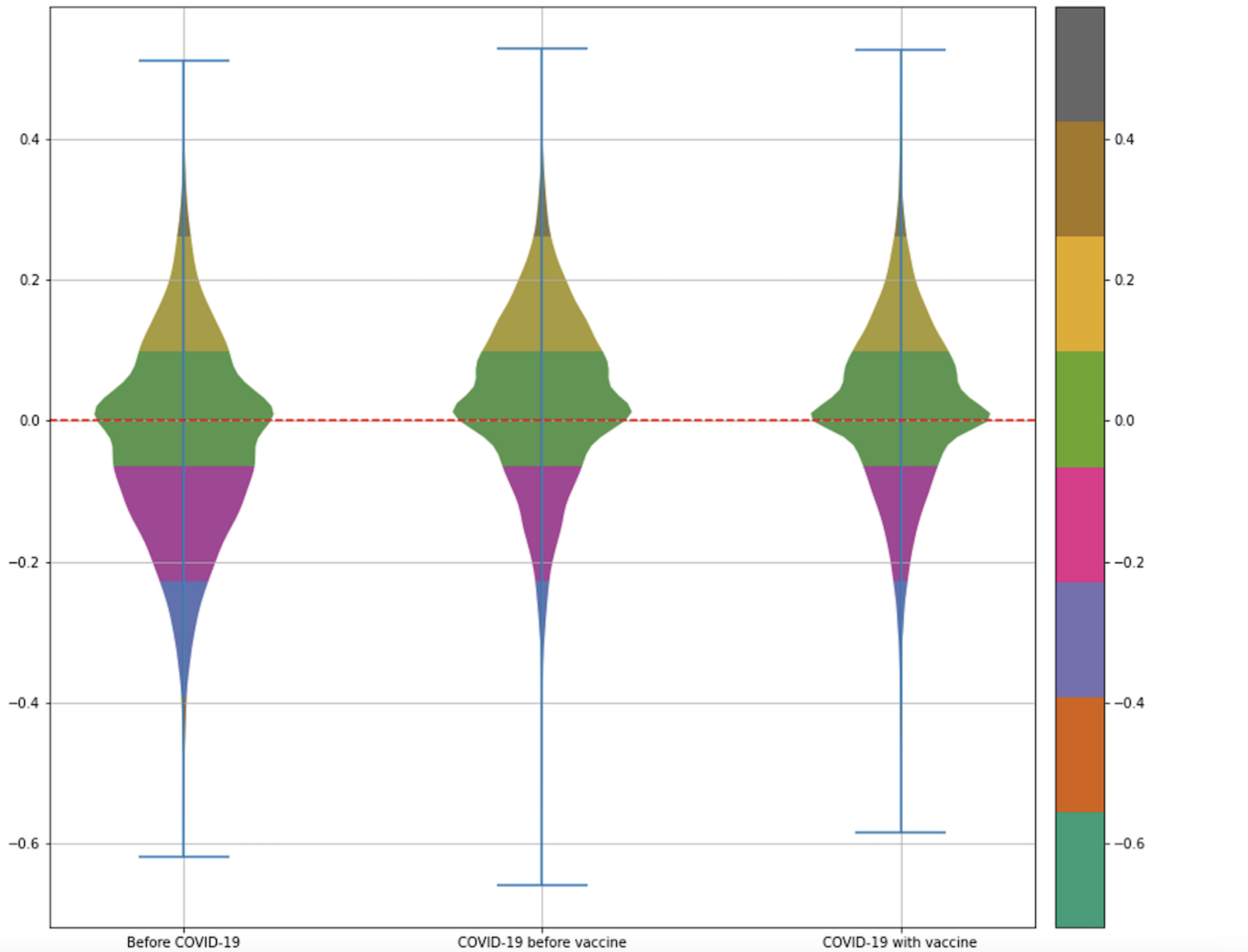
XSL•FO
RenderX

To investigate the difference in sentiment distribution between the 2 periods during the pandemic, we contrasted the topics and named entities mentioned in both periods. The period "Before the COVID-19 vaccine announcement" can largely be interpreted as the period in which all vaccines were under development, while "After the COVID-19 vaccine announcement" is the period in which some vaccines were rolled out and others were still under development. Although there is a difference between the periods before COVID-19 and after COVID-19, there was not a sizable sentiment discrepancy between the 2 periods during the pandemic (Figure 5).

We further investigated the topic polarization of the articles relating to the COVID-19 vaccine development and rollout. We extracted articles associated with 2 topics from Figure 3: "Vaccine development" and "Vaccine rollout." One could argue that "Vaccine production" (topic 10) should be merged with "Vaccine rollout" in line with our interpretation of the periods. However, we wanted to avoid manual intervention in topic annotations. The individual articles were extracted from the data giving 2 data sets of approximately the same size (846 and 814 headlines, respectively).

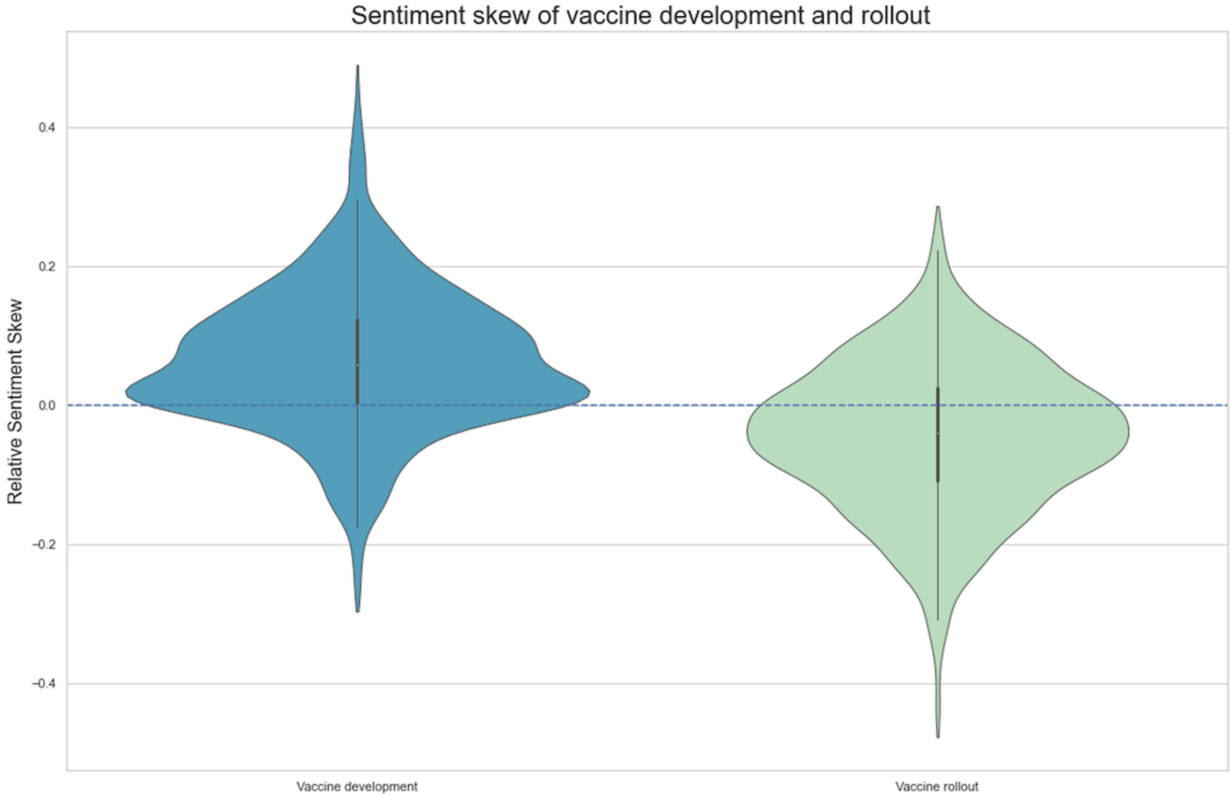
We assessed sentiment polarization of the topics "Vaccine development" and "Vaccine rollout." RSS of raw VADER sentiment for "Vaccine development" and "Vaccine rollout" is illustrated in Figure 6, which shows a change in vaccine sentiment between the development and trial phase and the rollout of the vaccines. Figure 6 illustrates that, for "Vaccine development," sentiment is overwhelmingly positive, with almost the entire interquartile range above the zero line. Of the headlines in "Vaccine development," 23% had negative RSS, while 77% had positive RSS. This is very different from "Vaccine rollout," for which 66% had negative RSS and only 34% had positive RSS. Additionally, the widest area lies above zero for "Vaccine development" and below zero for "Vaccine rollout." Therefore, the RSS with the highest frequency is positive for "Vaccine development" and negative for "Vaccine rollout." The largest and smallest RSS for the 2 topics are quite different: "Vaccine Development" lies in the range from –0.3 to just below 0.5, while "Vaccine rollout" lies in the range from –0.5 to 0.3; so, their RSS values are equally spread, but their ranges are differently situated. This suggests that the difference in sentiment distributions between the 2 COVID-19 periods could be attributed to more negative coverage during vaccine rollout.

Figure 5. Relative sentiment skew (y axis) of vaccine coverage in the 3 periods used in this study.



XSL•FO
RenderX

Figure 6. Relative sentiment skew for the topics "Vaccine development" and "Vaccine rollout" using the raw sentiment value.



Most Common Organizations Mentioned in the Context of COVID-19 Vaccines and Sentiment Toward Them

To gain more granular insight into sentiment polarization during the pandemic period, we investigated the top entities mentioned. We employed SpaCy to perform NER, and the 30 most frequently mentioned companies or organizations for all 3 periods are illustrated in Figure 7.

Unsurprisingly, the most common associations were between well-known COVID-19 vaccine manufacturers, namely "AstraZeneca" (in collaboration with Oxford), "Pfizer" (in collaboration with BioNTech), "BioNTech," "Moderna," "Oxford,", "Johnson & Johnson," and "Sputnik V." Though AstraZeneca and Oxford, as well as Pfizer and BioNTech, developed their vaccines as a partnership, they were frequently mentioned separately; thus, we opted to keep them as separate entities.

Of the 30 most frequent named entities, in both English and non-English headlines, 16 occurred in both data sets, colored green in Figure 7. The nonoverlapping entities were mainly attributed to national organizations or companies. For instance, "NHS" and "HHS" are the National Health Service and the Department of Health and Human Services from the United Kingdom and United States, respectively, and were solely found among the 30 most frequent English entities. "Rospotrebnadzor" is the Federal Service for Surveillance on Consumer Rights in Russia, and "RDIF" and "PAH" are also Russian and were found solely among the 30 most frequent non-English entities. Additionally, company names are the same across different languages, whereas some national organizations are not; for instance, the abbreviation for the World Health Organization is WHO in English, while in French, it is OMS.

The frequency at which vaccine manufacturers were mentioned within all news headlines increased from almost zero before COVID-19 to most frequently mentioned within the period after the vaccine announcement (Table 2). Therefore, vaccine manufacturers were assessed only within the COVID-19 pandemic.

The most common associations with vaccine manufacturers indicated progress in development and rollout and were health-related (eg, side effects). Detailed analysis of the n-grams for each vaccine developer are in Section 2 of Multimedia Appendix 1. Vaccines by Moderna and Pfizer were chiefly associated with n-grams indicating progress of clinical trials and their rollouts. By contrast, top n-grams associated with AstraZeneca and Johnson & Johnson were linked to side effect reporting (eg, unexplained illness, blood clot). Throughout the pandemic, Sputnik V was mentioned not in a medical context but rather frequently linked to Russia and Vladimir Putin, containing frequent n-grams like "Soviet Union," "President Vladimir Putin," and "Russia Soviet Union."

We investigated the extent to which the difference in the context of vaccine manufacturers influenced news article sentiment. In Figure 8, we plotted the proportion of negative and positive sentiments toward the vaccine manufacturer entities before and after the vaccine announcement. In the period before the COVID-19 vaccine announcement, entities appear to have similar negative polarizations, AstraZeneca and Johnson & Johnson being noted as slight outliers with more negative coverage. After the COVID-19 vaccine announcement,

XSL•FO
RenderX

AstraZeneca had a notably higher ratio of negative articles and a lower ratio of positive articles. Despite Johnson & Johnson being associated with side effects (as per our n-gram analysis), AstraZeneca received notably worse press. We removed AstraZeneca coverage from Figure 5 and Figure 6 to test whether the higher associated volume of negative news influenced the slightly more negative polarization in the phase after the COVID-19 vaccine announcement. In both cases, we did not find that AstraZeneca was the main driver in more negatively polarized articles in that period (please see Tables S1 and S2 in Multimedia Appendix 1).

Figure 7. The 30 most frequent entities (companies and organizations) found in the (A) English and (B) non-English data. The green names are the organizations and companies that were found in both English and non-English data.

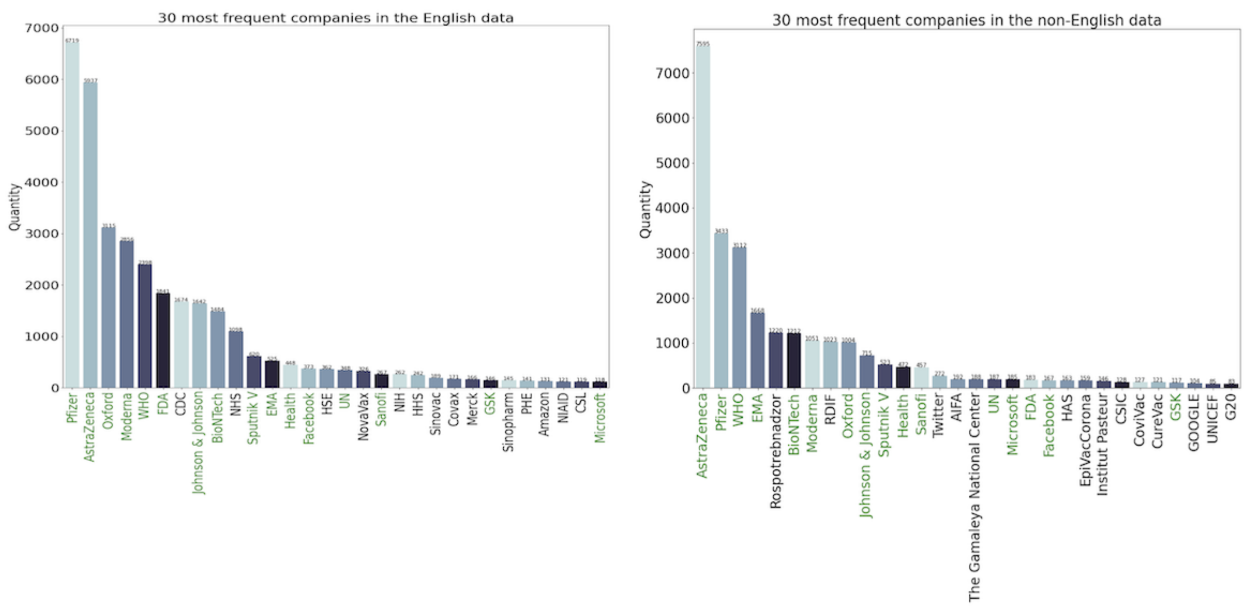
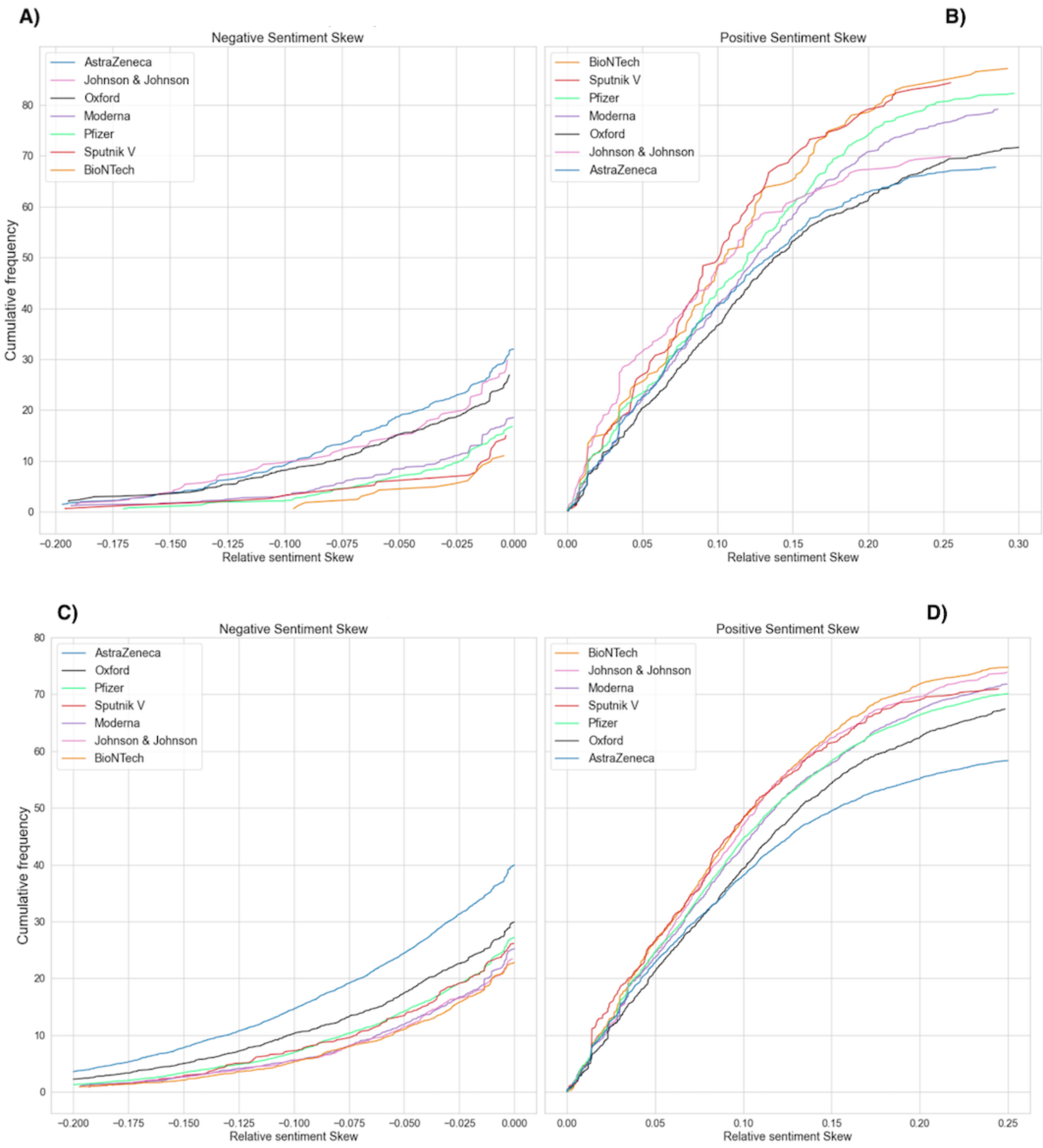


Table 2. The 21 different subsets created with respect to the different vaccines and periods, including the number of times each of the different manufacturers were mentioned within the news headlines in each subset (7 vaccine manufacturers in 3 periods).

Manufacturer	Before COVID-19	Before the COVID-19 vaccine announcement	After the COVID-19 vaccine announcement
AstraZeneca	3	747	5134
BioNTech	1	163	2118
Johnson & Johnson	17	332	1050
Moderna	3	647	2256
Oxford	3	1010	2288
Pfizer	27	513	6042
Sputnik V	0	153	700

XSL•FO
RenderX

Figure 8. Proportion of negative and positive sentiment polarization with respect to entities associated with vaccine manufacturing in the periods “Before the COVID-19 vaccine announcement” and “After the COVID-19 vaccine announcement”: (A) negative sentiment skew "Before the COVID-19 vaccine announcement," (B) positive sentiment skew "Before the COVID-19 vaccine announcement," (C) negative sentiment skew "After the COVID-19 vaccine announcement," (D) positive sentiment skew "Before the COVID-19 vaccine announcement".



Discussion

We used text mining to study vaccine reporting on the front pages of top national news outlets. We demonstrated that reporting on vaccines increased in volume from coverage of around 0.1% on front pages to almost 4% of all headlines during the pandemic. Despite reporting covering the vaccines’ side effects, overall coverage can be classified as positive, in line with previous studies of social media that reported positive polarization of vaccine-related tweets [7].

The news ecosystem accounts for 76% of the information people consume [31], which can affect people’s behavior, for instance making them more hesitant toward vaccines. This can be exacerbated by circulation of misinformation [21] and by vaccine reporting along partisan lines [19].

However, news is only one facet of the entire media ecosystem, and much information is communicated via social media [19-22]. Social media encourages active participation in the form of clicks, likes, retweets, and shares, which are then readily quantifiable by engagement. With news however, the

XSL•FO
RenderX

engagement is much more nuanced, especially because of more passive information consumption when people merely scan headlines. Nonetheless, traditional news is still vital in forming opinions and, in many cases, constitutes the initial discourse on other platforms.

We focused on analyzing headlines from a handful of western countries to provide a data-centric analysis of vaccine coverage across several countries. Similar studies have been conducted in individual countries (eg, Brazil [22]) or other regions (eg, Africa [20]). Our study encompasses countries that were among the first to manufacture and introduce the vaccine on a large scale (United States, Russia, Germany, United Kingdom). In these countries, policy makers had to navigate vaccine hesitancy and ongoing COVID-19 restrictions with sophisticated media coverage throughout the development and rollout phases.

We analyzed how front-page headline vaccine reporting evolved during the COVID-19 pandemic. For the analysis, we made a set of assumptions that are associated with certain limitations. Our focus on the headlines in predominantly developed western countries underrepresents the situation faced in other parts of the world that were also affected by COVID-19, where vaccine hesitancy is compounded by inequality in vaccine manufacturing and distribution [32,33]. We justify using headline information by virtue of normalizing heterogeneous long-form texts across different news sites and by capturing the behavior of passive scanning of headlines. However, this introduces a disconnect between the information in the full article that might not be reflected in an attention-attracting headline and thus leads to different information consumption by the reader. Within our data set, we opted for a keyword-based approach that was previously used to measure the extent of COVID-19 reporting [23]. The approach is designed to increase the precision of identified headlines, though at the expense of recall. For instance, the headline "UK measles outbreak: 500,000 British children don't have crucial jab - Daily Star. MORE than half a million children in the UK didn't receive a…" was not extracted for the English vaccine data set, as it does not contain any of the chosen key words given in Table 1, even though it clearly pertains to vaccination. Developing a more complex topic model would not guarantee better performance and comparability between different languages, as one would have to develop a suitable model that captures the same linguistic nuances. Therefore, we resorted to simple mentions of basic vaccine-derived keywords to aid comparison across countries.

Even though this approach underestimates the number of vaccine-related articles, COVID-19 vaccine reporting was still given central prominence, unlike before the outbreak when vaccines were covered only sporadically. Studying the volume of vaccine coverage motivated our division of the data into the 3 periods, before COVID-19, during COVID-19 but before the vaccines, and with COVID-19 vaccines. It is possible that our definitions of the second and third periods could have influenced our results. However, we found it reasonable to make these divisions according to the large rise in the relative frequency in vaccine headlines due to the Pfizer and BioNTech press release on November 9, 2020. This press release influenced all countries, while many of the other cornerstones in this period were more country-specific. For instance, the United Kingdom was the first country to approve the Pfizer-BioNTech vaccine on December 2, 2020, with the United States Food and Drug Administration approval of the Pfizer-BioNTech vaccine occurring on December 11, 2020.

Our topic modelling and sentiment analysis showed that COVID-19 increased the proportion of vaccine headlines by more than an order of magnitude, from a negligible 0.1% to a formidable 4% during vaccine rollout across 172 ONSs. Reporting on vaccines prior to COVID-19 was negatively polarized. By contrast, vaccine-related reporting during the pandemic is positively polarized. Though we note a discrepancy in sentiment polarization pre- and post-COVID-19, this could be attributed to sampling bias post-COVID-19, as there was significantly more vaccine coverage. Moreover, sentiment polarization in the headlines might not relate directly to vaccines but rather to tangential topics. We therefore also analyzed the tendences in sentiments relating to specific concepts or entities, such as vaccine development or vaccine manufacturers.

We performed in-depth sentiment analysis of the subtopic AstraZeneca, which received more negative coverage because of widely reported side effects and delivery issues. According to our analysis, however, such negative reporting was not significant enough to alter the overall positive narrative of vaccines in the news. Although The University of Oxford co-created the vaccine, it does not experience an equally large proportion of negative headlines as does AstraZeneca, which might be reflected in the media coverage frequency of the 2 with respect to vaccines. Although AstraZeneca is mentioned 5881 times during the pandemic, Oxford is mentioned 3298 times, mostly in the period before the COVID-19 vaccine announcement, while for AstraZeneca the majority is in the subsequent period. Therefore, AstraZeneca is more frequently connected with the vaccine in the media coverage than Oxford.

Our findings study the online news media's vaccine coverage and are also applicable more widely to general mistrust of authority and science. Although direct connections between news coverage and vaccine uptake are beyond the scope of this study, we have comprehensively characterized sentiment toward COVID-19 vaccination in the online news media. Future survey-based studies into vaccine hesitancy will hopefully benefit from our work, as it details the changing information landscape on which the public ultimately base their decisions. Our work is therefore also important for public health policy makers who require knowledge of the information that the public consumes when designing vaccine mandates.

Acknowledgments

DL and SB acknowledge joint center funding from the UK Medical Research Council and Department for International Development (grant MR/R015600/1). DL acknowledges funding from the Vaccine Efficacy Evaluation for Priority Emerging Diseases (VEEPED)

XSL•FO
RenderX

grant (ref. NIHR:PR-OD-1017-20002) from the National Institute for Health Research (NIHR). SB acknowledges funding from The Novo Nordisk Young Investigator Award (NNF20OC0059309), The Danish National Research Foundation via a chair position, and the NIHR Health Protection Research Unit in Modelling Methodology.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary information.
[PDF File (Adobe PDF File), 1476 KB - infodemiology_v2i2e35121_app1.pdf]

References

1. Neagu M. The bumpy road to achieve herd immunity in COVID-19. J Immunoassay Immunochem 2020 Nov 01;41(6):928-945 [FREE Full text] [doi: 10.1080/15321819.2020.1833919] [Medline: 33086932]
2. Voysey M, Clemens SAC, Madhi SA, Weckx LY, Folegatti PM, Aley PK, Oxford COVID Vaccine Trial Group. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet 2021 Jan 09;397(10269):99-111 [FREE Full text] [doi: 10.1016/S0140-6736(20)32661-1] [Medline: 33306989]
3. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, 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]
4. 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]
5. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger JA. Vaccine hesitancy: an overview. Hum Vaccin Immunother 2013 Aug 27;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
6. Fridman A, Gershon R, Gneezy A. COVID-19 and vaccine hesitancy: A longitudinal study. PLoS One 2021 Apr 16;16(4):e0250123 [FREE Full text] [doi: 10.1371/journal.pone.0250123] [Medline: 33861765]
7. 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]
8. Jemielniak D, Krempovych Y. An analysis of AstraZeneca COVID-19 vaccine misinformation and fear mongering on Twitter. Public Health 2021 Nov;200:4-6 [FREE Full text] [doi: 10.1016/j.puhe.2021.08.019] [Medline: 34628307]
9. Raciborski F, Jankowski M, Gujski M, Pinkas J, Samel-Kowalik P. Changes in attitudes towards the COVID-19 vaccine and the willingness to get vaccinated among adults in Poland: analysis of serial, cross-sectional, representative surveys, January-April 2021. Vaccines (Basel) 2021 Jul 29;9(8):A [FREE Full text] [doi: 10.3390/vaccines9080832] [Medline: 34451957]
10. Lyu H, Zheng Z, Luo J. Both rates of fake news and fact-based news on Twitter negatively correlate with the state-level COVID-19 vaccine uptake. arXiv. Preprint posted online February 17, 2022 2020 [FREE Full text]
11. 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]
12. Righetti N. The impact of the politicization of health on online misinformation and quality information on vaccines. Italian Sociological Review 2021;11(2):1 [FREE Full text] [doi: 10.13136/isr.v11i2.448]
13. Motta M. Republicans, not Democrats, are more likely to endorse anti-vaccine misinformation. American Politics Research 2021 Jun 14;49(5):428-438 [FREE Full text] [doi: 10.1177/1532673x211022639]
14. Baines A, Ittefaq M, Abwao M. #Scamdemic, #Plandemic, or #Scaredemic: what Parler social media platform tells us about COVID-19 vaccine. Vaccines (Basel) 2021 Apr 22;9(5):421 [FREE Full text] [doi: 10.3390/vaccines9050421] [Medline: 33922343]
15. Kalimeri KG, Beiró MG, Urbinati A, Bonanomi A, Rosina A. Human values and attitudes towards vaccination in social media. 2019 Presented at: WWW '19: The Web Conference; May 13-17, 2019; San Francisco, CA. [doi: 10.1145/3308560.3316489]
16. Jenkins MC, Moreno MA. Vaccination discussion among parents on social media: a content analysis of comments on parenting blogs. J Health Commun 2020 Mar 03;25(3):232-242. [doi: 10.1080/10810730.2020.1737761] [Medline: 32154770]
17. Chadwick A, Kaiser J, Vaccari C, Freeman D, Lambe S, Loe B, et al. Online social endorsement and Covid-19 vaccine hesitancy in the United Kingdom. Social Media + Society 2021 Apr 05;7(2):205630512110088 [FREE Full text] [doi: 10.1177/20563051211008817]
18. Harrison M, Lancaster K, Rhodes T. “A matter of time”: Evidence-making temporalities of vaccine development in the COVID-19 media landscape. Time & Society 2021 Jul 15;31(1):132-154 [FREE Full text] [doi: 10.1177/0961463x211032201]

XSL•FO
RenderX

19. Motta M, Stecula D. The influence of partisan media in the face of global pandemic: how news media influenced COVID-19 vaccine hesitancy. SocArXiv. Preprint posted online September 16, 2021 2021. [doi: 10.31235/osf.io/xj4nq]
20. Gbashi S, Adebo OA, Doorsamy W, Njobeh PB. Systematic delineation of media polarity on COVID-19 vaccines in Africa: computational linguistic modeling study. JMIR Med Inform 2021 Mar 16;9(3):e22916 [FREE Full text] [doi: 10.2196/22916] [Medline: 33667172]
21. Stecula D, Pickup M. How populism and conservative media fuel conspiracy beliefs about COVID-19 and what it means for COVID-19 behaviors. Research & Politics 2021 Feb 15;8(1):205316802199397 [FREE Full text] [doi: 10.1177/2053168021993979]
22. 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 Feb 10;7(2):e24585 [FREE Full text] [doi: 10.2196/24585] [Medline: 33480853]
23. Krawczyk K, Chelkowski T, Laydon DJ, Mishra S, Xifara D, Gibert B, et al. Quantifying online news media coverage of the COVID-19 pandemic: text mining study and resource. J Med Internet Res 2021 Jun 02;23(6):e28253 [FREE Full text] [doi: 10.2196/28253] [Medline: 33900934]
24. Wagner W. Steven Bird, Ewan Klein and Edward Loper: natural language processing with Python, analyzing text with the Natural Language Toolkit. Lang Resources & Evaluation 2010 May 27;44(4):421-424 [FREE Full text] [doi: 10.1007/s10579-010-9124-x]
25. Amri S, Zenkouar L. Amazigh POS Tagging Using TreeTagger: A Language Independant Model. In: Ezziyyani M, editor. Advanced Intelligent Systems for Sustainable Development (AI2SD'2018). AI2SD 2018. Advances in Intelligent Systems and Computing, vol 915. Cham, Switzerland: Springer International Publishing; 2019:622-632.
26. Grootendorst M. Bertopic: Leveraging bert and c-tf-idf to create easily interpretable topics. Version v0.7 2021 Apr 01:4. [doi: 10.1007/springerreference_179544]
27. Ebeling R, Córdova Sáenz CA, Nobre J, Becker K. The effect of political polarization on social distance stances in the Brazilian COVID-19 scenario. JIDM 2021 Aug 05;12(1):18 [FREE Full text] [doi: 10.5753/jidm.2021.1889]
28. Valensise C, Cinelli M, Nadini M, Galeazzi A, Peruzzi A, Etta G, et al. Lack of evidence for correlation between COVID-19 infodemic and vaccine acceptance. arXiv. Preprint posted online September 14, 2021 2020 [FREE Full text] [doi: 10.31219/osf.io/sxd5t]
29. Honnibal M, Montani I, Van LS, Boyd A. spaCy: Industrial-strength natural language processing in python. Zenodo. URL: https://spacy.io/ [accessed 2022-08-20]
30. Gilbert C, Hutto E. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. ERIC GILBERT. 2014. URL: http://eegilbert.org/papers/icwsm14.vader.hutto.pdf [accessed 2022-08-20]
31. Flaxman S, Goel S, Rao J. Filter bubbles, echo chambers, and online news consumption. PUBOPQ 2016 Mar 22;80(S1):298-320 [FREE Full text] [doi: 10.1093/poq/nfw006]
32. Mullard A. How COVID vaccines are being divvied up around the world. Nature. 2020 Nov 30. URL: https://www.nature.com/articles/d41586-020-03370-6 [accessed 2022-08-20]
33. Cioffi A, Cioffi F. COVID-19 vaccine: Risk of inequality and failure of public health strategies. Ethics Med Public Health 2021 Jun;17:100653 [FREE Full text] [doi: 10.1016/j.jemep.2021.100653] [Medline: 33758775]

Abbreviations

NER: named entity recognition
NIHR: National Institute for Health Research
ONS: online news source
RSS: relative sentiment skew
VEEPED: Vaccine Efficacy Evaluation for Priority Emerging Diseases
WHO: World Health Organization

Edited by T Mackey; submitted 22.11.21; peer-reviewed by C Williams, J Luo; comments to author 30.12.21; revised version received 25.02.22; accepted 19.07.22; published 20.09.22.

Please cite as:
Christensen B, Laydon D, Chelkowski T, Jemielniak D, Vollmer M, Bhatt S, Krawczyk K
Quantifying Changes in Vaccine Coverage in Mainstream Media as a Result of the COVID-19 Outbreak: Text Mining Study
JMIR Infodemiology 2022;2(2):e35121
URL: https://infodemiology.jmir.org/2022/2/e35121
doi:10.2196/35121
PMID:36348981

XSL•FO
RenderX

©Bente Christensen, Daniel Laydon, Tadeusz Chelkowski, Dariusz Jemielniak, Michaela Vollmer, Samir Bhatt, Konrad Krawczyk. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 20.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Asymmetric Influence of Emotion in the Sharing of COVID-19 Science on Social Media: Observational Study

Kai Luo[1], PhD; Yang Yang[2], PhD; Hock Hai Teo[1], PhD

[1]National University of Singapore, Singapore, Singapore

[2]University of Warwick, Coventry, United Kingdom

Corresponding Author:
Hock Hai Teo, PhD
National University of Singapore
15 Computing Drive
Singapore, 117418
Singapore
Phone: 65 65162979
Email: teohh@comp.nus.edu.sg

Abstract

Background: Unlike past pandemics, COVID-19 is different to the extent that there is an unprecedented surge in both peer-reviewed and preprint research publications, and important scientific conversations about it are rampant on online social networks, even among laypeople. Clearly, this new phenomenon of scientific discourse is not well understood in that we do not know the diffusion patterns of peer-reviewed publications vis-à-vis preprints and what makes them viral.

Objective: This paper aimed to examine how the emotionality of messages about preprint and peer-reviewed publications shapes their diffusion through online social networks in order to inform health science communicators' and policy makers' decisions on how to promote reliable sharing of crucial pandemic science on social media.

Methods: We collected a large sample of Twitter discussions of early (January to May 2020) COVID-19 medical research outputs, which were tracked by Altmetric, in both preprint servers and peer-reviewed journals, and conducted statistical analyses to examine emotional valence, specific emotions, and the role of scientists as content creators in influencing the retweet rate.

Results: Our large-scale analyses (n=243,567) revealed that scientific publication tweets with positive emotions were transmitted faster than those with negative emotions, especially for messages about preprints. Our results also showed that scientists' participation in social media as content creators could accentuate the positive emotion effects on the sharing of peer-reviewed publications.

Conclusions: Clear communication of critical science is crucial in the nascent stage of a pandemic. By revealing the emotional dynamics in the social media sharing of COVID-19 scientific outputs, our study offers scientists and policy makers an avenue to shape the discussion and diffusion of emerging scientific publications through manipulation of the emotionality of tweets. Scientists could use emotional language to promote the diffusion of more reliable peer-reviewed articles, while avoiding using too much positive emotional language in social media messages about preprints if they think that it is too early to widely communicate the preprint (not peer reviewed) data to the public.

(JMIR Infodemiology 2022;2(2):e37331) doi:10.2196/37331

KEYWORDS

COVID-19; science communication; emotion; COVID-19 science; online social networks; computational social science; social media

Introduction

Background

The COVID-19 pandemic has led to an unparalleled surge in global research publications on a single topic in documented history [1]. Research publications on COVID-19 accounted for roughly 8% of all PubMed research outputs in 2020 [1]. Such an incredible surge was seen in not only traditional scientific sources (eg, journals) but also preprint servers [1,2]. This uptake in research output coincides with the active social media engagement of COVID-19 science from the public [3]. The urgency and immediacy of pandemic information needs had

XSL•FO
RenderX

promoted the use of social media for science communication among the public to an unprecedented level [1,4-7]. Understandably, the communication of COVID-19 science on social media is of critical importance because it could influence people's behaviors and affect the effectiveness of government measures [5]. However, given the variance of scientific publications in terms of quality and the instantaneity of information transmission on social media, it is imperative for policy makers and scientists to understand what drives the diffusion of research publications on social media.

Communication of science to the public has traditionally relied on professionals (eg, journalists, scientists, and public health authorities) to meticulously translate scientific findings for public consumption [2]. Even in this professionally moderated communication context, prior studies have found that the virality of professionally articulated messages was strongly influenced by how they were framed [8]. For example, framing cancer research with an appropriate emotion can increase the public's understanding, quality perception, trust, and engagement with the findings [9]. It is noteworthy that communication through social media, being both unmediated and spontaneous, provides a fertile ground that could augment the impact of emotion on content virality, especially during a crisis [10]. Indeed, recent studies in the COVID-19 context have shown that emotion-laded communication on social media could influence a wide-range of pandemic-related issues, such as vaccine communication, public health compliance, and preventive behavior [11-14]. Thus, we sought to investigate how the emotionality present in the text of social media messages about scientific publications on COVID-19 would influence their virality.

Theoretical Background

Text-based emotions refer to the presence of fine-grained emotions, such as happy, sad, and angry, in human languages [15]. Prior research has found that text-based emotions in the form of emotion words or emotional framing of messages could affect people's cognitive processing of the information in the context of written communication [16]. There have been 2 mainstream theoretical perspectives on emotions in prior studies [17]. One is the dimensional perspective that posits dimensions, such as valence and arousal, are the basic elements of emotions [18], and the other is the discrete perspective that considers discrete entities, such as happy, sad, anger, and fear, as the basic elements of emotions [19]. Prior literature has investigated the role of text-based emotions in online content sharing from different perspectives [20-23], and has provided competing theoretical explanations of how emotion influences content sharing. First, in social media engagement, people exhibit a social tendency to present a positive self-image for altruistic reasons (eg, to help others) or self-enhancement [24]. People are motivated to share things that make them look good or help signal their desired identities. Indeed, it is found that people are more likely to share positive scientific findings [8], positive New York Times articles [25], and positive marketing content [26,27]. Second, contrary to self-enhancement, there is also a "negativity bias" explanation [28,29]. It argued that, due to its evolutionary advantages, information involving negative emotions is generally found to be detected, processed, and transmitted faster than information involving positive emotions [20-23]. Content that aroused negative emotions was found to spread faster, especially in the domain of social media news, politics, and science conspiracy [30-33]. The third and perhaps most widely used theoretical explanation suggests that it is high-arousal emotions, whether of positive or negative valence, that contribute to online virality [34-37]. This perspective argues that beyond valence, emotions also differ in the level of psychological arousal or activation [38], and the psychological arousal and activation (or deactivation) of the emotion influence the transmissibility of the content [25].

Given the plurality of the emotional dynamics in social media sharing, we aimed to first establish which of the 3 theoretical explanations mentioned above is most likely true in the context of social media sharing of COVID-19 scientific research. Although self-enhancement motivation has been established in the context of the interpersonal sharing of professionally mediated science communication [8], the science behind the emerging phenomenon of sharing scientific findings about a novel infectious disease through large online social networks could be much more complex. On the one hand, the heightened situational uncertainty induced by the pandemic [39] could potentially lead to even stronger "negativity bias." Recent studies found a heightened prevalence of negative emotions or a negative emotional climate on social media during the early months of the pandemic [10,40]. On the other hand, findings from early COVID-19 scientific research were arguably important information sources of pandemic news. Taking COVID-19 preprints as an example, although news media largely refrained from citing findings from preprints in their reports before the pandemic, the use of COVID-19 preprints became the new norm during the pandemic [2], and they were used in news articles at a rate almost 100 times that of non–COVID-19 preprints [41]. Would this "news-like" status combined with heightened situational uncertainty lead to more salient negativity bias in the diffusion of social media messages of COVID-19 science or would the emotional dynamics be dominated by high-arousal emotions, regardless of positive or negative emotions? More importantly, do the sources of the messages (eg, preprint servers vs peer-reviewed journals) lead to different emotional dynamics in their diffusion?

Peer-reviewed journal publications and preprints differ in their scientific uncertainty in that there is a possibility that the results may be invalidated by subsequent studies [42,43]. Although all studies carry some degree of scientific uncertainty, it is arguably much higher in preprints. A rigorous peer review and editorial process can help scrutinize and mitigate scientific uncertainty in most journal publications, but such a process is absent in preprints. This has led to heated debates over the virtue and danger of the use of preprints in science communication to the public [44-46]. However, partly due to the rare use of preprints in science communication to the public, it remains unknown whether social media messages about preprints exhibit a different pattern of diffusion from that of peer-reviewed journal publications. Moreover, to mitigate the influence of scientific uncertainty in the communication of any research, past studies have emphasized the moderator role of scientists [43]. Scientists are considered as important moderators in the communication of science to the public. Their expertise could facilitate better

XSL•FO
RenderX

articulation on the significance and implication of scientific findings while clarifying the potential scientific uncertainty [43]. Yet, we have limited understanding of how the identity and emotions of scientists jointly influence the diffusion of social media messages of scientific research. Thus, we also investigated the extent to which scientist participation in the social media sharing of COVID-19 science influences the emotional dynamics.

Research Questions

To address the above gaps in our knowledge, we collected all Twitter discussions of nearly 10,000 early (January to May 2020) COVID-19 English research articles in the life science and biomedical fields in both peer-reviewed journals and preprint servers from Altmetric. Altmetric provides quantification of the attention received online for an individual research article. It is increasingly being used as a research metric for science evaluation [47]. Using these data, we sought to address the following research questions:

1. What aspect of emotion (ie, positive valence, negative valence, or arousal) best explains the emotional dynamics in the social media sharing of COVID-19 scientific outputs?
2. Do the emotional dynamics of sharing have similar or divergent patterns between messages of preprint and peer-reviewed journal publications?
3. What are the emotional dynamics associated with the role of scientists as social media message creators in the sharing of COVID-19 science?

Methods

Data

To answer our research questions, we collected data from several sources. First, we obtained COVID-19–related medical English peer-reviewed journal publications, published prior to mid-May 2020, from the MEDLINE database (accessed through PubMed), where we retrieved each publication's unique digital object identifier (DOI). We then used the PubMed application programming interface (API) to further retrieve each publication's detailed metadata (ie, journal, title, category, authors, abstract, etc). Second, we extracted the DOIs of preprint medical publications in the same period from bioRxiv and medRxiv. We further used the bioRxiv API to extract all detailed metadata of each preprint. At the time of data collection, there were 6552 articles available on MEDLINE and 3725 articles from bioRxiv and medRxiv together. Third, social media mentions of all articles from the MEDLINE database and preprint servers were collected from Altmetric, a London-based commercial company that tracks, analyzes, and collects the online activity around scholarly outputs from a selection of online sources, such as blogs, Twitter, Facebook, Google+, mainstream news outlets, and media. We used a research fetch API to query the Altmetric database using DOIs. Fourth, because of Twitter's terms of use, Altmetric could only share the status ID of tweets through their API. We further retrieved the details of each tweet through a Twitter developer account using the REST API.

The Altmetric collection of tweets contains original tweets, retweets, quoted tweets, and replies. We used original tweets and their retweets, which yielded a raw sample of 268,003 original tweets created before June 1, 2020. We further removed tweets from nonhuman accounts (eg, organizational accounts or bots) through (1) manually checking and matching all official Twitter accounts of each publisher, journal, and preprint server, and (2) manually checking accounts with excessively high tweet volume (>200 tweets) in our data. This resulted in a final sample of 243,567 original tweets and 729,319 retweets. See Multimedia Appendix 1 for more information on the raw data and the data cleaning process mentioned above [48,49]. Lastly, due to the fast-changing COVID-19 situation worldwide in the early months, we sought to collect situational data related to COVID-19 to serve as controls. More specifically, we further collected (1) daily worldwide COVID-19–confirmed cases and confirmed fatality data from a verified source, OurWorldInData, which is operated by the University of Oxford, and (2) daily global COVID-19 Twitter data [48].

By focusing on the early months (January to May 2020) of the COVID-19 pandemic, we generated a large corpus of original tweets (n=243,567) for analysis. Accordingly, our data covered 8612 articles from 1161 peer-reviewed journals in the MEDLINE database and 2 preprint servers (ie, bioRxiv and medRxiv) in the life science and biomedical fields (see Multimedia Appendix 1 for more details). Each tweet had a valid URL reference to the article, which was identified by a unique DOI, on either the journal or preprint website. Using the DOI, we could identify whether the article referred in the tweet was a *preprint* research article, a *peer-reviewed* research article, or an *opinion/letter* piece published in a peer-reviewed journal. Opinion/letter pieces include editorials, correspondence, letters, and comments. They are published individual opinions from esteemed members of the scientific community rather than research articles. They do not go through a peer-review process, but they also have a unique DOI. Correspondingly, we further constructed 3 subgroups of original tweets mentioning these different article types. The distribution of original tweets among these 3 different types of scientific articles was as follows: 47,570 tweets for preprint articles; 97,769 for peer-reviewed journal research articles; and 98,228 for journal opinion/letter pieces.

Our raw tweet data contained many non-English tweets as Altmetric collected those tweets based on the presence of valid URLs to the DOI-referenced articles instead of text keywords. To process these data, we wrote and used a simple detect-then-translate program, using a Google Translate API, to translate all non-English tweet texts, user screen names, and user biographies (self-described text descriptions) to English. The translated tweet texts were then used to generate variables in this research. Specifically, to quantify the emotion in each tweet, we first used the previously validated Linguistic Inquiry and Word Count (LIWC) dictionaries [49] of the affective process to count the presence of both positive (eg, important, positive, and hope) and negative (eg, fatal, lower, and critical) emotional words in the tweet text. The positive and negative dictionary word counts were generated using licensed LIWC 2015 software.

XSL•FO
RenderX

As mentioned earlier, the discrete perspective is also a critical theoretical approach to investigate emotions [19]. Thus, in addition to the valence of tweets, we wanted to take into account the discrete entities of emotions as well to provide a more comprehensive and robust view on the impact of emotions in the social media sharing of COVID-19 scientific outputs. To this end, we used a state-of-the-art machine learning algorithm trained in the tweet context (CrystalFeel) to gauge which of the 4 specific emotions (ie, joy, anger, fear, and sadness) was most salient in the tweet [50,51]. We sent the translated text corpus to the authors of CrystalFeel who returned the predicted label. Example tweets are provided in Table 1. Although multiple discrete emotions could appear in the same text concurrently, the algorithm is designed to output the most salient one based on an independently calculated intensity score for each individual emotion.

Lastly, content sharing was measured by the number of retweets. Because our data covered a relatively long timespan (ie, 5 months), we counted the number of retweets within a fixed period (eg, the first 168 hours [a week]) after the time of the tweet to make the retweet count of different tweets comparable.

Answering our third research question required us to identify scientists in related fields (ie, medical doctors or academic researchers in the life science and biomedical fields) among tweet message creators. Unfortunately, there was no reliable existing method for us to identify the relevant scientists. To ensure cost-effectiveness and maintain a focused research scope, we developed (and pilot tested) a 2-step classification approach that relied on keyword identification and heuristic rules. This rule-based algorithm extracted formal job titles (eg, clinician, doctor, physician, and surgeon) and related medical terms (eg, cardiology and gastroenterology) from the user screen name along with their text biography and then differentiated scientists from nonscientists. Our manual verification coding validated a 95.5% F1 score for the classification performance. We acknowledge that this method is imperfect as it can lead to underidentification of scientists. We estimated 30%-50% underidentification through manual validation of our classification results on random samples (Multimedia Appendix 2 [52,53]). Underidentification may result in an underestimation of the effect of scientists' engagement. In other words, it may lead to more conservative estimation of the effect size; however, the direction of the estimated effect should be unbiased.

We further included a wide range of previously established control variables that capture the characteristics of the users, referenced articles, and COVID-19 pandemic situation. Table 2 provides descriptions of all variables used in this study, while Table 3 presents the summary statistics of all variables in the full sample as well as each subsample.

Table 1. Example tweets of each specific emotion.

Emotion	Tweet examples[a]
Joy	• "Some more good news - In this cohort of patients hospitalized for severe Covid-19 who were treated with compassionate-use [DRUG], clinical improvement was observed in [NUMBER] of [NUMBER] patients. #coronavirus #COVID-19" • "Good news. Large, retrospective [JOURNAL] study of n=[NUMBER]. [DRUG] did not increase risk of severe #COVID19." • "Some clinical important found about 2019-nCoV from [JOURNAL]. I picked up some important info and translate it Here."
Anger	• "Are you serious? The stranger this gets the more it screams bioweapon. #COVID19 coronavirus male infertility" • "The more vitamin D the less mortality from Coronavirus! The skin produces vitamin D with the sun. So why should we be locked up inside?" • "I don't expect politicians to know understand the detail of science. But you can't insult science when you don't like it and then suddenly insist on something that science can't give on demand."
Fear	• "Horrific read about allocation of scarce medical resources with #COVID19 by [AUTHORS] in @[JOURNAL] - This is very sad and distressing." • "Severe COVID-19 complications: [SYMPTOM] may be observed in the acute phase in severe cases. Long-term [SYMPTOM] has been observed." • "Horrifying. Social distancing in [LOCATION] is almost next to impossible."
Sadness	• "Reading this here left me with depression without enough meme." • "Sadly, this new covid fact will be totally ignored and causing so many lives." • "First time I see a political editorial at the [JOURNAL]. And it is about the disaster that is happening in [COUNTRY]. So sad."
Neutral	• "Clinical Characteristics and Results of [TEST & SUBJECT] With COVID19." • "The present study provides ten key recommendations for the management of COVID-19 infections in [DISEASE GROUP]: #COVID19" • "Here is the link of the last study on [DRUG]!"

[a]The URL has been removed.

XSL•FO
RenderX

Table 2. Descriptions of all variables.

Variable	Description
RT7D	# of retweets in the first 168 hours
preprint	=1 if the tweet source is a preprint article
peer	=1 if the tweet source is a peer-reviewed article
letter	=1 if the tweet source is a journal opinion/letter piece
scientist	=1 if the user is classified as a doctor or researcher in the life science and biomedical fields
liwc_positive	# of positive emotion dictionary words identified by LIWC[a] 2015
liwc_negative	# of negative emotion dictionary words identified by LIWC 2015
emotion: joy	=1 if the tweet text is predicted to have a salient emotion of joy
emotion: anger	=1 if the tweet text is predicted to have a salient emotion of anger
emotion: fear	=1 if the tweet text is predicted to have a salient emotion of fear
emotion: sadness	=1 if the tweet text is predicted to have a salient emotion of sadness
emotion: neutral	=1 if the tweet text is predicted to have no specific emotion
log_follower	(log) number of followers the user had
verified	=1 if the user is a verified user
length	# of words in the tweet text
hashtags	# of hashtags used in the tweet
mention	=1 if the tweet contains any mention of other users
title_length	# of words in the reference article in preprints or journal
title_liwc_pos	# of positive emotion words in the title identified by LIWC 2015
title_liwc_neg	# of negative emotion words in the title identified by LIWC 2015
log_cov_tweet	(log) rolling 7-day total number of global coronavirus tweets
log_cov_case	(log) rolling 7-day total number of global new confirmed COVID cases
log_cov_fatality	(log) rolling 7-day total number of global new confirmed COVID fatalities

[a]LIWC: Linguistic Inquiry and Word Count.

XSL•FO
RenderX

Table 3. Summary statistics of all variables.

Variable	Combined sample (N=243,567), mean (SD)	Preprint (N=47,570), mean (SD)	Peer-reviewed article (N=97,769), mean (SD)	Journal letter (N=98,228), mean (SD)
RT7D[a]	4.928 (85.873)	6.351 (75.654)	5.022 (87.606)	4.145 (88.729)
preprint	0.195 (0.396)	N/A[b]	N/A	N/A
peer	0.401 (0.490)	N/A	N/A	N/A
letter	0.403 (0.491)	N/A	N/A	N/A
scientist	0.183 (0.387)	0.156 (0.363)	0.179 (0.383)	0.201 (0.401)
liwc_positive	0.324 (0.634)	0.316 (0.619)	0.300 (0.614)	0.352 (0.661)
liwc_negative	0.206 (0.506)	0.208 (0.498)	0.191 (0.480)	0.221 (0.535)
emotion: joy	0.245 (0.430)	0.280 (0.449)	0.248 (0.432)	0.225 (0.417)
emotion: anger	0.050 (0.219)	0.045 (0.207)	0.034 (0.181)	0.070 (0.255)
emotion: fear	0.410 (0.492)	0.400 (0.490)	0.416 (0.493)	0.409 (0.492)
emotion: sadness	0.026 (0.159)	0.021 (0.143)	0.021 (0.145)	0.033 (0.179)
emotion: neutral	0.269 (0.443)	0.254 (0.435)	0.281 (0.449)	0.264 (0.441)
log_follower	6.367 (2.174)	6.345 (2.266)	6.329 (2.205)	6.415 (2.096)
verified	0.039 (0.194)	0.038 (0.191)	0.039 (0.194)	0.039 (0.194)
length	19.661 (12.893)	21.477 (13.021)	19.969 (12.807)	18.475 (12.796)
hashtags	0.648 (1.387)	0.647 (1.378)	0.667 (1.428)	0.630 (1.350)
mention	0.200 (0.400)	0.176 (0.381)	0.201 (0.401)	0.211 (0.408)
title_length	11.060 (4.733)	13.051 (5.063)	12.511 (4.303)	8.652 (3.859)
title_liwc_pos	0.101 (0.322)	0.074 (0.280)	0.090 (0.301)	0.125 (0.359)
title_liwc_neg	0.087 (0.289)	0.087 (0.290)	0.103 (0.311)	0.070 (0.262)
log_cov_tweet	15.829 (0.166)	15.817 (0.138)	15.831 (0.176)	15.834 (0.168)
log_cov_case	12.507 (1.283)	12.530 (1.300)	12.504 (1.346)	12.498 (1.209)
log_cov_fatality	9.690 (1.530)	9.740 (1.562)	9.681 (1.611)	9.674 (1.428)

[a]RT7D: number of retweets in the first 168 hours.

[b]N/A: not applicable.

Statistical Analysis

To answer each of our research questions, we examined (1) the impacts of positive versus negative emotional language; (2) the impacts of specific emotions, such as joy, anger, fear, and sadness; and (3) the role of scientists as social media message creators in sharing about COVID-19 medical scientific papers through statistical analysis. We referred to the collective findings from answering these questions as the emotional dynamics in sharing COVID-19 science on social media. Because the distribution of the retweet count was highly skewed (see Table 3), we fitted a negative binomial regression with a maximum likelihood estimator, which is the most appropriate for data with overdispersion. This method is consistent with prior studies using Twitter data [37]. To further ensure that we obtained an unbiased standard error for statistical inference, we used clustered robust standard error [54] at the article level to account for and correct potential intracluster error correlation.

Consistent with prior studies [8], we estimated models both with and without article-level fixed effects. Models without fixed effects capture the between-article comparison, while models with fixed effects provide within-article comparison. The article-level fixed effect, or within-article effect, results were obtained using unconditional fixed effect negative binomial estimators [55]. More specifically, article dummies were included in the regression model to obtain the unconditional fixed effect results. Lastly, we assessed the robustness of our results under 2 criteria: (1) an alternative window for counting retweets (eg, 48 hours after the original tweet rather than a week), and (2) an alternative statistical model, that is, a zero-inflated negative binomial model, to account for the excessive presence of zeros in the retweet count. We showed that our key findings were highly robust under these criteria. More details are discussed and reported in Multimedia Appendix 3 [56,57].

Lastly, to buttress any findings from the statistical analysis on the effect of positive and negative emotion words in tweet text, we further conducted explorative analyses using a word cloud plot. We created 4 text corpuses along the emotion dimension (ie, positive vs negative) and tweet source dimension (ie, preprint vs peer reviewed). For example, if a *positive* dictionary word identified using LIWC 2015 appeared in tweet or retweet

XSL•FO
RenderX

text (the text in the retweet was exactly the text in the original tweet being retweeted) about a preprint, this word was added to the *positivepreprint* text corpus. Then, each word in the 4 text corpuses was processed to keep only the word stem and the term frequency-inversed document frequency weight for each word in the text corpuses to create the word cloud. More details on the text processing and word cloud creation process are provided in Multimedia Appendix 1.

Ethical Considerations

This paper uses only secondary public data from an authorized Twitter commercial data vendor in compliance with Twitter privacy policy. Apart from the public Twitter handle, our data do not contain any individual identifier.

Results

Positive Versus Negative Language

We started with positive and negative emotional language. In the combined sample of all original tweets, our regression analysis (see Multimedia Appendix 4) revealed a significant main effect of positive emotional language on retweet rate (incidence rate ratio [IRR] 1.075, 95% CI 1.027-1.125; P=.002) but not for negative emotional language (IRR 1.015, 95% CI 0.953-1.082; P=.64). The results implied that one additional positive emotional word in a tweet mentioning a COVID-19 research article was associated with, on average, a 7.5% higher retweet rate, while a negative emotional word had a neutral impact. It highlighted that positivity spreads faster than negativity in the Twitter sharing of COVID-19 research, implying the existence of a "positivity bias" rather than a "negativity bias," where positive emotion was found to spread faster. Further, the moderation test between LIWC emotional dictionary word counts and tweet source indicators revealed a positive interaction effect between the positive emotional word count and preprint indicator (IRR 1.129, 95% CI 1.034-1.233; P=.007), implying that an additional positive emotional word would increase the retweet rate difference between tweets mentioning preprint research and peer-reviewed research by 12%, while all other interactions remained insignificant. This points to a differential effect of the presence of emotion in tweets about different scientific sources. Thus, we next examined the effects of positive and negative emotional language separately on each subgroup to check if this pattern persisted in all 3 subgroups of tweets mentioning different types of articles (see Models 1-3 in Table 4).

The above results suggested that the "positivity bias" was only prevalent and visible in tweets that mentioned COVID-19 preprints. To further check the findings' robustness, we also analyzed the within-article effects following a past study on the interpersonal sharing of science to the public [8]. Specifically, we used fixed effects to control for the articles' influence on retweet count. As shown in Models 4-6 in Table 4, the within-article effects were largely consistent with the previously observed pattern. Only the positive word count in the preprint subgroup was found to significantly increase the retweet count. All other estimated coefficients of positive and negative emotional words remained insignificant.

Our results implied that there were divergent patterns among these 3 subgroups. More specifically, the "positivity bias" was only present in tweets mentioning preprints, which predicted that one additional positive emotional word was associated with a 17.7% increase in the retweet rate (IRR 1.177, 95% CI 1.089-1.272; P<.001), while the effect of a negative word was neutral (IRR 0.980, 95% CI 0.883-1.088; P=.70; see Figure 1 for a graphical illustration). In tweets mentioning either research articles or opinion/letter pieces in peer-reviewed journals, neither positive emotional words (research article: IRR 1.048, 95% CI 0.990-1.110; P=.11; opinion/letter pieces: IRR 1.043, 95% CI 0.952-1.143; P=.37) nor negative emotional words (research article: IRR 1.033, 95% CI 0.944-1.131; P=.47; opinion/letter pieces: IRR 1.041, 95% CI 0.936-1.158; P=.45) had statistically significant effects on the retweet rate.

Although the results of the statistical analyses implied the existence of a "positivity bias," they cannot explain why it exists. Hence, we sought to further provide some explorative insights. Using word cloud plots (Figure 2), we showed that the positive words in tweets about preprints had a higher concentration of words like "hope," "support," and "promise" than tweets about peer-reviewed research (see Multimedia Appendix 5 for the exact weight difference). According to the psychological meaning of words [49], besides the positive affective process, the other categories shared by at least two of these three words were "verb," "cognitive process," and "present focus." Qualitatively, these aspects could further elicit a sense of action alongside positivity, which could be a key positivity aspect that people seek under adverse circumstances, such as the COVID-19 crisis.

XSL•FO
RenderX

Table 4. Negative binomial estimation results using the Linguistic Inquiry and Word Count emotional dictionary word counts in subgroups.

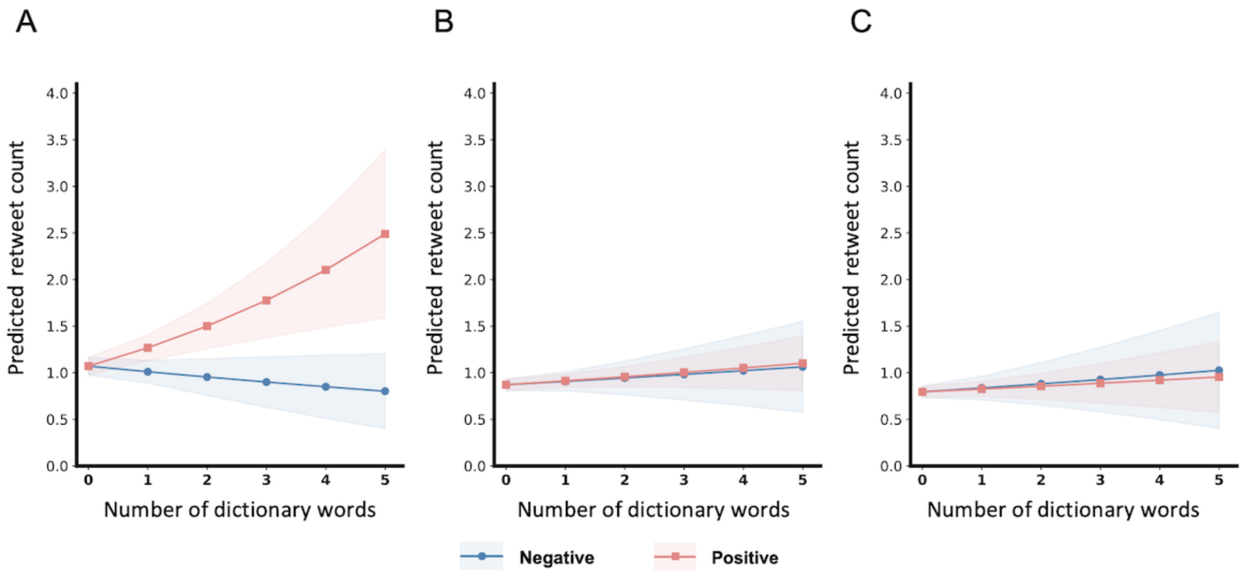
Variable	Model 1 (preprint; N=47,570)[a,b]	Model 2 (peer reviewed; N=97,769)[a,b]	Model 3 (journal letter; N=98,228)[a,b]	Model 4 (preprint; N=47,570)[a,c]	Model 5 (peer reviewed; N=97,769)[a,c]	Model 6 (journal letter; N=98,228)[a,c]
liwc_positive						
IRR[d]	1.177[e]	1.048	1.043	1.084[f]	1.048[g]	1.029
SE[h]	0.047	0.030	0.049	0.036	0.027	0.025
liwc_negative						
IRR	0.980	1.033	1.041	1.031	1.030	1.032
SE	0.052	0.047	0.056	0.040	0.043	0.032
log_follower						
IRR	1.785[e]	1.879[e]	1.891[e]	1.930[e]	1.915[e]	1.933[e]
SE	0.059	0.027	0.026	0.025	0.022	0.024
verified						
IRR	2.040[e]	1.865[e]	1.465[e]	2.003[e]	2.032[e]	1.822[e]
SE	0.283	0.223	0.202	0.196	0.210	0.275
length						
IRR	1.049[e]	1.050[e]	1.051[e]	1.055[e]	1.053[e]	1.049[e]
SE	0.004	0.002	0.004	0.002	0.002	0.002
hashtags						
IRR	1.042[f]	1.037[e]	1.012	1.064[e]	1.032[e]	1.018
SE	0.017	0.013	0.016	0.015	0.011	0.018
mention						
IRR	1.944[e]	1.604[e]	1.703[e]	1.601[e]	1.469[e]	1.632[e]
SE	0.149	0.083	0.090	0.137	0.067	0.083
title_length						
IRR	0.992	0.979[e]	1.018[f]	N/A[i]	N/A	N/A
SE	0.007	0.006	0.009	N/A	N/A	N/A
title_liwc_pos						
IRR	1.051	1.056	1.001	N/A	N/A	N/A
SE	0.124	0.112	0.077	N/A	N/A	N/A
title_liwc_neg						
IRR	0.914	1.083	1.007	N/A	N/A	N/A
SE	0.082	0.092	0.077	N/A	N/A	N/A
log_cov_tweet						
IRR	0.861	0.904	1.201	0.636[f]	0.842	0.874
SE	0.143	0.218	0.299	0.142	0.178	0.213
log_cov_case						
IRR	0.864	0.778[g]	0.846	0.717[g]	0.762[f]	0.728[f]
SE	0.155	0.100	0.146	0.144	0.095	0.091
log_cov_fatality						
IRR	1.144	1.254[f]	1.149	0.977	1.127	1.075

XSL•FO
RenderX

Variable	Model 1 (preprint; N=47,570)[a,b]	Model 2 (peer reviewed; N=97,769)[a,b]	Model 3 (journal letter; N=98,228)[a,b]	Model 4 (preprint; N=47,570)[a,c]	Model 5 (peer reviewed; N=97,769)[a,c]	Model 6 (journal letter; N=98,228)[a,c]
SE	0.178	0.136	0.175	0.168	0.119	0.131
ln(alpha)						
IRR	4.596[e]	4.369[e]	4.172[e]	3.593[e]	3.711[e]	3.367[e]
SE	0.151	0.120	0.143	0.165	0.119	0.113
constant						
IRR	0.188	0.089	0.000[f]	N/A	N/A	N/A
SE	0.503	0.337	0.002	N/A	N/A	N/A

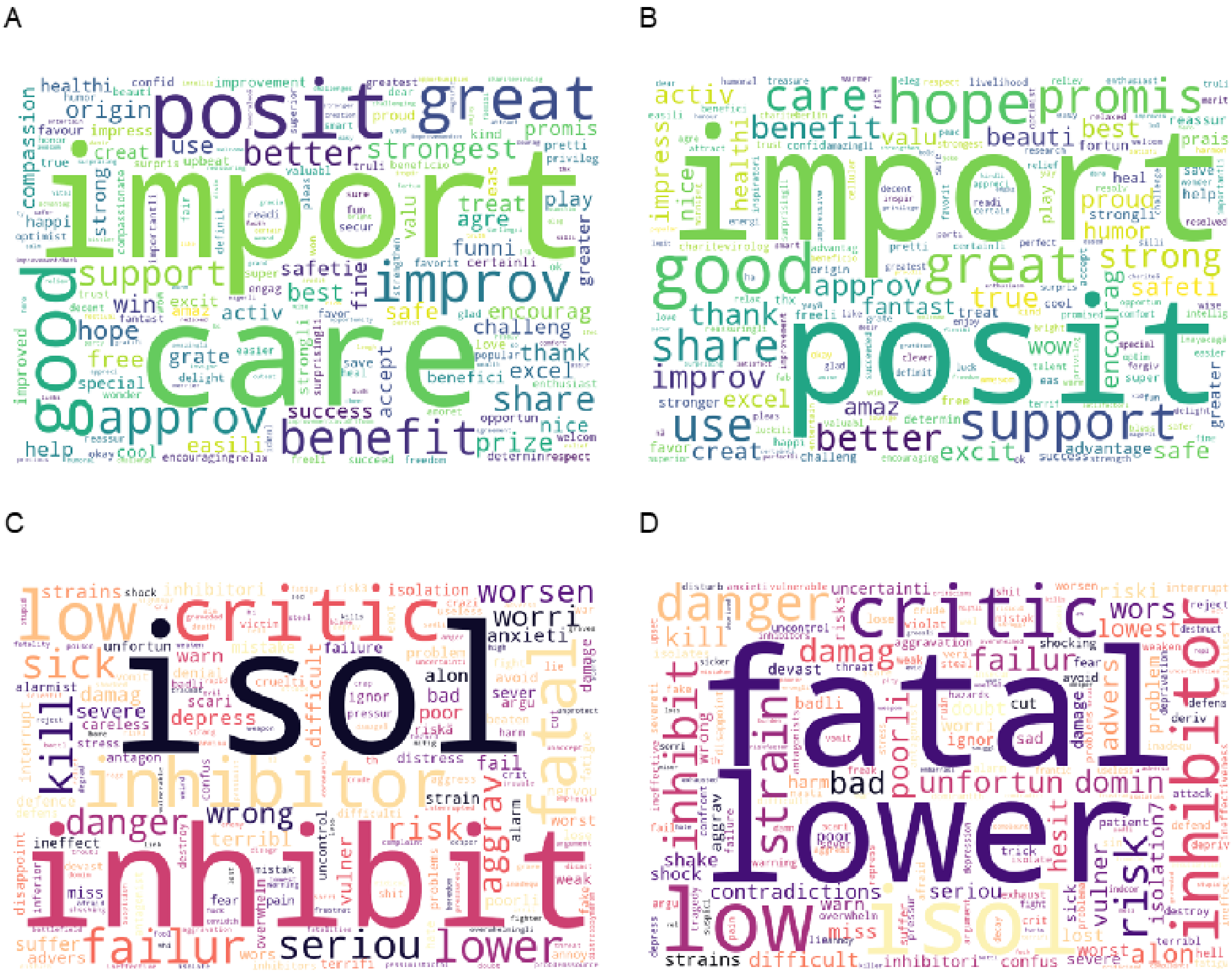
[a]Dependent variable: retweets in the first 168 hours.
[b]No fixed effect.
[c]Fixed effect.
[d]IRR: incidence rate ratio.
[e]*P*<.01.
[f]*P*<.05.
[g]*P*<.10.
[h]Robust standard error clustered by article.
[i]N/A: not applicable.

Figure 1. Prediction of the retweet count for (A) preprints, (B) peer-reviewed articles, and (C) journal letters. Positive emotion Linguistic Inquiry and Word Count dictionary words in tweets about preprints predict the highest retweet count. Bands indicate the 95% CIs.



XSL•FO
RenderX

Figure 2. Word cloud plot of all positive/negative emotional words in tweets about preprints and peer-reviewed articles (word size indicates the term frequency-inversed document frequency weight). (A) positive–peer-reviewed articles; (B) positive–preprints; (C) negative–peer-reviewed articles; (D) negative–preprints.



Specific Emotion

Next, we examined the impact of a specific emotion on retweet count. In this analysis, we used a machine learning approach that was developed for tweet text analysis [51] rather than a general word count–based method. The algorithm classified the emotion in each tweet into 4 categories: joy (happiness), anger, fear, and sadness, as well as a neutral (no specific emotion) condition. For analytical purpose, we focused on these 4 basic emotions as they are the most commonly studied ones in the computational and evolutionary models of emotion [58,59]. Among the classified emotions of the combined tweet sample, 24.5% (59,674/243,567) involved joy, 5.0% (12,178/243,567) involved anger, 41.0% (99,862/243,567) involved fear, and 2.6% (6,333/243,576) involved sadness. This left 26.9% (65,520/243,567) of tweets that had no specific emotion. We have further provided details on the distribution of these specific emotions in all 3 subgroups in Table 3. The results of this classification were largely consistent with the findings of recent studies that have profiled public emotions on social media during the COVID-19 pandemic [10,40], where the authors also found a prevalence of negative emotions such as fear.

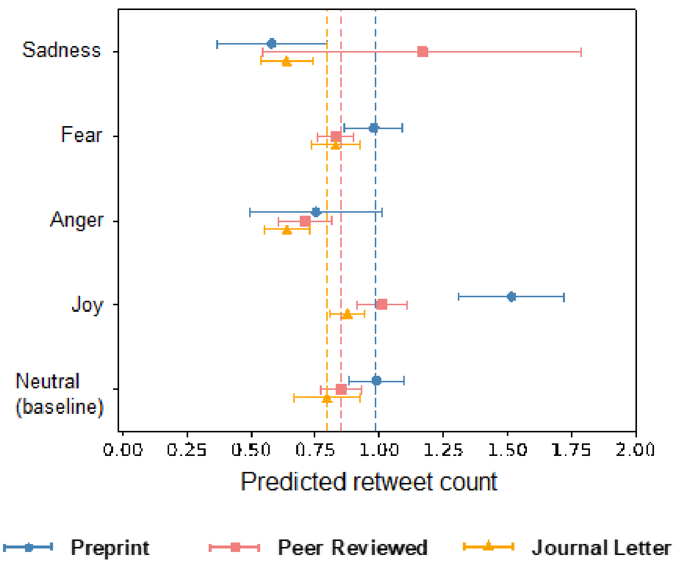
The regression analysis on the combined sample (see Multimedia Appendix 6) revealed that, compared with the presence of no specific emotion, joy was associated with a 25.6% increase in retweet count (IRR 1.256, 95% CI 1.158-1.362; *P*<.001), anger was associated with a 20.4% decrease in retweet count (IRR 0.796, 95% CI 0.702-0.901; *P*<.001), and both fear (IRR 0.998, 95% CI 0.908-1.097; *P*=.97) and sadness (IRR 0.946, 95% CI 0.723-1.237; *P*=.68) had no effect on retweet count. These results confirmed the general existence of a "positivity bias," and only the positive emotion of joy contributed to content sharing. More importantly, high-arousal negative emotions, such as anger and fear, were found to have either a negative or neutral impact on content sharing.

With further analysis, we again observed that the "positivity bias" was most prevalent in tweets mentioning preprints. In the combined sample (see Multimedia Appendix 6), the analysis revealed that the interaction between the preprint subgroup indicator and the joy indicator was significantly positive (IRR 1.290, 95% CI 1.092-1.524; *P*=.003). The interaction between the preprint subgroup indicator and the sadness indicator was significantly negative (IRR 0.429, 95% CI 0.334-0.524; *P*=.009). This difference was also observed in subgroup analysis (see Figure 3 and Models 1-3 in Table 5). More specifically, in the preprint subgroup, joy predicted a 50.3% increase in retweet count (IRR 1.503, 95% CI 1.324-1.707; *P*<.001) and sadness

XSL•FO
RenderX

predicted a 41.0% decrease in retweet count (IRR 0.590, 95% CI 0.417-0.834; *P*=.003). Both high-arousal negative emotions (anger and fear) had neutral impacts on retweet count. In comparison, joy had a smaller but significant positive impact on retweet count (IRR 1.186, 95% CI 1.073-1.310; *P*=.001) in the journal research article subgroup but not in the opinion/letter subgroup. Similarly, anger was associated with less retweets (IRR 0.843, 95% CI 0.725-0.980; *P*=.03) in the journal research article subgroup but not in the opinion/letter subgroup. Sadness had negative effects on retweet count (IRR 0.810, 95% CI 0.671-0.977; *P*=.03) in the journal opinion/letter subgroup but not in the journal research article subgroup. Lastly, fear did not appear to have any effects across all subgroups. Additional results from fixed effect analysis of the within-article effects were again largely consistent (see Models 4-6 in Table 5). Thus, overall, our results showed that a positive-valence emotion, rather than a negative-valence emotion or high-arousal emotion, contributes to higher content sharing of social media messages about COVID-19 scientific research.

Figure 3. Prediction of retweet count according to emotion. Joy in tweets about preprints predicts the highest retweet count. Error bars indicate 95% CIs.



XSL•FO
RenderX

Table 5. Negative binomial estimation results using a specific emotion in subgroups.

Variable	Model 1 (preprint; N=47,570)[a,b]	Model 2 (peer reviewed; N=97,769)[a,b]	Model 3 (journal letter; N=98,228)[a,b]	Model 4 (preprint; N=47,570)[a,c]	Model 5 (peer reviewed; N=97,769)[a,c]	Model 6 (journal letter; N=98,228)[a,c]
joy						
IRR[d]	1.503[e]	1.186[e]	1.100	1.317[e]	1.117[f]	1.110[g]
SE[h]	0.098	0.060	0.086	0.094	0.056	0.061
anger						
IRR	0.777	0.843[f]	0.835	0.809[f]	0.883	0.968
SE	0.140	0.065	0.095	0.067	0.067	0.103
fear						
IRR	0.998	0.985	1.023	1.152	0.971	0.992
SE	0.079	0.056	0.106	0.109	0.055	0.064
sadness						
IRR	0.590[e]	1.440	0.810[f]	0.619[e]	1.128	0.905
SE	0.104	0.393	0.078	0.091	0.196	0.084
log_follower						
IRR	1.786[e]	1.874[e]	1.893[e]	1.929[e]	1.914[e]	1.935[e]
SE	0.059	0.024	0.026	0.025	0.022	0.024
verified						
IRR	2.149[e]	1.892[e]	1.485[e]	2.029[e]	2.030[e]	1.821[e]
SE	0.315	0.222	0.212	0.203	0.208	0.273
length						
IRR	1.053[e]	1.052[e]	1.052[e]	1.056[e]	1.055[e]	1.050[e]
SE	0.003	0.002	0.004	0.003	0.002	0.002
hashtags						
IRR	1.037[f]	1.038[e]	1.010	1.063[e]	1.032[e]	1.017
SE	0.017	0.013	0.017	0.014	0.011	0.018
mention						
IRR	1.849[e]	1.593[e]	1.689[e]	1.596[e]	1.466[e]	1.618[e]
SE	0.136	0.084	0.086	0.121	0.067	0.082
title_length						
IRR	0.994	0.979[e]	1.015[g]	N/A[i]	N/A	N/A
SE	0.006	0.006	0.009	N/A	N/A	N/A
title_liwc_pos						
IRR	1.090	1.064	1.029	N/A	N/A	N/A
SE	0.141	0.111	0.075	N/A	N/A	N/A
title_liwc_neg						
IRR	0.934	1.096	1.047	N/A	N/A	N/A
SE	0.077	0.099	0.072	N/A	N/A	N/A
log_cov_tweet						
IRR	0.867	0.864	1.216	0.639[f]	0.845	0.879

XSL•FO
RenderX

Variable	Model 1 (preprint; N=47,570)[a,b]	Model 2 (peer reviewed; N=97,769)[a,b]	Model 3 (journal letter; N=98,228)[a,b]	Model 4 (preprint; N=47,570)[a,c]	Model 5 (peer reviewed; N=97,769)[a,c]	Model 6 (journal letter; N=98,228)[a,c]
SE	0.142	0.176	0.309	0.143	0.177	0.215
log_cov_case						
IRR	0.851	0.791[g]	0.853	0.723	0.759[f]	0.728[f]
SE	0.152	0.100	0.142	0.145	0.094	0.091
log_cov_fatality						
IRR	1.155	1.238[f]	1.144	0.970	1.132	1.075
SE	0.179	0.133	0.169	0.166	0.119	0.131
ln(alpha)						
IRR	4.536[e]	4.356[e]	4.167[e]	3.574[e]	3.708[e]	3.366[e]
SE	0.158	0.116	0.142	0.164	0.118	0.113
constant						
IRR	0.157	0.160	0.000[f]	N/A	N/A	N/A
SE	0.417	0.519	0.001	N/A	N/A	N/A

[a]Dependent variable: retweets in the first 168 hours.
[b]No fixed effect.
[c]Fixed effect.
[d]IRR: incidence rate ratio.
[e]*P*<.01.
[f]*P*<.05.
[g]*P*<.10.
[h]Robust standard error clustered by article.
[i]N/A: not applicable.

Role of Scientists as Social Media Message Creators

We compared the difference in the retweet rate between tweets from scientists and nonscientists. The distributional differences of specific emotions between scientists and nonscientists in each subgroup are reported in Multimedia Appendix 7. In all subgroups (see Models 1-3 in Table 6), we observed a baseline “toning up” effect of scientists’ participation, where their tweets were associated with, on average, a 40%-60% higher retweet count than tweets from nonscientists (preprint: IRR 1.618, 95% CI 1.358-1.928; *P*<.001; journal research article: IRR 1.434, 95% CI 1.260-1.632; *P*<.001; journal opinion/letter pieces: IRR 1.513, 95% CI 1.204-1.901; *P*<.001). However, we only observed significant interaction effects between the scientist indicator and the emotion indicators for joy (IRR 1.235, 95% CI 1.031-1.479; *P*=.02), anger (IRR 1.767, 95% CI 1.262-2.474; *P*=.001), and fear (IRR 1.339, 95% CI 1.124-1.594; *P*=.001) in the journal research article subgroup. All other interaction terms were not significant (Figure 4). Further within-article effect analysis using fixed effects revealed consistent results (see Models 4-6 in Table 6).

These results highlighted that scientists’ participation could alter the emotional dynamics in the social media sharing of messages of preprints, as their expressed positive emotions (ie, joy) and high-arousal negative emotions (ie, anger and fear) could enhance sharing. In comparison, the indifferences in the emotional dynamics between scientists’ tweets and nonscientists’ tweets about preprints may suggest that it is the emotion elicited by the messages about preprints, rather than who expressed it, that influences content sharing.

XSL•FO
RenderX

Table 6. Negative binomial estimation results using interactions between the scientist indicator and the specific emotion indicators in subgroups.

Variable	Model 1 (preprint; N=47,570)[a,b]	Model 2 (peer reviewed; N=97,769)[a,b]	Model 3 (journal letter; N=98,228)[a,b]	Model 4 (preprint; N=47,570)[a,c]	Model 5 (peer reviewed; N=97,769)[a,c]	Model 6 (journal letter; N=98,228)[a,c]
scientist						
IRR[d]	1.618[e]	1.434[e]	1.513[e]	1.667[e]	1.393[e]	1.601[e]
SE[f]	0.145	0.095	0.176	0.143	0.085	0.134
joy						
IRR	1.540[e]	1.110[g]	1.060	1.344[e]	1.044	1.102
SE	0.123	0.068	0.112	0.113	0.060	0.074
anger						
IRR	0.762	0.751[e]	0.811	0.819[h]	0.772[e]	0.975
SE	0.154	0.063	0.116	0.079	0.060	0.129
fear						
IRR	1.020	0.893[g]	1.008	1.194	0.890[g]	0.984
SE	0.097	0.059	0.140	0.130	0.057	0.080
sadness						
IRR	0.618[h]	1.516	0.783[h]	0.623[e]	1.057	0.883
SE	0.126	0.490	0.085	0.107	0.224	0.092
scientist × joy						
IRR	0.887	1.235[h]	1.048	0.914	1.236[h]	0.971
SE	0.119	0.114	0.139	0.116	0.106	0.099
scientist × anger						
IRR	1.194	1.767[e]	1.258	1.106	1.913[e]	1.034
SE	0.333	0.304	0.222	0.231	0.369	0.150
scientist × fear						
IRR	0.916	1.339[e]	1.019	0.875	1.330[e]	1.000
SE	0.105	0.119	0.151	0.103	0.101	0.096
scientist × sadness						
IRR	0.869	0.789	1.168	1.127	1.372	1.137
SE	0.317	0.281	0.213	0.376	0.374	0.187
log_follower						
IRR	1.772[e]	1.858[e]	1.878[e]	1.914[e]	1.895[e]	1.922[e]
SE	0.058	0.024	0.026	0.026	0.022	0.024
verified						
IRR	2.091[e]	1.733[e]	1.482[h]	1.986[e]	1.929[e]	1.800[e]
SE	0.315	0.210	0.228	0.207	0.207	0.284
length						
IRR	1.052[e]	1.052[e]	1.053[e]	1.055[e]	1.055[e]	1.050[e]
SE	0.004	0.002	0.004	0.003	0.002	0.002
hashtags						
IRR	1.038[h]	1.035[e]	1.011	1.066[e]	1.030[e]	1.020

XSL•FO
RenderX

Variable	Model 1 (preprint; N=47,570)[a,b]	Model 2 (peer re-viewed; N=97,769)[a,b]	Model 3 (journal letter; N=98,228)[a,b]	Model 4 (preprint; N=47,570)[a,c]	Model 5 (peer re-viewed; N=97,769)[a,c]	Model 6 (journal letter; N=98,228)[a,c]
SE	0.017	0.012	0.017	0.015	0.011	0.018
mention						
IRR	1.847[e]	1.560[e]	1.643[e]	1.602[e]	1.447[e]	1.599[e]
SE	0.140	0.082	0.086	0.125	0.066	0.081
title_length						
IRR	0.993	0.978[e]	1.011	N/A[i]	N/A	N/A
SE	0.006	0.006	0.008	N/A	N/A	N/A
title_liwc_pos						
IRR	1.073	1.026	1.026	N/A	N/A	N/A
SE	0.142	0.097	0.074	N/A	N/A	N/A
title_liwc_neg						
IRR	0.922	1.103	1.034	N/A	N/A	N/A
SE	0.075	0.097	0.071	N/A	N/A	N/A
log_cov_tweet						
IRR	0.863	0.878	1.207	0.636[h]	0.895	0.898
SE	0.140	0.176	0.309	0.145	0.191	0.225
log_cov_case						
IRR	0.848	0.826	0.863	0.719[g]	0.759[h]	0.740[h]
SE	0.153	0.106	0.147	0.144	0.097	0.091
log_cov_fatality						
IRR	1.151	1.187	1.131	0.977	1.140	1.074
SE	0.180	0.127	0.171	0.167	0.123	0.129
ln(alpha)						
IRR	4.489[e]	4.250[e]	4.098[e]	3.539[e]	3.624[e]	3.308[e]
SE	0.161	0.116	0.149	0.166	0.116	0.116
constant						
IRR	0.178	0.110	0.000[h]	N/A	N/A	N/A
SE	0.466	0.353	0.002	N/A	N/A	N/A

[a]Dependent variable: retweets in the first 168 hours.

[b]No fixed effect.

[c]Fixed effect.

[d]IRR: incidence rate ratio.

[e]*P*<.01.

[f]Robust standard error clustered by article.

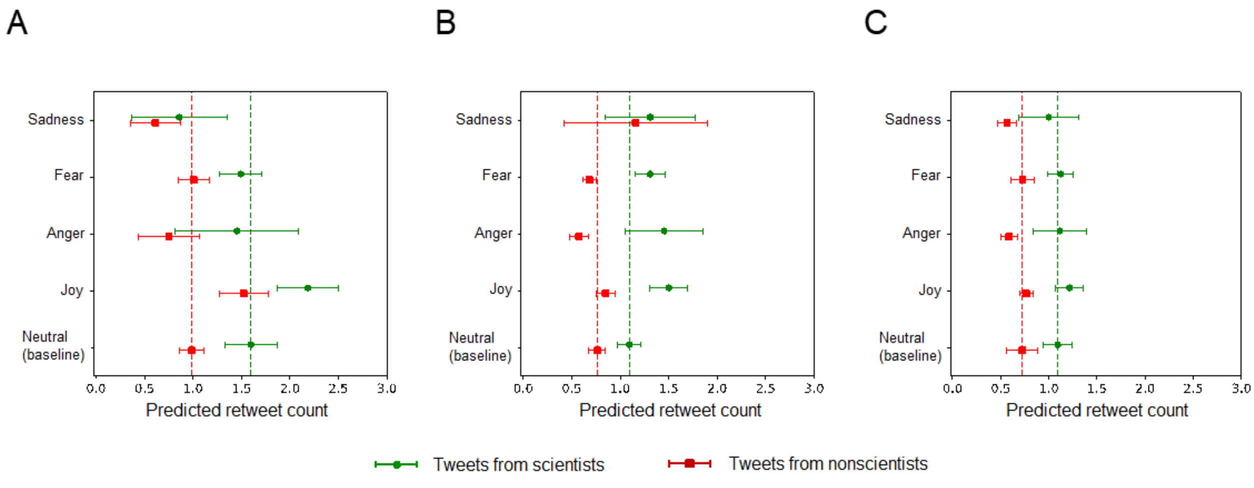
[g]*P*<.10.

[h]*P*<.05.

[i]N/A: not applicable.

XSL•FO
RenderX

Figure 4. The retweet rate difference between tweets from scientists and nonscientists for (A) preprints, (B) peer-reviewed articles, and (C) journal letters. Error bars indicate 95% CIs.



Discussion

Principal Findings

The COVID-19 crisis may have already created a lasting change to the scientific communication process [60], leading this process to become more immediate and transparent as exemplified by explosive use and sharing of preprints. Should we be worried? Using 243,567 original tweets, which generated 729,319 retweets, about 8612 COVID-19 articles from medical peer-reviewed journals and preprint servers in the early months of the pandemic, we shed light on this question by investigating the emotional dynamics of social media sharing of COVID-19 scientific outputs. Our quantitative analyses revealed 3 key findings.

First, we observed a positivity bias. A positive-valence emotion, rather than a negative-valence emotion or high-arousal emotion, contributed to the sharing. Even though the pandemic has given COVID-19 research a heightened "news-like" status, the dissemination of this research on social media did not exhibit a pattern mimicking social media news. Instead, it implied that social media users' sharing of COVID-19 science may be motivated by altruistic reasons or self-enhancement, which was consistent with previous studies on the sharing of science to the public in interpersonal communication settings [8]. However, to the best of our knowledge, the observed differential emotional dynamics of content sharing in messages that mentioned different sources (ie, preprints, peer-reviewed journal research, and journal opinion/letter pieces) have not been demonstrated previously.

Second, the "positivity bias" was most salient in messages of preprints than messages of articles in peer-reviewed journals. What drives this observed difference in emotional dynamics, especially between tweets about preprints and peer-reviewed research? One possibility could be the nature of preprints, as preprints involve nonvetted findings. The peer-review process helps scrutinize and mitigate the scientific uncertainty of a scientific manuscript, and the process often leads to tone-downed findings and conclusions [61]. Without undergoing this "toning down" process, the raw findings in preprints are more likely to be novel, eye-catching, and political [62], which could boost the effect of emotion on content sharing.

Given the self-enhancement explanation behind the "positivity bias," it is also possible that tweets about preprints possess higher self-enhancement potential. Findings in preprints may be perceived by social media users to have higher self-enhancement value because they may be perceived as more novel and impactful [62]. Our explorative analysis using word cloud visualization could provide support for this conjecture as it implied that the positive language in tweets about preprints tends to contain more action-oriented positive words than tweets about peer-reviewed articles. This potential action-positivity perspective also aligns with a self-enhancement explanation, as self-enhancement is linked to not only a positive mindset and stress resistance, but also action orientation [63]. Future research efforts could expand on this conjecture to conduct more in-depth investigations.

Finally, we showed that scientists' participation in the social media sharing of COVID-19 science exhibited differential emotional dynamics in tweets about different scientific sources. Specifically, scientists played a moderating role in the sharing of social media messages about peer-reviewed research, as their expressive positive emotions (ie, joy) and high-arousal negative emotions (ie, anger and fear) further enhanced sharing. However, the same pattern was not observed in messages about preprints. Given that peer-reviewed journal research contains arguably much more reliable findings than preprints, the presence of enhancing and neutral effects of scientists' emotions in tweets about peer-reviewed research and preprints, respectively, could imply a moderated emotional communication process by scientists on social media, selectively promoting more reliable findings. Therefore, our study highlights the instrumental role of scientists in moderating science communication to the public on social media, echoing recent calls for promoting more effective science communication from both the scientific community [64] and the public [65] during crises.

XSL•FO
RenderX

Limitations

Our focus on studying the messages that explicitly referenced COVID-19 research (ie, with a valid URL reference), however, limited us from examining other messages that may have contained scientific research information but did not provide a valid reference. Lack of a valid reference or source ambiguity is a key factor leading to rumor mongering [66] or differentiating science from science conspiracy on social media [67]. Examining the emotional dynamics in these types of messages would be an interesting future research direction. Would the “positivity bias” still exist or would a “negativity bias” prevail instead? Examining these questions would provide insights on social media management, especially the importance of a valid source reference in online messaging. Further, our study design could not fully explicate the causal relationship between the emotion present in tweet text and the subsequent diffusion (retweet). Studies that aim to examine such a causal relationship may consider a randomized study design using either a laboratory experiment or a large-scale field experiment. A future study could also expand on our study to examine the social media sharing of a broader range of scientific outputs beyond COVID-19. Additionally, we detected and translated non-English tweets using only the Google Translate API. Future studies may consider cross-validating this process with human verification or alternative approaches.

Conclusions

Notwithstanding these limitations, our study provides useful implications that add to the ongoing debate regarding the virtue and danger in the use of preprints in science communication to the public [44-46]. Distorted social media dissemination of science could potentially resemble that of misinformation or scientific conspiracy. For instance, in a direct comparison of the online spread of scientific and conspiracy-theory content, a recent study showed that a negative emotion was more likely to enhance the engagement and virality of conspiracy content [30]. We provided evidence that, at least from the perspective of emotional dynamics, social media sharing of COVID-19 science did not exhibit such a distorted pattern that overtly promotes negative emotional messages. On the contrary, positive emotional messages were found to transmit faster, especially in preprints. However, the extent to which such positive but unverified findings of preprints are widely shared on social media was beyond the scope of this study. Practically, our findings highlighted the instrumental role played by scientists in promoting the dissemination of more reliable findings, which can have important implications for social media platform governance in terms of public discourse, especially during crises. Scientists could infuse messages about peer-reviewed articles with positive and high-arousal emotions but try to tone down the emotionality of messages about preprints to reduce the scientific uncertainty in communication. Scientists’ strategic use of emotions in social media sharing could help promote organized and orderly social media sharing of science without relying on explicit and centralized controls on the accessibility of preprints to the public.

Acknowledgments

We sincerely thank Altmetric for providing us academic access to their database for collecting the data.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Information on data processing and cleaning.

[DOCX File , 471 KB - infodemiology_v2i2e37331_app1.docx]

Multimedia Appendix 2

Information on the method for classifying scientist Twitter users.

[DOCX File , 29 KB - infodemiology_v2i2e37331_app2.docx]

Multimedia Appendix 3

Information on robustness tests.

[DOCX File , 55 KB - infodemiology_v2i2e37331_app3.docx]

Multimedia Appendix 4

Negative binomial estimation results using Linguistic Inquiry and Word Count emotional word counts in the combined sample.

[DOCX File , 28 KB - infodemiology_v2i2e37331_app4.docx]

Multimedia Appendix 5

Top 15 positive or negative emotion word stems (Linguistic Inquiry and Word Count 2015) in each text corpus.

[DOCX File , 26 KB - infodemiology_v2i2e37331_app5.docx]

XSL•FO
RenderX

Multimedia Appendix 6

Negative binomial estimation results using specific emotion indicators in the combined sample.

[DOCX File , 28 KB - infodemiology_v2i2e37331_app6.docx]

Multimedia Appendix 7

Kolmogorov-Smirnov test statistics on the distribution of specific emotions between tweets from scientists and tweets from nonscientists in each subgroup.

[DOCX File , 25 KB - infodemiology_v2i2e37331_app7.docx]

References

1. Lever J, Altman RB. Analyzing the vast coronavirus literature with CoronaCentral. Proc Natl Acad Sci U S A 2021 Jun 08;118(23):e2100766118 [FREE Full text] [doi: 10.1073/pnas.2100766118] [Medline: 34016708]
2. Fleerackers A, Riedlinger M, Moorhead L, Ahmed R, Alperin JP. Communicating Scientific Uncertainty in an Age of COVID-19: An Investigation into the Use of Preprints by Digital Media Outlets. Health Commun 2022 May 03;37(6):726-738. [doi: 10.1080/10410236.2020.1864892] [Medline: 33390033]
3. Foell J. Social media science communication is a nonstop academic conference for all. Nat Hum Behav 2021 Jul 26;5(7):812-812. [doi: 10.1038/s41562-021-01138-0] [Medline: 34040172]
4. Koerber A. Is It Fake News or Is It Open Science? Science Communication in the COVID-19 Pandemic. Journal of Business and Technical Communication 2020 Sep 22;35(1):22-27. [doi: 10.1177/1050651920958506]
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. Narayan KMV, Curran JW, Foege WH. The COVID-19 Pandemic as an Opportunity to Ensure a More Successful Future for Science and Public Health. JAMA 2021 Feb 09;325(6):525-526. [doi: 10.1001/jama.2020.23479] [Medline: 33560328]
7. Majumder MS, Mandl KD. Early in the epidemic: impact of preprints on global discourse about COVID-19 transmissibility. The Lancet Global Health 2020 May;8(5):e627-e630. [doi: 10.1016/s2214-109x(20)30113-3]
8. Milkman KL, Berger J. The science of sharing and the sharing of science. Proc Natl Acad Sci U S A 2014 Sep 16;111 Suppl 4(Suppl 4):13642-13649 [FREE Full text] [doi: 10.1073/pnas.1317511111] [Medline: 25225360]
9. Drummond C, Fischhoff B. Emotion and judgments of scientific research. Public Underst Sci 2020 Apr 26;29(3):319-334. [doi: 10.1177/0963662520906797] [Medline: 32098582]
10. Steinert S. Corona and value change. The role of social media and emotional contagion. Ethics Inf Technol 2020 Jul 21;23(Suppl 1):1-10 [FREE Full text] [doi: 10.1007/s10676-020-09545-z] [Medline: 32837288]
11. Chou WS, Budenz A. Considering Emotion in COVID-19 Vaccine Communication: Addressing Vaccine Hesitancy and Fostering Vaccine Confidence. Health Commun 2020 Dec 30;35(14):1718-1722. [doi: 10.1080/10410236.2020.1838096] [Medline: 33124475]
12. Heffner J, Vives M, FeldmanHall O. Emotional responses to prosocial messages increase willingness to self-isolate during the COVID-19 pandemic. Pers Individ Dif 2021 Feb 15;170:110420 [FREE Full text] [doi: 10.1016/j.paid.2020.110420] [Medline: 33082614]
13. Harper CA, Satchell LP, Fido D, Latzman RD. Functional Fear Predicts Public Health Compliance in the COVID-19 Pandemic. Int J Ment Health Addict 2021 Apr 27;19(5):1875-1888 [FREE Full text] [doi: 10.1007/s11469-020-00281-5] [Medline: 32346359]
14. Shen F, Min C, Lu Y, Chu Y. The effect of cognition and affect on preventive behaviors during the COVID-19 pandemic: a cross-sectional study in China. BMC Public Health 2021 Apr 14;21(1):722 [FREE Full text] [doi: 10.1186/s12889-021-10784-y] [Medline: 33853555]
15. Acheampong FA, Wenyu C, Nunoo‐Mensah H. Text‐based emotion detection: Advances, challenges, and opportunities. Engineering Reports 2020 May 28;2(7):e12189. [doi: 10.1002/eng2.12189]
16. Bayer M, Sommer W, Schacht A. Font size matters--emotion and attention in cortical responses to written words. PLoS One 2012 May 9;7(5):e36042 [FREE Full text] [doi: 10.1371/journal.pone.0036042] [Medline: 22590518]
17. Harmon-Jones E, Harmon-Jones C, Summerell E. On the Importance of Both Dimensional and Discrete Models of Emotion. Behav Sci (Basel) 2017 Sep 29;7(4):66 [FREE Full text] [doi: 10.3390/bs7040066] [Medline: 28961185]
18. Russell JA. A circumplex model of affect. Journal of Personality and Social Psychology 1980 Dec;39(6):1161-1178. [doi: 10.1037/h0077714]
19. Izard CE. Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. Perspect Psychol Sci 2007 Sep 24;2(3):260-280. [doi: 10.1111/j.1745-6916.2007.00044.x] [Medline: 26151969]
20. Bebbington K, MacLeod C, Ellison TM, Fay N. The sky is falling: evidence of a negativity bias in the social transmission of information. Evolution and Human Behavior 2017 Jan;38(1):92-101. [doi: 10.1016/j.evolhumbehav.2016.07.004]
21. Dijksterhuis A, Aarts H. On wildebeests and humans: the preferential detection of negative stimuli. Psychol Sci 2003 Jan 06;14(1):14-18. [doi: 10.1111/1467-9280.t01-1-01412] [Medline: 12564748]

XSL•FO
RenderX

22. Peeters G, Czapinski J. Positive-Negative Asymmetry in Evaluations: The Distinction Between Affective and Informational Negativity Effects. European Review of Social Psychology 1990 Jan;1(1):33-60. [doi: 10.1080/14792779108401856]
23. Carver LJ, Vaccaro BG. 12-month-old infants allocate increased neural resources to stimuli associated with negative adult emotion. Dev Psychol 2007 Jan;43(1):54-69 [FREE Full text] [doi: 10.1037/0012-1649.43.1.54] [Medline: 17201508]
24. Sedikides C. Assessment, enhancement, and verification determinants of the self-evaluation process. Journal of Personality and Social Psychology 1993 Aug;65(2):317-338. [doi: 10.1037/0022-3514.65.2.317]
25. Berger J, Milkman KL. What Makes Online Content Viral? Journal of Marketing Research 2012 Apr 01;49(2):192-205. [doi: 10.1509/jmr.10.0353]
26. Tellis GJ, MacInnis DJ, Tirunillai S, Zhang Y. What Drives Virality (Sharing) of Online Digital Content? The Critical Role of Information, Emotion, and Brand Prominence. Journal of Marketing 2019 Apr 24;83(4):1-20. [doi: 10.1177/0022242919841034]
27. Taylor DG, Strutton D, Thompson K. Self-Enhancement as a Motivation for Sharing Online Advertising. Journal of Interactive Advertising 2012 Mar;12(2):13-28. [doi: 10.1080/15252019.2012.10722193]
28. Rozin P, Royzman EB. Negativity Bias, Negativity Dominance, and Contagion. Pers Soc Psychol Rev 2016 Dec 21;5(4):296-320. [doi: 10.1207/s15327957pspr0504_2]
29. Baumeister RF, Bratslavsky E, Finkenauer C, Vohs KD. Bad is Stronger than Good. Review of General Psychology 2001 Dec 01;5(4):323-370. [doi: 10.1037/1089-2680.5.4.323]
30. 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]
31. 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]
32. Rathje S, Van Bavel JJ, van der Linden S. Out-group animosity drives engagement on social media. Proc Natl Acad Sci U S A 2021 Jun 29;118(26):e2024292118 [FREE Full text] [doi: 10.1073/pnas.2024292118] [Medline: 34162706]
33. Hansen LK, Arvidsson A, Nielsen FA, Colleoni E, Etter M. Good Friends, Bad News - Affect and Virality in Twitter. In: Park JJ, Yang LT, Lee C, editors. Future Information Technology. Communications in Computer and Information Science, vol 185. Berlin, Heidelberg: Springer; 2011:34-43.
34. Guadagno RE, Rempala DM, Murphy S, Okdie BM. What makes a video go viral? An analysis of emotional contagion and Internet memes. Computers in Human Behavior 2013 Nov;29(6):2312-2319. [doi: 10.1016/j.chb.2013.04.016]
35. Berger J. Arousal increases social transmission of information. Psychol Sci 2011 Jul 20;22(7):891-893. [doi: 10.1177/0956797611413294] [Medline: 21690315]
36. Crockett MJ. Moral outrage in the digital age. Nat Hum Behav 2017 Nov 18;1(11):769-771. [doi: 10.1038/s41562-017-0213-3] [Medline: 31024117]
37. 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]
38. Smith CA, Ellsworth PC. Patterns of cognitive appraisal in emotion. Journal of Personality and Social Psychology 1985;48(4):813-838. [doi: 10.1037/0022-3514.48.4.813]
39. Rutter H, Wolpert M, Greenhalgh T. Managing uncertainty in the COVID-19 era. BMJ 2020 Sep 01;370:m3349. [doi: 10.1136/bmj.m3349] [Medline: 32873549]
40. Vemprala N, Bhatt P, Valecha R, Rao HR. Emotions During the COVID-19 Crisis: A Health Versus Economy Analysis of Public Responses. American Behavioral Scientist 2021 Oct 25;65(14):1972-1989. [doi: 10.1177/00027642211050901]
41. Fraser N, Brierley L, Dey G, Polka JK, Pálfy M, Nanni F, et al. The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. PLoS Biol 2021 Apr 2;19(4):e3000959 [FREE Full text] [doi: 10.1371/journal.pbio.3000959] [Medline: 33798194]
42. Schneider S. Communicating Uncertainty: A Challenge for Science Communication. In: Drake J, Kontar Y, Eichelberger J, Rupp T, Taylor K, editors. Communicating Climate-Change and Natural Hazard Risk and Cultivating Resilience. Advances in Natural and Technological Hazards Research, vol 45. Cham: Springer; 2016:267-278.
43. Fischhoff B, Davis AL. Communicating scientific uncertainty. Proc Natl Acad Sci U S A 2014 Sep 16;111 Suppl 4(Suppl 4):13664-13671 [FREE Full text] [doi: 10.1073/pnas.1317504111] [Medline: 25225390]
44. Sheldon T. Preprints could promote confusion and distortion. Nature 2018 Jul 24;559(7715):445-445. [doi: 10.1038/d41586-018-05789-4] [Medline: 30042547]
45. Fry N, Marshall H, Mellins-Cohen T. In praise of preprints. Microb Genom 2019 Apr;5(4):e000259 [FREE Full text] [doi: 10.1099/mgen.0.000259] [Medline: 30938670]
46. Sarabipour S. Preprints are good for science and good for the public. Nature 2018 Aug 29;560(7720):553-553. [doi: 10.1038/d41586-018-06054-4] [Medline: 30158619]
47. Kwok R. Research impact: Altmetrics make their mark. Nature 2013 Aug 22;500(7463):491-493. [doi: 10.1038/nj7463-491a] [Medline: 23977678]

XSL•FO
RenderX

48. 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]
49. 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]
50. Mohammad SM, Bravo-Marquez F. WASSA-2017 Shared Task on Emotion Intensity. arXiv. 2017. URL: https://arxiv.org/abs/1708.03700 [accessed 2022-11-24]
51. Gupta RK, Yang Y. CrystalFeel at SemEval-2018 Task 1: Understanding and Detecting Emotion Intensity using Affective Lexicons. In: Proceedings of the 12th International Workshop on Semantic Evaluation. 2018 Presented at: 12th International Workshop on Semantic Evaluation; June 2018; New Orleans, Louisiana. [doi: 10.18653/v1/s18-1038]
52. Zhang Z, Bors G. “Less is more”. OIR 2019 Dec 17;44(1):213-237. [doi: 10.1108/OIR-05-2019-0143]
53. Ke Q, Ahn YY, Sugimoto CR. A systematic identification and analysis of scientists on Twitter. PLoS One 2017;12(4):e0175368 [FREE Full text] [doi: 10.1371/journal.pone.0175368] [Medline: 28399145]
54. Cameron AC, Miller DL. A Practitioner’s Guide to Cluster-Robust Inference. J. Human Resources 2015;50(2):317-372. [doi: 10.3368/jhr.50.2.317]
55. Allison PD, Waterman RP. Fixed-Effects Negative Binomial Regression Models. Sociological Methodology 2016 Jun 23;32(1):247-265. [doi: 10.1111/1467-9531.00117]
56. Staub KE, Winkelmann R. Consistent estimation of zero-inflated count models. Health Econ 2013 Jun;22(6):673-686. [doi: 10.1002/hec.2844] [Medline: 22623339]
57. Sheu ML, Hu TW, Keeler TE, Ong M, Sung HY. The effect of a major cigarette price change on smoking behavior in california: a zero-inflated negative binomial model. Health Econ 2004 Aug;13(8):781-791. [doi: 10.1002/hec.849] [Medline: 15322990]
58. Gadanho SC, Hallam J. Exploring the role of emotions in autonomous robot learning. Association for the Advancement of Artificial Intelligence (AAAI). 1998. URL: https://www.aaai.org/Papers/Symposia/Fall/1998/FS-98-03/FS98-03-017.pdf [accessed 2022-11-24]
59. Pacella D, Ponticorvo M, Gigliotta O, Miglino O. Basic emotions and adaptation. A computational and evolutionary model. PLoS One 2017 Nov 6;12(11):e0187463 [FREE Full text] [doi: 10.1371/journal.pone.0187463] [Medline: 29107988]
60. Haseltine WA. How COVID Changed Science. Scientific American. URL: https://www.scientificamerican.com/article/how-covid-changed-science/ [accessed 2022-11-24]
61. Horbach S. No time for that now! Qualitative changes in manuscript peer review during the Covid-19 pandemic. Research Evaluation 2021;30(3):231-239. [doi: 10.1093/reseval/rvaa037]
62. Heimstädt M. Between fast science and fake news: Preprint servers are political. The London School of Economics and Political Science. 2020. URL: https://blogs.lse.ac.uk/impactofsocialsciences/2020/04/03/between-fast-science-and-fake-news-preprint-servers-are-political/ [accessed 2022-11-24]
63. Dufner M, Gebauer JE, Sedikides C, Denissen JJA. Self-Enhancement and Psychological Adjustment: A Meta-Analytic Review. Pers Soc Psychol Rev 2019 Feb 13;23(1):48-72. [doi: 10.1177/1088868318756467] [Medline: 29534642]
64. Saitz R, Schwitzer G. Communicating Science in the Time of a Pandemic. JAMA 2020 Aug 04;324(5):443-444. [doi: 10.1001/jama.2020.12535] [Medline: 32749498]
65. What the pandemic has taught us about science communication. World Economic Forum. 2021. URL: https://www.weforum.org/agenda/2021/06/lessons-for-science-communication-from-the-covid-19-pandemic/ [accessed 2022-11-24]
66. Oh O, Agrawal M, Rao HR. Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets During Social Crises. MIS Quarterly 2013;37(2):407-426. [doi: 10.25300/misq/2013/37.2.05]
67. 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]

Abbreviations

API: application programming interface
DOI: digital object identifier
IRR: incidence rate ratio
LIWC: Linguistic Inquiry and Word Count

XSL•FO
RenderX

Edited by T Mackey; submitted 15.02.22; peer-reviewed by K Reuter, H Alhoori, L Popova; comments to author 30.04.22; revised version received 31.05.22; accepted 05.11.22; published 08.12.22.

Please cite as:
Luo K, Yang Y, Teo HH
The Asymmetric Influence of Emotion in the Sharing of COVID-19 Science on Social Media: Observational Study
JMIR Infodemiology 2022;2(2):e37331
URL: https://infodemiology.jmir.org/2022/2/e37331
doi:10.2196/37331
PMID:36536762

©Kai Luo, Yang Yang, Hock Hai Teo. Originally published in JMIR Infodemiology (https://infodemiology.jmir.org), 08.12.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Infodemiology, is properly cited. The complete bibliographic information, a link to the original publication on https://infodemiology.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX